# Group-based Estimation of Missing Hydrological Data

By

Amin Elshorbagy

A Thesis
Submitted to the Faculty of Graduate Studies
In Partial Fulfillment of the Requirements
For the Degree of

Doctor of Philosophy

Department of Civil and Geological Engineering
University of Manitoba
Winnipeg, Manitoba

© May, 2001

National Library of Canada
Acquisitions and Bibliographic Services
395 Wellington Street
Ottawa ON K1A 0N4
Canada

Bibliothèque nationale du Canada
Acquisitions et services bibliographiques
395, rue Wellington
Ottawa ON K1A 0N4
Canada

*Your file Votre référence*

*Our file Notre référence*

The author has granted a non-exclusive licence allowing the National Library of Canada to reproduce, loan, distribute or sell copies of this thesis in microform, paper or electronic formats.

The author retains ownership of the copyright in this thesis. Neither the thesis nor substantial extracts from it may be printed or otherwise reproduced without the author's permission.

L'auteur a accordé une licence non exclusive permettant à la Bibliothèque nationale du Canada de reproduire, prêter, distribuer ou vendre des copies de cette thèse sous la forme de microfiche/film, de reproduction sur papier ou sur format électronique.

L'auteur conserve la propriété du droit d'auteur qui protège cette thèse. Ni la thèse ni des extraits substantiels de celle-ci ne doivent être imprimés ou autrement reproduits sans son autorisation.

0-612-62635-0

Canada

**THE UNIVERSITY OF MANITOBA**

**FACULTY OF GRADUATE STUDIES**

*****

**COPYRIGHT PERMISSION**

# GROUP-BASED ESTIMATION OF MISSING HYDROLOGICAL DATA

**BY**

**AMIN ELSHORBAGY**

**A Thesis/Practicum submitted to the Faculty of Graduate Studies of The University of Manitoba in partial fulfillment of the requirement of the degree**

**of**

**DOCTOR OF PHILOSOPHY**

**AMIN ELSHORBAGY © 2001**

**Permission has been granted to the Library of the University of Manitoba to lend or sell copies of this thesis/practicum, to the National Library of Canada to microfilm this thesis and to lend or sell copies of the film, and to University Microfilms Inc. to publish an abstract of this thesis/practicum.**

**This reproduction or copy of this thesis has been made available by authority of the copyright owner solely for the purpose of private study and research, and may only be reproduced and copied as permitted by copyright laws or with express written authorization from the copyright owner.**

To

My late father, my mother,

my wife, and my daughter.

**ABSTRACT**

Water resources planning and management require complete data sets of many variables, such as rainfall, streamflow, and temperature. Unfortunately, records of hydrologic processes are usually short and often have missing observations. Attracted by the importance of estimating missing data, hydrologic researchers have adopted and developed various models and techniques to in-fill missing data. The diversity of the adopted techniques does not necessarily indicate diversity in the approach. A major commonality exists in most of the applications of these techniques; that is, any hydrologic time series record is perceived as a sequence of single-valued observations irrespective of the time scale of the data or their underlying structure.

In this research, the group approach, different from the traditional single-valued approach, is proposed. The approach perceives the periodic hydrologic data as sequence of groups rather than single-valued observations. The techniques suggested to handle the group approach, after modification, are regression, time series analysis, partitioning modeling, and artificial neural networks. Various models representing these four techniques are briefly presented and applied to single series and bi-series cases, respectively. Also group time series models are developed in this thesis for the same purpose. It turns out that the group approach is highly useful for estimating consecutive missing values, and possibly other applications, such as long-term forecast.

On the other hand, in non-periodic data (e.g., daily flows) where seasonality does not play a major role and a definite number of repetitive low dimensional groups of

observations cannot be found in the geophysical year, another approach of identifying and modeling groups is sought. The nonlinearity and dynamic behavior of non-periodic hydrologic data sets have been indicated in water resources literature as issues that influence the performance of modeling tools that ignore nonlinearity and dynamics inherent in the data structure. Consecutive missing streamflows are estimated, using the principles of chaos theory, in two steps. First, the existence of chaotic behavior in daily flows of the river is investigated. Second, the analysis of chaos is used to configure two models employed to estimate missing data: artificial neural networks and K-nearest neighbors. Also, another local linear model is applied for comparison purposes. The results highlight the utility of using the analysis of chaos for configuring the models.

In an unprecedented trial, in the chaos literature in water resources, the effect of the chaotic behavior on the analysis of two cross-correlated time series is investigated. The effect of both nonlinearity and dynamics is shown through application to daily streamflows. Other issues such as noise reduction and the reliability of its application to hydrologic time series are discussed. It is recommended that current noise reduction algorithms should be applied with caution and used for better estimation of chaotic invariants. The raw data should always be the basis for any further hydrologic analysis. After decades of adopting stochastic hydrology, chaos analysis, which has been recently introduced to hydrology, provides challenges and opportunities in hydrologic research. It has the potential to change the way in which hydrologic, and other real, processes are perceived, analyzed, and interpreted. The phenomenon that used to be treated as random may turn out to be nonlinear deterministic (chaotic) process.

# Acknowledgment

It is deep-rooted in the Islamic traditions and conducts that the "*one who does not express thanks to people, cannot be grateful to Allah (The Lord)*". Therefore, after praising and thanking Allah, Glorified and Exalted is He, I would like to use this little space to express my gratitude to the people without their support, this work would have not been realized. First, I would like to thank Prof. U. S. Panu, my research supervisor, for his continuous technical and financial support throughout my Ph.D. program. I appreciate the time and effort that he has devoted to me. His sincere help and insightful assistance and guidance have been extremely helpful to me.

I would also like to extend my thanks to Prof. S. P. Simonovic, my research co-supervisor, for his support, advices, and critical reading and review of my work. He has been always teaching me, directly and indirectly, that success is fruit provided by one tree; working hard.

I would also like to acknowledge Prof. M. Pawlak, the member of my advisory committee, for his support. Some of the ideas of this research are based on concepts I have learnt in his "Pattern recognition" course.

I would like also to thank the Natural Sciences and Engineering Research Council (NSERC), Canada, for its financial support through the research grant No. OGP-0004404 given to my supervisor (1997-2001) and I would also like to extend my thanks to the

University of Manitoba for the graduate fellowship (UMGF) awarded to me (1999-2001) and Lakehead University for facilitating my research during the period (1998-1999).

Special thanks are also expressed to Dr. Abu-Zeid, the Minister of Public works and Water Resources (MPWWR) of Egypt and Dr. M. El-Qadi, Chairman of the National Water Research Center (NWRC), Egypt, for giving me the chance to come to the University of Manitoba in pursuit of my Ph.D. program.

To all my colleagues at former Facility for intelligent Decision Support (FIDS), I acknowledge your support and kind companionship.

Last but not the least, I am extremely grateful to my mother, brother, and sisters for their support and prayers for me. Shahira, my wife, and Habiba, my daughter, have been always to me like two candles when it was so dark. The care of Shahira and the smile of Habiba defeated any feeling of depression.

# Contents

**Abstract** .....i

**Acknowledgment** .....iii

**List of Figures** .....ix

**List of Tables** .....xii

**1. Introduction**

1.1 Background .....2

1.2 Area of interest .....3

1.3 Problem recognition .....5

1.4 Research objectives .....7

1.5 Synopsis of the thesis .....7

**2. Estimation of missing data: a review**

2.1 Introduction .....10

2.2 Single-valued approach .....11

2.2.1 Regression analysis .....11

2.2.2 Time series analysis .....15

2.2.3 Interpolation analysis .....17

2.2.4 Pattern recognition .....18

2.2.5 Artificial neural networks .....19

2.3 The group approach .....20

**3. Perception of groups in periodic data**

3.1 Introduction .....27

3.2 Proposed approach and methodology .....31

3.3 Data segmentation .....37

3.4 Difference between group approach and existing models .....42

3.5 Proposed group-based techniques and models for data in-filling .....43

**4. Application of group approach to periodic hydrologic data**

4.1 Introduction .....50

4.2 Application to streamflow records .....50

4.2.1 Segmentation .....51

4.2.2 Identifying inter-group relationships towards model selection .....59

4.2.3 Application to the single series case .....63

4.2.4 Application to the bi-series case .....68

4.2.5 Estimation of partially missing segment .....70

4.3 Indicators of model performance .....72

4.4 The pooled mean squared error (PMSE) .....74

4.5 Discussion .....76

4.6 Concluding remarks .....79

**5. Perception of groups in non-periodic data: the chaos approach**

5.1 Introduction .....83

5.2 What is chaos? .....84

5.2.1 Reconstruction of the phase-space .....87

5.2.2 Correlation dimension .....88

5.2.3 Other methods for detecting chaos .....90

5.3 Review of studies investigating the concepts of chaos in hydrology .....94

5.4 Estimation of missing streamflow data .....106

5.5 Summary .....110

**6. Analysis of chaos: applications to daily streamflows**

6.1 Introduction .....113

6.2 Application to single time series .....113

6.2.1 Correlation dimension .....114

6.2.2 Lyapunov exponent .....116

6.2.3 Entropy .....116

6.2.4 Method of Surrogate data .....117

6.2.5 Results and analysis .....118

6.2.6. Concluding remarks .....121

6.3 Application to cross-correlated time series .....122

6.3.1 Case study .....124

6.3.2 results and analysis of estimation of missing data .....126

6.3.3 Discussion .....126

6.3.4 Concluding remarks .....129

6.4 Analysis of noise .....130

6.4.1 Noise in hydrologic data .....132

6.4.2 Effect of noise .....135

6.4.3 Noise reduction .....137

6.4.4 Noise reduction in hydrology .....139

6.4.5 Case study 1: A single time series .....144

6.4.6 Comparison between two noise reduction algorithms .....165

6.4.7 case study 2: two cross-correlated series .....170

6.5 Summary and concluding remarks .....176

**7. Conclusions**

7.1 Summary of the thesis .....179

7.2 Research contribution.....183

7.2.1 level-1 contribution .....183

7.2.2 level-2 contribution .....184

7.2.3 level-3 contribution .....185

7.3 Possible research extension .....187

7.4 Study limitations .....188

**References**.....190

**Appendices**.....206

# List of Figures

3.1 Schematic diagram for group in-filling in a single time series .....34

3.2 Schematic diagram for group in-filling in a single time series .....35

3.3 Schematic diagram representing a single series in the group approach as a multivariate problem .....45

4.1 Monthly distance from the series mean flow of the English River .....53

4.2 Box-plots of the monthly flows of the English River at Umferville .....54

4.3 Box-plot of the English River. a) pattern class 1, b) pattern class 2, and c) pattern class 3 .....55

4.4 Correlation structure of data groups (English River). a) autocorrelation function and b) partial autocorrelation function .....63

4.5 Autocorrelation function of data groups (Little River) .....64

4.6 Autocorrelation function of data groups (Reed Creek) .....64

4.7 Cross-correlation function of data groups (Little & Reed Creek) .....65

4.8 Estimated and observed values of the English River flows .....68

4.9 Estimated and observed values of the Little River flows .....70

4.10 Modified estimates of the missing segments (English River) .....71

4.11 Modified estimates of the missing segments (Little River) .....71

5.1 Phase-space reconstruction of a signal X(t) .....89

5.2 Flowchart of the different stages of modeling a chaotic time series .....95

5.3 Structure of the reconstructed time series data .....108

6.1 Variation of daily flows of the English River .....114

6.2 Correlation integral of the English River data .....115

6.3 Correlation exponent of the English River data .....115

6.4 Estimate of correlation entropy of the English River data .....117

6.5 Correlation dimensions of the English River data set and its surrogate .....118

6.6 Actual and estimated data using ANN technique .....119

6.7 Actual and estimated data using K-nn technique .....120

6.8 Actual and estimated data using ANN and K-nn techniques .....120

6.9 Correlation dimension of the time series, a) Little River and b) Reed Creek ...125

6.10 Actual and estimated data of the Reed Creek raw data using a) linear regression and b) artificial neural networks .....127

6.11 Time series of (a) raw SOI data and noise-reduced data by (b) moving average, (c) low-pass filter and (d) nonlinear smoothing. [*after Kawamura et al. (1998)*] .....145

6.12 Change in prediction accuracy using Local Linear Model (LLM) .....151

6.13 Change in prediction accuracy using Artificial Neural Networks (ANNs) ....151

6.14 Change in prediction accuracy using ANNs and LLM (10 days are used to compute MSE) .....152

6.15. Change in prediction accuracy using ANNs and LLM (100 days are used to compute MSE) .....153

6.16 Autocorrelation function of the noise time series of (a) 2% noise, (b) 4% noise, (c) 10% noise, and (d) 50% noise .....155

6.17 Partial autocorrelation function of the noise time series of 10% noise level ...156

6.18 Actual and predicted values of the noise series using AR(1) .....156

6.19 Phase-space diagram of sea surface temperature (S Index) data.(a) raw data, (b) after noise removal. [after Jayawardena and Gurung (2000)] .....158

6.20 Correlation exponent of the raw and noise-reduced data of the English River .....159

6.21 Entropy of the English River data. (a) raw data and (b) 10% noise-reduced data .....161

6.22 Predicted data of the English River using LLM and ANNs. (a) predicted versus observed data and (b) predicted versus noise-reduced data .....164

6.23 Autocorrelation function of the removed noise, a) method of Schreiber et al. (1993) and b) method of Grassberger et al. (1993) .....168

6.24 Partial autocorrelation function of the removed noise, a) method of Schreiber et al. (1993) and b) method of Grassberger et al. (1993) .....169

6.25 Correlation structure of the 18.7 % (q=5) noise removed from the Little River data, a) autocorrelation function and b) partial autocorrelation function .....172

6.26 Correlation structure of the 3.7 % (q=15) noise removed from the Little River data, a) autocorrelation function and b) partial autocorrelation function .....173

6.27 Actual and estimated data of the Reed Creek noise-reduced data using a) linear regression and b) artificial neural networks .....174

# List of Tables

4.1 Basic information on the three rivers under consideration .....50

4.2 Correlation matrix of the monthly flows of English river at Umferville .....57

4.3 Percentage of classification/misclassification using Euclidean and Mahalanobis distances for the segments of different classes of the three rivers .....59

4.4 Actual and predicted flows, residuals, and relative errors (a hypothetical case) .....76

4.5 Summary of model performance for the English River case .....78

4.6 Summary of model performance for the Little River case .....79

6.1 Mean squared and mean relative error of estimated data .....118

6.2 The mean squared and the mean relative errors of the estimated values .....128

6.3 Noise removed and estimated by the methods of Schreiber (1993) and Schouten et al. (1994) .....147

6.4 Prediction accuracy of the raw and noise reduced data of the English River .....163

6.5 Estimation accuracy of the raw and noise-reduced data of the English River .....166

6.6 Different levels of reduced noise and the corresponding cross-correlation values .....170

6.7 The mean squared and the mean relative errors of the estimated values .....175

# Chapter 1

# 1. INTRODUCTION

1.1 Background

1.2 Area of interest

1.3 Problem recognition

1.4 Research objectives

1.5 Synopsis of the thesis

---

*" Verily We have granted thee a manifest opening:*

*That ALLAH may forgive thee thy faults of the past and those to follow; fulfill His favor to thee; and guide thee on the straight way"*

*The Glorious Qura'n (48: 1-2)*

## 1.1 Background

Water resources planning and management require complete data sets of many variables, such as rainfall, streamflow, evapo-transpiration, and temperature. Unfortunately, records of hydrologic processes are usually short and often have missing observations. The existence of data gaps might be attributed to a number of reasons, such as interruption of measurements because of equipment failure, effects of extreme natural phenomena such as hurricanes or landslides or of human-induced factors such as wars and civil unrest, mishandling of observed records by field personnel, or accidental loss of data files in the computer system. Deficiencies in hydrologic data series vary from 5 to 10 percent in the case of runoff data (Correll et al. 1999) and up to 25 percent in the case of oceanic storm surges (Zhang et al. 1997). Most hydrologic models do not tolerate missing observations and thus data in-filling techniques have evolved to deal with incomplete data sets.

Attracted by the importance of estimating missing data, hydrologic researchers have adopted and developed various models and techniques to deal with the problem. Efforts are not only devoted to extending short records by adding lengthy segments of estimated data, but attention is also given to the gaps of short duration. One can find in the literature cases where sophisticated techniques are used for estimation of a single missing observation (e.g., Griffith et al. 1985). In water resources, the commonly used techniques for estimation of missing data are based on regression analysis, time series analysis, artificial neural networks, and interpolation techniques. The diversity of the above-mentioned techniques does not necessarily indicate diversity in the approach. A major

commonality exists in most of the applications of these techniques; that is, any hydrologic time series record is perceived as a sequence of single-valued observations irrespective of the time scale of the data or their underlying structure. In hydrologic data (e.g., streamflows), it can be noted that annual or seasonal data might be independent while monthly or weekly data of the same river have significant levels of autocorrelations. Different techniques may be employed for modeling annual and monthly data; however, in both cases (year 1, year 2,...or month 1, month 2,..) observations are treated, in the literature, as single-valued entities that have inter-relations modeled at one stage.

## 1.2 Area of interest

Hydrologic time series data can be roughly classified into periodic and non-periodic data. Periodicity of monthly streamflows can be easily noticed even by visual inspection where high-flow and low-flow seasons can be identified. The maximum number of seasons in monthly data cannot exceed 12. Therefore, it is possible to have seasonal models with 12 average flows and 12 variances to handle the periodicity. In the case of daily flows, one may argue that data are still periodic. But such a claim can be marginalized for two reasons; first, analysis of autocorrelation function may indicate the insignificance of periodicity in daily streamflows, and second, treating daily data as periodic implies that 365 models should be involved in modeling the data set. Therefore, daily data are usually treated as non-periodic data and this research is no exception.

Missing data can be categorized according to Beale and Little (1975) into three groups: (1) Data of trivial importance are missing (e.g., a few sparsely distributed, not consecutive, missing observations in a long historical record). In this case a simple in-filling method such as in-filling by using series average or simple interpolation could be satisfactory; however, more sophisticated techniques can be used for such an in-filling process (e.g., Griffith et al. 1985). Peak or extreme values should not be encountered in the data gaps to be in-filled by the simple in-filling methods, (2) fundamental data are missing (e.g., lengthy segments or many intermittent observations) where data patterns or structure cannot be recognized from the remaining record. Beale and Little (1975) further stated that in this case, any attempt for in-filling such a record may be unreliable and therefore, the whole record should be dropped according to the current available state of knowledge and techniques, and (3) significant data are missing (e.g., a segment of consecutive observations). In the latter case, the missing data are considered important enough (quantitatively or qualitatively) to deserve developing a technique that estimates them as accurately as possible. At the same time, the data gaps are too short to have significant damaging effect on the patterns and structure of the whole record (Elshorbagy et al. 1999).

Since missing values under the third category occur more often, this category is the focus of this research. This case could happen due to interruption of measurements because of prolonged equipment failure, stopping the measurements at some stations due to any reasons (e.g., budget limitations) and resuming them after some time, and accidental loss of data files. Further, the available literature on estimation of missing hydrologic data can

be classified into single-valued approach (such as regression, time series analysis, interpolation, and artificial neural networks) and group approach. Groups are considered to possess certain characteristics that distinguish them uniquely as separate entities. In other words, the elements that form a group represent a certain uniqueness of the group. In time series, time dependent observations that form groups collectively reflect the unique characteristics of such groups. This research focuses on the group behavior of the periodic and non-periodic data sets with no restriction regarding the modeling techniques, whether linear or nonlinear techniques.

## 1.3 Problem recognition

Estimating missing consecutive observations is worth investigating because available traditional techniques will have to use an estimated observation to estimate the successive ones. In order to propose a methodology that can avoid such a shortcoming, identifying data groups and modeling them as vectors of consecutive values rather than single-valued observations are required and such considerations are the focus of this research. Within each group, the use of data values with the same characteristics ensures group homogeneity. This group homogeneity enhances the extraction of satisfactory information for improved accuracy of data estimation. It is in this vein that there are some advantages to forming seasonal groups towards the development of data infilling models, such as the following: (i) the confidence boundary over the group duration is stable; and (ii) parameters of interest such as correlation coefficient are obtained over the homogeneous subset rather than across the entire set of data.

In this research, it is argued that periodic hydrologic data can have a hierarchical type of structure that needs to be modeled at more than one stage. For example, when different months and different seasons are recognized in the same time series, months within each season can be modeled at the first stage and inter-seasons relations can be modeled at a second stage. This approach is identified as "group approach" because of that inter-season (group) structure. The underlying difference between the traditional single-valued approach and the group approach is that in the latter one, data are perceived as groups that have inter and intra relations. When the monthly flows are autocorrelated and the seasonal flows are independent, the flows can be described as independent groups (seasons) of correlated elements (internal single-valued monthly observations). In the same manner, one can have independent groups of independent elements or correlated groups of correlated elements. It is worth mentioning that available techniques such as regression and time series analysis can be modified to cope with the group approach.

On the other hand, in non-periodic data (e.g., daily flows) where seasonality does not play a major role and a definite number of repeating low dimensional groups of observations cannot be found in the geophysical year, other ways of identifying and modeling groups should be sought. The nonlinearity and dynamic behavior of some hydrologic data sets have been indicated in water resources literature as issues that influence the performance of any modeling tool that ignores the inherent properties of nonlinearity and dynamics of hydrologic data structure. Such features, sometimes, get swept away on the aggregate level. For example, the effect of nonlinear dynamics could be significant in daily data, but is less significant on the monthly level and probably of

much less significance in the annual data sets. One of the appropriate approaches to handle the nonlinear dynamics in non-periodic data is through the use of principles of chaos theory (Kantz and Schreiber, 1997). Investigating the existence of chaos in daily data can provide information about group structure in the data through some chaotic invariants, which will be explained later.

## 1.4 Research objectives

Towards the target of estimating missing groups of data, the objectives of this research can be stated as follows:

1. To identify the group behavior of periodic hydrologic data.
2. To model periodic hydrologic data as groups and demonstrate the utility of the group approach.
3. To investigate the existence of nonlinear deterministic dynamics (chaos) in some non-periodic hydrologic data and link that characteristic to the group approach defined earlier.
4. To model non-periodic hydrologic data as groups using principles of the chaos theory.

## 1.5 Synopsis of the thesis

The different parts of this thesis are presented as follows:

Chapter 2: Key publications and reports from water resources literature pertinent to estimation of missing data are reviewed. A clear distinction between traditional single-valued approach's literature and that related to group approach is also given.

Chapter 3: In this chapter, the perception of groups in periodic hydrologic data is explained and the general methodology for grouping and modeling groups is presented.

Chapter 4: Streamflow data are used as illustrative examples to show the applicability of the proposed methodology to periodic data.

Chapter 5: Perception of data grouping in non-periodic hydrologic data is presented through the principles of chaos theory. Pertinent chaos literature is reviewed in this chapter.

Chapter 6: Applications of chaos theory to daily streamflows are provided to show the application of the proposed methodology to non-periodic data. Also, other important issues related to time series analysis of chaotic data are discussed and critically reviewed.

Chapter 7: A brief summary of the thesis is given. The different levels of contribution of this thesis are highlighted. Also, the scope for further studies and research work and some implied limitations of this research are concisely discussed.

# Chapter 2

# 2. ESTIMATION OF MISSING DATA: A REVIEW


2.1 Introduction

2.2 Single-valued approach

2.2.1 Regression analysis

2.2.2 Time series analysis

2.2.3 Interpolation analysis

2.2.4 Pattern recognition

2.2.5 Artificial neural networks

2.3 The group approach


---

*" There is, in their stories, instruction for people endued with understanding. It is not a tale invented, but a confirmation of what went before it, a detailed exposition of all things, and a guide and a mercy to any such as believe." The Glorious Qura'n (12: 111)*

## 2.1 Introduction

Researchers have been tackling the problem of missing data in different ways and from different perspectives. Even their definitions of "missing data" and the expressions that they have used to describe the in-filling process are no less diversified than the different techniques that they have used. A group of researchers tackled the problem of intermediate missing data where data or observations before and after the missing observations are available (e.g., Beale and Little, 1975; Lettenmaier, 1980; Griffith et al. 1985; Dax, 1985; Panu, 1992; Mott et al. 1994; Gyau-Boakye and Schultz, 1994; and Bennis et al. 1997). The word "***patching***" (e.g., Hughes and Smakhtin, 1996; Makhuvha et al. 1997a; and Pegram, 1997) or "***filling***" (e.g., Panu, 1992; and Gyau-Boakye and Schultz, 1994) is used to express the in-filling in these cases. Others consider the cases in which data are available only from one side of the gap or the gap is so lengthy that the data set is considered bounded from one side only. Generally, the in-filling process in this case is called "***extension***" (e.g., Hirsch, 1982; Alley and Burns, 1983; and Hughes and Smakhtin, 1996), "***synthesizing***" (e.g., Beauchamp et al. 1989 and Simonovic, 1995), or "***estimation***", which is also used to indicate patching type of in-filling (e.g., Griffith et al. 1985; Ben-Zvi and Kesler, 1986; Berkowitz et al. 1992; Bennis et al. 1997; and Knotters and van Walsum, 1997). The word "***augmenting***" is used especially when estimating the model parameters becomes the overriding objective and the estimation of the missing data comes as a direct result of applying the developed model (e.g., Fiering, 1962; Gilroy, 1971; Moran, 1974; and Vogel and Stedinger, 1985). It should be noted that other words such as "***reconstruction***" (e.g., Hirsch, 1979) and "***completing***" (e.g., Dax, 1985) are also

employed to denote estimating missing data. In the above-mentioned cases, single variable and multi-variable applications are found in the hydrologic literature. This research explores the first type: "in-filling" or "patching" where data before and after the gap are available. The available literature on estimation of missing hydrologic data can be classified into single-valued approach (such as regression, time series analysis, interpolation, and artificial neural networks) and group approach.

## 2.2 Single-valued approach

### *2.2.1 Regression analysis*

Augmenting hydrological data using multivariate linear regression has attracted many researchers since early sixties. A short data set is augmented using other data sets by Fiering (1962), when he augmented a short data set using other p data sets. Estimates of the mean and variance, which have greater reliability than those resulting from using the short record only, have been obtained. Then, the mean squared error of the variance for p= 1 and p= 2 has been derived. Since this approach yields a biased estimate of the variance, an unbiased variance estimate has been proposed by Matalas and Jacobs (1964) by adding a noise term to the regression estimates of the dependent variable. In parallel to the formulation in the work of Matalas and Jacobs, Gilroy (1971) has derived a procedure for estimating the sample variance of the estimate of the variance for the general case of p +1 variables. The whole process was reviewed and maximum likelihood estimators, corrected for bias, have been presented by Moran (1974).

A design for a comprehensive data in-filling procedure, which uses linear regression equations for multivariate data, has been given by Young et al. (1970). The procedure employs criteria for truncating data sets or deciding how many data gathering stations are necessary. The data sets of different sites have been arranged in a way that facilitates their use in a step by step filling procedure. The most complete data series is used to infill the following one, and so on. A noise term has been added to the regression equations to avoid variance reduction due to data infilling. Also, the linear regression technique has been satisfactorily used to generate missing monthly relative humidity by Shih and Cheng (1989). In their application they have tried to adhere to the assumptions of linear regression. They have used it only for the stations that have records proved to be statistically white noise. When data showed a significant degree of periodicity, they recommended the use of time series analysis.

Hirsch (1979) has employed linear regression and log-log regression for record extension. Seasonality in the data has not been handled explicitly in that early work. The models have been formulated as if all of the variability in flows were random rather than partly cyclic and partly random. Regular regression (REG) method and regression plus noise (RPN) term added method were considered in a work done by Hirsch (1982). Regular regression (REG) and regression plus noise (RPN) have been compared with two other versions of the Maintenance of Variance Extension (MOVE1) and (MOVE2) for extending short streamflow records. It was concluded that RPN and MOVE2 are better than the other two methods in terms of preserving the variance, but it was noted that they slightly decrease variability in more extreme events.

Motivated by the work of Matalas and Jacobs (1964), Vogel and Stedinger (1985) have improved the variance estimate at a short-record gage. They have also been able to obtain a unique extended record and to preserve the autocorrelation structure by developing the new estimators (MOVE3) and (MOVE4). A regression-based approach which selects a base station among several in a region for in-filling of missing data is presented by *Alley and Burns (1983)*. The authors have, sometimes, used the word infilling to describe the procedures while it was rather extension. The data gaps were lengthy enough to be considered as open at one end. The approach differs from others in that different stations can be selected as the base station to fill-in different values. The approach has also provided a decision rule for using only flow values from the same month or all flow values in developing the extension equation used for estimating a particular value. A deficiency in this approach is that the use of different stations for estimating different values at one site is not justified on solid statistical or physical grounds. The same shortcoming can be reported concerning the unjustified use of cyclic and non-cyclic models for estimating values at the same site. These shortcomings are admitted by the authors when they have concluded that "*...an alternative, and perhaps preferred, application of the methodology is to simply use the mixed-station approach to identify separate periods during which a particular station will be used for data fill-in and to determine whether to use cyclic or noncyclic extension equations...*"

Alternative methods for analyzing multivariate data with missing observations from one or more independent variables are presented by Beale and Little (1975). Their objective is mainly to develop a model that describes the data set. They have considered the case

when some observations are missing from one or more independent variables. Many methods have been presented in their work including the methods that assume multivariate normal distribution of the data and a method that does not require the assumption of normality. The objective was to estimate the parameters of the distribution that lead to the regression coefficients and hence to estimating missing observations. In-filling, extension and generation are considered by Simonovic (1995) and approached by Multiple Nonlinear Standardized Correlation (MNSC) analysis. Although the technique incorporates the nonlinearity, it is mentioned here because it is based on transforming data to their logarithms, replacing the logarithms by their empirical probabilities, and then replacing the probabilities by their standardized variables. This means that data are transformed to be linear before proceeding further with regression. The author has been successful in improving the results of the statistical model by considering some physical parameters.

A direct example of patching missing hydrologic data is conducted by Makhuvha et al. (1997a,b). The techniques used have been confined to regression methods. Three methods have been classified under the approach of "best subset selection" which concerns partitioning the multi-site data into a subset of control sites to be used in the regression, ignoring the remainder. The second approach is the Expectation Maximization (EM) algorithm, which iteratively computes the maximum likelihood estimates of parameters of a model when some observations are missing. The authors have used annual rainfall data and suggested, in case of monthly data, to group the data into a small number of homogeneous seasons. In their second paper, in a set of three

papers, comparisons of accuracy, bias, and efficiency among different methods have been conducted. It has been concluded that the pseudo-EM method, which introduces some modifications to the EM algorithm, was preferred to the rest of the tested methods. A missing or bad block of climate data is filled or replaced using linear regression technique by Mott et al. (1994). The point of interest in their research is that the authors have thought of every observation as a member of a group, but groups and grouping concepts as exist in literature are not invoked in their methods.

*2.2.2 Time series analysis*

Following the forecasting approach, forward and backward estimates are derived for each missing value by Lettenmaier (1980). A single value out of the two estimates is calculated. The importance of considering both time and space has been put into focus by Griffith et al. (1985). Space-time interdependencies are assumed to be characterized by a first-order Markov structure. The cases of single missing value and cluster of missing values are addressed in their paper. Another example of considering a time series model in a multi-site problem has been given by Kuczera (1987). Maximum likelihood estimates (MLE) for the multi-site lag-one streamflow model for incomplete data cases have been provided.

An effort for improving single variable and multivariable techniques for estimating missing hydrological data has been made by Bennis et al. (1997). The authors have discussed the non-stationarity of the time series and the autocorrelation found in the residuals of the regular time series models. They have proposed the use of Kalman Filter

(KF), combined with autoregressive (AR) model, instead of ordinary least squares (OLS) technique to allow for having time variant model parameters. Estimation accuracy of missing data has been increased but the problem of using the estimated value for estimating the next missing one was not solved in case of single-variable model. The time and space inter-relation has received further attention by Van Geer and Zuur (1997). They have extended Box-Jenkins transfer/noise models, known to be single-output models, to a multiple output model. The purpose of doing this was to model several groundwater head series and to enable a spatial interpolation in terms of time series analysis. This extension has been achieved by taking the spatial correlation of the white noise into account and spatially interpolating the parameters of the transfer and noise models. The information transfer from one series to another series has been through the covariance structure of both noise series. Considering the parameters and the white noise of the single transfer/noise models to be spatially random process that is continuous in space, the models have been extended to unmeasured locations. Thus, unmeasured values have been estimated.

Regression and time series methods have been compared by Beauchamp et al. (1989) regarding their ability to synthesize missing streamflow records. Although regression approach has been credited for explaining a large percentage of the variation in the observed data, it has been criticized by the authors for not accounting for the autocorrelation that exists among the original observations. Family of ARIMA models and transfer function-noise models have been used to represent the time series analysis technique. According to the authors, time series analysis is accredited for flexibility and

accounting for autocorrelation, seasonality, and introduction of intervention variable. It is believed that the argument presented in their work is valid if it is assumed that regression approach is applied regardless of the structure of the original observations. It seems that the authors overlooked the fact that seasonality and autocorrelation of the original data should be checked carefully before applying the regression approach.

### *2.2.3 Interpolation*

When the observed phenomenon is characterized by spatial configuration, like the case of a groundwater wells network, a simple function of the space using the historical observations can be formulated (Ben-Zvi and Kesler, 1986). Their model represents a space surface fluctuating parallel to itself on the time axis and it has been used to estimate any missing data. Finally they have brought the attention to the fact that employing periodical configurations may reduce the errors of estimating the missing data. The constraint of parallel fluctuations is removed by the method proposed by Berkowitz et al. (1992). The adopted method is based on integration of two sources of information: prior estimates assuming stochastic behavior of the system and online estimates assuming probability distribution function of field data conditioned on the system state. Another form of using interpolation technique for estimating missing observations is presented by Hughes and Smakhtin (1996). It is based on the use of 1-day flow duration curves for each month of the year and on the assumption that flows occurring simultaneously at sites in reasonably close proximity to each other correspond to similar percentage points on their respective duration curves.

A crude method of grouping is noticed in the work of Dax (1985). The author has built an interpolation technique for estimating missing data on the assumption that each year is composed of two seasons, rainy season with water levels monotonically increasing and dry season with water levels monotonically decreasing. Estimated data are supposed to follow the pre-assumed pattern, though data within each season are treated as single-valued observations. Combining time series and interpolation techniques, Terry et al. (1986) have used time series models (ARMA) to generate time series data similar to the historically observed data. Episodes of missing values were artificially introduced and then filled by linear interpolation of the two nearest values scheme. The authors have noted that the occurrence of missing values did not cause major errors in estimating the parameters for the low order models such as ARMA(2,0) and ARMA(2,1).

*2.2.4 Pattern recognition*

An example of using pattern recognition technique with single-valued data approach is the work of Zhang and Berndtsson (1991). They have used pattern recognition as a useful technique when patterns in space and time need to be evaluated simultaneously. They handle the issue of soil water dynamics by introducing both a time lag and a spatial lag, depending on the depth, in the calculation of the covariances between consecutive pairwise combined time series of soil water content. It is possible to trace a pulse of infiltrated water through a correlation matrix representing the time and space that is used directly for interpolation and estimation of values at unmeasured locations. Outside the area of water resources, pattern recognition techniques were employed for dealing with data sets with some missing observations. The efforts were devoted to classification

process or density estimation from incomplete data (Kittler, 1978; Titterington and Mill, 1983; Flick and Jones, 1986; Queiros and Gelsema, 1988; Titterington and Sedransk, 1989; Zhu, 1990; and Pawlak, 1993). Also, simple methods for imputation or in-filling of missing data have been proposed by Little and Rubin (1987) with the objective of conducting statistical analysis on the completed record.

*2.2.5 Artificial Neural Networks (ANNs)*

The last decade has witnessed many applications of ANNs in water resources. These include, among others: rainfall forecasting (French et al. 1992), multivariate modeling of water resources time series (Raman and Sunlikumar, 1995), modeling of rainfall-runoff process (Hsu et al. 1995), and flow forecasting (Zealand et al. 1999 and Elshorbagy et al. 2000c). The promising results due to the use of ANNs in water resources make them a feasible technique to be employed in this research. The literature related to the use of ANN with missing or incomplete data is limited. Karunanithi et al. (1994) have used the ANN as a synthesizer to predict the river flows and compared the results with the corresponding results of the regular parametric models. The results are in favor of the ANN technique. Kuligowski and Barros (1998) have compared the performance of ANNs with simple techniques, such as regression and weighted averages for estimation of missing rainfall data. The tests performed on a sample of rain gauges show that ANN technique produces results that compare favorably to the other tested techniques. Outside the water resources domain ANNs have been used by Gupta and Lam (1996) to estimate missing values in a multivariate data set. The accuracy of the results is superior to that obtained by iterative regression analysis. It should be noted that some contributions have

been also made to improve the capabilities of the ANN for learning from incomplete training data and to handle inputs with missing attributes (Ishibuchi et al. 1993; and Pedreira and Parente, 1995).

The foregoing literature presented above follows the traditional methods of hydrologic data analysis, which is based on information contained in individual data. These methods ignore information contained in and among groups of data. A few published reports that are available in literature and which use the concept of extracting information from data groups are briefly presented below.

## 2.3 The group approach

A few published reports can be found in the area of water resources where data sets are dealt with as sequences of recognized groups. Panu et al. 1978; and Panu and Unny, 1980a have based their research on the fact that sequences of hydrologic data corresponding to daily, weekly, or monthly measurements occur in well-defined groups. They have considered groups of data on hydrologic phenomena observed at regular time intervals as objects. Each object is described by hydrologic patterns that, in turn, are collective properties or features. The authors have provided the first trial of using pattern recognition techniques in hydrology and solidifying the idea of handling data as groups. The objective of their work was to generate hydrologic data while maintaining their group structure. The procedure followed can be summarized as follows:

A pattern recognition system is formulated by segmenting the time-series data into N segments. Each segment is divided in to K sections or seasons. A pattern vector is formed by taking a set of n measurements on the $k^{th}$ season. All the hydrologic patterns belonging to a season are collected together to form a pattern class. Pattern vectors can be reduced to feature vectors by considering the geometry of each object. Pattern vectors or feature vectors in pattern classes are further grouped into sub- classes. The centre of each sub-class is identified as a reference vector that will be used for classifying future vectors. Classification is based on the distance between the reference vector and any arbitrary feature vector.

The intrastatistical structure in the data groups is indicated by the probability distribution of the features around each element of a reference vector representing them. A univariate normal distribution for each element in the feature vector is assumed sufficient in their work. The interstatistical structure deals with the occurrence of the reference vectors in successive seasons. Reference vectors are assumed to occur in pure random order.

To synthesize N years of record, reference vectors are generated for each season such that these vectors in each season preserve the a priori probability distribution determined empirically from historical data. A set of criteria is considered for rejecting or accepting each generated feature vector. The authors have applied their adopted technique to South Saskatchewan River streamflow data. Finally, the authors concluded that the approach adopted in their work provides a satisfactory procedure for hydrologic data synthesis.

Both inter-group and intra-group structures are claimed to be provided by Panu and Unny (1980b). They have assumed a Markovian model to describe the inter-groups relations while they divide each group type into sub-groups. This further division helps in their objective of generating synthetic data but cannot be considered as intra-group modeling. No model is specified in their work to describe the underlying mathematical relation that links the individual observations within each group. Their model structure would allow them to deal with a group as a whole but not with a partial group.

Also, the research mentioned earlier (Dax, 1985 and Mott et al. 1994) indicates, implicitly, the existence and significance of groups in the hydrologic data. Dax (1985) has assumed that each year was composed of two seasons, a rainy season with water levels monotonically increasing and dry season with water levels monotonically decreasing. So it was desired that the estimated data would follow this seasonal pattern. This concept can lend a support to the fact that each observation is a member of a specific group of observations. The proposed interpolation method was based on a sequence of data that has the same seasonal pattern as the original data. This sequence of data was created by a simple additive model that depended on the mean of the series and the seasonal deviations from the mean, or by standard forecasting process of an ARIMA model. It is noted that the concept of grouping is not extensively applied, only one observation before and one after the observation under consideration have been used to check the trend of the segment.

Missing or bad blocks of climate data have been filled or replaced using linear regression technique by Mott et al. (1994). The objective of their research was to determine an effective method of estimating missing or replacing bad climate data based on information from surrounding weather stations. The point of interest in their work is that the authors thought of every observation as a member of a group. Although the authors never mentioned the word group explicitly, there was a sort of unjustified grouping in their two proposed methods. In the first method, they have developed a linear regression equation for every day of the year based on a window of 30 days before and 30 days after that day of the year. Three years of complete data sets have been used to develop the equations. The length of the window is not justified and groups became like moving windows rather than distinct groups. In the second method, they have found out that a specific block of data before the missing block of the current year was providing estimated values with the same level of accuracy when the whole data set was used. This result can support the argument of data groups adopted in this thesis. The results could be attributed to the fact that the block that precedes the missing block is forming the group within which the missing values are lying. This group might have captured information, about the missing block, equal to or more than that contained in the three years of historical data. Khalil et al. (1998) have indicated the possibility of in-filling missing data, using ANN, based on the group approach. However, questions of how to employ the group approach for in-filling missing data and how to adapt the ANN to perform it are not clearly answered in their work.

Three group types are identified in hourly water demand data by Shvarster et al. (1993). They have used time series analysis technique to model the intra-group structure. However, the data within each group are treated as single-valued observations. Three repeating segments are observed in the daily demand pattern: "rising", "osciliating", "falling", then "rising" again the following day. These have been called "states" of the demand curve and have been defined as successive states of a Markov process. The transition probabilities between states have been learned and low-order ARIMA models fitted to each segment. The model is then used to forecast hourly demands for a period of one to several days ahead. The analysis and forecasting contain the following steps: assuming, learning, recognition, and forecasting. Posterior probabilities of each state at all times are produced using an estimation algorithm. The sequence of values of such estimations can be viewed as sets of time series. These time series have been seen to have a fixed period of 24 hours, and modelled by ARIMA models. After selecting the state with highest probability, at each time, demands are forecasted using the forecasted states. Finally, the authors concluded that their adopted approach can also serve other purposes than forecasting.

Perceiving data as groups can give insight into the data structure and affect the results of further analysis using the data under consideration. For a long time, annual flood peaks have been assumed and treated as serially independent. Lately, it has been shown by Booy and Morgan (1985) that flood peaks on the Red River at Winnipeg, Canada, show clustering characteristics. As a result, flood risk has been shown to be substantially higher than that estimated by conventional methods that assume serial independence of the peak

flows. This supports the grouping found by Panu and Unny (1980a) in some annual streamflow data and the long-term persistence in annual streamflows shown by Hurst (1951 and 1956).

The research conducted in the area of estimation of missing hydrological data can be classified in several ways under different categories. However, different applications are classified in this thesis under two main categories; single-valued approach and group approach. Even though the foregoing literature is diversified with regard to the techniques used in the analysis, most of the work comes under the single-valued approach, where each observation is perceived as an individual entity. Considering different data structures, other than single-valued approach, has not been given a significant attention in the past, with the exception of a few publications. The group approach, which is a different way of perceiving the data structure, is under consideration in this research. Therefore, it is presented and explained in the ensuing chapter.

# Chapter 3

# 3. PERCEPTION OF GROUPS IN PERIODIC DATA

3.1 Introduction

3.2 Proposed approach and methodology

3.3 Data segmentation

3.4 Difference between group approach and existing models

3.5 Proposed group-based techniques and models for data in-filling

---

*" Thus do We explain the signs by various ways that they may say, thou hast learnt this from somebody, and that We may make the matter clear to those who know"* *The Glorious Qura'n (6: 105)*

## 3.1 Introduction

Perceiving the hydrologic data as groups can be considered different from the single-valued approach in the underlying algorithm and the modeling mathematics. The concept of data groups has been always recognized in the traditional single-valued approach. For example, arbitrarily assuming that the streamflows can be discretized into hourly units (assuming that the hour is the smallest unit and flows are constant over the one hour duration), then 24 observations per day and 720 observations per month can be obtained. In hydrologic models dealing with monthly flows, the 720 values of each month are aggregated in one value, which is the average of these values. So, when a monthly value is predicted or forecasted based on the previous monthly value, in essence, it is a group of 720 hourly values that are forecasted based on the previous group of 720 observations. Condensing the whole group in one value is viewed here as a way of simplifying the complexity of the needed mathematics when the multi-dimensionality of groups is maintained. In this regard, it is argued that the concept of data groups is widely recognized but traded for simplicity in stochastic hydrology.

Another way of recognizing groups in hydrologic literature is through the application of disaggregation models. As a first step, data are treated as aggregate values and in a subsequent step, these aggregate values are disaggregated into multi-dimensional elements (Valencia and Schaake, 1973). The mathematics needed to deal with the group vectors is simplified by dividing the modeling procedures into the above-mentioned two steps. On the other hand, if disaggregation model is used to estimate a missing season (group of values), an aggregate value of the season has to be estimated first and then

disaggregated into its original elements. Performing these steps entails two stages of parameter estimation and hence more complications. Further, daily or short time period data does not lend itself well to disaggregation (Salas et al. 1980).

A clear diversion from the traditional single-valued approach in the stochastic hydrology literature is a work by Panu et al. (1978) based on concepts of pattern recognition. Their proposition was made for the sake of generating synthetic hydrological sequences. In generation, only the statistical properties of the series are important, and they are the parameters that should be preserved. The generated sequences could be far from the original data in terms of error measure (e.g., mean squared error), yet the generation process can be successful as long as the statistical properties (mean, variance, skewness,..) are preserved. Where in-filling or estimating missing data is considered, minimizing the squared error (difference between estimated and true values), such as the case of this research, becomes the overriding objective. The proposition of Panu et al., which is briefed and clarified here, can be modified for in-filling of missing data.

The hydrologic time wave form (HTW) is a plot of a random variable, such as rainfall or discharge, versus time. Any section of this HTW, for example, a section corresponding to a time duration of one season in a geophysical year, is an object. Thus the HTW consists of many objects. A collection of all such objects is considered to constitute a universe. The simplest example of a universe is that containing hydrologic objects from two categories, namely, the dry season and the wet season, describing all the objects of the HTW. The HTW or any observed hydrologic sequences of data or observations can be

converted into a hydrologic "signature" by joining the data points by a continuous line representing simple geometric figures. The characteristics of such a signature can be evaluated in terms of several different specified features. These features and the collection of such features can be considered equivalent to letters and words, respectively, in a written language. This interesting analogy with written language is made by Panu et al. (1978) because it is there that pattern recognition has found wide application.

*"....Written language is an expression of human thought in signature. Such a signature contains a proper set of words arranged according to a specified grammar in order to provide proper meaning. It has been observed that any sequence of letters, as well as any sequence of words in a written language, is stochastic in character. It is groups of letters that provide a meaning to the stochastic written language. For example, individual letters, such as W, I, N, D, and O, carry no meaning. However, if they are considered as a group of letters (WINDO), the group not only provides a sense but also indicates that it is an incomplete group and the next most probable letter to complete the group is the letter W. In this case the occurrence of the letter W depends on the previous sequence of letters...".*

This is the argument provided by Panu et al. (1978). In this research this argument is accepted and also extended as follows (Elshorbagy et al. 2000a):

If (WIN) is available, then the process of guessing whether it is complete or not is not an easy task. If it is not complete then one letter could be missing. That letter might be D, E, G, S, or K. If more letters than one are assumed missing then the number of possibilities increases and might become an unmanageable process. Two major criteria can be considered to limit the number of possibilities and to facilitate the task. The first criterion

is the group length and the second criterion is the existence of groups (words) that come before and after the group (word) under consideration. If we know that the group length is four units long and that only one letter is missing, then we need to consider the letters that come before and after the word (WIN). Assuming the following letters are forming the series:

**TODAYISCOLDBUTTHE*WIN*-ISCALM** (3.1)

If the grammar is learnt from the previous letters, sentences, and as well from other related texts, this sequence of letters might be segmented and written as follows:

**TODAY IS COLD BUT THE WIN- IS CALM** (3.2)

In this case, the missing letter is, most probably, the letter ***D***. The estimated letter depends on: first, segmenting the sequences into meaningful groups (words), second, knowing the length of the group where some data are missing, and third, defining the structure of the whole sentence before and after the group with missing data. These three steps cannot be conducted without learning the grammar of the language. Returning to our hydrological application on monthly streamflow data for example, the analogy is made as follows: the letters represent the monthly values, the words might represent seasonal values, and the sentence might represent the whole year. If the HTW is regarded as a meaningful text, it is conceptualized that monthly observations are linked in a specific way to form the seasons, seasons are linked in a specific way to form the year,

and also years might be linked in a specific way to form the whole series. The problem in hydrologic series is that such a language of the time series and its grammar is neither well defined nor understood. Grammar, in this case, has to be deciphered and learnt from the whole text (time series under consideration) and also from any other available related texts (related time series).

## 3.2 Proposed approach and methodology

Every hydrological observation in a time series record is conceived as a member of a season, segment, or a group of observations. Such a concept requires the hydrological data record to be segmented into parts. These parts might be named segments, groups or seasons. Each season consists of a number of consecutive observations. For example, let us consider a monthly streamflow record for 50 years: the entire data set includes (12*50=600) observations. Plotting these observations on a graph, one can observe that well-defined patterns, which can be proved by further investigations, exist in the data.

So, the **first step** in our proposed methodology is to divide the data into segments – let us assume 150 segments are obtained out of the 600 observations. These segments need not to be of equal lengths (i.e., equal number of months in each segment is not required). Also the names of the months (Jan., Feb.,..etc..) are disregarded such that the same month can be associated with different group members (months) every year. Implicitly this means that one month might fall in a wet season in one year, semi-wet in another year, dry in a third year, and so on. At the same time, having segments of different lengths

means that the hydrological seasons are of different durations. The second assumption seems to be feasible and close to the reality while the first one is a problematic assumption due to several reasons. The first reason is that the names of the months can be considered as hydrological attributes because historical observations show that the same month has common properties, every year, concerning any hydrological parameters (e.g., precipitation, temperature,..etc..). Therefore, considering every month to be a member of a specific fixed group repeating over the years not only sounds logical but is also a realistic assumption from a hydrologic viewpoint. The second reason relates to the concept of modeling itself, where generalization and pre-determined pattern structure are sought and modeled to mimic the reality as closely as possible. If a small percentage of the observations is not following the assumed pattern, they are considered as exceptions and are included in the error probability of the model. The third reason is guided by the considerations of practicality. Consideration of the exceptions or the minority of the observations in formulating the model structure suggests that they may constitute a significant pattern in the data set. Referring to our example of 50 years of monthly data, many sets or replicates of 50 years data may be required in order to track the patterns of the minority. In other words, hundreds and probably thousands of years of data will be needed for the analysis, which imposes an obstacle that hinders any progress in the modeling approach.

Since the objective of modeling in general, including the approach proposed in this thesis, is to reassure that simplicity and closeness-to-reality can coexist in hydrological modeling, the assumptions related to the first step of our proposed methodology are

formulated as follows: (1) the segments are allowed to be of different lengths, (2) definite types of segments are recognized and assumed to be repeating over the years. It should be noted that repeating segments share common statistical properties, and not identical internal values.

After segmenting the data as a first step, the **second step** in the proposed methodology is to prepare the segmented data for further analysis. The analyses or the models that adopt the single-valued approach in stochastic hydrology require that the data satisfy some basic statistical characteristics such as normality, trend, seasonality, and correlation structure. What has to be achieved in order to apply the proposed group approach is the identification of the statistical properties of the data groups rather than the individual observations. Stationarity, multivariate normality within the groups, seasonality, and independence or correlation functions among the groups should be tested before proceeding further with any analysis. To help understand the terminology used in this thesis, a brief explanation of all such terms is given in appendix A.

The **third step** of our proposed methodology is to model the data as a sequence of groups in order to in-fill the missing segments. Two cases of applications are addressed in this research. These cases are explained below.

*Case 1:* single series case, where only one time series is available for the analysis. It is the time series with data gaps that need to be filled. From now on, this time series record is named as the *target time series*. In cases of streamflow data, it can be called *target*

*river*. Many scenarios of data gaps can occur in such a record (Panu, 1992). The cases of entirely and partially missing segment are assumed in this research. A schematic diagram of the single series case is shown in Figure 3.1.

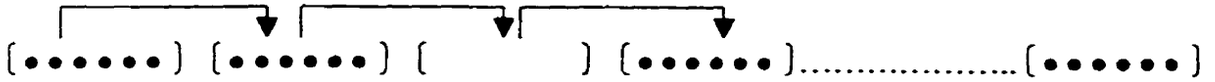


**Fig. 3.1 Schematic diagram for group in-filling in a single time series.**

***Case 2***: multiple series case: where time series other than the target one are also available. These series are referred to as *reference time series*, or *reference rivers* in case of streamflow data. They should be cross-correlated to the target series. The number of the reference time series could be one or more, and they need not to be restricted to the same hydrological variable of the target series (e.g., flows, precipitation, temperature,..etc..). The bi-series case, which is a special case of the multiple series, is considered in this thesis. A schematic diagram of the bi-series case is shown in Figure 3.2

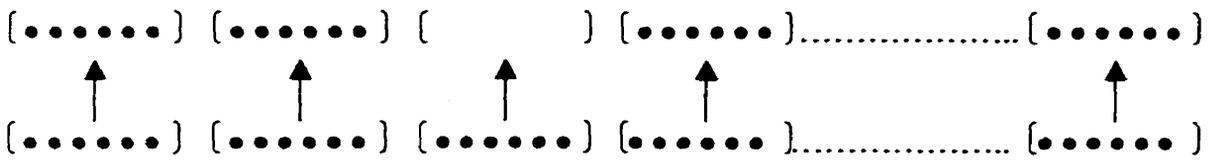


**Fig. 3.2 Schematic diagram for group in-filling in a bi-series case.**

The analogy between the traditional single-valued approach and the group approach is depicted in Figures 3.1 and 3.2. In the proposed group approach, the segment of consecutive observations replaces the single-valued observations. Therefore, techniques that are capable of modeling hydrologic data according to the single-valued approach can be modified to cope with the group approach. Moreover, the techniques and the models that have been developed to deal with groups rather than single-valued data can also be employed for analysis of groups. As mentioned earlier, two levels of data modeling (intra-group and inter-group) may be required to in-fill data gaps.

In the proposed methodology, each group is viewed as an object or a rigid segment of multiple elements (observations). These elements behave as a group in relation to other groups. In this regard, the intra-group structure needs not to be modeled separately from the inter-group model. Development of a model that treats the data as a sequence of multi-dimensional segments is considered satisfactory to handle inter and intra structures of data groups. At the same time this group character is utilized in handling partial groups as will be shown in chapter 4.

Based on the previous discussion, statistical analyses of data groups play the major role in selecting the most appropriate technique for the analysis. In the single series case, the time series analysis can be employed after modifying the formulations to deal with groups. It should be noted that time series models can be applied to the single series case when autocorrelations among groups are considered significant. The artificial neural networks (ANNs) and multivariate partitioning modeling (MPM) technique (Johnson and Wichern, 1988) are also used for the single series case. The ANN technique can easily deal with groups with no required modifications, and the MPM is mainly designed to handle data vectors (groups). The details of these techniques are provided in section 3.5. It is worth mentioning that ANNs can handle both autocorrelated and independent groups. The conditions of significant cross-correlation and insignificant auto-correlation among groups are assumed for application of the proposed methodology in case of multiple series analysis. As in the single series case, the techniques of ANNs and MPM are utilized for the multiple series case. Further, the regression analysis technique lends itself to the analysis when cross-correlation between the *reference* and *target* time series is significant and groups within each of the participating time series are considered to be serially independent. According to the group approach, regression technique should be modified to handle the situation of multiplicity in both of predictors and response variables. It is found that multivariate multiple regression (MMR) technique (Johnson and Wichern, 1988) can be employed to deal with group regression.

## 3.3 Data segmentation

As the first and most significant step in the group approach, data segmentation is further elaborated in this section. The analysis pertaining to groups is perhaps best accomplished through pattern recognition (Panu and Unny, 1980a). The analogy, discussed earlier, between groups of hydrologic observations and formal language makes pattern recognition a good candidate for segmentation and recognition of a definite number of classes in the hydrologic data.

Segmenting the data into a definite number of repeating groups entails answering three questions. The first question is what an appropriate number of groups is, the second is where and how the boundaries between different groups can be drawn, and the third is how to test or verify the performance of the segmenting technique? In the hydrologic literature, the first two questions are addressed through the use of the periodogram as a tool for determining number of seasons or harmonics in a year (Panu et al. 1978 and Khalil et al. 1998). The segmenting of hourly water demand data is done by only observation in Shvarster et al. (1993). Additional elaborations concerning the grouping process are not considered in their work. The statistical method of bi-plot is used by Pegram (1997) to check the consistency of multiple sites or months to form a group. The third question of verifying the performance of the segmenting technique has not been given enough attention in the hydrology literature.

The pattern recognition domain is one of the appropriate areas for finding some answers to, or at least addressing, the above-noted questions. The problem can be described in the pattern recognition context by considering each segment type as a pattern class (PC). The segments from the same type can be treated as pattern vectors (PV). So in the case of monthly streamflows, a season (e.g., Jan, Feb, Mar, and Apr) is a PV where every month is an attribute of that vector. Pattern recognition may be regarded as a technique of discriminating the input data, not between individual patterns but between populations, via the search for features or invariant attributes among members of a population (Tou and Gonzalez, 1974). The features that represent differences between pattern classes are called inter-class features while the features that characterize attributes common to all patterns belonging to one class are called intra-class features.

Although the analogy between the pattern recognition field and the problem at hand is feasible, one should bear in mind two important points. First, considering each month as an attribute of the PV implies the case of multivariate or multi-dimension, but all attributes are of the same unit of measurement (e.g., flow: volume of water/time). Second, a learning set of data where the patterns are already of known classes is not available. Therefore, our hydrological problem will come under the umbrella of the unsupervised learning technique. Suppose that we are given a set of pattern vectors $\{X_1, X_2, \ldots\ldots, X_N\}$ of unknown classification: a good way to let them cluster themselves in groups and get separated from each other is by calculating their distance from a fixed reference. The motivation for using distance functions as a clustering and/or classification tool follows naturally from the fact that the most obvious way of establishing a measure

of similarity between pattern vectors, which are also considered as points in the Euclidean space, is by determining their proximity (Tou and Gonzalez, 1974). In some cases of hydrological data, observations with higher mean values tend to be of higher variance. If the flow value and the variance are taken as two criteria for segmenting the data, then streamflow data can be clustered according to their values and associated variance. Other methods of classifying hydrological patterns were given by Kojiri et al. (1994). It is in this sense that seasonal mean flow and variance can be used as inter-class features. By calculating the distance between every month and the time series mean value as a reference, observations might get clustered (e.g., Jan, Feb,...etc..as PC1, May, Jun,...etc as PC2, and so on) for the N years. Hence the pattern vectors can be denoted as follows:

$$\mathbf{PV}_i^{(j)} = \left[\mathbf{x}_1, \mathbf{x}_2, \ldots\ldots\ldots\ldots \mathbf{x}_k\right] \tag{3.3}$$

where the superscript (j) refers to pattern class, subscript (i) refers to the year, and k is the number of months in the class.

Considering the question of how many classes should be used, one can say that outside of the hydrology domain, this problem has been recognized by some researchers. The difficulty inherent in estimating the number of classes is demonstrated by Dubes (1987). Many methods were proposed for this purpose, some of which depend on minimizing the squared error by calculating the distance from each PV to the centroid of the class (PC).

Criteria were developed for stopping rules or preventing more splitting of classes (Duda and Hart, 1973).

Nonhierarchical clustering techniques (e.g., k-means), which are designed to group items into a collection of k clusters, can be investigated. The k-means algorithm is based on starting by partitioning the items into k initial classes or by specifying k initial centroids. Two problems will be encountered in applying this algorithm to the problem at hand. First, k has to be determined in advance, which is to be avoided in this investigation. Second, the existence of an outlier might produce at least one group with widely dispersed items (Johnson and Wichern, 1988). Therefore the k-means algorithm is not considered in this research. In this hydrologic application the problem seems to be unique because the required segmenting also entails deciding on the dimensionality of the pattern vectors and the Euclidean space. By segmenting the monthly data, the dimensionality of the Euclidean space, which may be different for each pattern class, is determined. This relates to the earlier proposition of calculating the distances from each month to the mean of the whole series. Plotting these distances and their variance can help easily segment them into classes.

After segmenting the time series into 'm' pattern classes, PC1, PC2,...PCm, these classes need to be checked in order to evaluate the segmentation process. In the segmentation process, the proximity or the distance measure is used; therefore, it is convenient to use the distances between patterns as a measure of similarity. In this case, the Euclidean distance measure (Eq.3.4) readily lends itself to the process because of its familiar

interpretation as a measure of proximity. The Euclidean distance between two pattern vectors $X_1$ and $X_2$ is given by

$$d_E = \| X_1 - X_2 \| = \sqrt{(X_1 - X_2)'(X_1 - X_2)} \tag{3.4}$$

where $(X_1 - X_2)'$ denotes the transpose of $(X_1 - X_2)$.

The centroid $\mu_i$ for each pattern class is calculated as follows:

$$\mu_{il}^{(j)} = \frac{1}{N} \sum_{l=1}^{N} x_{il} \qquad \forall i \text{ and } \forall j \tag{3.5}$$

where j denotes the pattern class, $l$ is the year, and N is the number of pattern vectors (segments) in the class, which is the number of years (in case of monthly data).

For calculating the similarity measure within the class, it might be argued that the Mahalanobis distance can be a better representative because it includes the class covariance matrix. It is believed that the Mahalanobis distance (Eq. 3.6), differs from Euclidean by weighing vectors according to their variance. The months of higher variance are given lower weight.

$$d_M = (X - \underline{\mu})' C^{-1} (X - \underline{\mu}) \tag{3.6}$$

where X represents a pattern vector (segment), $\underline{\mu}$ is the mean vector, and C is the covariance matrix of a pattern population. In our segmentation process, the similarity of variances was considered, visually, in clustering the months together in groups. Therefore it is assumed that the differential effect of the monthly variances is insignificant within each class. Yet the Mahalanobis distance is used in this thesis, and its results are compared to Euclidean distance.

## 3.4 Difference between group approach and existing models

It may be worthwhile to clarify the essential difference between the proposed group approach and the seasonal time series analysis models (e.g., periodic autoregressive moving average, PARMA model), especially when the word "season" may be used interchangeably with "group type". The PARMA model has been proposed to overcome the problem of seasonality in the auto-correlation by developing a model with different parameters for each season (Salas, 1992). In that sense, there is a difference between PARMA and the group approach. In PARMA, the word "season" represents one value rather than a vector, and adjacent observations are not considered as a group. For example, in order to apply PARMA model to monthly data where a season corresponds to one month, then 12 models are required. On the other hand, in the group approach, the number of models is significantly reduced (e.g., 2 models when a group corresponds to 6 months is assumed). Therefore, the PARMA is a member of the traditional single-valued approach.

## 3.5 Proposed group-based techniques and models for data in-filling

Models of the time series analysis, regression, ANNs, and MPM techniques are employed to model the data groups and to estimate the missing segments. A description of each type of models follows.

*a) Multivariate time series modeling*

Multivariate analysis of time series started to receive attention as early as 1969 (Gnanadesikan, 1977). Afterwards, a number of multivariate models have been proposed in water resource literature. Different multivariate ARMA models were classified and reviewed by Salas et al. (1985). Useful notes on the use and abuse of multivariate time series models in hydrology were presented by Camacho et al. (1987). The autoregressive model AR(p) with constant parameters (Salas et al. 1980) is given in Equation (3.7) below.

$$y_t = \mu + \phi_1 (y_{t-1} - \mu) + \ldots\ldots\ldots + \phi_p (y_{t-p} - \mu) + \varepsilon_t \qquad (3.7)$$

where $y_t$ is the time dependent variable, $\varepsilon_t$ is the time independent series (error) which is independent of $y_t$ and normally distributed with mean zero and variance $\sigma_\varepsilon^2$, $\mu$ is the average value of the time series, and $\phi$ is the autoregressive coefficient. The group approach presented in this thesis treats the single series as a multivariate problem. Therefore, the multivariate AR model can be the most appropriate form of AR models to

be modified and employed to handle data groups. The multivariate AR(1) model is given in Equations (3.8) and (3.9).

$$Y_t = \underline{\mu} + \underline{\sigma}\, Z_t \tag{3.8}$$

$$Z_t = A_1\, Z_{t-1} + B\underline{\varepsilon}_t \tag{3.9}$$

where $\underline{\mu}$ is a (k*1) matrix, $\underline{\sigma}$ is a (k*k) diagonal matrix, and $Y_t$, $Z_t$, and $Z_{t-1}$ are (k*1) matrices assuming that we have 'k' series. A and B are (k*k) parameter matrices. In our proposed group approach, a single time series with consecutive segments, each of which consists of 'k' elements, is treated as 'k' series. Figure 3.3 provides a schematic diagram for representing a single series of monthly flows using the group approach. Three segment types of four elements each are assumed for illustration. In this case the problem of a single time series is viewed as similar to the case of multivariate AR model with four time series. The data are arranged in a fashion that facilitates calculation of different correlation coefficients. This model is named the *autoregressive group* (ARG) model. In the same way of designing a periodic multivariate AR model (Salas et al. 1980), a periodic ARG model (PARG) for each segment type is developed and applied in this thesis.

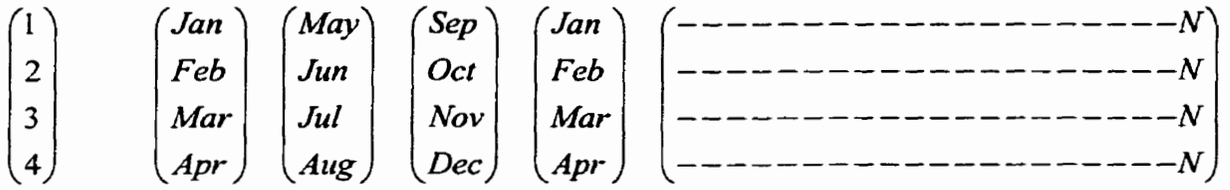



**Fig. 3.3 Schematic diagram representing a single series in the group approach as a multivariate problem**

*b) Multivariate partitioning model (MPM) technique*

Assuming the case of single or multiple time series with two segment types (classes) $X^{(1)}$ and $X^{(2)}$, let $X = \begin{bmatrix} X^{(1)} \\ \hline X^{(2)} \end{bmatrix}$ be distributed as $N_p(\underline{\mu}, C)$ with $\underline{\mu} = \begin{bmatrix} \underline{\mu}^{(1)} \\ \hline \underline{\mu}^{(2)} \end{bmatrix}$, $C = \left[\begin{array}{c|c} C_{11} & C_{12} \\ \hline C_{21} & C_{22} \end{array}\right]$, and $|C_{22}| > 0$. Then the conditional distribution of $X^{(1)}$, given $X^{(2)} = X_2$, is normal with

$$Mean = \underline{\mu}^{(1)} + C_{12} C_{22}^{-1} \left( X_2 - \underline{\mu}^{(2)} \right) \tag{3.10}$$

There can be more than two groups. For example, the case of three groups will be given in chapter 4. The result provided by Equation (3.10) is given by Wichern and Johnson (1988) as a property of multivariate normal distribution. Note that the vertical and horizontal lines shown inside a matrix notation indicate that it is a partitioned matrix. The resultant mean of Equation (3.10) is used as an estimate of the missing segment in this research for both cases of single and bi-series. In the case of a single time series, $X_1$ is considered as the missing segment of "target river" and $X_2$ is the previous segment assuming lag-1 transition model. For the bi-series case, $X_1$ is the missing segment of the "target river" and $X_2$ is the concurrent segment of the "reference river".

*c) Multivariate multiple regression (MMR) technique*

Simple and multiple regression techniques are widely used in hydrology. In both cases of applications, the response variable is always a single dependent variable. The group approach requires a regression model that can handle multiple response variables as well as multiple predictors. The MMR technique is the statistical linear method for estimating values of one or more dependent variables from a collection of predictor (independent) variables (Johnson and Wichern, 1988). One such multivariate MMR model is presented as follows:

$$\underset{(N*k)}{Y} = \underset{(N*(r+1))}{Z} \quad \underset{((r+1)*k)}{\underline{\beta}} + \underset{(N*k)}{\underline{\varepsilon}} \tag{3.11}$$

Where $\underline{\varepsilon}$ is the residual term, whose sum of squares and cross product is obtained as follows.

$$\underline{\hat{\varepsilon}}'\underline{\hat{\varepsilon}} = Y'Y - \hat{Y}'\hat{Y} = Y'Y - \underline{\beta}\hat{}'Z'Z\underline{\hat{\beta}} \tag{3.12}$$

Where k is the number of elements in each estimated response vectors Y, N is the number of observations, (r+1) is the rank of the design matrix Z, and the unknown parameter matrix, $\beta$, is estimated as follows

$$\underline{\hat{\beta}} = (Z'Z)^{-1}Z'Y \tag{3.13}$$

Where Y is the estimated vector and has the following form:

$$\hat{Y} = Z\underline{\hat{\beta}} = Z(Z'Z)^{-1}Z'Y \qquad (3.14)$$

It is noted that the multivariate multiple regression equation is based on the assumptions that the expectation of errors is equal to zero, the errors are statistically independent, the variance of errors is constant, and the errors are normally distributed.

*d) Artificial neural network (ANN) technique*

ANN is a computing paradigm that may have various types of configurations employing different learning algorithms. The feed-forward neural networks with back propagation (BP) learning algorithm are the most widely used neural networks (Anderson, 1995; Freeman and Skapura, 1991). This research employs networks with three layers. The configuration of a neural network includes determining the number of hidden layers, the number of nodes in each of the hidden layers, and the connection weights. Details on the ANNs and the BP training algorithm (to estimate the network parameters) can be found in (Dhar and Stein, 1997; Anderson, 1995; Freeman and Skapura, 1991; and Maren et al. 1990). It is worth mentioning that the ability of the ANN technique to establish a functional relationship between multiple inputs and outputs makes it a self-justified candidate for the group approach proposed in the present research.

In this chapter, the group approach has been presented and explained. The implicit and explicit consideration of groups in the hydrologic literature have been discussed in light of the difference between the group approach, proposed in this thesis, and the existing models. An analogy between the written languages and the hydrologic time series has been provided to clarify the theoretical utility of the group approach in handling the periodic hydrologic data. The steps of implementing the group approach to periodic data have been outlined. An autoregressive-groups (ARG) model and periodic autoregressive-groups (PARG) model have been proposed for time series analysis of data groups. Other existing techniques such as multivariate multiple regression, artificial neural networks, and multivariate partitioning model have been modified to handle the group approach. What is needed now is to implement the developed and modified models to real hydrologic case studies to verify the utility and the applicability of the proposed group approach. In the following chapter, the group approach is validated through the application of the models, proposed in this chapter, to two case studies; first, the case of a single time series and second, the case of multiple time series. In-filling the missing data is used, in this thesis, as a commonly encountered hydrologic application.

# Chapter 4

# 4. APPLICATION OF GROUP APPROACH TO PERIODIC HYDROLOGIC DATA

4.1 Introduction

4.2 Application to streamflow records

4.3 Indicators of model performance

4.4 The pooled mean squared error (PMSE)

4.5 Discussion

4.6 Concluding remarks

---

*" ALLAH disdains not to use the similitude of things, even of a gnat as well as anything above it. Those who believe know that it is the truth from their Lord; but those who reject Faith say: "What means Allah by this similitude?" By it He causes many to stray, and many He leads into the right path; but He causes not to stray, except those who forsake the path"* *The Glorious Qura'n (2: 26)*

## 4.1 Introduction

The approach introduced in the previous chapter is utilized for estimating the missing data. The aim of this chapter is to demonstrate the application of the techniques presented earlier, using the group approach. The application consists of two cases: a) single series case, where only one time series is available for the analysis and b) bi-series case, where two cross-correlated time series are available, the reference series with complete record and the target series that has missing observations.

## 4.2 Application of the group approach to streamflow records

Different models representing four techniques are applied to monthly streamflows of three rivers. These rivers are the English River at Umferville, Ontario, Canada; and the Little River and Reed Creek, Virginia, USA. Table 4.1 summarizes information such as drainage area, length of records used for calibration and verification of the different models, and lag-1 autocorrelation (AC) coefficient for each river.

**Table 4.1. Basic information on the three rivers under consideration.**

| River | Drainage area ($Km^2$) | Calibration | Verification | AC |
|---|---|---|---|---|
| English | 6230 | 1926- 1985 | 1986- 1991 | 0.65 |
| Little | 777 | 1929- 1979 | 1980- 1989 | 0.46 |
| Reed Creek | 640 | 1929- 1979 | 1980- 1989 | 0.45 |

Prior to segmenting the time series records into groups, the data are pre-analyzed as single-valued observations. Absence of outliers and other indications of major interventions in the natural flows need to be verified. In spite of the importance of such pre-analysis, it is a step that is usually overlooked by analysts (Salas et al. 1980). For the case of the English River, flow record at another station of the same river, Sioux Lookout, is used to verify the flows at Umferville. The Little River and Reed Creek are used together for the bi-series analysis. Therefore, they can be used to verify each other. These two rivers are from the same basin with cross-correlation coefficient equal to 0.78. Historical monthly data and mass curve are plotted for each river. Also, for each pair of comparative records (Little River versus Reed Creek and English River at Umferville versus Sioux Lookout), double mass curves and historical flows are plotted. These graphs help detect the existence of extreme values, outliers, or any significant slope change in the curves. The need for deletion or correction of observations is not detected in the preliminary investigations on the data sets. Examples of graphs of data pre-analysis are given in appendix B.

### *4.2.1 Segmentation*

The English River is a challenging example for any segmenting or clustering technique because the annual and monthly values of the mean flows and the variances are larger than normally expected. The mean value of the whole series is calculated to be 56.9 $m^3/s$ and is taken as a reference point from which distances to all months are calculated. Distances are allowed to be of either positive or negative values. For simplicity of illustration, the average values of the distances are plotted (Figure 4.1). By visual

inspection of Figure 4.1, one can draw vertical lines in two places, first between April and May and second between July and August. At these two places there are jumps in the calculated distances, which are large enough to justify the partitioning. So, tentatively we have three pattern classes (PCs) as follows: PC1 {Jan, Feb, Mar, Apr}; PC2 {May, Jun, Jul}; and PC3 {Aug, Sep, Oct, Nov, Dec}. Each pattern class consists of one segment type with N realizations; where N is the number of years. For PC1 each segment can be considered as a point in the 4-dimensional Euclidean space. Likewise, PC2 and PC3 are considered to be of three and five dimensions, respectively, in the Euclidean space.

One of the useful ways of representing the data to help segment them and also to judge the resulting segments is by plotting them on a Box-plot graph type. The Box-plot for the monthly flows of the English River (Figure 4.2) shows the interquartile, the median, the range, and the outliers (shown as asterisks) for every month. It should be noted that *outlier* is defined here as the value more than 1.5 box-length from the $25^{th}$ or $75^{th}$ percentiles. The monthly Box-plots provide further support for data partitioning into three classes. It can be noted from Figure 4.2 that the first four months form a class, the following three or four months are close to each other, forming another class, and the last four months can be considered as forming a third class. At the same time one can expect the third class to be a source of analytical difficulties, compared to the other two classes, due to the existence of a large number of outliers (Figure 4.2). Any claims that one can reduce the variability of the third class by splitting it into more classes can be refuted by a simple fact: the variability is not concentrated in one month, but rather exists in the last four months constituting the third class. Figures 4.3a, 4.3b, and 4.3c are presenting Box-

plots for every class separately. In Figure 4.3a the location of the median, length of the box, and the interquartile show clearly that there is no significant skewness in the distribution. The few outliers are not of great influence on the class. The second pattern class (Figure 4.3b) has some positive skewness. In the third class, the existence of many outliers is not introduced because of the clustering process. The individual attributes experience similar variabilities while it can be noticed that the positive skewness of the attributes is slightly smoothed out and improved by bringing them together into one class.

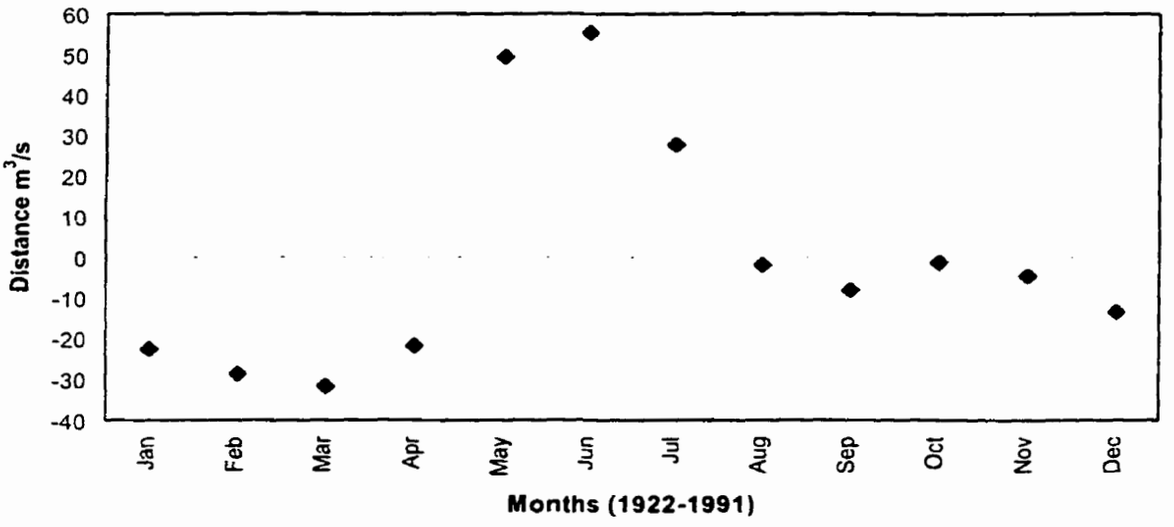



**Fig. 4.1 Monthly distance from the series mean flow of the English River.**

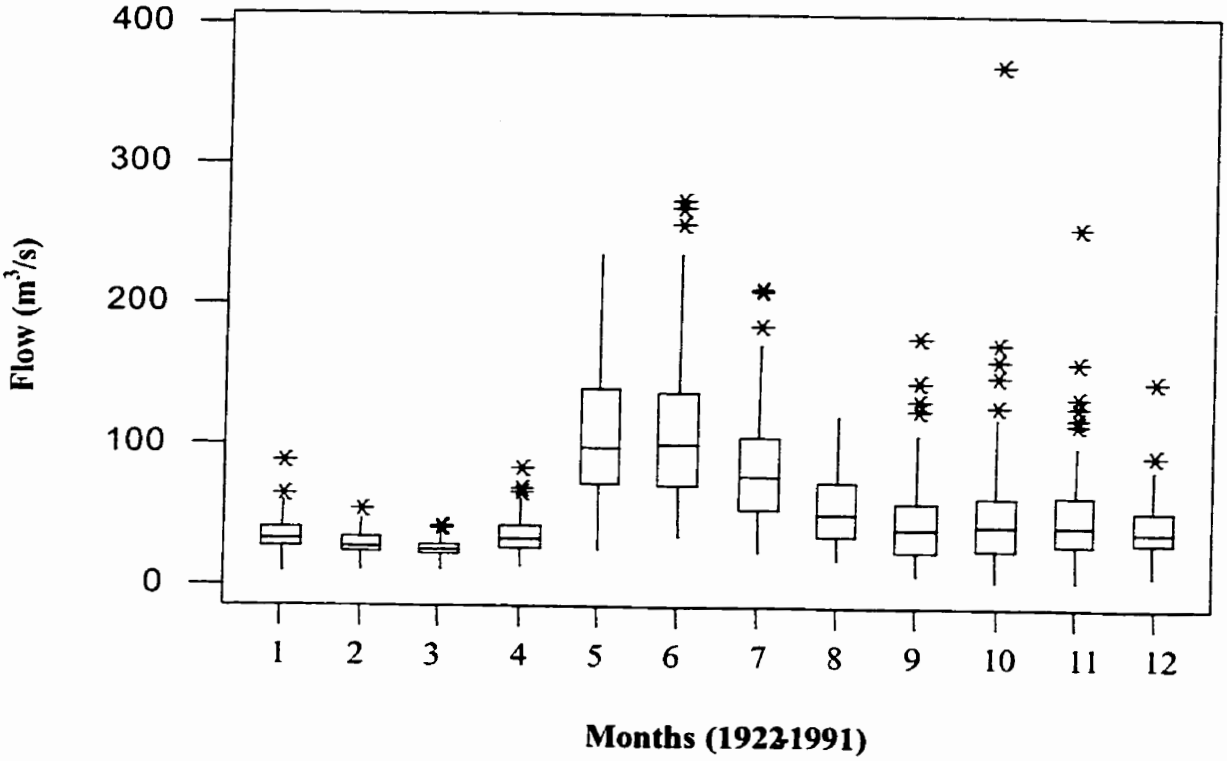
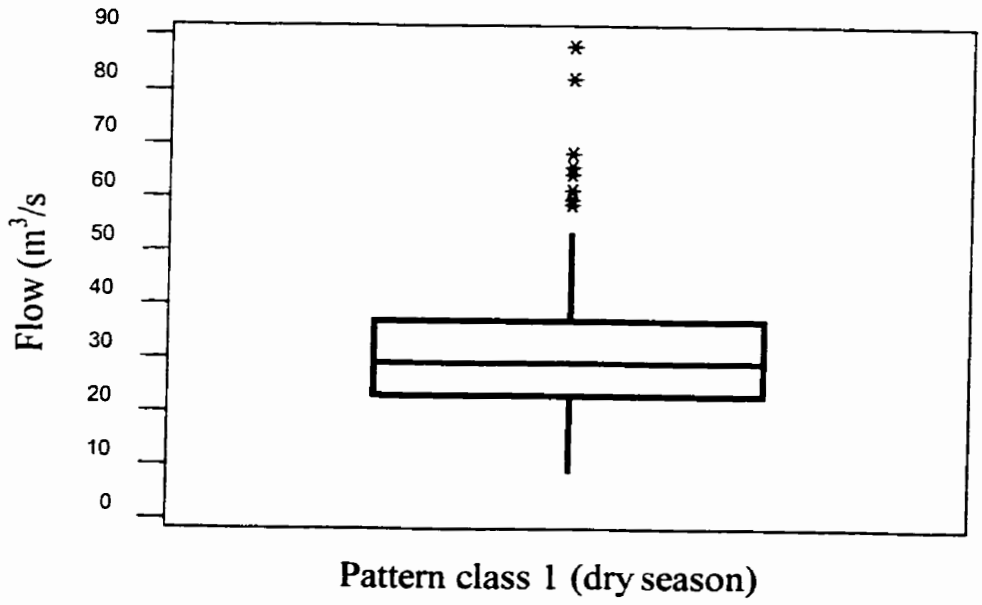
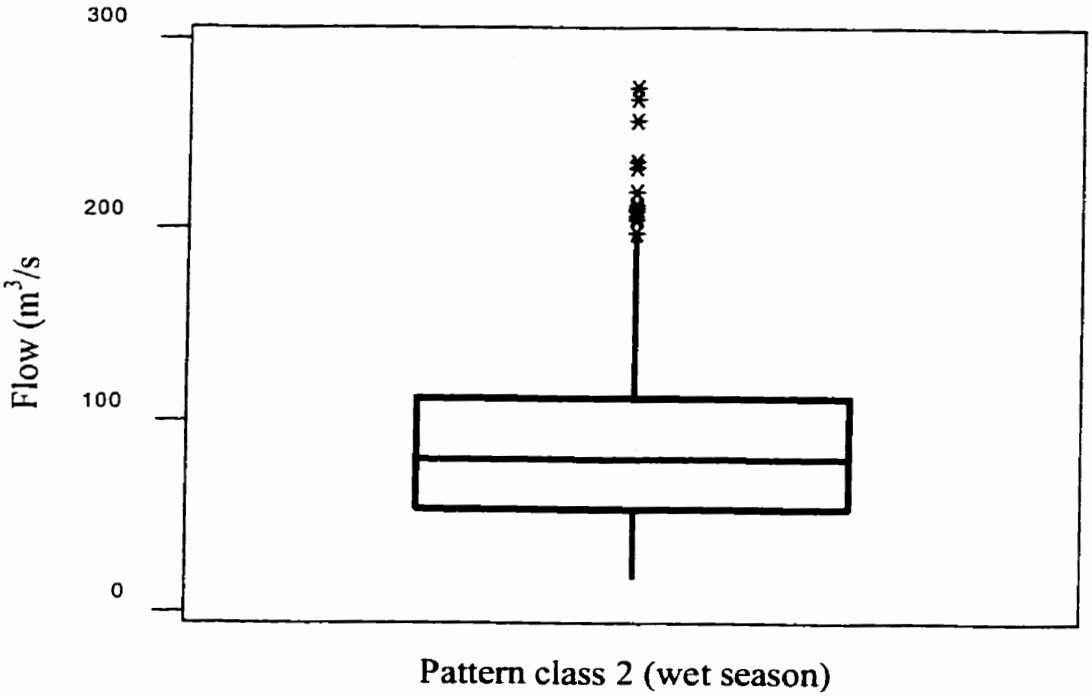


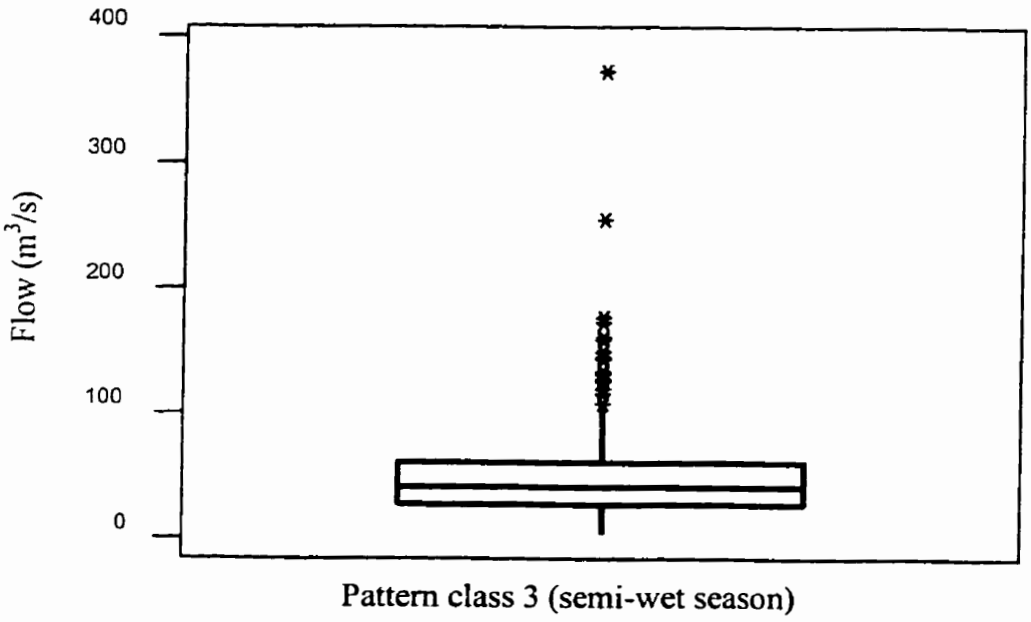

**Fig. 4.2 Box-plots of the monthly flows of the English River at Umferville.**




a)




b)




c)

**Fig. 4.3 Box-plot of the English River. a) pattern class 1, b) pattern class 2, and c) pattern class 3.**

The correlation matrix of the monthly flow data of the English River is depicted in Table 4.2. Moving horizontally from each month can give an idea about the trend of the correlation among months. Assuming a threshold correlation coefficient value of 0.5, below which the two variables (months) could be considered insignificantly correlated, the data can be clustered into classes. A vertical boundary can be drawn when the value before is more than 0.5 and the value after is less than 0.5. It is clear from Table 4.2 that a vertical boundary between April and May can be drawn. One can argue that the boundary can be drawn between March and April, but this can be refuted by presenting the difficulty of associating April with May because of their weak correlation. So April is better classified with March. The more difficult situation arises in the vicinity of August and September. Drawing the boundary after September means taking out from October, November, and December, a month that is significantly correlated with all of them. The line could be drawn after either July or August. In both cases there is no clear cut boundary, and the trade-off should be balanced and judged by the analyst. In this research, the boundary is considered after August to make the case easier by having three seasons of equal lengths (Elshorbagy et al. 2000b).

Two important observations can be derived from the correlation matrix. The first is that within each class lag-1 and lag-2 autocorrelation, for example, could exist while jumping from one class to the other might not be statistically justified (e.g., lag-1 correlation between April and May <0.5). Secondly, in case of applying the traditional ACF and PACF to the data, an AR(1) or AR(2) time series model might be recommended. It occurs due to the process of averaging the lag-k autocorrelation coefficients over the

**Table 4.2. Correlation matrix of the monthly flows of English river at Umferville.**

| Month | JAN | FEB | MAR | APR | MAY | JUN | JUL | AUG | SEP | OCT | NOV | DEC |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| JAN | 1.00 | 0.96 | 0.92 | 0.40 | 0.33 | 0.33 | 0.23 | 0.17 | 0.08 | 0.11 | 0.13 | 0.16 |
| FEB | | 1.00 | 0.98 | 0.48 | 0.43 | 0.42 | 0.29 | 0.20 | 0.07 | 0.12 | 0.11 | 0.16 |
| MAR | | | 1.00 | 0.56 | 0.47 | 0.41 | 0.29 | 0.18 | 0.08 | 0.13 | 0.12 | 0.18 |
| APR | | | | 1.00 | 0.47 | 0.14 | 0.14 | 0.07 | 0.10 | 0.04 | 0.12 | 0.15 |
| MAY | | | | | 1.00 | 0.63 | 0.53 | 0.37 | 0.05 | 0.03 | 0.18 | 0.22 |
| JUN | | | | | | 1.00 | 0.82 | 0.52 | 0.27 | 0.16 | 0.27 | 0.30 |
| JUL | | | | | | | 1.00 | 0.81 | 0.50 | 0.30 | 0.33 | 0.36 |
| AUG | | | | | | | | 1.00 | 0.68 | 0.35 | 0.29 | 0.33 |
| SEP | | | | | | | | | 1.00 | 0.74 | 0.54 | 0.54 |
| OCT | | | | | | | | | | 1.00 | 0.78 | 0.71 |
| NOV | | | | | | | | | | | 1.00 | 0.97 |
| DEC | | | | | | | | | | | | 1.00 |

entire data set. Therefore, care must be exercised in the case of applying traditional time series analysis, which ignores the concept of groups, on such a data set. It is believed that these two observations lend a strong support to the grouping approach proposed in this thesis. Finally, it is worth to mention that the use of the correlation matrix for segmenting the data is viewed as an auxiliary tool to verify the process.

The same procedure is repeated with the Little River and Reed Creek – except the correlation matrix (because of the weakness of the autocorrelations for both of the two rivers). Two pattern classes PC1 and PC2, each of six months duration, are identified for the two rivers. Table 4.3 summarizes the percentage of correct and mis-classified rivers segments. Both distance measures of the Euclidean and the Mahalanobis are used. Graphs showing the segments of the Little River and the Reed Creek are provided in appendix C.

It should be noted that the Mahalanobis distance includes the variance of the segments. If the variance is neglected, both measures, the Euclidean distance ($d_E$) and the Mahalanobis distance ($d_M$), should provide similar results. Deterministically speaking, a segment belonging to PC1 of the English River has low probability of being mis-classified while it is statistically merging with PC3. Opposite comments can be made with respect to PC3.

As a concluding remark for the process of data clustering or grouping, one can say that grouping is not only a necessary step for further analysis, but it also helps provide an insightful understanding into the data at hand. Searching the data for a structure of natural

groupings is an important explanatory technique. Grouping can provide an informal means for assessing dimensionality, identifying outliers, and suggesting interesting hypotheses concerning relationships (Johnson and Wichern, 1988).

**Table 4.3. Percentage of classification/misclassification using Euclidean and Mahalanobis distances for the segments of different classes of the three rivers.**

| Rivers | Class | Euclidean Distance ($d_E$) | | | Mahalanobis Distance ($d_M$) | | |
|---|---|---|---|---|---|---|---|
| | | PC1 | PC2 | PC3 | PC1 | PC2 | PC3 |
| English River | PC1 | 93% | 0% | 7% | 47% | 0% | 53% |
| | PC2 | 7% | 68% | 25% | 0% | 97% | 3% |
| | PC3 | 41% | 22% | 37% | 0% | 30% | 70% |
| Little River | PC1 | 69% | 31% | N/A | 87% | 13% | N/A |
| | PC2 | 15% | 85% | N/A | 7% | 93% | N/A |
| Reed Creek | PC1 | 84% | 16% | N/A | 100% | 0% | N/A |
| | PC2 | 2% | 98% | N/A | 28% | 72% | N/A |

*4.2.2 Identifying inter-group relationships towards model selection*

Having the data set segmented as consecutive groups, one should proceed with studying the inter-group structure. The traditional or single-valued approach is based on understanding the time series as a sequence of individual observations. Therefore, identification of the structure linking the individual observations is required to model the

time series in the single-valued approach (whether in regression or time series analysis). In the proposed approach, the model entities are groups of observations rather than individuals, and the links sought are supposed to represent the relations among vectors rather than single-valued observations. Characteristics such as trend, seasonality, independence, autocorrelation and cross-correlation coefficients, and normality or multivariate normality should be checked in relation to data groups.

*Autocorrelation and cross-correlation*

Autocorrelation and cross-correlation coefficients among groups are ideally represented in a form of (k*k) matrix where k is the number of elements in each segment. For English River, the autocorrelation coefficient of any lag is (4*4) matrix while auto and cross-correlation coefficients are represented by (6*6) matrices in cases of Little River and Reed Creek. Such matrices are difficult to interpret and cannot be represented on 2-dimensional plots as autocorrelation function (ACF) and partial autocorrelation function (PACF). For the purpose of this research, it is considered that a mean value of each segment can satisfactorily represent that subtle characteristic of the segment in studying the correlation structure. The ACF and PACF of the English River data groups are respectively plotted as Figures 4.4a and 4.4b. The slow decay of the ACF and the cutoff in the PACF after lag-1 demonstrate that a lag-1 autoregressive group, ARG(1), model can be successful in modeling the English River data as groups. The independence of the Little River and Reed Creek data, represented by the ACFs in Figures 4.5 and 4.6, as well as the significant group cross-correlation (Figure 4.7) refutes the utility of time series analysis in this case (bi-series case). It should be noted that the lag-1 autocorrelation may

look significant according to Figure 4.5 but there is no gradual decay that indicates the existence of serial dependence. Regression technique lends itself to studying the Little River and Reed Creek as a bi-series case. Two other techniques, ANNs and MPM, are used along with regression and time series analysis for comparison purposes. Information about autocorrelation or independence is not required for application of ANN technique. The details of calculating the autocorrelation and cross-correlation of groups are given in appendix D.

*Deterministic components*

Trends, shifts, and cyclicity in the mean and standard deviation can be removed before modeling the data groups. Perceiving the data as groups does not affect the definition of these deterministic components. Removing these components can be done in the same way as in the traditional single-valued approach. Details of such procedures are provided by Salas (1992).

*Multivariate normality*

It should be noted that time series analysis and MPM techniques require data to follow multivariate normal distribution given in equation (4.1),

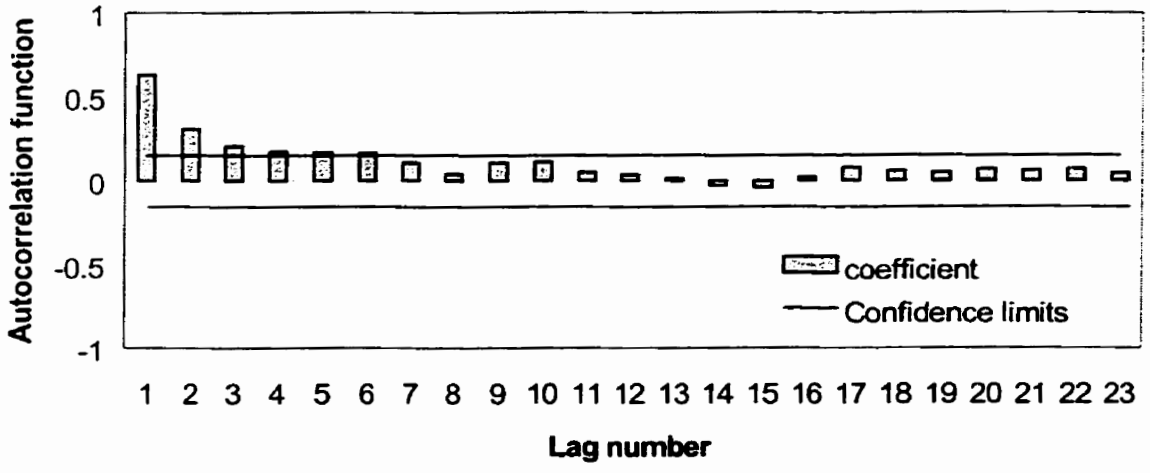
$$f(X)=\frac{1}{(2\pi)^{k/2}|C|^{1/2}}e^{-1/2(X-\underline{\mu})'C^{-1}(X-\underline{\mu})} \tag{4.1}$$

where k is the dimension of the group, which is equal to the number of elements in each segment, and C is the covariance matrix of the pattern class. In regression technique, the

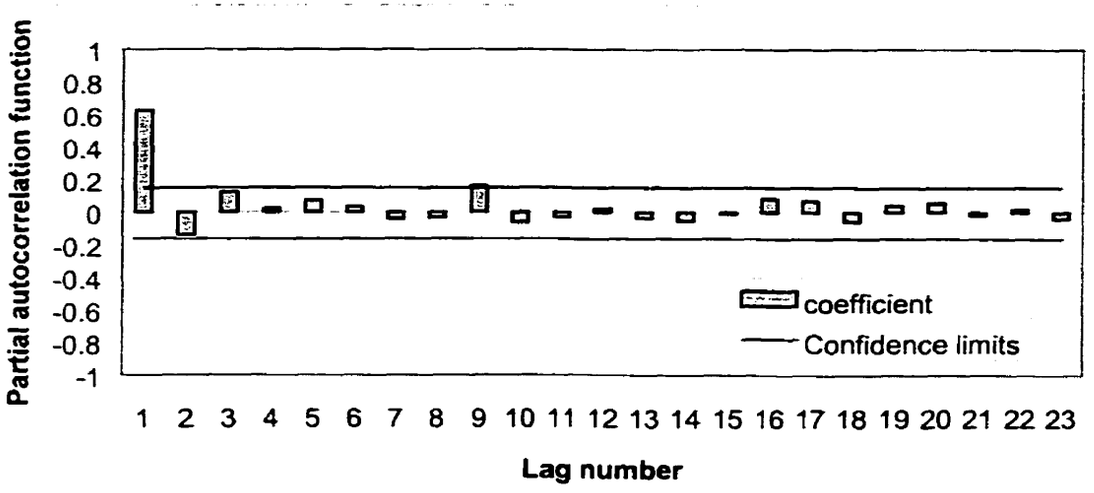
residuals have to be normally distributed while the variables themselves need not follow normal distribution (Makhuvha et al. 1997a). To check the multivariate normality, the following relation should hold for a k-dimensional normal distribution:

$$\left(X-\underline{\mu}\right)' C^{-1}\left(X-\underline{\mu}\right) \leq \chi_k^2(\alpha) \tag{4.2}$$

where $\chi_k^2(\alpha)$ is upper (100 )th percentile of a chi-square distribution with k degrees of freedom (Johnson and Wichern, 1988). For the three rivers involved in this research, it is found that log-transformation is sufficient to bring the data close to multivariate normal distribution. Other methods of transforming the data to multivariate normal may lead to better transformation and improvement of final results, but only log-transformation is applied in this research. Normality tests are shown in appendix E.




a)




b)

**Fig. 4.4 Correlation structure of data groups (English River). a) autocorrelation function and b) partial autocorrelation function.**

*4.2.3 Application to the single series case (English River)*

In the following, three techniques for estimating missing data, namely, time series modeling, multivariate partitioning modeling, and artificial neural networks are presented.

*a) Time series autoregressive groups ARG(p) model*

Analyzing the ACF and PACF of the data groups reveals that ARG(1) is worth testing. In general AR(1) and AR(2) are the most commonly used models of the AR(p) type models in hydrology. Needless to say, perceiving the data as groups can further weaken the autocorrelation links among data constituents (groups rather than single-valued observations) due to the local autocorrelation within each group. ARG(1) or even

independent groups type of modeling might be the most dominant structure in modeling monthly streamflows using the group approach. The model represented by equations (3.8) and (3.9) is used with the calibration data to estimate the model parameters. The parameter matrices $A_1$ and B are estimated as follows (Salas et al. 1980).

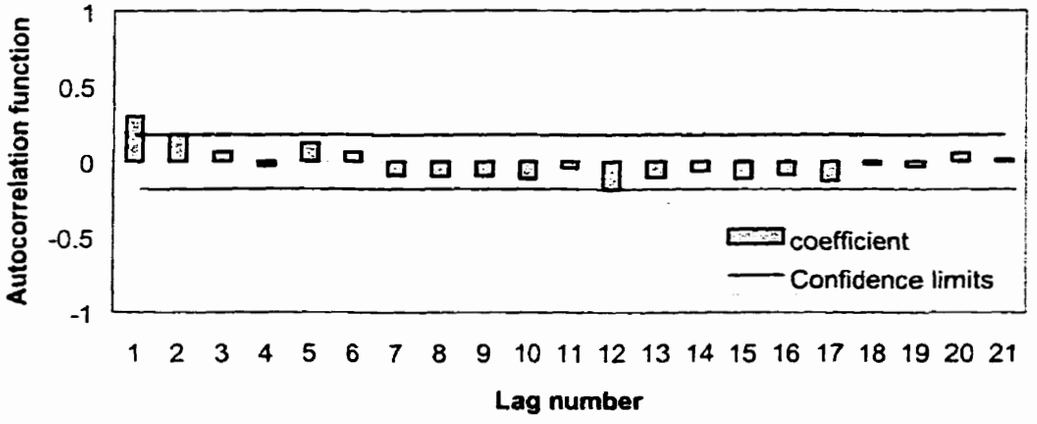



**Fig. 4.5 Autocorrelation function of data groups (Little River).**

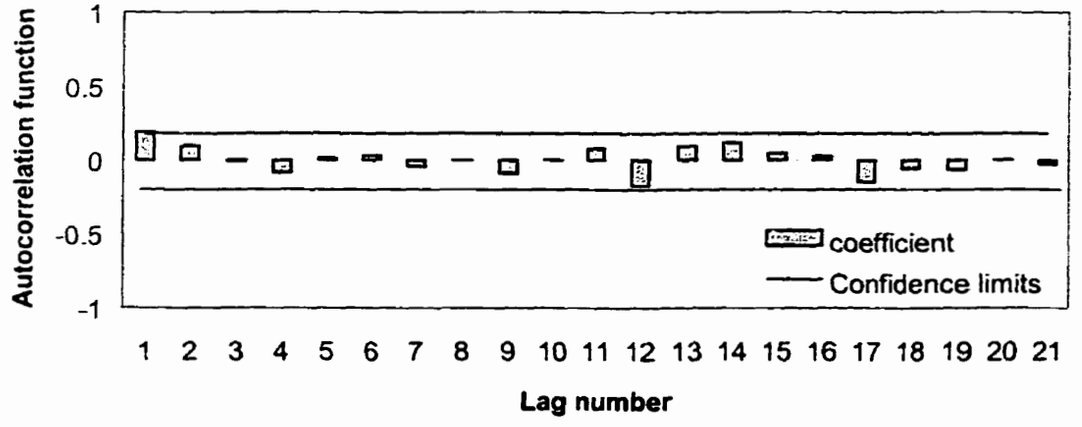



**Fig. 4.6 Autocorrelation function of data groups (Reed Creek).**

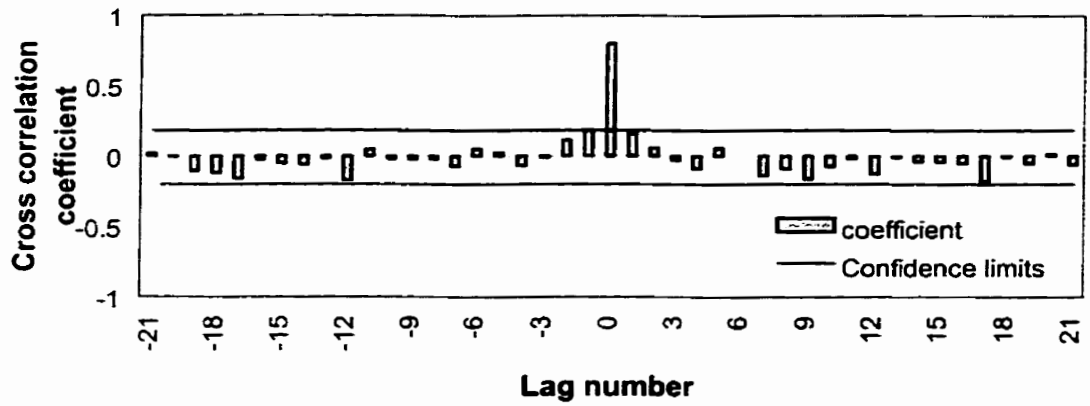



**Fig. 4.7 Cross-correlation function of data groups (Little & Reed Creek)**

$$\hat{A}_1 = \hat{M}_1 \hat{M}_0^{-1} \tag{4.3}$$

$$\hat{B}\,\hat{B}^T = \hat{M}_0 - \hat{A}_1 \hat{M}_1^{-1} \tag{4.4}$$

$$\hat{M}_k = E\left[Z_t\, Z_{t-k}^T\right] = \begin{bmatrix} \rho_k^{11}, \rho_k^{12}, \ldots\ldots\ldots\ldots \rho_k^{1n} \\ .. \\ .. \\ .. \\ \rho_k^{n1}, \qquad\qquad \rho_k^{nn} \end{bmatrix} \tag{4.5}$$

where T denotes the transpose of the matrix, $\rho_k^{ij}$ = corr $[Z_t^{(i)}, Z_{t-k}^{(j)}]$ is the lag-k cross-correlation coefficient between $Z_t^{(i)}$ and $Z_{t-k}^{(j)}$ for $i \neq j$. As a model verification step, the random vector of the English River is checked. It is uncorrelated in time and space, that is, $E[\varepsilon_t^{(i)}\varepsilon_t^{(j)}]=0$ for $i \neq j$.

*b) Periodic autoregressive groups PARG(1) model*

To check the influence of the periodicity of the autocorrelation among the three classes of the English River data, a periodic model given in equation (4.6) is employed,

$$Z_{\nu,\tau} = A_{1,\tau}\, Z_{\nu,\tau-1} + B_{\tau}\, \underline{\varepsilon}_{\nu,\tau} \tag{4.6}$$

where three (4*4) periodic coefficient matrices represented by $A_{1,\tau}$ and three (4*4) matrices represented by $B_{\tau}$ are estimated. Model verification procedures similar to those mentioned earlier are conducted. Appendix F shows details of developing ARG(1) and PARG(1) models.

*c) Multivariate Partitioning modeling (MPM) technique*

The form of MPM given in equation (3.10) is rewritten to handle the case of three classes for application to the English River. The mean values of classes 1, 2, and 3 are given below.

$$Mean\,1 = \underline{\mu}^{(1)} + C_{13}\, C_{33}^{-1} \left( X3 - \underline{\mu}^{(3)} \right) \tag{4.7}$$

$$Mean\,2 = \underline{\mu}^{(2)} + C_{21}\, C_{11}^{-1} \left( X1 - \underline{\mu}^{(1)} \right) \tag{4.8}$$

$$Mean\,3 = \underline{\mu}^{(3)} + C_{32}\, C_{22}^{-1} \left( X2 - \underline{\mu}^{(2)} \right) \tag{4.9}$$

where

$$X = \begin{bmatrix} X1 \\ X2 \\ X3 \end{bmatrix} \quad \text{and} \quad C = \begin{bmatrix} C_{11} & C_{12} & C_{13} \\ C_{21} & C_{22} & C_{23} \\ C_{31} & C_{32} & C_{33} \end{bmatrix}$$

For simplicity, each equation of the three above-mentioned equations is considered representing the mean of the conditional distribution of a specific class, given the values of the preceding class and independently of the third class (Elshorbagy et al 2000b). These equations are used to estimate the missing segments of the English River data at Umferville.

*d) Artificial neural network (ANN) technique*

Multilayer feed forward ANN with back propagation learning algorithm is applied to the English River data. After many trials, network configuration that estimates the missing data in the best way is ANN(7-6-4) with seven input nodes, six hidden nodes, and four output nodes. The reason for having three extra input nodes is the use of the monthly mean, annual mean, and monthly variance as inputs. The trials showed that these additional inputs help accelerate the convergence process of ANNs. One can easily notice that the three additional inputs represent three components that are recognized during the pre-analysis of the time series. These components are the trend, monthly average value, and monthly variance. A lag-1 transition model is assumed such that any segment of four observations is estimated based on the previous segment. Therefore, the data are arranged in a way that makes a (180*7) matrix of the input data and a (180*4) matrix of the output data of the ANN model. The 180 rows represent 60 years of data with three segments each.

The estimated values of the English River corresponding to the missing data using the four models are given in Figure 4.8 along with the observed values for comparison. It should be noted that the estimated values are calculated as segments of four consecutive values. Each segment is estimated in one step at a time (i.e., the estimated value within the segment is not used for estimating the successive one).

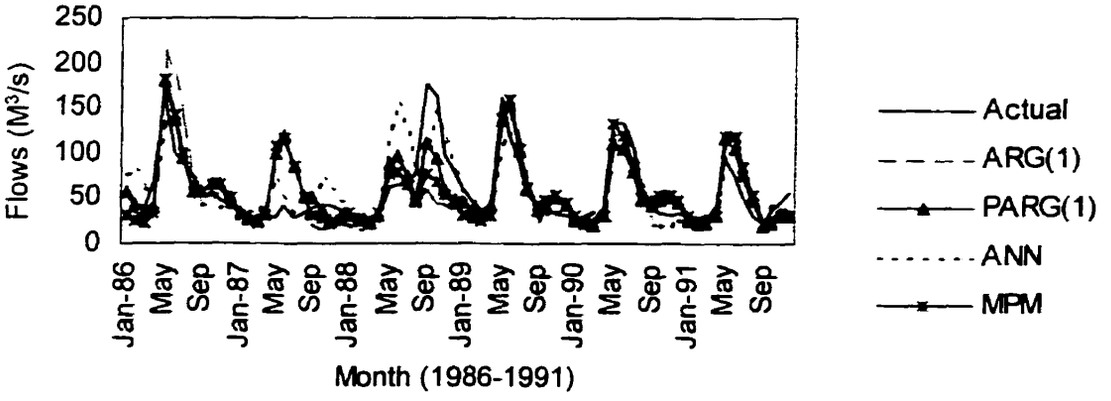



**Fig. 4.8 Estimated and observed values of the English River flows.**

*4.2.4 Application to the bi-series case*

In the following, three techniques, namely, multivariate multiple regression, multivariate partitioning modeling, and artificial neural networks are presented.

*a) Multivariate multiple regression (MMR) technique*

The model equation and the parameter estimation technique are given in the previous chapter. Y is a (102*6) data matrix of the *target river* (Little River), and Z is a (102*6)

matrix of the *reference river* (Reed Creek). The 102 rows represent 51 years of data with two segments each. The model is verified by checking the independence of the errors and their distribution. The errors are found to be independent and their distribution is close to normal.

*b) MPM technique*

Equation (3.10) is ideal for the bi-series case where X2 is the segment known from the *reference river* (Reed Creek) and X1 is the segment to be estimated for the *target river* (Little River). The data of both rivers are arranged in a way that collates the data from each of the two concurrent classes of both rivers in one matrix. The (102*12) data matrix can be used to calculate the main covariance matrix, C. The constituent components of this matrix are $C_{11}$, $C_{12}$, $C_{21}$, and $C_{22}$.

*c) ANN technique*

Data in 102 rows are arranged such that there are six inputs to the network and six outputs. The six inputs are the number of elements in each class of the Reed Creek while the six outputs are the elements in each class of the Little River. It should be noted that the 102 refers to the number of data rows used for training the ANN model. ANN (6-8-6) is found to provide the best estimate of the missing segments for the Little River.

The estimated values of the Little River missing data using the three models, for the bi-series case, are shown in Figure 4.9 along with the observed values. It should be noted

that estimated values are calculated as segments of six consecutive values (one segment at a time) in one step rather than one by one single-valued outputs.

*4.2.5 Estimation of partially missing segment*

Sometimes, one might face the case of having only part of a segment missing. Due to adoption of the group approach in this research, each segment is considered as a single object. Thus, a segment is estimated as an inseparable set of consecutive values. When missing values constitute only a part of a segment, a solution should be found in a way that integrates explicitly the information available from the known part of the segment. In this research, it is assumed that the first value of the missing segment is known and the rest are to be estimated.

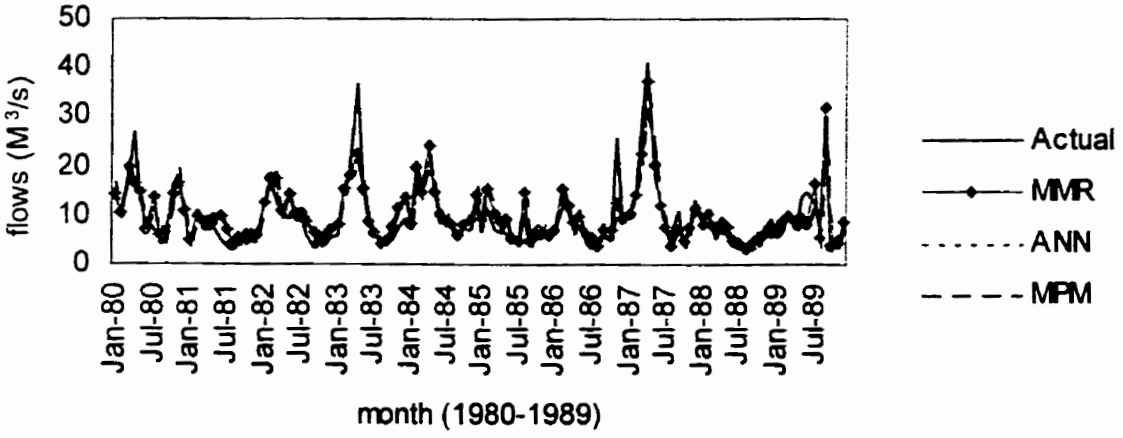



**Fig. 4.9 Estimated and observed values of the Little River flows.**

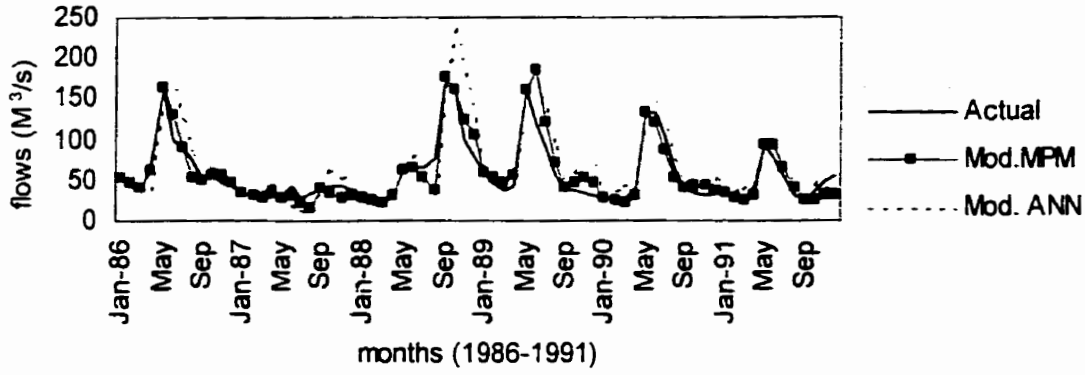



**Fig. 4.10 Modified estimates of the missing segments (English River).**

In other words, only the last three and the last five observations out of the segments are missing in the English River and Little River respectively. Similarly, other varieties of partial segments are possible, but for demonstration purposes, only the first value of the segment is assumed to be known.

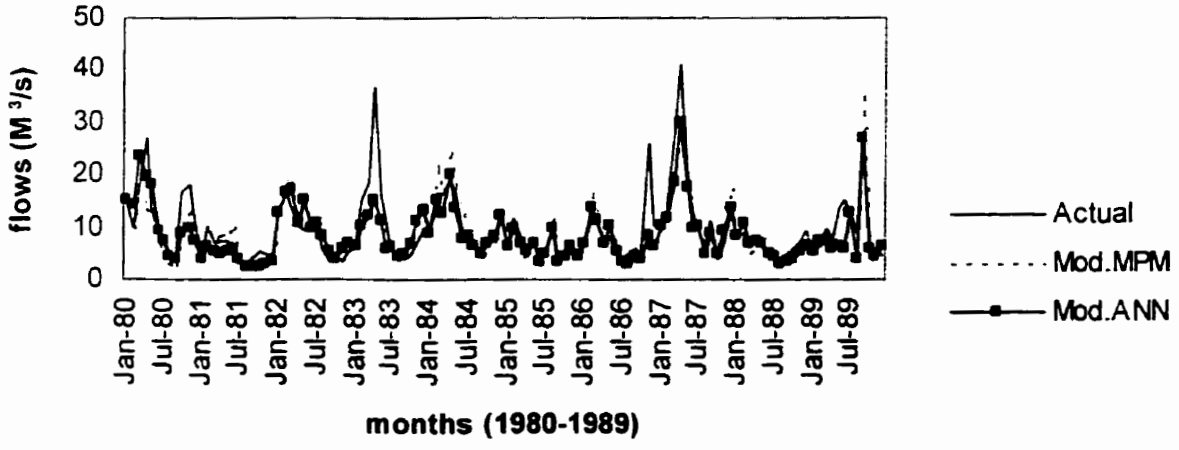



**Fig. 4.11 Modified estimates of the missing segments (Little River).**

Keeping the modeling procedures unchanged, first, the whole segment (including the first known observation) is assumed missing and estimated in one step as done before. Second, the percent error of the first value is calculated as follows.

$$p_e = \frac{d_{el} - d_{kl}}{d_{kl}} \tag{4.10}$$

Where, $d_{el}$ is the estimated value of the first observation of the segment, $d_{kl}$ is the known value of the first observation. The third step is modifying the rest of the estimated values within each segment according to the ($p_e$) of the first known value using the following equation.

$$d_{mei} = (1 - p_e) * d_{ei} \tag{4.11}$$

Where, $d_{mei}$ is the modified value of any estimated value in the missing segment, and $d_{ei}$ is the estimated value of any observation after the first one in the missing segment. Figures 4.10 and 4.11 show the modified estimated values of the English River and Little River assuming that the first observation of the missing segment is known.

## 4.3 Indicators of model performance

A mean squared error (MSE) is one of the most commonly used performance measures in hydrological modeling. Many researchers used MSE or its root (RMSE) as an accuracy

measure, (e.g., Bastarache et al. 1997; Shamseldin, 1997; and Carpenter and Bartheleny, 1994). Others, like Karunanithi et al. (1994), use both MSE and mean relative error (MRE) to try to fill some of the gaps left by considering only MSE, and state that the squared error and the relative error provide different types of information about model predictive capabilities. Formulas for calculating MSE and MRE are given below:

$$MSE = \frac{1}{n}\sum_{i=1}^{n}\left(\hat{D}_i - D_i\right)^2 \tag{4.12}$$

$$MRE = \frac{1}{n}\sum_{i=1}^{n}\left|\frac{\left(\hat{D}_i - D_i\right)}{D_i}\right| \tag{4.13}$$

where $D_i$ stands for the measured (actual) value, $\hat{D}_i$ is the estimated value, and n is the number of observations. Using the MRE to rank different models according to their performance may result in a different rank from that obtained using MSE. To be more comprehensive in the evaluation of different aspects of model performance, Hsu et al. (1995), use RMSE, percent error in volume (% VE), percent error in matching the maximum flow (% MF), and the correlation (CORR) between the observed and simulated flows. Those who use only the MSE to assess the performance of different models fail to prove the appropriateness of MSE as a sole measure of model performance. Meanwhile, others who use more than one measure do not provide guidelines for choosing the model when different measures favor different models. In some cases, preference of one model to another is based on a small difference in the MSE. The MSE is an absolute value and not referenced to any benchmark point, nor comprehended according to a pre-designed

threshold to indicate the significance of the MSE value or even the significance of the difference between two calculated MSEs of different models. It is apparent that a statistical measure capable of overcoming the above-discussed difficulties is required.

## 4.4 The pooled mean squared error (PMSE)

Assuming a common hydrological problem of predicting streamflow using two different models A and B, a different statistic is proposed to achieve the following objectives: (a) to combine the effect of the MSE and MRE in one measure, called pooled mean squared error (PMSE); and (b) to increase the effect of the significant residuals which are above a pre-determined threshold, and to reduce the differential effect of the insignificant residuals which are below the pre-determined threshold (Elshorbagy et al. 2000c).

A hypothetical numerical example is provided to show the steps followed in the development of the PMSE. Table 4.4 shows the actual, estimated flows, residuals, and relative errors for the hypothetical example. The steps are as follows:

1. A threshold is determined and only the residuals above its value are considered significant. A threshold can be any value, but this research considers 5% of the actual values.
2. Relative error values of the outputs from the two models are pooled together to form one set of size N (N = 20 in this hypothetical example, $n_1 = n_2 = N/2$).
3. If the relative error is $\leq 5\%$, the predicted value is given a constant rank. In our example, predicted values with residuals less than or equal to 5% of the actual

values are ranked (1).

4. Predicted values with relative errors higher than the threshold 5% are ranked so that the highest residual is given the rank (N), the next highest is given the rank (N-1) and so on.

5. The PMSE for each model is calculated using the following relationship:

$$PMSE = \frac{\sum_{i=1}^{N/2} \left(\hat{D}_i - D_i\right)^2 \cdot k_i}{\sum_{i=1}^{N} k_i} \qquad (4.14)$$

where $k_i$ is the rank of the residual error . The ideal value of MSE, MRE, and PMSE is 0.0. In our example MSE of model A = 25 and MRE (A) equal 0.048 while MSE(B) = 20.8 and the MRE(B) = 0.0281. One can favor model B over model A while the values in Table 4.5 clearly show that all predicted values of model A can be accepted because they lie within the pre-defined threshold. Two values out of ten (20%) predicted by model B are in clear violation of the threshold, which indicates low reliability of model B. It can be concluded that neither of the two regular measures able to interpret or explain the results of our example. On the other hand, calculating the PMSE results in values of 9.72 and 43.42 for models A and B respectively. These values are more reasonable according to the concept of satisfying or violating the threshold. It should be noted that other authors establish measures of errors for different purposes. The closest example is Sorooshian et al. (1983 and 1993), who use a maximum likelihood estimator function for the heteroscedastic error case (HMLE). The main point of their approach weighs the

squared residuals by the actual observations. In other words, they implicitly combine the squared error and the relative error.

**Table 4.4 Actual and predicted flows, residuals, and relative errors (a hypothetical case)**

| Actual flow | Model (A) | | | | Model (B) | | | |
|---|---|---|---|---|---|---|---|---|
| | Predicted Flow | Residual values | Relative error | | Predicted Flow | Residual values | Relative error | |
| 100 | 105 | 5 | 0.05 | (1) | 101 | 1 | 0.01 | (1) |
| 110 | 115 | 5 | 0.045 | (1) | 109 | -1 | 0.009 | (1) |
| 90 | 85 | -5 | 0.056 | (18) | 100 | 10 | 0.111 | (20) |
| 100 | 105 | 5 | 0.05 | (1) | 101 | 1 | 0.01 | (1) |
| 105 | 110 | 5 | 0.048 | (1) | 104 | -1 | 0.0095 | (1) |
| 113 | 108 | -5 | 0.044 | (1) | 114 | 1 | 0.0088 | (1) |
| 95 | 100 | 5 | 0.053 | (17) | 96 | 1 | 0.0105 | (1) |
| 120 | 125 | 5 | 0.043 | (1) | 119 | -1 | 0.0083 | (1) |
| 105 | 100 | -5 | 0.048 | (1) | 115 | 10 | 0.095 | (19) |
| 117 | 122 | 5 | 0.043 | (1) | 118 | 1 | 0.0085 | (1) |

*** : The values in parenthesis are the pooled rank**

## 4.5 Discussion

The results of the single series case, presented in Table 4.5, reflect the superiority of the ANN model in estimating the missing values when the whole segment is missing. On the other hand, the ANN model becomes inferior to the rest of the models when the first observation of the missing segment becomes known. The results of the ANN model do

not improve with the acquisition of additional information (first observation in the segment). This is because the modification process conducted using equations 4.10 and 4.11 is based on the group concept, and the whole segment is assumed to have consistent signs (+ve or -ve) of the relative errors for all elements within the segment. It seems that the ANN model, during its search to minimize the squared error, tends to violate the group-structure of the data. Perhaps it treats each element within the segment differently. Therefore, the apparent superiority of the ANN model when the whole segment is missing should not be generalized to include all cases of missing data (Elshorbagy et al. 2000b).

Similar to Table 4.5, Table 4.6 summarizes the results for the bi-series case (the Little River). The ANN model tends to provide better estimates for the missing values when the whole segment is missing. The last three columns of Table 4.6 indicate that none of the adopted techniques is able to benefit from the availability of the first element of the missing segment. It is noted that the differences between the MSE values of different models in Table 4.6 are not big; however, the less the MSE value, the better the model performance. Also the use of PMSE helps address this problem. Keeping in mind that the three techniques follow different routes to approach the problem solution, the utility of the grouping becomes questionable in this bi-series case. One explanation, supported by the relative success of the group approach in the English River case and by the improper results in the Little River case, is the role of the autocorrelation. The idea of tying each group of elements in one segment to behave as a single object seems to require significant autocorrelation among the single-valued elements. Therefore, the significant

autocorrelation of the English River, unlike that of the Little River, plays a major role in supporting the group approach.

Another point of interest is worth mentioning in light of that modification process described by equations 4.10 and 4.11. The significant improvement in the accuracy of estimating the values of the partially-missing segments of the English River diverts attention toward the utility of the group approach for forecasting purposes.

**Table 4.5 Summary of model performance for the English River case.**

| Models | Whole missing segment | | | Partial missing segment | | |
|---|---|---|---|---|---|---|
| | MSE | MRE | PMSE | MSE | MRE | PMSE |
| ARG(1) | 988 | 0.345 | 457 | 247 | 0.154 | 138 |
| PARG(1) | 569 | 0.315 | 235 | 183 | 0.152 | 100 |
| MPM | 773 | 0.358 | 347 | 201 | 0.157 | 109 |
| ANN | 457 | 0.25 | 195 | 468 | 0.241 | 281 |

**Table 4.6 Summary of model performance for the Little River case.**

| Models | Whole missing segment | | | Partial missing segment | | |
|---|---|---|---|---|---|---|
| | MSE | MRE | PMSE | MSE | MRE | PMSE |
| MMR | 9.4 | 0.248 | 5.6 | 14.1 | 0.213 | 9.5 |
| MPM | 10.5 | 0.210 | 6.5 | 15.5 | 0.198 | 10.5 |
| ANN | 8.7 | 0.207 | 5.2 | 15.2 | 0.209 | 10.2 |

A complete segment of four elements (monthly values) can be forecasted as a first step. Then, when the first value becomes available, a significant improvement in the following three monthly values can be achieved without remodeling or re-estimation of the model parameters. This finding not only saves the time and effort of re-estimating model parameters, but also provides satisfactory estimates for multiple steps ahead (3-steps in our example).

## 4.6 Concluding remarks

This chapter demonstrates that monthly (periodic) streamflow data can be clustered into a definite number of classes according to the group approach. The flow values and their variances are reasonable characteristics that can be used for data segmentation. The results of the application of different techniques for estimation of missing segments indicate that the autocorrelation among elements of each segment plays a major role in the success of the group approach. In both, single series and bi-series cases, artificial

neural network model shows superiority over other models, applied in this research, in estimating missing segments using the group approach. However, care must be exercised when ANN models are applied to cases of partially-missing segments. ANNs tend to violate the group structure of the data to minimize the overall errors of the estimates.

Other techniques represented by group time series models (ARG(1) and PARG(1)) and MPM are more consistent in handling data groups. Their accuracy of missing data estimation improves when a part of the missing segment becomes available. The discussion given earlier on the utility of the group approach for forecasting applications indicates that the group approach may have potential for other applications in water resources. The success of the group approach is contingent on the segmenting process and the selection of the models that can handle groups. Segmentation procedures and models used in this thesis are not claimed to be exhaustive. Improvement in these two aspects of the group approach reflects on the final results of the application. Finally, the group approach is not proposed as a replacement for the traditional single-valued approach but rather as a sound alternative when the hydrologic data show clustering characteristics and as a strong candidate when consecutive (group of) observations are missing (Elshorbagy et al. 2000b).

In this chapter, the concept of the group approach has been tested and applied to monthly streamflow records as representatives of periodic hydrologic data. A definite number of groups (seasons), with definite and small number of elements (months), in the geophysical year have been identified. The processes of segmenting the data into groups

and modeling the group structure is mainly based on the fact that the time series shows a significant periodicity. Thus, the applicability of the same approach and methodology to hydrologic data with different time scale (e.g., weekly or daily flows) is not certain. The daily streamflows are considered as an example of non-periodic hydrologic data. The characteristics of the daily hydrologic data might pose the necessity for adopting different concepts, techniques, and methodologies to perceive the data structure and model the time series accordingly. In the following chapter, the concept of the chaos theory is investigated and its suitability to explain the behavior of the daily streamflows is explored.

# Chapter 5

# 5. PERCEPTION OF GROUPS IN NON-PERIODIC DATA: THE CHAOS APPROACH

5.1 Introduction

5.2 What is chaos?

5.3 Review of studies investigating the concepts of chaos in hydrology

5.4 Estimation of missing streamflow data

5.5 Summary

---

*"The Fools among the people will say: What hath turned them from the direction to which they were used? Say: To Allah belong both East and West. He guideth whom He wills to a way that is straight"*

*The Glorious Qura'n (2: 142)*

## 5.1 Introduction

The methodology explained in the previous chapters, utilizing the information contained in groups of periodic data might not be feasible for non-periodic data. The daily streamflows are considered in this thesis as an example of non-periodic hydrologic data. The lack of the apparent deterministic component of periodicity, within the year, in the daily streamflows makes the task of segmenting the data, into a definite number of repeating groups, difficult. If the segmentation procedures, presented earlier, were adopted for chopping daily flows, one might end up with one of the following results: (i) three or four group types for every geophysical year with 90 or 100 elements for each group or (ii) large number of group types (e.g., 90 or 100 groups) with three or four elements each. Both results pose difficulty and impracticality due to the unmanageable increase in the number of parameters and models needed for the data. Therefore, another way of looking for group structure, or that window encompassing observations of similar characteristics, in the daily data should be sought.

On the other hand, recent studies have indicated the existence of nonlinear deterministic dynamics in some hydrological data (e.g., among others, Rodriguez-Iturbe et al. 1989). Such a characteristic of the non-periodic data diverts the attention from regular stochastic modeling towards the principles of chaos theory. In chaos theory, as will be shown later, a specific data structure is searched through identifying some chaotic invariants. One of these invariants is a dimension of a certain sub-space of the data, which indicates the width of a window entailing a specific number of observations. Even though the word

group or grouping is not invoked in chaos-related literature, such type of structure is envisioned in this research as a support for the group approach adopted in this thesis. The basic principles and definitions of chaos theory are given hereafter.

## 5.2 What is chaos?

As there is no commonly accepted definition of chaos, one can only characterize chaos as a seemingly random and irregular signal generated by deterministic process with some additional properties (Embrechts, 1994). The first property is that the deterministic process that generates a chaotic signal must show sensitivity to initial conditions. A second property is that a chaotic signal should be associated with a strange attractor (i.e., that the phase space shows a figure or distribution of points characterized by a fractal or noninteger dimension). In order to understand such expressions like strange attractor and phase space, we need to know what a fractal is.

*(a) What is a fractal?*

Benoit Mandelbrot introduced fractals in the 1970s to describe irregular shapes, such as mountains, clouds, wiggly lines, and coagulations of points. Since then, several books have been published that carefully define what exactly constitutes a fractal (Barnsley, 1988; Mandelbrot, 1983; Feder, 1989; and Peters 1991, 1994).

Geometric structures such as points, lines, and surfaces can be categorized according to their topological dimension as follows: 0 for a point, 1 for a line, and 2 for a surface.

Shapes such as wiggly lines or circles have a topological dimension of 1, because they can be stretched into a straight line. Mathematicians have defined several other dimensions. One type is the Hausdorff dimension, which has an interesting characteristic in that it can also take noninteger, or fractional, values. The figures that Mandelbrot studied can be assigned such noninteger Hausdorff dimensions. Fractals are geometric figures with a noninteger Hausdorff dimension. A fractal is a set for which the Hausdorff dimension strictly exceeds the topological dimension (Mandelbrot, 1983). The Hausdorff dimension can be approximated by the fractal dimension or correlation dimension.

A fractal dimension is a measure of how much an m-dimensional space is occupied by an object. For example, as given by Embrechts (1994): "*a cotton ball has three physical dimensions. If one squeezes it with a press, it tries to occupy two dimensions but fails. It still has bulk in the third dimension. So the flattened cotton ball has a dimension somewhere between 2 and 3. Similarly, a point has one dimension, and a straight line has two dimensions. A wiggly line, which is a straight line squeezed together, has a dimension between 1 and 2.*"

*(b) Methods for computing fractal dimensions*

There are a few distinct methods for computing fractal dimensions, one of which is the correlation analysis. To estimate the fractal dimension of a time series, the concept of the correlation dimension is useful and often applied (Kantz and Schreiber, 1997). Before computing the correlation dimension, some definitions, regarding the chaotic behavior are given.

*(c) Definitions*

Chaotic time series analysis determines whether an irregular and seemingly random set of data points resulted from a chaotic process. A set of discrete data points that represent values of a signal as a function of time is a time series. A chaotic time series, which represents a dynamic system, can be described by a phase space diagram that describes the evolution of the system from some initial state. The chaotic time series has a phase space trajectory that converges to a sub-space regardless of the initial conditions. This sub-space is called an "attractor". An attractor can lie in an m-dimensional phase space and can be multi-dimensional with dimensions less than m. Deterministic systems for which long term predictability is possible have attractors of integer dimension (Embrechts, 1994). When the dynamic system is sensitive to initial conditions, the attractors have non-integer or "fractal" dimensions. Such attractors are called "strange attractors" and systems containing them are called "chaotic dynamic systems" (Jayawardena and Lai, 1994). So, a strange attractor is nothing more than a phase space trajectory that stays on a fractal set. Chaotic time series analysis finds the dimension of this attractor. The higher the dimension, the more complex the underlying process that generated the time series. An example of a chaotic system is given in appendix G.

In principle, a chaotic system can be modeled by a number of nonlinear first-order differential equations. The minimum number of differential equations is equal to the integer that embeds (i.e., is just larger than) the fractal dimension. The dimension of the phase space that spans this minimal number of differential equations is called the

embedding dimension. The embedding dimension will always be larger than the fractal dimension of the strange attractor.

*5.2.1 Reconstruction of the phase-space*

The first step in the process of analysis of chaos is that of attempting to reconstruct the dynamics in phase-space. The most popular analysis technique for a chaotic time series is the time delay method (Takens, 1981). In this method, a series of equally spaced data is considered. Takens showed that a topological picture of the attractor in phase-space can be reconstructed by the time delay method. Such a topological picture bears a reasonable resemblance to the original attractor, but certain directions have been stretched and deformed. The time delay method consists of choosing a proper delay time and reconstructing a set of m-dimensional vectors or m-tuples, where m is not known a priori. The components of the m-tuples form the directions of the new (topologically equivalent) phase-space in which the attractor will be reconstructed by plotting the points corresponding to the m-tuples in this new phase-space.

The dynamics of a scalar time series $\{x_1, x_2, \ldots\ldots x_n\}$ are embedded in the m-dimensional phase-space (m>d, where d is the dimension of the attractor). The phase-space is defined by:

$$Y_t = \{x_t, x_{t-\tau}, x_{t-2\tau}, \ldots\ldots\ldots\ldots x_{t-(m-1)\tau}\} \qquad (5.1)$$

where $\tau$ is the time delay. Usually, the choice of $\tau$ is made with the help of the auto-correlation function or the mutual information content (Fraser and Swinney, 1986). It may be chosen as the lag time at which the autocorrelation becomes zero (Kantz and Schreiber, 1997) or at which the autocorrelation function falls below a threshold value commonly defined as 1/e (Tsonis and Elsner, 1988). However, it has been reported that considering various values of $\tau$ demonstrates that the results do not show a strong dependence on the actual value chosen (Porporato and Ridolfi, 1997). In this thesis, as will be explained later (section 5.4), $\tau$ is used also as the maximum length of missing segment of observations that can be estimated in one step (i.e., estimation of a missing value does not depend on the previous missing one) by the proposed technique.

### *5.2.2 Correlation dimension*

The method of correlation dimension, as explained by Embrechts (1994), consists of centering a hyper sphere around a point in hyperspace or phase-space, letting the radius (r) of the hyper sphere grow until all points are enclosed, and keeping track of the number of data points that are enclosed by the hyper sphere. The slope of the line on a double logarithmic plot will be an estimate of the fractal dimension of the set of data points. A phase-space is shown in Figure 5.1. The fractal dimension is a measure of the relation between the variance of the signal and the time scale, so it provides a measure of the homogeneity of the signal.

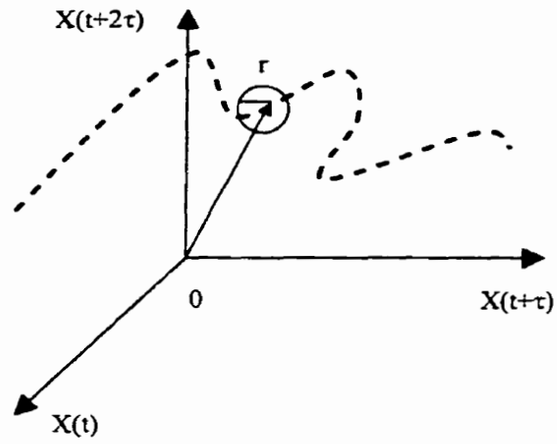



**Fig. 5.1 Phase-space reconstruction of a signal X(t)**

For an m-dimensional phase-space, the correlation integral C(r) given by Theiler (1986):

$$C(r)=\lim_{N\to\infty}\frac{2}{N(N-1)}\sum_{\substack{i,j\\1\le i\le j\le N}}H\left(r-\left|Y_i-Y_j\right|\right) \tag{5.2}$$

where H is the Heaviside step function, with $H(u)=1$ for $u>0$, and $H(u)=0$ for $u\le 0$; N is the number of points on the reconstructed attractor, r is the radius of the sphere centered on $Y_i$ or $Y_j$. The most commonly used norms for $\left|Y_i-Y_j\right|$ are the maximum norm and the standard Euclidean norm. The maximum norm is the maximum absolute difference between the elements of $Y_i$ and $Y_j$. The Euclidean norm, which is the distance between two points in the space, is adopted in this research, and it is also used

by others (e.g., Jayawardena and Lai, 1994). If the phenomenon is chaotic, for a large number of points, beyond a certain m the correlation integral follows the power law:

$$C(r) \underset{\substack{r\to 0 \\ N\to\infty}}{\sim} \alpha\, r^{\nu} \tag{5.3}$$

where $\alpha$ is constant; and $\nu$ is the correlation dimension, which represents generally good estimate of the fractal dimension d of the attractor. The slope of the log C(r) versus log r plot is given by:

$$\nu = \lim_{\substack{r\to 0 \\ N\to\infty}} \frac{\log C(r)}{\log r} \tag{5.4}$$

For a random process, $\nu$ varies linearly with increasing m, without reaching a saturation value, whereas for a deterministic process the value of $\nu$ saturates (levels off) after a certain m. The saturation value, d, is the fractal dimension of the attractor or the time series.

### *5.2.3 Other methods for detecting chaos*

There are also other additional methods available in literature that help detect chaotic behaviour and indicate its existence in the system. Hereafter, Lyapunov exponents, entropy, and the method of surrogate data are briefly explained.

*(a) Lyapunov exponents*

If the fractal dimension of a signal is known and it has been established that there is evidence of chaos through strange attractors in the phase space, then it is possible to estimate the predictability of the signal (Embrechts, 1994). In other words, if chaotic attributes are found in a time series, then according to Takens (1981) it is possible to extrapolate the dynamics of the series within a range specified by the Lyapunov exponents. However, these exponents have not been used in hydrology to determine predictability because of the difficulty of interpreting the computed values in a practical way.

Lyapunov exponents are useful for diagnosis of nonlinear dynamic systems. They describe the rate at which close points in the phase-space diverge. There is one exponent for each dimension. If one or more Lyapunov exponents are positive, the system is chaotic (Frison, 1994). Therefore, one needs to compute only the maximal Lyapunov exponent. The Lyapunov exponents are invariants with respect to initial conditions. Therefore, they form a supportive tool for identifying a chaotic system. However, in water resources literature, Lyapunov exponents have been ignored by researchers such as Sivakumar et al. 1998; 1999c; and Porporato and Ridolfi 1997, as a necessary indication of chaotic behavior.

Consider the representation of the time series data as a trajectory in the embedding space. Assuming that one observes a close return $s_{n'}$ to a previously visited point $s_n$, then one can consider the distance $\Delta_o = s_n - s_{n'}$ as a small perturbation, which should grow

exponentially in time. Its future can be read from the time series: $\Delta_l = s_{n+l} - s_{n'+l}$. If one finds that $|\Delta_l| \cong \Delta_o e^{\lambda l}$ then $\lambda$ is the maximum Lyapunov exponent (Hegger et al. 1999). The routine given by Kantz and Schreiber (1997) is used to calculate the maximal Lyapunov exponent in this research.

*(b) Kolmogorov entropy*

Entropy is a thermodynamic quantity describing the amount of disorder in the system. It can characterize the amount of information needed to predict the next measurement with a certain precision. The most popular one is the Kolmogorov entropy. The Kolmogorov entropy of a time series gives a lower bound to the sum of the positive Lyapunov exponents. An estimate of the Kolmogorov entropy (K) is $K_2$ (Jayawardena and Lai, 1994).

$$K_2(m) \cong \lim_{r \to 0} \left( \frac{1}{\Delta t} \{ \log[C_m(r)] - \log[C_{m+1}(r)] \} \right) \tag{5.5}$$

$$K_2 \cong \lim_{m \to \infty} [K_2(m)] \tag{5.6}$$

where $\Delta t$ is the time interval between two successive observations, $K_2$ is supposed to be zero for regular systems (e.g., periodic), positive and finite for chaotic systems and infinite for stochastic process.

*(c) Method of surrogate data*

Another way of supporting the argument that a specific data set is coming from a nonlinear deterministic system might be by rejecting the hypothesis that it is coming from a linear process. The method of surrogate data (Theiler et al. 1992) makes use of the substitute data generated in accordance with the probabilistic structure underlying the original data. The null hypothesis consists of a candidate linear process and the objective is to reject the hypothesis that the original data have come from a linear stochastic process. A null hypothesis is formulated, for example, that the data have been created by a stationary Gaussian linear process. Then, it is attempted to reject this hypothesis by comparing results for the data to appropriate realizations of the null hypothesis. Sivakumar et al. (1999c) follow the algorithm provided by Theiler et al. (1992) and the significance of a discriminating statistic obtained for surrogate data is judged. In Theiler et al. (1992), it is mentioned that a value of ~ 2 of the statistic cannot be considered significant whereas a value of ~ 10 is highly significant. It is not explained how the authors made an inference from these values and on what statistical basis they derived the conclusion.

Since the primary interest is to give a support to the identification of chaos in time series, provided by the previous invariants, one of the discussed invariants will be used for visual inspection and comparison. In this thesis, the correlation dimension for different m-values of both original and surrogate data will be plotted and inspected. Since the null assumption leaves room for free parameters, the process of generating the surrogate data has to take these into account (Hegger et al. 1999). One approach is to construct constrained realizations of the null hypothesis. This approach of constrained realizations

is adopted, for generating the surrogate data, using the algorithm provided by Kantz and Schreiber (1997). Constrained realizations are obtained by randomizing the data subject to the constraint that an appropriate set of parameters remains fixed (e.g., random data with a given periodogram can be made). More details of that technique can be found in Kantz and Schreiber (1997) and Hegger et al. (1999).

To summarize the procedures of diagnosis and modeling of a chaotic time series, the different stages are shown in Figure 5.2. These stages are identified as identification, modeling, and verification.

## 5.3 Review of studies investigating chaos in hydrology

During the past few decades, hydrologists have been conducting comprehensive research regarding the appropriate type of analysis for the hydrologic data. Statements such as linear versus nonlinear, deterministic versus stochastic, black box versus conceptual models, parametric versus nonparametric became part of the common hydrologic vocabulary. Influenced by the fast-advancing research on chaotic behavior in the physics field, Rodriguez-Iturbe et al. (1989) has indirectly introduced to hydrologists a new topic for comparative analysis. The last few years have witnessed the birth of the topic of stochastic versus chaotic time series analysis and the most recent example is provided by Sivakumar et al. (2001). The literature on chaos in water resources is limited and research is still in its infancy. However, various levels of research and applications are found among the available publications. Only the question of whether data are chaotic or not is

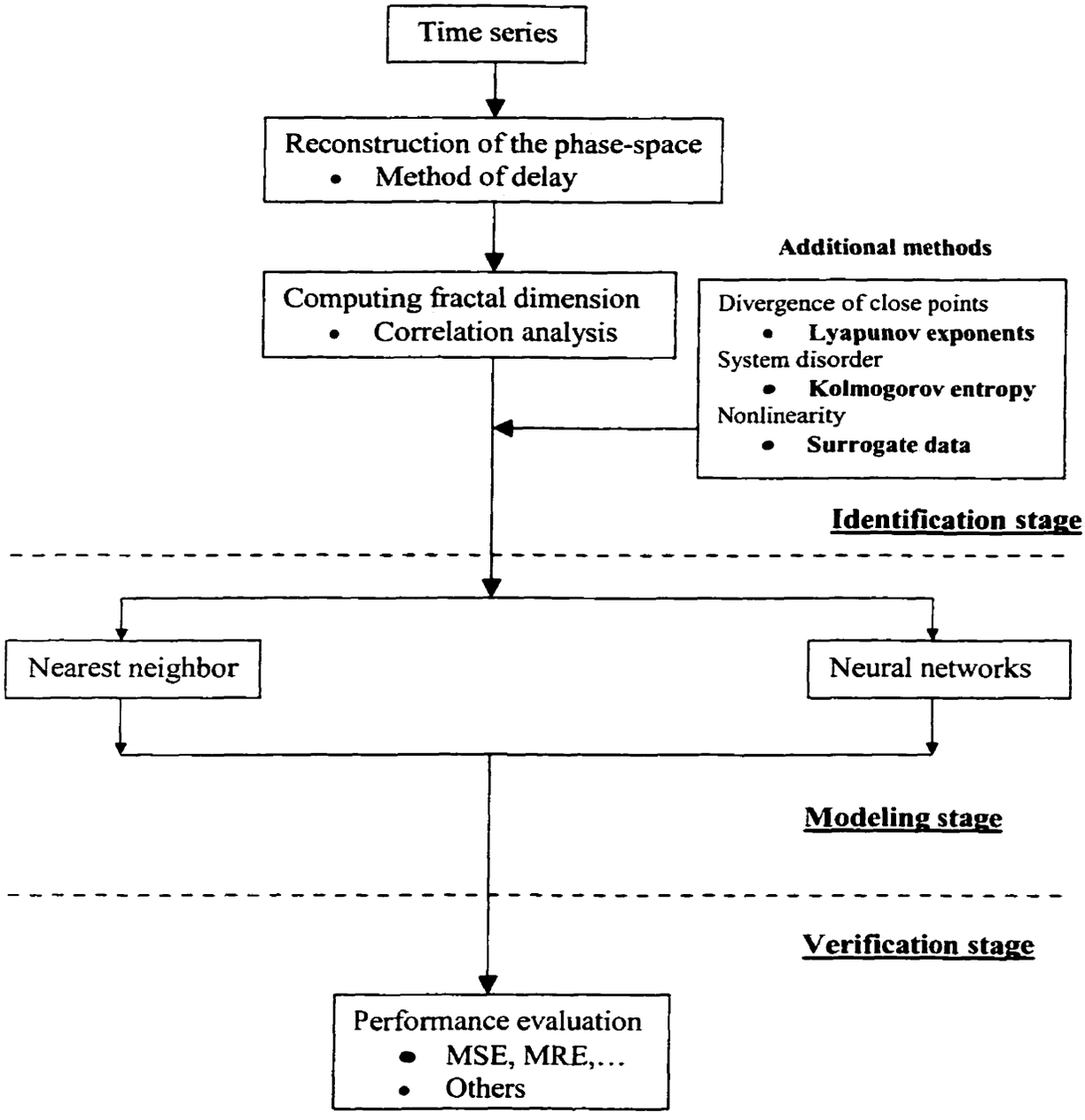



**Fig. 5.2 Flow chart of the different stages of modeling a chaotic time series**

addressed by Rodriguez-Iturbe et al. (1989); Angelbeck and Minkara (1994) and Sivakumar et al. (1998). Others (e.g., Jayawardena and Lai, 1994; Porporato and Ridolfi, 1997) have taken a step forward by trying to predict future values of the variable under consideration.

In many applications, nonlinear modeling tools have provided better results when used in hydrological time series analysis. Few examples, among others, are the superiority of nearest neighbors technique over ARMA models for predicting streamflows (Jayawardena and Lai, 1994), ANNs over ARMA (e.g., Hsu et al. 1995), ANNs and nonlinear regression over linear regression. Further, the superiority of ANNs over nonlinear regression in predicting river flows has been attributed to the possible existence of nonlinear dynamics, which are not captured by the regression technique (Elshorbagy et al. 2000c). Daily rainfall and streamflows might show large dispersions from a mean motion similar to those exhibited by a stochastic process. This behavior might result either from a random probabilistic structure in the data or from a nonlinear deterministic system highly sensitive to the initial conditions (Rodriguez-Iturbe et al. 1989). In the latter case, the dynamics are deterministic although the appearance is similar to that of a stochastic process. Chaotic systems cannot be distinguished from stochastic processes using conventional statistical tools. Systems are said to be chaotic if they are nonperiodic, sensitive to initial conditions, and long predictability is lost (Grassberger and Procaccia, 1983; and Proccacia, 1988).

It is worth mentioning that some doubts have been raised about the existence of chaos in hydrologic data (Chilardi and Rosso, 1990). However, the importance of evaluating

hydrologic data from the viewpoint of nonlinear dynamics is stressed by others (Sivakumar et al. 1999c). In reality, describing a time series as either a totally linear stochastic process or fully nonlinear deterministic chaos is not a practical approach. Any time series might have components of both systems. The analyst has to decide whether the process to be modeled is linear stochastic or deterministic chaos (Kantz and Schreiber, 1997). Based on the impossibility of long-term prediction in chaotic time series, the chaotic applications in water resources literature handled only short-term prediction. Cases where missing lengthy records or consecutive observations, which are common in hydrologic data, have not been addressed. More detailed review of the chaos-related studies in hydrology is given below.

Fraedrich (1986) has estimated the dimensions of attractors from phase-space trajectories of observed weather and climate variables (local surface pressure, relative sunshine duration, and zonal wave amplitude). The author has observed a low fractal dimensionality between 3 and 4 for the weather attractor, if inter-annual variability and seasonal changes were eliminated. The climate variable has also revealed a low dimensionality (between 4 and 5) of the climate attractor. Further study of weather and climate attractors has been done by Fraedrich (1987). Predictability of weather and climate variables has been deduced from phase-space trajectories and estimated by the cumulative distance distributions of expanding pairs of points on the single variable trajectory. It is simply a method of estimating predictability that combines the concepts of correlation dimension and Lyapunov exponents. As a first estimate, the author has obtained a predictability time scale of about two weeks for the weather variable and 10-15 thousand years for the climate variable.

Time series describing the state of the Southern Oscillation (SO) have been analyzed by Hense (1987) using the correlation dimension method to find the dimensionality of the strange attractor. From the results obtained for the precipitation and sea surface temperature time series, the author hypothesized the existence of strange attractor with a fractal dimension between 2.5 and 6.0. It is worth mentioning here that the monthly time series of the Southern Oscillation index (SOI) has also been analyzed by Kawamura et al. (1998) to examine its chaotic characteristics. Three schemes, moving average, low-pass filter and nonlinear smoothing, have been used to reduce noise and enhance chaotic properties. Autocorrelation and spectral characteristics, as well as three chaos-oriented properties – phase space trajectory, the largest Lyapunov exponent, and correlation dimension – have then been examined. No significant signs of chaotic behavior have been found for either the noise-reduced SOI time series or the raw one. The SOI time series was considered to be stochastic rather than chaotic from the viewpoint of dynamical systems theory.

More comprehensive study on obtaining attractor dimensions from meteorological time series has been conducted by Henderson and Wells (1988). They have used the correlation dimension and the limit capacity techniques for estimating the dimension of the attractor of meteorological time series. The results have suggested a low-order attractor for the large-scale flow of size 5.0-6.5, while the gust front data showed a local attractor of 4.0-5.5. Forecasts have been made using a model that simulates the attractor by reconstructing it using time delay method. The prediction function has been gained by tabulating the future value of the lagged series element as a function of position in the

phase-space. The global properties and local structure of the weather attractor over Western Europe have been investigated by Keppenne and Nicolis (1989), who have characterized the dynamics based on the computation of the dimensions of manifolds on which the systems evolve. Several embedding techniques have been used and compared using the data at nine different meteorological stations. All methods have given similar results, namely, that the data in different stations derive from a single deterministic dynamical system spanning a relatively low-dimensional manifold embedded in a low-dimensional phase-space (~7.0). The average time scale of the "growth of errors" associated with divergence of nearby initial conditions has been found to be a few weeks. It has been found that the predictability time in the regions of the attractor, which correspond to low values of the observed time series data, is slightly above one month decreasing to about two weeks for high values of the time series observations.

Mohan et al. (1989) have analyzed the oxygen isotope ratio data from deep-sea cores as a representative of climate data. The authors have questioned the meaningfulness of the correlation dimension calculated for weather and climate attractors by others. The two issues related to the effect of noise and the natural sampling time at which the data are available have been addressed by the authors. They have adopted a different method, called BK (Broomhead and King, 1986) method, for analyzing the time series and have shown that the previously estimated correlation dimension value of 3.1 is inconclusive. Further, they added that, due to noise and large sampling time, it is not possible to embed the data in more than three dimensions and obtain meaningful results. Such a conclusion is in disagreement with study results of Zeng et al. (1992), who have analyzed the fractal

dimension, Lyapunov exponent spectrum, Kolmogorov entropy, and predictability for chaotic attractors in the atmosphere. The time series of daily surface temperature and pressure, over several regions of the United States and the North Atlantic Ocean with different climatic signal-to-noise ratios, have been used for the analysis. Though the total number of data points is large (from about 13,800 to about 36,500), the authors have indicated that it is still too small to obtain a reliable estimate of the correlation dimension. However, it has been shown that this dimension is greater than 8 and that the previous estimate of low fractal dimensions in the atmosphere is spurious. The predictability time is longer over regions with a higher climatic signal-to-noise ratio. It has also been mentioned that the computation of the Lyapunov exponents is slightly sensitive to the selection of the time delay. Another example of the discrepancies in the dimensions estimate of weather attractors is the study of Wang (1995). The author has attempted to detect the presence of low-dimensional deterministic chaos in temperature data by estimating the correlation dimension with the Hill estimate that has been developed by Mikosch and Wang (1995). There has been no convincing evidence of low dimensionality with either global or local temperature data sets. Further, the author has concluded that any apparent reduction in the dimension estimates appears to be due to effects of statistical bias, but neither is it a purely random stochastic process. The dimension of the climatic attractor, according to Wang (1995), may be significantly larger than 10.

Rodriguez-Iturbe et al. (1989) has posed the question regarding possible chaotic dynamics in rainfall of storm events. It has been found that both characteristics of the

correlation integral and the Lyapunov exponents of the historical data give preliminary support to the presence of chaotic dynamics with a strange attractor. A time series of 1990 points of 15-s rainfall has been used for the analysis. Similarly, Sharifi et al. (1990) have analyzed the rainfall data of three storms. The data sets consist of the discrete time series of the inter arrival times of one-hundredth of a millimeter rainfall amounts. In all cases, the number of data points in the data sets was at least 3300. Evidences of the existence of low-dimensional attractors, with correlation dimensions less than 4, have been found in the storm data. The authors indicated that the key to a successful research effort in this area appears to be the availability of very high temporal resolution samples of storm rainfall.

Islam et al. (1993) have made a trial to explain the low correlation dimension estimates for the atmospheric data. They have reviewed weather and climate attractors. The authors have suggested that the physical constraints on variables may be an explanation for the seemingly low estimate of correlation dimension. A good and low estimate may simply mean that a low number of variables may capture certain dynamical aspects of the analyzed time series rather than the entire underlying dynamical system. Along the same way of detecting the existence of chaotic behavior in water resources systems, Angelbeck and Minkara (1994) have investigated the complexity of wastewater flow phenomena using chaos analysis. The authors have used power spectrum, correlation exponent, and Lyapunov exponent analyses in their study. Flow time series of three wastewater treatment plants with number of points ranging from 1,200 to 3300 have been investigated. The analyses have provided a qualitative description of the irregular

behavior of wastewater flows and indicated the existence of chaos in the data sets. Finally, it has been concluded that wastewater flow data can be very complex and that accurate prediction using stochastic models may be invalid.

In addition to the traditional analysis of investigating the existence of chaos in the time series data, Jayawardena and Lai (1994) have moved one step forward by predicting future values. Variations of local approximation methods, which are based on the nearest neighbor technique, are employed for nonlinear prediction of rainfall and streamflow values. Correlation dimension, Lyapunov exponent, and Entropy have been used for the diagnostic analysis. The authors have found convincing statistical evidence to believe that the rainfall and streamflow data series could be better modeled by time delay embedding method than by the traditional linear ARMA approach.

Sangoyomi et al. (1996) have gone back to stress the issue of dimension estimation of nonlinear systems with a different application. They have studied the possibility that variations in the volume of the Great Salt Lake may be described as a low dimensional nonlinear system. Three methods (correlation dimension, nearest neighbor dimension, and false neighbor dimension) of estimating attractor dimension have been applied and compared. The authors have highlighted some factors that may influence dimension calculations, such as number of data points considered in the analysis and whether the underlying nonlinear dynamics has been sampled in a representative manner. A direct continuation of Sangayomi et al.'s study has been made by Lall et al. (1996). The time series data set has been considered as the outcome of a finite-dimensional nonlinear

dynamical system. Short-term forecasting has been made using nonparametric regression to recover the nonlinear autoregressive skeleton of the underlying dynamics. The authors have demonstrated the utility of the proposed approach with synthetic periodic data, data from low-dimensional chaotic dynamical systems, and time series data of the Great Salt Lake.

Puente and Obregon (1996) have reported the existence of chaos in storm events observed in Boston by analyzing a time series of 1990 points using the correlation dimension method, the Kolomogorov entropy method, the false neighbor algorithm, and the Lyapunov exponent method. They have presented a deterministic fractal-multifractal (FM) approach for modeling the storm event. A detailed comparison of the real and fitted time series revealed the possibility of using a deterministic FM approach for a faithful representation of the Boston storm event, and led the authors to hint that a stochastic framework for rainfall modeling might be bypassed. An application that is slightly different from rainfall and streamflow-related applications has been provided by Jeong and Rao (1996). The characteristics of tree rings have been studied to determine the long-term climatic behavior of a region. The authors have analyzed 13 tree ring series from the Salt and Verde River basins in Arizona to determine their chaos characteristics. Two methods of analysis have been used in the study: Grassberger-Procaccia algorithm and the method proposed by Scargle (1990). No evidence of chaotic behavior in these tree ring series has been indicated.

Concepts from nonlinear dynamics (attractors, spectral analysis, phase-space reconstruction, and time delay embedding) and their application to modeling of soil-water processes have been presented by Doscher (1997). A review of nonlinear prediction models based on time series data has been provided. The author has concluded that discovering the nonlinear dynamics in hydrological processes can help direct the modeling efforts.

Porporato and Ridolfi (1997) have given particular emphasis to the problem of noise in the time series and also to nonlinear prediction. The authors have stressed the significance of differencing the time series to obtain the first difference. This process is an important indication about the existence of chaotic component. Then, an interpolation function using the nearest neighbor technique has been estimated for forecasting future values of the time series. The noise level has been reduced, using the algorithm proposed by Grassberger et al. (1993). Details of noise reduction procedures are deferred to the next chapter. Finally, the authors concluded that remarkable predictive results have been attained with the nonlinear method. Further studies regarding the existence of deterministic chaos in streamflow data have been carried out by Wang and Gan (1998), who have estimated the correlation dimensions of the unregulated streamflow data of six rivers in the Canadian prairies to be about 3.0. However, based on their observation of the consistent underestimation of the correlation dimension for the randomly re-sampled data by an amount of 4-6, they have interpreted that the actual dimensions of the streamflow data should be between 7 and 9. The authors have used two methods for calculating the correlation dimension: the Hill estimate and Grassberger-Procaccia algorithms.

Sivakumar et al. (1998, 1999c) have investigated the daily rainfall data of different record lengths observed from each of six stations in Singapore using the correlation dimension method and the nonlinear prediction method, and have provided convincing evidence regarding the existence of chaos. They have also employed the surrogate data method, which indicated the absence of linearity in the rainfall time series. Subsequently, Sivakumar et al. (1999a,b) have studied the influence of presence of noise (measurement error) on the correlation dimension and prediction accuracy estimates, by proposing a systematic approach for noise reduction, coupling a noise level determination method and a noise reduction method. The outcomes have provided additional support regarding the existence of deterministic component in the rainfall phenomenon and possible reasons for the low prediction accuracy estimates achieved in their earlier study (Sivakumar et al. 1999c). Pasternack (1999) has raised serious concerns about the correlation integral analysis (CIA) as a method for calculating the dimension of the attractor. The author has stated that no attractor with dimension less than 5.0 can be found for daily streamflow data. A few published reports have been used to support the claims made by the author that most of the studies that reported attractors of low dimension are faulty. However, it is believed that the streamflow data set used by the author, which turns out to be of dimension more than 5.0 cannot be considered sufficient for reliable and general conclusion.

The number and diversity of applications of chaos theory in hydrology deserve to be reviewed and summarized to highlight the important issues pertinent to hydrologic applications. A pace toward this objective has been made by Sivakumar (2000) who has

briefly reviewed some of the past studies investigating chaos in hydrological processes. The author has addressed some of the important issues in the application of chaos theory in hydrology and provided possible interpretations to the results reported in these studies. According to Sivakumar (2000), an insight into the studies has revealed that most of the problems such as data size, noise, and delay time have been addressed by past studies, and caution has been taken in the application of the methods and interpretation of the results. The study has also revealed that the problem of data size is not as severe as it was assumed to be, whereas the presence of noise seems to have more influence on the nonlinear prediction method than the correlation dimension method. Finally, the study has suggested that the hypothesis of deterministic chaos for hydrological processes is valid and has great practical potential.

There exists an apparent similarity between time dependent chaotic (nonlinear deterministic dynamic) analysis and the domain of nonlinear system dynamics because both deal with the change of behavior over time. However, a distinction between them exists in the sense that the former merely handles a response variable of a system whereas the latter describes the system as a whole.

## 5.4 Estimation of missing streamflow data

The problem of estimation of missing segments (consecutive observations) in streamflow records has been reported recently (Panu et al. 2000; and Elshorbagy et al. 2000a,b). This approach, which is similar in concept to prediction of flows for few steps ahead,

contradicts the concept of chaotic behavior. When a time series is investigated and proved to be chaotic, such type of long-term prediction or estimation of missing segments is not expected to be valid. In this section, a trial is made to combine both of the group approach and analysis of chaos. This combination is achieved by restructuring the data as explained below.

Using the technique of reconstruction of attractor, the time series is presented as shown in Figure 5.3. A segment of missing consecutive data can be confined to the last column, the $m^{th}$. The data set will be structured so that each observation in column m can be estimated based on the previous m-1 columns. The missing segment, indicated in Figure 5.3 inside a rectangle, can be as lengthy as $\tau$. In this way, the missing $\tau$ observations can be estimated using the values from $x_k$ till $x_{(m-1)\tau+k-1}$, which do not include any of the observations from $x_{(m-1)\tau+k}$ till $x_{(m-1)\tau+k+\tau-1}$ (that are supposed to be missing).

In this research, ANNs, as global approximators, are used for estimating the missing data. The number of input nodes will be m-1, output nodes will be equal to one, and the hidden nodes will be obtained using trial and error as explained in the next chapter. A conceptual advantage of this configuration is that the information obtained about m from the correlation dimension computation is used for configuring the ANN model. Therefore, the correlation dimension is not only used as an invariant indicator of chaos but also the length of the groups. The theoretical possibility of predicting the $m^{th}$ dimension, using the previous m-1 dimensions, has been indicated by Kantz and Schreiber (1997). Even

though the ANN model configuration has one output, the way in which the problem is structured allows for estimation of ten outputs in one step (Elshorbagy et al. 2001a).

Another technique for making nonlinear prediction or estimation of missing data is the K-nearest neighbor (K-nn) algorithm. It is a representative of the local approximation method, which uses information of only nearby states to make prediction. The local approximators are always considered to provide good results in chaotic time series. Therefore, they are widely used in chaos literature (e.g., in water resources, Porporato and Ridolfi, 1997; and Sivakumar et al. 1999c).

| | 1 | 2 .. .. | m-1 | m |
|---|---|---|---|---|
| 1 | $x_1$ | $x_{1+\tau}$ | | $x_{(m-1)\tau+1}$ |
| 2 | $x_2$ | $x_{2+\tau}$ | | $x_{(m-1)\tau+2}$ |
| . | | | | |
| . | | | | |
| . | | | | |
| k | $x_k$ | | | $x_{(m-1)\tau+k}$ |
| . | | | | .. |
| . | | | | .. |
| $k+\tau-1$ | $x_{k+\tau-1}$ | | $x_{(m-1)\tau+k-1}$ | $x_{(m-1)\tau+k+\tau-1}$ |
| . | | | | |
| . | | | | |
| $n-(m-1)\tau$ | $x_{n-(m-1)\tau}$ | $x_{n-(m-2)\tau}$ | $x_{n-\tau}$ | $x_n$ |

**Fig. 5.3 Structure of the reconstructed time series data.**

To estimate $x_{i+t}$ based on $Y_i$ (m-dimensional vector) and historical observations, K nearest neighbors of $Y_i$ are found based on the minimum distances $|Y_i - Y_j|$. For K number of neighbors, the estimation of $x_{i+t}$ can be taken as the average of the K values of $x_j$.

Considering the data structure shown in Figure 5.3, the subspace estimation method is employed here. The m-dimensional points are projected on the first, completely known, m-1 dimensions. The nearest neighbors to the vectors in the rows from (k) to $(k + \tau - 1)$ can be calculated using Euclidean norm. Note that the $m^{th}$ dimension of the vectors from (k) to $(k + \tau - 1)$ is the missing segment. A similar technique is used in pattern recognition problems and is known as subspace classification (Queiros and Gelsema, 1988). In pattern recognition, the objective is only to classify objects, therefore the problem is projected on m-1 dimensions (features) and objects are classified based on the complete features regardless of the incomplete one. In our application, once the nearest neighbors are identified, their $m^{th}$ dimension (which is known) is used for the estimation of the missing $m^{th}$ dimension of the vectors from (k) to $(k + \tau - 1)$.

It is reported by Casdagli (1991) that when smaller number of neighbors give the most accurate estimation or short-term prediction compared to that of larger number of neighbors, then this may be considered as a strong evidence for low-dimensional chaos in the data.

## 5.5 Summary

The last decade has witnessed the introduction of the principles of chaos theory to the field of water resources. The nonlinear deterministic dynamics (chaotic behavior) of climatic and hydrologic time series has been investigated and reported in hydrology-related literature. In this chapter, the principles of chaos have been presented and their application to real time series have been explained. Studies investigating chaos in hydrology have been briefly reviewed to show the validity of the concepts of chaos to hydrologic time series and also to shed some light on the current state of knowledge in this relatively new field. The procedures of diagnosing and modeling chaotic time series have been also identified and illustrated in this chapter.

Since this thesis focuses on the process of estimation of missing consecutive hydrologic observations, the problem of missing consecutive observations has been reformulated to fit the context of chaos analysis. The methodological procedures of using the characteristics of the nonlinear dynamics and the chaotic invariants to estimate missing data have been developed and explained. A length equal to the embedding dimension has been recognized as the zone of influence required to estimate or predict the subsequent group of observations. The estimated group of observations should be of a length that is equal to or less than the time lag ($\tau$), used to reconstruct the dynamics of the time series.

In this chapter, the concepts of chaos and analysis of chaotic time series have been explained with regard to the validity of their application in hydrology. When a time series

is investigated and proved to be chaotic rather than stochastic, the analysis of chaos becomes a strong candidate to handle the time series. Having the approach of statistical and stochastic hydrology settled in the minds of analysts for long time, many questions can be raised regarding the analysis of chaos such as: What type of techniques and models will be able to handle chaotic data? Knowing that data are chaotic, what will be the effect on the modeling approach? How applicable is the idea of using the concepts of chaos in formulating and configuring the models? Giving the nature of sensitivity to initial conditions and other chaos characteristics, can global modeling approach such as ANNs be used to model chaotic time series? Even if modeling a single chaotic time series is possible, what is the potential for modeling multiple chaotic time series? And finally, how can the random error term, well perceived in stochastic hydrology, be tackled in the context of chaos analysis?

The following chapter focuses on the above-mentioned questions. Endeavors of addressing these questions through applications of the concepts of chaos to streamflow data will be pursued and explained in chapter 6.

# Chapter 6

# 6. ANALYSIS OF CHAOS: APPLICATIONS TO DAILY STREAMFLOWS

6.1 Introduction

6.2 Application to single time series

6.3 Application to cross-correlated time series

6.4 Analysis of noise

6.5 Summary and concluding remarks

---

*" They said: O Noah! Thou hast disputed with us, and much hast thou prolonged the dispute with us: now bring upon us what thou threatenest us with, if thou speakest the truth!? He said: Truly, Allah will bring it on you if He wills,- and then, ye will not be able to frustrate it!*

*It was revealed to Noah: None of thy people will believe except those who have believed already! So grieve no longer over their deeds"*

*The Glorious Qura'n (11:32-33, 36)*

## 6.1 Introduction

In this chapter, few issues are addressed: first, estimation of missing consecutive observations of a single chaotic time series, second, the use of chaotic invariants or dimensions for configuring the hydrologic models, third, the possibility of using global modeling techniques for modeling chaotic time series, which are commonly handled using local modeling techniques, fourth, the possible extension of the analysis of chaotic time series to handle multiple time series, and fifth, the issue of the data noise and its effect on modeling the time series. The reliability and the utility of the noise reduction approach are addressed with regard to its applicability to the analysis of chaotic time series in hydrology.

## 6.2 Application to single time series

In this section, the streamflow data from English River at Umferville, Ontario, Canada is used and analyzed to investigate the possible existence of chaotic behavior. The average daily flow is 124.3 $m^3/s$ and the standard deviation is 92.3 $m^3/s$. The issue related to length of the data record that is considered sufficient for chaos analysis has been addressed by many authors. While Frison (1994) suggests a minimum of ~ 10,000 data points for reliable results, others accepted a number of points as low as 1,200 observations (Jayawardena and Lai, 1994) and 1,500 observations (Sivakumar et al. 1999c). In this application, 11,000 observations of the daily flows are used. The daily streamflows of the English River are shown in Figure 6.1.

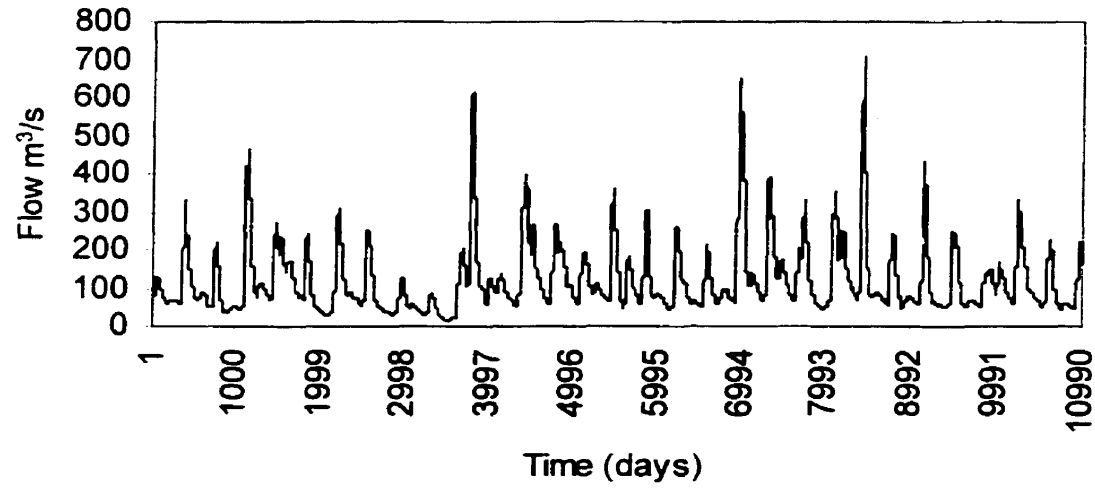



**Fig. 6.1 Variation of daily flows of the English River.**

*6.2.1 Correlation dimension*

Daily flows comprising of 11,000 observations of the English River are considered for investigating the existence of chaotic behavior in the streamflow. The value of lag time ten (i.e., $\tau$ =10 days) is considered in the analysis with the purpose of estimating (in-filling) ten consecutive missing observations. The correlation integral (sum) C(r) and the correlation exponent $\nu$ are computed, as explained earlier, from the data set. The relationship between the correlation integral C(r) and the radius r for various values of embedding dimensions m is shown in Figure 6.2. The correlation exponent increases with the increase in the embedding dimension up to a certain point (m = 8) and saturates beyond that point (Figure 6.3). The saturation value of the correlation exponent (dimension) is ~ 2.4.

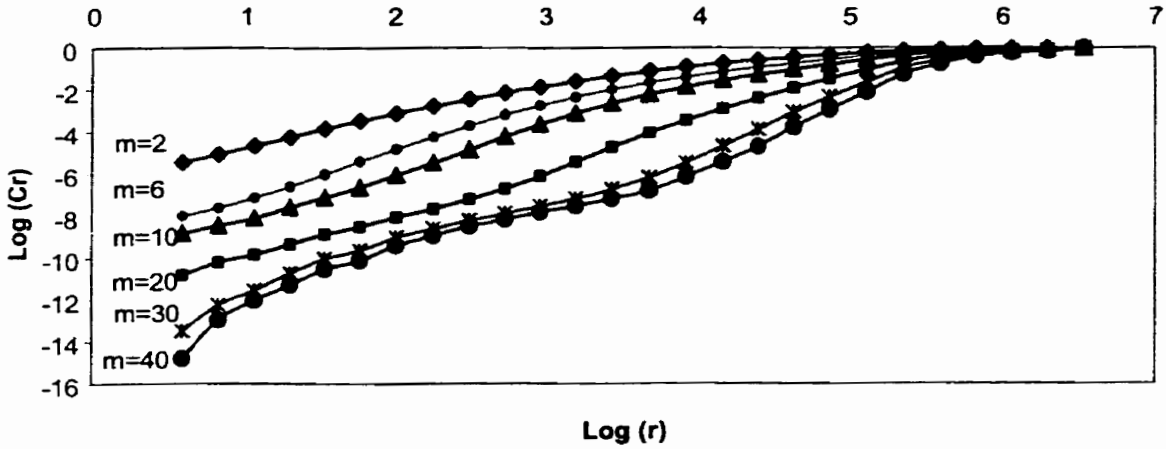



**Fig. 6.2 Correlation integral of the English River data.**

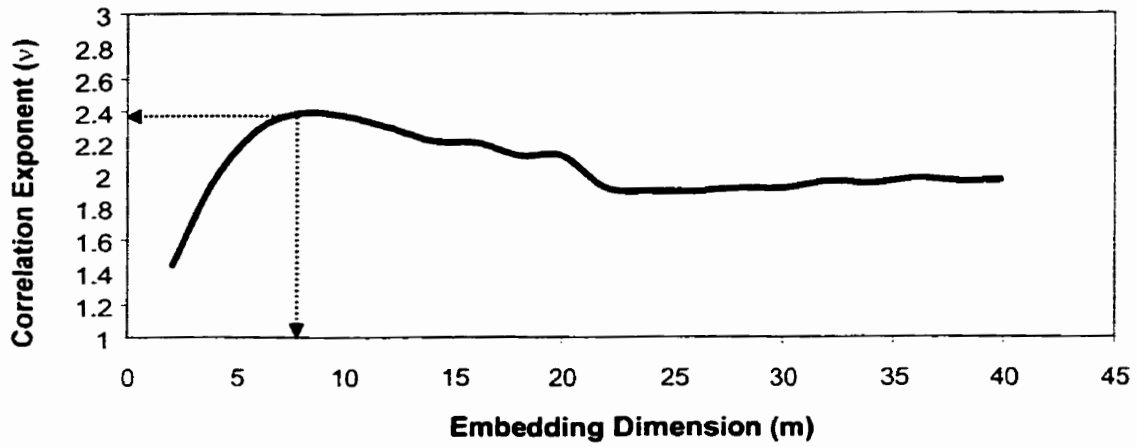



**Fig. 6.3 Correlation exponent of the English River data.**

The nearest integer above the correlation dimension value (d = 3) is taken as the minimum dimension of the phase-space that can embed the attractor. The value of m at the saturation point (m = 8) is supposed to provide the sufficient number of variables to describe the dynamics of the attractor.

### *6.2.2 Lyapunov exponent*

Using the algorithm of Kantz and Schreiber (1997), the largest Lyapunov exponent is found to be positive ($9.1x\ 10^{-3}$). It should be noted that it can be positive also for some random and ARMA processes as observed by Jayawardena and Lai (1994) and Rodriguez-Iturbe et al. (1989), when calculated using the algorithm of Wolf et al. (1985). Therefore, results of Lyapunov exponent computation are not recommended to be taken as a sole indication of chaotic behavior.

### *6.2.3 Entropy*

A positive finite $K_2$ value can be found for the data set under consideration. Figure 6.4 shows the relationship between the correlation entropy and log (r) for various m. A plateau can be observed at $K_2$ value ~ 0.2. This type of visual inspection of $K_2$ entropy is used and reported by Kantz and Schreiber (1997).

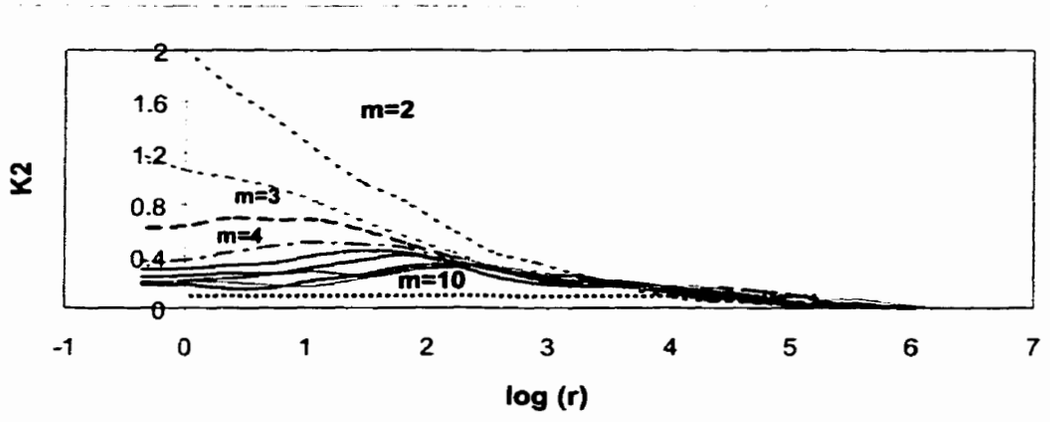



**Fig. 6.4 Estimate of correlation entropy of the English River data.**

*6.2.4 Method of Surrogate data*

Several realizations of surrogate data sets can be generated according to the probabilistic structure underlying the original data and also according to the null hypothesis as discussed earlier. The major aim of this step is to detect nonlinearity. The purpose of this process is to demonstrate the difference in behavior between the two data sets. In the case of the original data of the English River, the correlation exponent curve levels off at a certain point (i.e., m = 8) whereas the correlation exponents computed for the surrogate data continue increasing monotonically with the increase in embedding dimension (Figure 6.5). In this thesis, the visual difference between the two data sets is considered sufficient to indicate that original data might not come from a linear stochastic process.

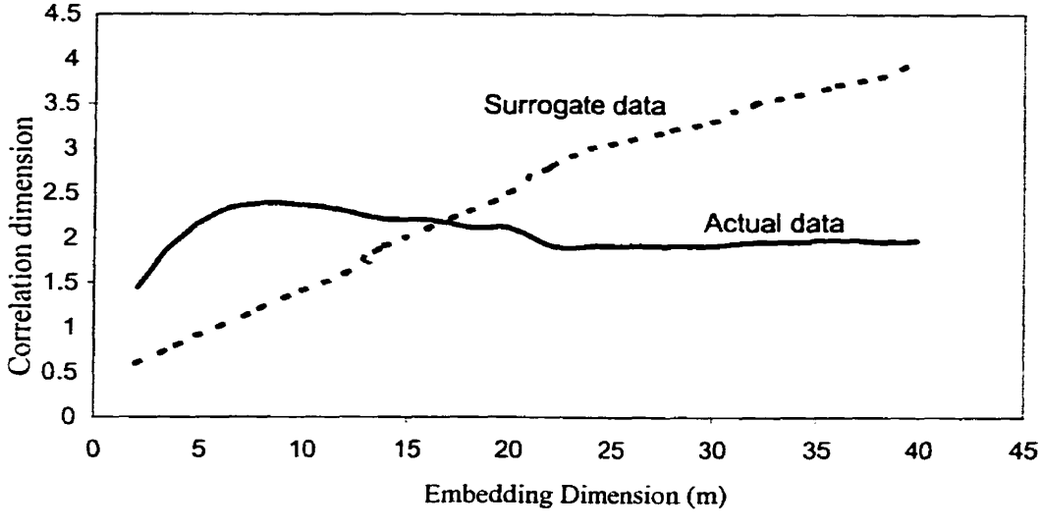



**Fig. 6.5 Correlation dimension of the English River data set and its surrogate.**

*6.2.5 Results and analysis*

The methodology and models, presented in section 5.4 of the previous chapter, for estimation of ten consecutive missing observations are applied to the streamflow data in this section. The ANNs and the K-nearest neighbors techniques are employed. The results shown in Figures 6.6 and 6.7 are summarized in Table 6.1. This table indicates that the lower the K-value, the better is the estimate of the missing observations (lower mean squared error and mean relative error).

**Table 6.1 Mean squared and mean relative error of estimated data**

| Technique | Mean squared error (MSE) | Mean relative error (MRE) |
|---|---|---|
| ANNs | 80.2 | 0.06 |
| K-nn (K=1) | 255.5 | 0.08 |
| K-nn (K=5) | 365.9 | 0.12 |
| K-nn (K=10) | 441.5 | 0.14 |

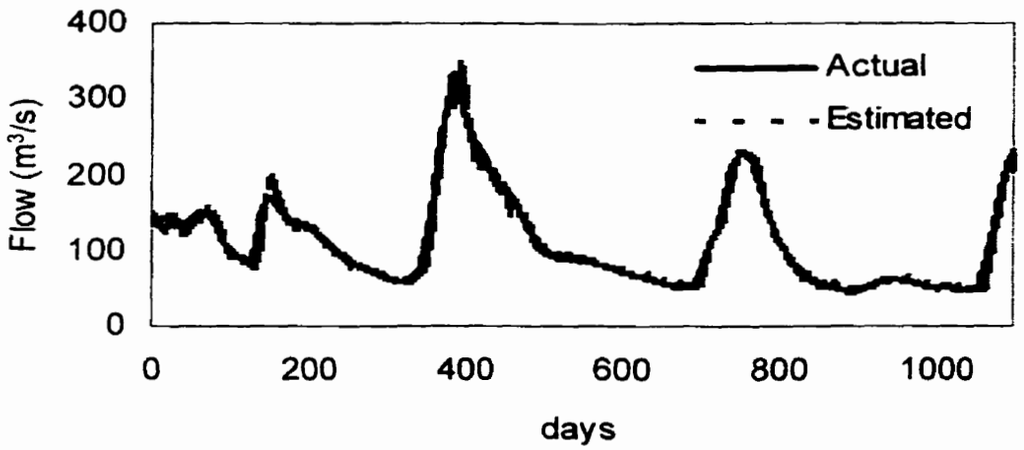



**Fig. 6.6 Actual and estimated data using ANN technique.**

This result may give another support to the previously discussed invariants (e.g., correlation dimension) that indicates the existence of chaos. Figure 6.8 shows the actual streamflows and those estimated by ANN and K-nn (using K=1) models. Both, Table 6.1 and Figure 6.8 indicate that ANNs are superior to K-nn for estimating the missing observations. Such a superior performance of ANN may be problem-related and needs extensive applications on various data sets before generalization. However, one can say that the superiority of ANNs might be attributed to the ability of ANNs to capture the nonlinear dynamics of the data. Such a characteristic of ANNs is indicated by others (Panu et al. 2000; and Elshorbagy et al. 2001c). Furthermore, the way ANNs are used in this research makes use of the identified embedding dimension, which means that the attractor is modeled directly using its first seven dimensions. It seems that the ANNs are

able to generalize the structure of the attractor throughout the whole data set in a more homogeneous fashion than K-nn model.

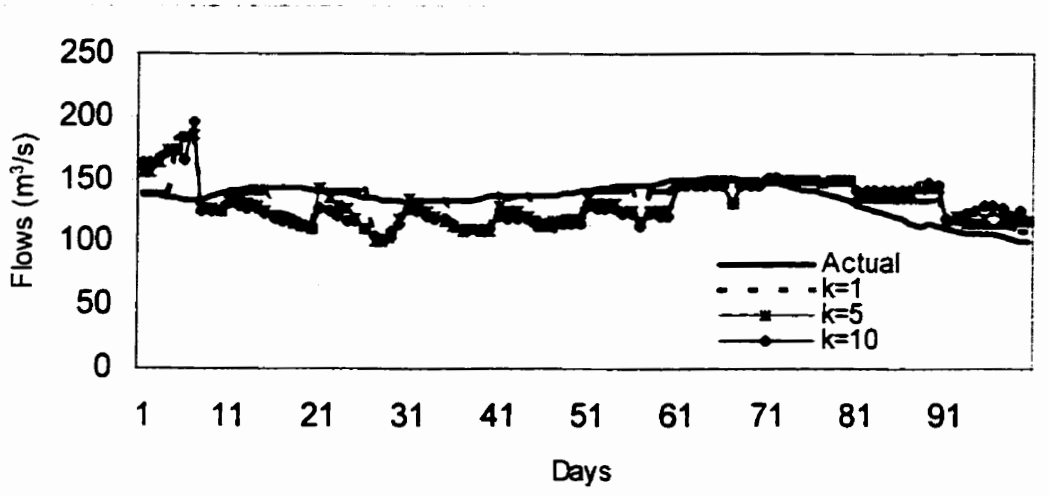



**Fig. 6.7 Actual and estimated data using K-nn technique.**

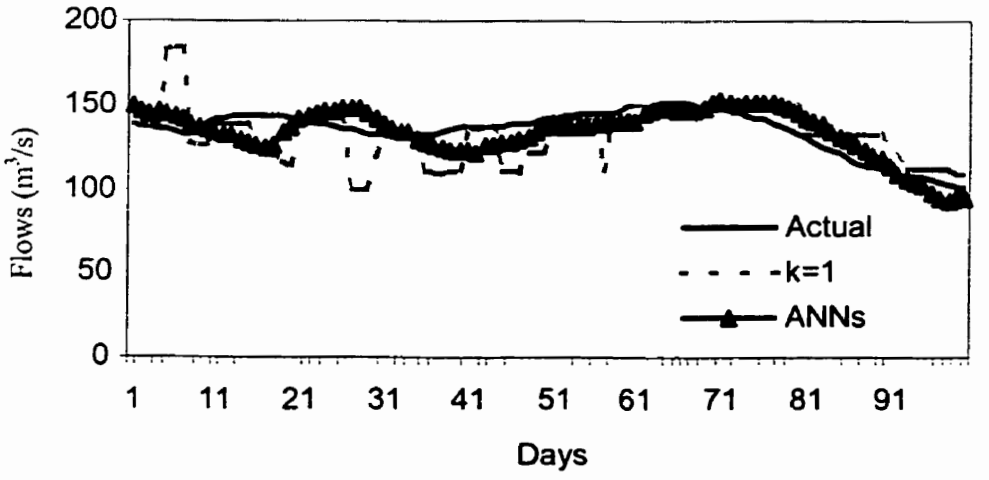



**Fig. 6.8 Actual and estimated data using ANN and K-nn techniques.**

On the other hand, the subspace technique used with the K-nn could not benefit significantly from the information related to the embedding dimension (m=8). This could be due to the aggregation that happens when the first seven dimensions of the attractor are represented by one number (Euclidean norm).

It should be noted that some recent studies have indicated that the noise that exists in hydrologic data may limit the performance of many modeling technique. Some methods have been proposed to reduce the level of noise in the data set (Sivakumar et al. 1999b; and 2000), which may lead to improvement in the accuracy of the estimation of missing data. The effect of noise on the accuracy of the estimation process is deferred to section 6.4.

*6.2.6. Concluding remarks*

In this section, the existence of chaotic behavior (nonlinear deterministic dynamics) in daily flows of the English River has been investigated. The correlation dimension, the Lyapunov exponent, the Kolmogorov entropy, and the method of surrogate data have been used in the analysis. There are sufficient indications to believe that the streamflows have some chaos and the data could be modeled by the time delay embedding method. Based on the attractor dimension, the minimum number of variables essential to model the dynamics of daily flows of the English River is identified as 3 and the number of sufficient variables is 8. The sufficient number of variables is used to configure the ANN model. Seven input nodes (first 7 dimensions) are used to estimate the output ($8^{th}$ dimension). The data are structured in a way that facilitates the estimation of 10

consecutive missing observations in one step (previously estimated value is not used for estimating the following one). Also, the K-nearest neighbor technique is used, with K=1, 5, and 10, to estimate the missing data. A subspace modeling approach is adopted by projecting the reconstructed attractor on a lower dimension scale (i.e., projecting the 8-dimensional attractor on its first known 7-dimensional space). The $8^{th}$ dimension is estimated using the previous seven dimensions. The ANN model shows superiority in the accuracy of estimating the missing data, which is attributed to the capability of the ANNs to capture the nonlinear dynamics and generalize the structure of the attractor on the whole data set. Finally, this work is considered as an endeavor towards establishing hydrologic chaotic modeling by using the chaos indicators (correlation dimension) directly in the process of modeling or configuring the data model (Elshorbagy et al. 2001a).

## 6.3 Application to cross-correlated time series

The application presented in the previous section and other reported case studies focus mainly on the univariate case where a single time series is always under consideration. The growing belief in the applicability of the chaos theory to the field of water resources along with the indications that some of the hydrologic time series might be chaotic rather than stochastic raise important questions to be addressed. In stochastic hydrology, multiple time series are considered using multivariate time series analysis approach (Salas et al. 1985). Data generation, augmentation, and extension of short records are always treated in the context of statistical and stochastic analysis where statistical

properties such as mean and variance are parameters to be preserved. When hydrologic data are proved to be chaotic (behavior of nonlinear deterministic dynamics), then analysis of multiple time series has to be reformulated to fit the context of chaos analysis.

In this section, two cross-correlated streamflow time series are analyzed with regard to the existence of chaos. One of the time series, named as the *reference river*, is used to estimate missing data of the other series, named here as the *target river*. The main objective of the analysis conducted in this section is to study the effect of the existence of nonlinear deterministic dynamics (chaotic behavior) on the process of analyzing cross-correlated hydrologic time series for data in-filling purposes.

Methodologies or techniques that help improve the accuracy of estimating missing observations are highly significant and useful for the process of hydrologic modeling, which requires complete data sets. In this section, a complete set of streamflow data is used as a *reference* river and another concurrent set of complete streamflow data, that is cross correlated with the *reference* river, is used as a *target* river. The data sets are divided into training (calibration) and testing (verification) sets. Few patches of consecutive observations are removed from the data set of the *target* river to test the estimation procedures.

Two techniques are used to estimate the missing data. First, linear regression (LR) technique and second the artificial neural networks (ANNs) as a representative tool of the

nonlinear approximation methods. The feed forward ANNs that employ the back propagation (BP) training technique are selected for the problem under consideration.

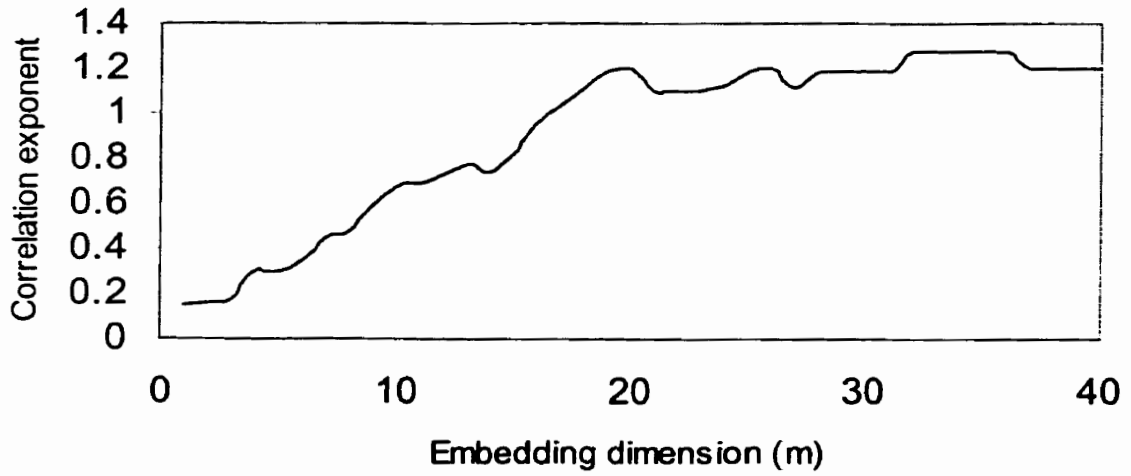
*6.3.1 Case study*

The daily streamflow data of the Little River (USGS # 03170000) and Reed Creek (USGS # 03167000), Virginia, USA are used in this section. The cross-correlation coefficient between the two time series is found to be equal to 0.76. A length of 20,000 concurrent observations of the two rivers is selected for the analysis. The Little River and the Reed Creek are used as the *reference* and *target* rivers respectively. Seven patches, each is of ten consecutive observations, are considered missing from the *target river*. The whole data set is divided into seven sections and one patch is randomly selected from each section so that the missing data may be representative of the entire data range.

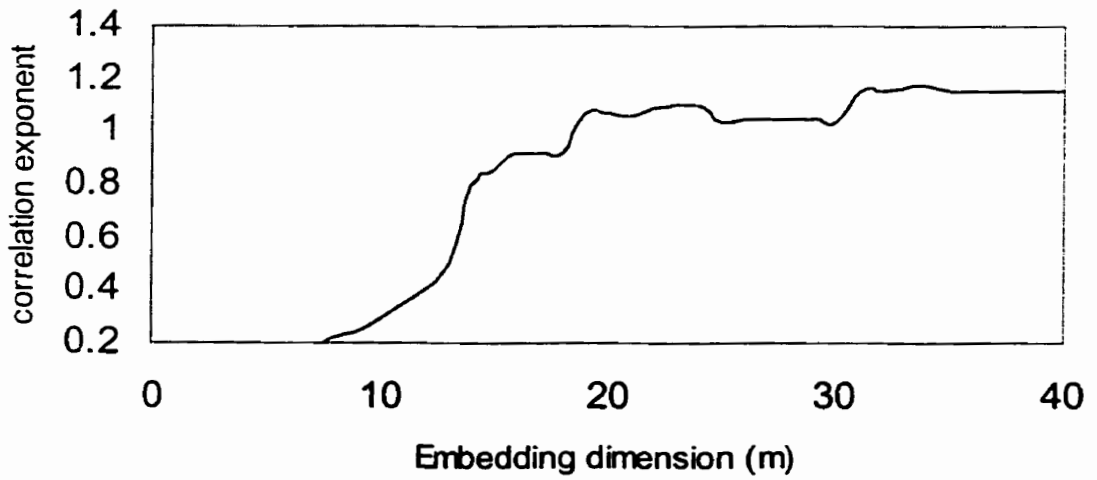
*Correlation dimension*

A value of lag time ten (i.e., 10 days) is considered in the analysis of the correlation dimension. The correlation integral (sum) C(r) and the correlation exponent $\nu$ are computed, as explained earlier, from the data set. Figure 6.9 shows that the correlation exponent (dimension) increases with the increase in the embedding dimension up to a certain point (m = 19) and saturates beyond that point. The saturation values of the correlation exponent (dimension) are ~ 1.19 and 1.07 for the Little River and the Reed Creek respectively. The nearest integer above the correlation dimension value (d = 2) is the minimum dimension of the phase-space that can embed the attractor. The value of m

at the saturation point (m = 19) is supposed to provide the sufficient number of variables to describe the dynamics of the attractor (Kantz and Schreiber, 1997).




a)




b)

**Fig. 6.9 Correlation dimension of the time series, a) Little River and b) Reed Creek**

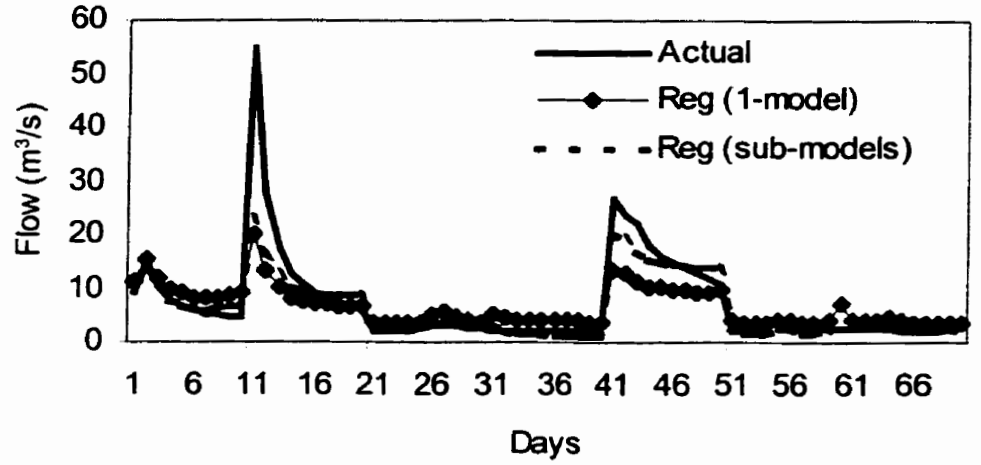
*6.3.2 Results and analysis of estimation of missing data*

Two experiments are conducted in this section to test the ability of both local and global modeling approaches to handle chaotic time series. Two standard techniques, linear regression (LR) and artificial neural networks (ANNs), are used in the analysis. The two experiments are designed as follows: (1) The training data are used to train one single model, which is verified using the testing data. The testing data are unseen during the training and are assumed to be missing. It is also worth mentioning that the testing data are selected to represent seven sections that are covering the entire data set. (2) Seven different local models are trained to represent the different sections of the data. Each model is verified on a test data set that represents the section of the data within which the model is trained. The mean squared error (MSE) and the mean relative error (MRE) are calculated as measures of accuracy of estimating the missing data.

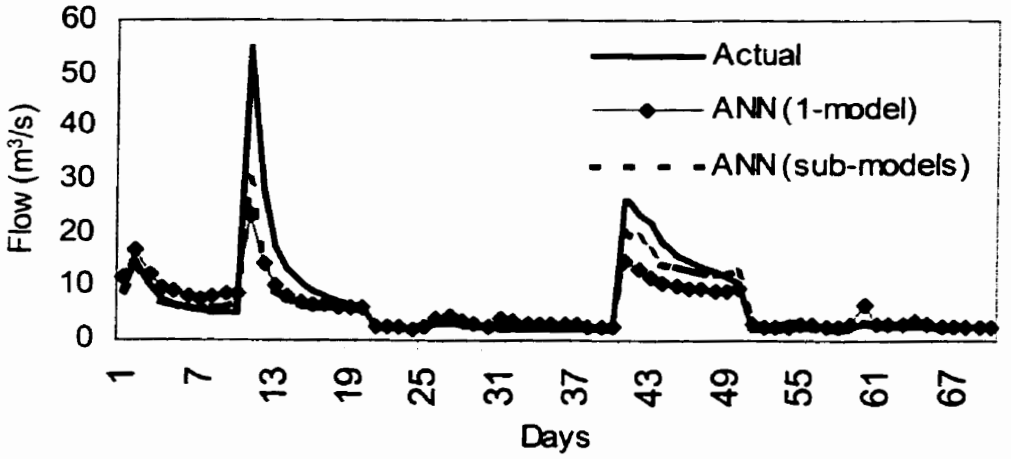
Other experiments can be designed as well but the two mentioned earlier are considered to be sufficient in this research to present the idea of the effect of local and global modeling on the analysis of chaotic time series. The actual and estimated missing data, using LR and ANN models, of the Reed Creek are shown in Figure 6.10. One can observe that using multiple local models (sub-models) can improve the accuracy of estimating the missing data.

*6.3.3 Discussion*

The MSE and the MRE of estimating the missing data resulting from different experiments are given in Table 6.2. It is evident that ANN models are superior to the LR




a)




b)

**Fig. 6.10 Actual and estimated data of the Reed Creek raw data using a) linear regression and b) artificial neural networks**

models given that the experiment type is fixed (i.e., either one global model or several local sub-models are considered). The nonlinearity in the data, which is not captured by the linear regression models, is better modeled using a nonlinear technique such as the ANNs. More importantly, one can observe the superiority of local models over the global models irrespective of the adopted technique, LR or ANNs. The nonlinearity that is inferred by the superiority of the ANNs over LR is complemented by the improvement of modeling accuracy when local models are employed. It is believed that the superiority of nonlinear local models in estimating the missing data can be attributed to the existence of chaos in the two cross-correlated time series considered in this analysis. This result supports the argument provided by others that local models can simulate chaotic time series better than global ones (e.g., Porporato and Ridolfi, 1997).

**Table 6.2. The mean squared and the mean relative errors of the estimated values.**

| Experiment | Raw data | |
|---|---|---|
| | MSE | MRE |
| **LR (global model)** | 32.30 | 0.55 |
| **LR (local models)** | 19.64 | 0.23 |
| **ANN (global model)** | 26.79 | 0.31 |
| **ANN (local models)** | 15.29 | 0.17 |

In a study of stochastic models of streamflows, Mujumdar and Kumar (1990) have shown that an autoregressive model, AR(1), is superior to other ARMA models that have more

parameters, in terms of the MSE of forecast values. In their paper, the authors have considered the results to be interestingly contrary to the common belief that models with a larger number of parameters give better forecasts. In light of the recent revealed chaotic behavior of the hydrologic time series, the results of Mujumdar and Kumar (1990) can be interpreted differently. For example, although both AR(1) and AR(10) are global models, when forecast is done using AR(1) only the value of the most recent observation is used in the forecast. In case of AR(10), a lengthier stretch of observations, 10 values, is utilized. That segment of observations may be longer than the window that constitutes the zone of influence for the future values to be forecast. This argument may justify why AR(1) resulted in lower MSE than models that include more parameters. The comment of Mujumdar and Kumar (1990) that "the simplest model is sufficient" may be reasonable if simple and more complicated models result in similar values of MSE. But the deterioration that occurs and shown in their study, when more parameters are included, can be attributed to the existence of chaos (Elshorbagy et al. 2001c).

### *6.3.4 Concluding remarks*

In this section, a trial has been made to model two cross-correlated rivers as chaotic time series. The existence of chaos in the daily flows of the Little River and the Reed Creek, Virginia, has been investigated. An indication of the chaotic behavior is shown using the correlation analysis. Segments of missing data in the Reed Creek are estimated using linear regression and artificial neural network (ANN) models. Two experiments are conducted with both of the modeling techniques: (1) a single global model is fitted to

estimate the missing data and (2) multiple local sub-models are fitted; each is in the direct vicinity of the missing data.

In general, ANN models show superiority over LR models in estimating the missing data. More significantly, the performance of local models over global models in terms of the MSE and MRE criteria is attributed to the chaotic behavior of the time series. The significant role that the chaotic behavior plays in this analysis highlights the importance of investigating similar effects of chaoticity in other hydrologic applications such as record extension, data generation, and multivariate analysis. When a time series is proved to be chaotic rather than stochastic, every application in stochastic hydrology should find its analogy in chaos analysis, which poses a research challenge for hydrologists.

## 6.4 Analysis of noise

In the mid eighties and early nineties when meteorologists and hydrologists (e.g., Fraedrich, 1987 and Rodriguez-Iturbe et al., 1989) imported the principles of chaos theory from physics to apply them to hydrologic time series, a caution in both analysis and reporting of the results was exercised. In general, one can find statements indicating that describing a geophysical time series as a fully chaotic system might not be feasible. For example, Kantz and Schreiber (1997) mentioned that any time series might have components of both systems, chaotic and stochastic. The analyst has to decide whether the process to be modeled is linear stochastic or deterministic chaotic. Such a lack of certainty and blurring of the boundaries between the two systems have been expressed in

the final remarks of Rodriguez-Iturbe et al. (1989). They raised the question of whether or not the complex multidimensional patterns of a rainfall storm can be usefully described by a single attractor in a single state space. Over a period of a decade, a few case studies applying the principles of chaos to hydrologic time series have been reported (e.g., among others, Jayawardena and Lai, 1994; Lall et al. 1996; Porporato and Ridolfi, 1997; and Wang and Gan, 1998). Until 1999 questions such as whether rainfall is chaotic or not were raised by Sivakumar et al. (1999c). The overriding objective of their work was to ensure the existence of some chaos in the time series.

Driven by the application-oriented approach of the hydrologists, Jayawardena and Lai (1994) advanced one step further by applying a local nonlinear approximation method for short-term prediction of daily rainfall and streamflow data. Similarly, a local nonlinear prediction was applied by Sivakumar et al. (1999c) on a case study of Singapore rainfall. Porporato and Ridolfi (1997) came up with the idea of importing another concept from the field of Physics and apply it to hydrologic time series. They adopted a technique proposed by Schreiber and Grassberger (1991) for reducing the noise that is contaminating the time series, and therefore improving the accuracy of prediction. After reducing the noise, the clean signal is considered for further analysis (i.e., prediction). This approach, which apparently improved the accuracy of prediction has encouraged others to build on the same idea of improving the prediction by cleaning the noisy signal. Three cases have been reported by Kawamura et al. (1998); Sivakumar et al. (1999b); and Jayawardena and Gurung (2000). It is acknowledged and agreed upon in the published reports that neither the clean signal nor the true underlying dynamics of the hydrologic

systems is known. However, the clean signals, after noise reduction, have been taken for granted as representatives of the underlying systems. As will be shown later, the perception of noise and noise reduction are dependent on the prediction tool. In this section, some concerns are raised and questions, regarding the reliability of the noise reduction approach, are addressed with regard to its applicability to chaotic time series analysis in hydrology.

### *6.4.1 Noise in hydrologic data*

#### *a) Noise in stochastic analysis*

The concepts of random variables and random phenomena have been used in the field of hydrology since early 20$^{th}$ Century. About the same time, statistics and probability theories were being applied to the analysis of river flow sequences. However, it was not until the beginning of the 1960's that the formal development of stochastic modeling began with the introduction of autoregressive (Markovian) models to seasonal and annual hydrologic time series (Salas et al. 1980). For generating synthetic sequences, Matalas (1967) has employed the lag-one Markov process, which includes a random component ($\varepsilon$) that has zero mean and unit variance and is independent of the signal ($x_t$). This component was sometimes called error term (e.g., Young et al. 1970) or independent stochastic component that has zero autocorrelation coefficient (Yevjevich, 1987). D'Astous and Hipel (1979) called it white noise when they introduced a stochastic intervention model where they differentiated between noise and white noise. They restricted the use of the term noise to the component that can be modeled by a stochastic

model. In general, the white noise component in stochastic models is considered to include any measurement error and/or components that are not captured by the model.

*b) Noise in chaos analysis*

Adopting the chaos analysis approach, the time series is considered to be seemingly irregular data that are generated by a nonlinear deterministic dynamic system. Such a system is referred to as chaotic system. The chaos analysis of a time series is dependent on the computation of some invariants (e.g., correlation dimension and Lyapunov exponents). The correlation dimension, as one of the straightforward invariants to be computed, is widely used in chaotic analysis of hydrologic, and any other real, data. One of the main problems in dimension estimation from time series has to do with the fact that time signals from natural phenomena are corrupted by noise. The measured time signal will always contain some noise due to random influences and inaccuracies that can never be eliminated completely (Schouten et al. 1994). Apparently, hydrologic time series are no exception. The inaccuracies may be caused by the measuring device and analog-to-digital conversion.

Considering the system to be chaotic, one has to make an important distinction between two types of noise, *measurement noise* and *dynamical noise*. *Measurement noise*, sometimes called additive noise, refers to the corruption of observations by errors, which are independent of the dynamics (Kantz and Schreiber, 1997). The dynamics satisfy:

$$X_{t+1} = F(X_t) \tag{6.1}$$

but we measure scalars:

$$y_t = y(X_t) + \eta_t \tag{6.2}$$

where $y(X_t)$ is a smooth function that maps points on the attractor to real numbers, and the $\{\eta_t\}$ are random numbers. The series $\{\eta_t\}$ is referred to as the measurement noise. Included in this series are real errors of measurement and other dynamics simultaneously present and superimpose themselves on the signal under investigation altering the measurement (Porporato and Ridolfi, 1997). In most practical cases, measurement noise is considered to be Gaussian, especially when it is due to a combination of various sources of independent random fluctuations. However, in some cases, signals may also be contaminated by measurement noise that is strictly bounded in magnitude. *Dynamical noise*, in contrast, is a feedback process wherein the system is perturbed by a small random amount at each time step (Kantz and Schreiber 1997):

$$X_{t+1} = F(X_t + \eta_t) \tag{6.3}$$

It might also be the case that dynamic phenomena or systems are influenced by intrinsic events taking place at random. In other words, dynamical noise implies that the system itself, and not simply the measurement, is disturbed at each instant. Therefore, noise may become an essential part of the dynamics and a clean signal (trajectory) does not exist a priori. The problem of "dynamical noise" is much more complex than the "additive noise" (Porporato and Ridolfi, 1997 and Schouten et al. 1994) and removing it from the

signal may not be justified, if at all possible. What is discussed in this thesis is the issue of removal or reduction of the additive (measurement) noise, in relation to the published case studies of noise reduction in hydrology.

### *6.4.2 Effect of noise*

It is worth mentioning that the computations of the above mentioned invariants (correlation dimension, entropy, and Lyapunov exponents) are based on the assumption that the data are noise free. The existence of noise complicates the process of computing these invariants (Kostelich and Schreiber, 1993). When hydrologists started adopting analysis of chaotic time series, the main effort has been focused on proving that chaos exist in hydrologic phenomena (e.g., rainfall and streamflow). Prediction imposed itself as an important hydrologic application where local nonlinear prediction models have been applied to the raw data after proving the existence of chaos (Jayawardena and Lai, 1994; Lall et al. 1996; Liu et al. 1998; and Sivakumar et al. 1999c). As reported in Sivakumar et al. (1999a and 2000), prediction accuracy has not been highly satisfactory. It is widely recognized that in chaotic systems, only short-term prediction is possible, unlike long-term prediction, due to the rapid evolution of the system with time. Since short and long-term are not strictly defined, one can find in the hydrologic applications of chaos predictions with 1 – 20 days ahead for daily streamflows (Liu et al. 1998), 1 – 10 days (Jayawardena and Lai, 1994), one-step ahead (Lall et al. 1998), 12 - 48 hours (Porporato and Ridolfi, 1997), 25 data points (Jayawardena and Gurung, 2000), and 100 days (Sivakumar et al. 1999b).

The unsatisfactory accuracy of prediction obtained in the above-mentioned hydrologic case studies has been attributed to the existence of noise contaminating the signal. Surprisingly there has been no referral to the possibility that low accuracy of prediction could have been caused by the inefficiency of the selected prediction model. Only a single local prediction model has been always employed in Porporato and Ridolfi (1997), Sivakumar et al. (1999b), and Jayawardena and Gurung (2000). Consequently, the noise reduction approach has been adopted, and the accompanied improvement in the prediction accuracy has been taken for granted as an indication of the existence of noise in the raw data.

In order to have the noise component removed, reduced, or separated from the main signal, one of the two following approaches has to be adopted. The first, is to fit an empirical model to the original data and separate the component that is not obeying the model (similar to ARMA type of models), and in this case any prediction using that model should be compared to the original data. Second, either the clean signal is known and the noise can be simply removed from the data or the true underlying dynamics governing the system is known and can be perfectly modeled. In this case also, the perfect model can be used to eliminate the noise. Following the second approach implies that any prediction process can be conducted using the clean signal. In hydrological signals (rainfall, streamflow, tree rings,..etc..), The second approach cannot be adopted because neither the clean signal nor the underlying dynamics of the system is known. On the other hand, the first approach might appeal to hydrologists (or any real data analysts). In stochastic hydrology, modelers adhere to the first approach by comparing the predicted

values to the observed ones, and also ensure that residuals (noise) are independent, and do not carry any structure that can be further modeled. In the reported applications of noise reduction in chaos-related hydrologic literature, the second approach has been apparently followed raising doubts about both components, the clean (modeled) signal and the left out (alleged noise) component. In light of the above discussion the reported publications with regard to noise reduction in hydrology are critically investigated. In section 6.4.5, the daily data of the English River at Umferville, Ontario, Canada are used to illustrate the viewpoint adopted in this thesis.

### *6.4.3 Noise reduction*

Good surveys of noise reduction methods that have been proposed to deal with time series data whose underlying dynamical behavior is characterized as low-dimensional chaos can be found in Kostelich and Schreiber (1993) and Davies (1994). Mostly, two methods have been adopted in hydrologic applications. First, the method proposed by Schreiber and Grassbeger (1991), which has been used by Porporato and Ridolfi (1997) and second, the method of Schreiber (1993), which has been employed by Kawamura et al. (1998), Sivakumar et al. (1999b). Jayawardena and Gurung (2000) have used also the method of Grassberger et al. (1993) in addition to the previous two methods. In order to understand the meaning of noise reduction, the methods of Schreiber (1993) and Schreiber and Grassbeger (1991) are briefly described.

*Schreiber and Grassbeger (1991):*

Backward and forward information are considered locally, reconstructing the dynamics in a $(\alpha+1+\beta)$ dimensional phase-space with $\alpha$-delay coordinates and $\beta$-forward coordinates.

$$\underline{y_i} = \left(y_{i-\alpha}, y_{i-\alpha+1}, \ldots\ldots, y_i, \ldots\ldots y_{i+\beta-1}, y_{i+\beta}\right) \tag{6.4}$$

For each point $\underline{y_i}$, The dynamics can be approximated locally in a $(\alpha+1+\beta)$-dimensional boundary $U^{(i)}$ with a linear model. The corrected value (noise reduced) can be calculated using the following formula:

$$y_i^{corr} = a^{(i)}\,\underline{y_i} + b^{(i)} \tag{6.5}$$

where $a^{(i)}$ *and* $b^{(i)}$ are coefficients dependent on the location in phase-space. For more details on obtaining the coefficients, the reader is referred to the original reference.

*Schreiber (1993):*

The basic idea is almost the same as the previous one, but using a simpler formulation for correcting the observations. Each measurement $y_i$ is replaced by the average value of this coordinate over points in the chosen neighborhood.

$$y_i^{corr} = \frac{1}{\left|U^{(i)}\right|} \sum_{U^{(i)}} y_i$$

From the above explanation and formulations, one can conclude that noise redu the analysis of chaotic time series means "smoothing" the signal rather than rem separating the noise component. A more representative name of this proc "suppressing", as used by Kantz and Schreiber (1997). It is useful to mention h the way noise is separated in ARMA modeling is different. ARMA model is fitte data and the residual is considered as the noise component, which is closer to the of "filtering" or "removing" than that pursued in the nonlinear noise reduction pr by Schreiber (1993) and Schreiber and Grassberger (1991). Therefore, the no noise reduction methods may be distorting the original data, and this is one of the that are raised in this research about noise reduction in chaotic hydrology. In the section, five articles reporting noise reduction in hydrology, are briefed and c reviewed with regard to the noise reduction process.

### *6.4.4 Noise reduction in hydrology*

*Porporato and Ridolfi (1997):*

The authors employed the method of Schreiber and Grassberger (1991) f reduction and they reported improvements in the estimation of the correlation di and prediction. The noise level removed was close to 6%. Unexpectedly, the co coefficient between the 24-hour forecasts and the observed ones for the noisy ti

is 0.95, which is very high, raising a doubt about the component that is perceived and removed as noise.

*Sivakumar et al. (1999a)*

In a comment on the work of Porporato and Ridolfi (1997), Sivakumar et al. (1999a), have indicated few problems with the noise reduction approach adopted by Porporato and Ridolfi. First, 6% noise level is too high to obtain that excellent accuracy of prediction, which was reported in the criticized work when noisy data were used. Second, after removing that considerable level of noise (6%), only marginal improvement in the prediction accuracy has been obtained, which is raising another doubt about the estimated and removed noise. Third, the number of iterations (200 iterations), used to obtain the parameters of the noise reduction algorithm that helped to achieve the reported improvements, is fairly large leading to over-correction of the observations. Overcorrection may have happened due to the use of inappropriate stopping criterion for the iteration procedure. Fourth, some of the absolute values removed as noise from observations are as high as 100 $m^3/s$ out of 700 $m^3/s$ and 50 $m^3/s$ out of 200 $m^3/s$, which is hard to explain in case of mean daily flow values. The comments of Sivakumar et al. on the work of Porporato and Ridolfi appear to be justified concerns. Finally, Sivakumar et al. (1999a) have proposed what they call a systematic noise reduction approach to handle the problems highlighted in the work of Porporato and Ridolfi (1997). A brief presentation and critique of their approach is given below.

*Sivakumar et al. (1999b):*

Based on the assumption that the optimal values of the parameters involved in the noise reduction methods depend on the level and nature of noise, Sivakumar et al. (1999b) proposed a approach for systematic noise reduction. Briefly, their approach includes additional two steps for stopping the iterations and noise removal. First, obtaining an initial estimate of the noise level using the method of Schouten et al. (1994), available within the RRCHAOS Software (Schouten and van den Bleek, 1994). Second, because of the underestimation of the noise level by the method of Schouten et al. (1994), the authors have tried certain lower and higher levels of noise around that initial estimate and continued removing different levels of noise, using the method of Schreiber (1993). They suggested the use of the improvement in the accuracy of prediction as a diagnostic tool and stopping the iterations. No more removal of noise is done when the prediction accuracy starts to deteriorate. The last estimated noise level is taken as the noise level in the data. Some doubts concerning the approach of Sivakumar et al. (1999b) are raised here and can be summarized as follows:

(a) The comment pertaining to over-correction that Sivakumar et al. (1999a) made on Porporato and Ridolfi (1997) has not been validated. Sivakumar et al. (1999a) mentioned that when a higher than the actual noise level is removed, the prediction accuracy decreases because of the removal of deterministic component. In Porporato and Ridolfi's case, they reported higher prediction accuracy after removing 6% noise. However, Sivakumar et al. (1999a) raised a doubt that over-correction might have happened in the work of Porporato and Ridolfi (1997).

(b) The initial estimate of noise level has been presented as one of the differences between the approach of Sivakumar et al. and that adopted by Porporato and Ridolfi. However, the final estimate of the noise level ends up being 2-3 times the initial estimate. The authors even admitted towards the end that having that initial estimate is not that significant.

(c) Consideration of the prediction accuracy as a diagnostic tool for verifying the removed noise is questionable. Such a diagnostic step implies that the prediction model is perfectly modeling the underlying dynamics. The prediction model used is the one suggested by Farmer and Sidorowich (1987), which yields nothing but the average values of the nearest neighbors to the point after which prediction is to be made. This averaging is again assuming a smooth function to be representing points in a certain neighborhood. Such a simplified model has been employed as a single prediction tool by Sivakumar et al. (1999b). In order to use prediction for verifying the noise reduction, either a variety of nonlinear models should be used or a single model that is known to be perfectly modeling the whole phenomenon (i.e., rainfall). One can easily conclude that what is confirmed to be correctly removed by the approach of Sivakumar et al. is the component that is not modeled by the specific model employed. In other words, what is removed can be considered noise to a single model, not the absolute noise.

(d) The way in which noise is removed and predictions are made indicates that such natural phenomena (e.g., rainfall) are expected to be highly deterministic, no much irregularities, and obeying smooth functions as close as possible. Applying this concept to hydrologic data for any further analysis (e.g., data extension, reservoir operation,

design of water structure) is highly risky, considering the lack of knowledge about the true signal.

(e) Elaborate analysis on the removed component has not been reported by Sivakumar et al. (1999b). Proving that the noise component has no structure, to be further modeled, is required. Autocorrelation function of the noise component and cross-correlation function between the noise and the "cleaned" signal might be good representatives of the internal structure of the noise series.

*Kawamura et al. (1998):*

Three different algorithms for reducing the noise have been used. Two of them are linear filters, moving average and low-pass filter, and the third is the nonlinear noise reduction method suggested by Schreiber (1993) and used by Sivakumar et al. (1999b). Graphs of the filtered series using the three methods along with the raw data are included from Kawamura et al. (1998) in this chapter to support the argument presented earlier. Even a visual check of Figures 6.11b and 6.11c versus 6.11a strongly supports the statement made by Schreiber and Grassberger (1991) that linear noise reduction techniques remove significant parts of the nonlinear systems. More importantly, the Schreiber's nonlinear noise reduction algorithm has reduced the variability of the data and thus has created less likely-to-happen straight-line segments in the data (Figure 6.11d). The authors have not used the filtered series for any further analysis, but if it is used like the case of Sivakumar et al. (1999b) and Porporato and Ridolfi (1997), then doubts raised about the reliability of the noise reduction in hydrology are legitimate, given the current state of knowledge about noise removal.

*Jayawardena and Gurung (2000):*

Two differences can be noticed in the work of Jayawardena and Gurung (2000). First, the authors tried three methods of Schreiber and Grassberger (1991), Grassberger et al. (1993), and Schreiber (1993) for filtering the data. But using different methods does not affect much of the argument regarding caution when applying such methods. According to Kantz and Schreiber (1997), "*most methods will reduce noise by a similar amount and their performance will not differ dramatically. Instead, the major criteria for the preferred algorithm will be robustness, ease of use and implementation, as well as the resources needed*". Second, the authors provided a comparison among different results when the system is described as chaotic versus stochastic. Chaos analysis has been applied to both raw and noise-reduced data whereas stochastic modeling has been applied to raw data only. If real additive noise exists in the data, why not have it removed also before modeling it stochastically? And if noise is only defined relative to the modeling approach – a component that cannot be captured by that specific model – then, what is the utility of comparing the predicted values of actual time series versus the predicted from the artificially noise-reduced time series?

*6.4.5 Case study 1: A single time series*

The daily data set of the English River at Umferville, Ontario, Canada, has been investigated and strong indications of the existence of chaos in its time series has been given in section 6.2. The same river is used in this section to support the argument concerning the noise reduction issue. The delay method of Takens (1981) has been used to reconstruct the dynamics in the phase-space.

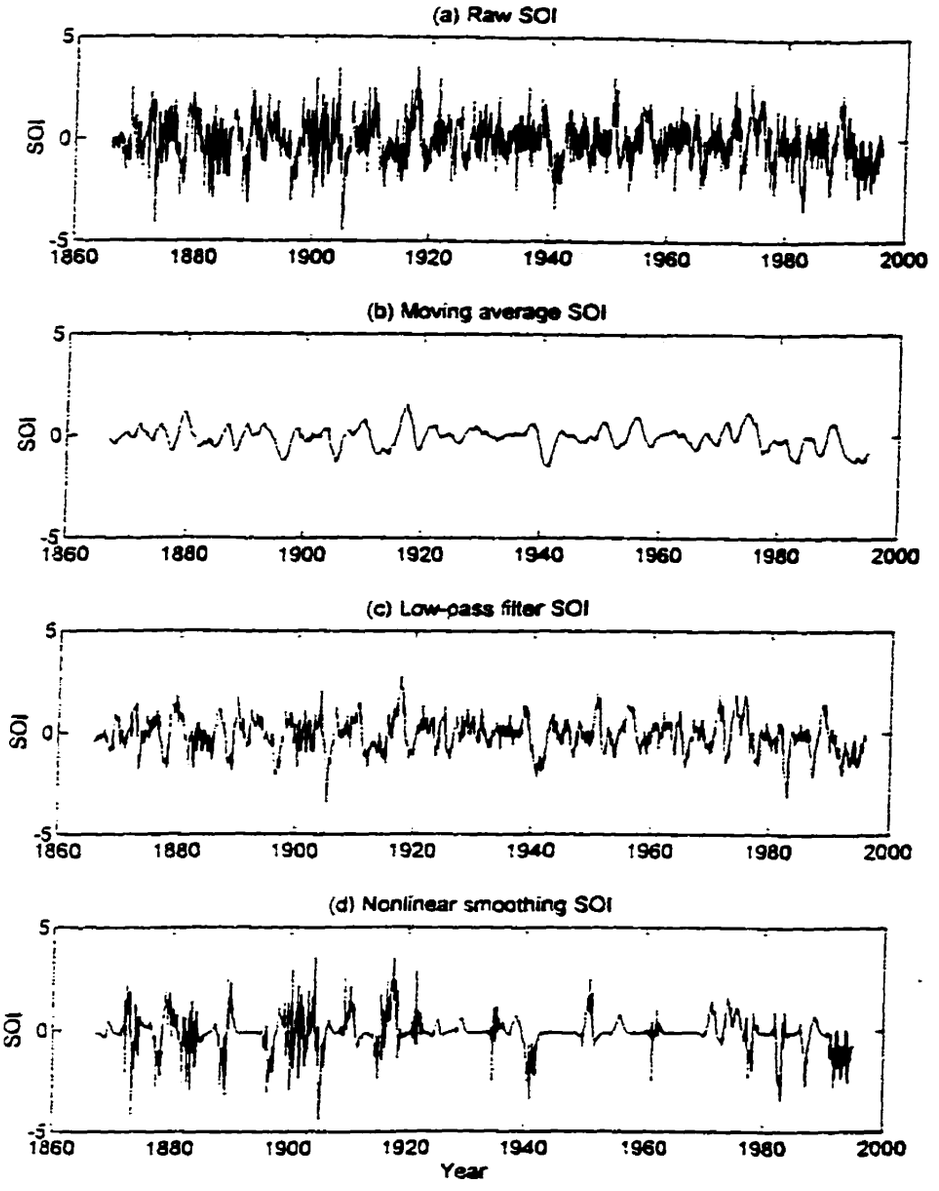



**Fig. 6.11 Time series of (a) raw SOI data and noise-reduced data by (b) moving average,(c) low-pass filter and (d) nonlinear smoothing [*after Kawamura et al. (1998)*].**

The dynamics of the scalar time series of the English River is found to be embedded in an m-dimensional phase space where m = 8 (Figure 6.3). A time-delay, $\tau$ , equal 10 is considered for the analysis.

*a) Noise reduction*

According to Kantz and Schreiber (1997), two methods of nonlinear noise reduction have been found to be reliable and effective on a broad variety of data sets. These methods are the simple algorithm by Schreiber (1993) and the locally projective scheme by Sauer (1992). In this section, the simple method of Schreiber (1993) is used through the TISEAN Software (Hegger et al. 1999). TISEAN allows for removing noise using a parameter (v) as a percentage of the standard deviation of the raw data*. The parameter (v) determines the neighborhood size for noise removal. The values of the parameter v considered in this research are 2%, 4%, 6%, 8%, 10%, and 50%. For consistency with the reported case studies of noise reduction in hydrology, the amount of noise removed is computed in a way that copes with the definition of noise given by Sivakumar et al. (1999b), as the ratio of standard deviation of the noise component to the standard deviation of the cleaned signal. Accordingly, the levels of noise removed in this case study from the English River data are: 0.2%, 1%, 2.1%, 4%, 5.5%, and 210%. The RRCHAOS, which employs the method of Schouten et al. (1994) has been used to get an initial estimate of the noise level in the raw data and also after each trial of cleaning the signal. According to Sivakumar et al. (1999b), the actual noise level is expected to be within 2-3 multiples of the RRCHAOS estimation. Table 6.3 summarizes the findings

* Personal communication with Holger Kantz, a co-author of TISEAN Software.

regarding the initial estimate of noise by RRCHAOS and also the inconsistency of estimating the noise level among different methods – what is removed as noise by the method of Schreiber (1993), is not accounted for as noise by the method of Schouten et al. (1994).

**Table 6.3. Noise removed and estimated by the methods of Schreiber (1993) and Schouten et al. (1994)**

| parameter (v) | 0.0 | 2% | 4% | 6% | 8% | 10% | 50% |
|---|---|---|---|---|---|---|---|
| Actual noise reduced | 0.0 | 0.2% | 1% | 2.1% | 4% | 5.5% | 210% |
| RRCHAOS[1] | 1.3% | 1.5% | 1.4% | 1.76% | 1.47% | 1.31% | 0.0% |

[1] estimation of noise level remaining in the data using RRCHAOS software.

*b) Nonlinear prediction*

In section 6.2, the K- nearest neighbor (K-nn) technique, as a local approximator, has been applied to estimate the missing data of the English River, and compared to the artificial neural network (ANN) model, as a global approximator. The ANN model has shown better performance than the K-nn technique. Since the K-nn, with various number of neighbors, is similar to the "$0^{th}$ order" model proposed by Farmer and Sidorowich (1987), another local approximator model is employed in this section. A Local Linear Model (LLM) approach is used and compared to the ANN model for predicting 10-days lead-time for the English River data.

*LLM approach (local fit)*

If it is assumed that the relation $x_{t+1} = f(\underline{x_t})$ is fulfilled by the data in good approximation for some unknown $f$, and that $f$ is smooth, predictions can be improved by fitting local linear models. According to Hegger et al. (1999), these models can be determined by minimizing:

$$\sigma^2 = \sum_{x_j \in U_t} \left(x_{j+1} - \underline{a_t}\,\underline{x_j} - b_t\right)^2 \tag{6.7}$$

with respect to $\underline{a_t}$ and $b_t$, where $U_t$ is the $\varepsilon$-neighborhood of $x_t$, excluding $x_t$ itself. Then, the prediction is:

$$\hat{x}_{t+1} = \underline{a_t}\,\underline{x_t} + b_t \tag{6.8}$$

where $\underline{a_t}$ and $\underline{x_t}$ are vectors while $b_t$ is a scalar value.

*Artificial neural networks (global fit)*

Feed forward neural networks that employ the back propagation technique, for training the network, are used. The structure of the input and output data of the ANNs is arranged in a way that makes use of the delay-time $\tau$. Since $\tau$ is taken equal to 10-days in this analysis, (m-1) inputs are used to predict the $m^{th}$ value and therefore, 10 consecutive days ahead can be predicted without using any of the predicted values for predicting the successive one. A similar technique for infilling consecutive missing data has been

employed in Elshorbagy et al. (2001a). Figure 5.3 of the previous chapter depicts the structure of the inputs and outputs to the ANN model. The data shown inside the rectangle represent the 10-days values to be predicted.

One of the concerns about the work of Sivakumar et al. (1999b) and Jayawardena and Gurung (2000) is the way that the authors evaluated the improvement in the prediction accuracy. Jayawardena and Gurung (2000) stated explicitly that only the last 25 data points of the time series were used for the prediction. So, the prediction accuracy is based on the average over these 25 observations only. Sivakumar et al. (1999b) did not explicitly mention their way of considering the prediction accuracy but it seems that they have used only that one window of predicted 100-data points for calculating the error. If the removed component is truly noise, then there will be consistency in the prediction accuracy, which is taken for granted by Sivakumar et al. (1999a) and Jayawardena and Gurung (2000) without verification.

In this thesis, the accuracy of the 10-days prediction is averaged over 10, 30, 60, and 100 days to test whether there is a real improvement in the prediction accuracy or not. This procedure has been repeated using the predictions obtained by both the LLM and ANN model. In this way, two shortcomings of the reported noise-reduction case studies in hydrology are avoided. The first one is the use of a single model for predicting and verifying the noise reduction scheme, and the second is estimating the prediction error over a single run of the prediction model. For convenience, the values of the parameter (v) of TISEAN Software are used as an indication of the noise level, in both discussion

and graphs. It should be noted that 0.0%, 2.0%, 4.0%, 6.0%, 8.0%, 10.0%, and 50.0% values of the parameter (v) correspond to 0.0%, 0.2%, 1.0%, 2.1%, 4.0%, 5.5%, and 210.0% of actual removed noise respectively. Figure 6.12 shows the change in the perception of the prediction accuracy when different numbers of data points are used to calculate the error (predicted values are obtained using LLM). The mean squared error (MSE) is used as a measure for the prediction accuracy. For example, there is an unjustified deterioration in the prediction accuracy when 6% noise level is removed (error is averaged over 10 and 30 days) and similarly at 4% noise level when error is averaged over 60 days.

Similarly, Figure 6.13 is demonstrating the same process using the ANN model. All ANN models have seven input nodes (= m-1) and one output nodes ($m^{th}$ coordinate of the embedding dimension). The number of hidden nodes are found, after many trials, to be as follows: 3, 10, 10, 8, 4, 7, and 5 for reduction of noise levels of 0%, 2%, 4%, 6%, 8%, 10%, and 50% respectively. Monotonic improvement in the prediction accuracy can be claimed only when 10 days are used with ANN model and 100 days are used with the LLM. It is believed that more reliability of the estimation of prediction error can be obtained when the error is averaged over more number of data points. Therefore the average over 100 days is considered in this research for assessing the prediction accuracy. Accordingly, one can say that there is a false deterioration of the prediction accuracy when 4% of the noise level is reduced. But improvement continues monotonically with more reduction of noise. The assertion of Sivakumar et al. (1999b) that prediction

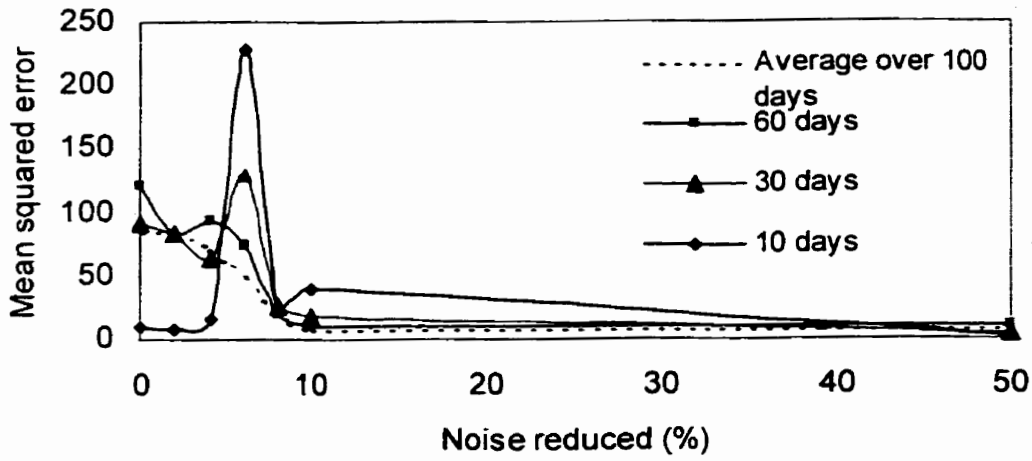



**Fig. 6.12 Change in prediction accuracy using Local Linear Model (LLM)**

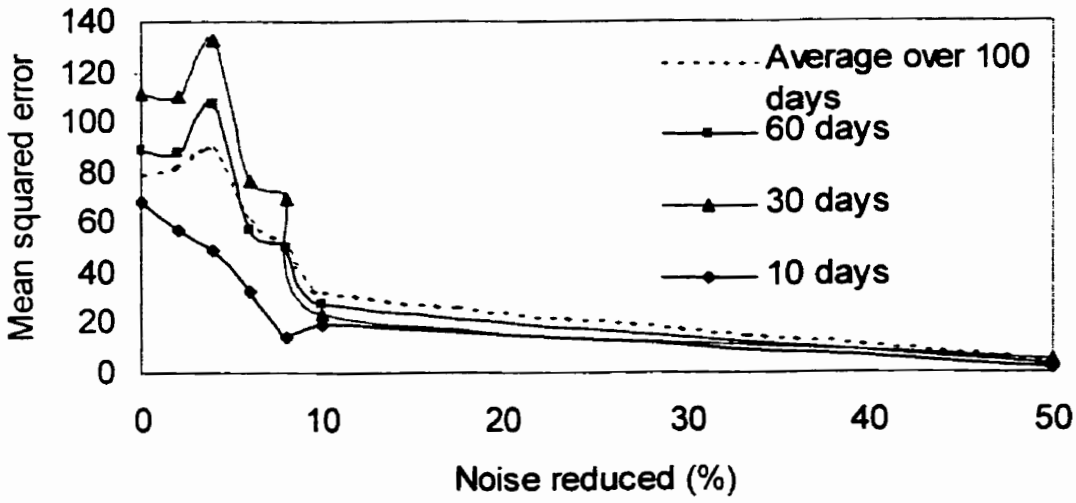



**Fig. 6.13 Change in prediction accuracy using Artificial Neural Networks (ANNs)**

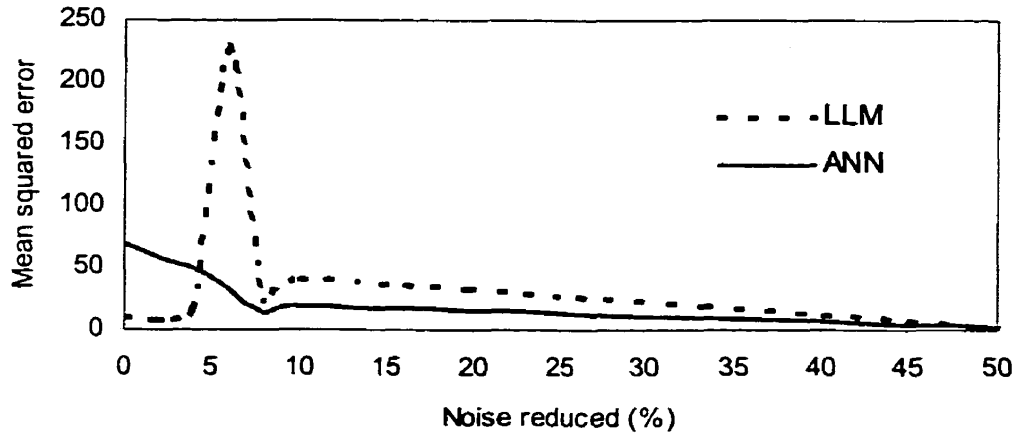



**Fig. 6.14 Change in prediction accuracy using ANNs and LLM (10 days are used to compute MSE)**

accuracy will deteriorate if over-correction occurred could not be validated in the case study of this thesis. Improvement continues until the actual level of reduced noise reached 210%. Figure 6.14 supports the argument regarding the relative perception of noise by different models and its relation to the period over which the error is averaged. What is considered as noise by LLM is not so according to the ANN model when 10 days are used to calculate the MSE.

According to the argument of noise reduction given by Porporato and Ridolfi (1997), the noise level in the data can be taken as (v = 10%), which is equivalent to actual 5.5% reduction of noise. The improvement in the prediction accuracy after this level is of considerably slower rate and significant improvement is achieved without removing much more of the original signal (Figure 6.15).

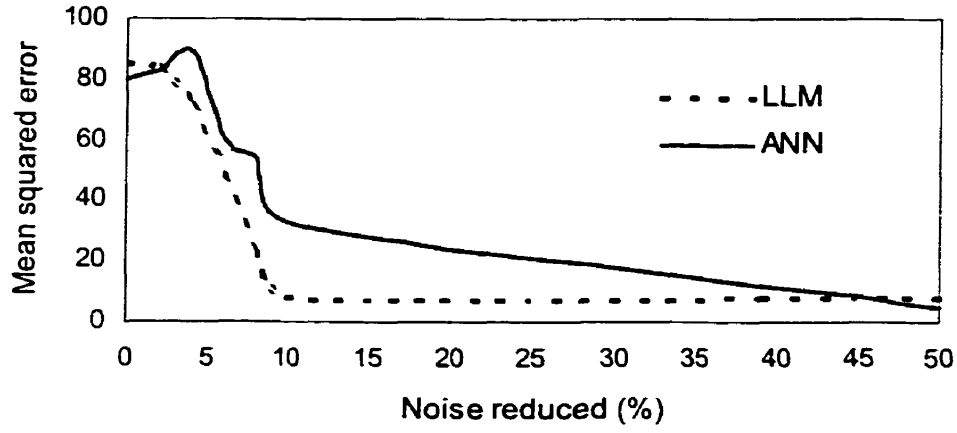
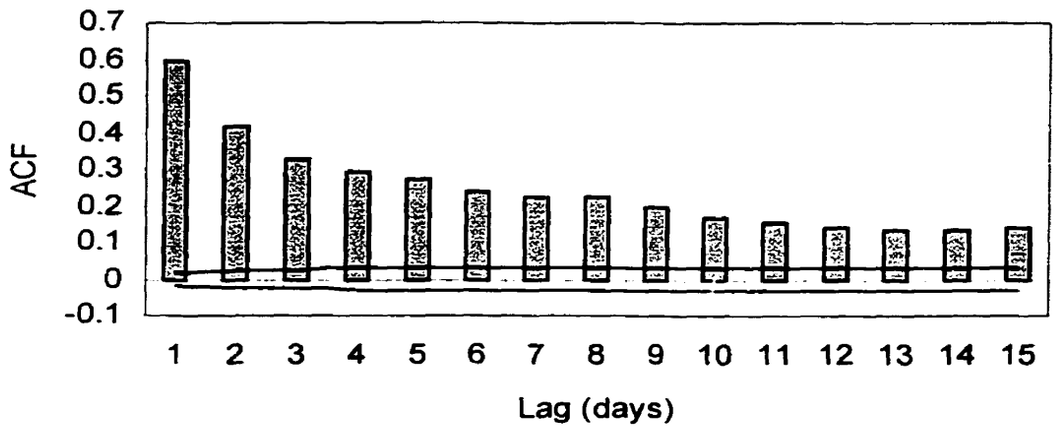



**Fig. 6.15. Change in prediction accuracy using ANNs and LLM (100 days are used to compute MSE)**

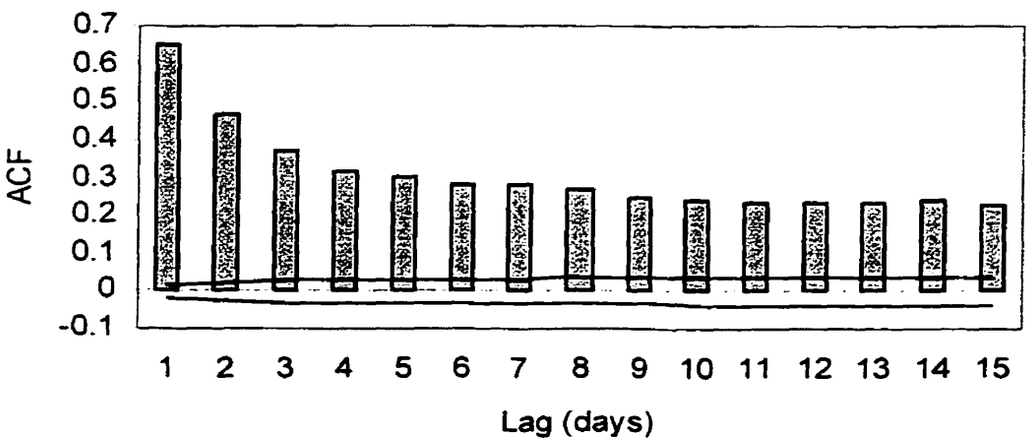
*c) How noisy is the removed noise?*

The noise reduction algorithm has to be verified by analyzing the removed noise component to ensure the absence of any further structure of the alleged noise. The autocorrelation function of the removed component (noise) is shown in Figure 6.16 for four different levels of noise. The figures indicate that the removed noise series are strongly autocorrelated. For 10% level of noise ($\nu$=10% and actual noise = 5.5%), for example, it is also found that there is a cutoff in the partial autocorrelation function (PACF) after lag-one suggesting a colored noise type of structure (Figure 6.17 shows the PACF of the time series of 10% noise level). Being a colored noise, treated as a regular time series, a linear time series model (ARMA) can be fitted to model the noise. The noise is modeled using autoregressive model AR(1). The actual removed noise and the

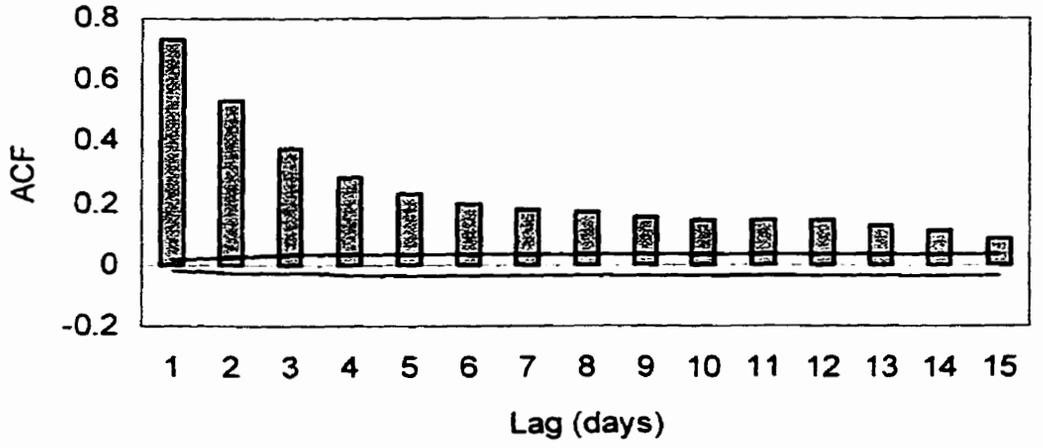
modeled one are shown in Figure 6.18. Therefore, a strong doubt is raised, in this thesis, about the structure of the component removed by the nonlinear noise reduction method of Schreiber (1993).



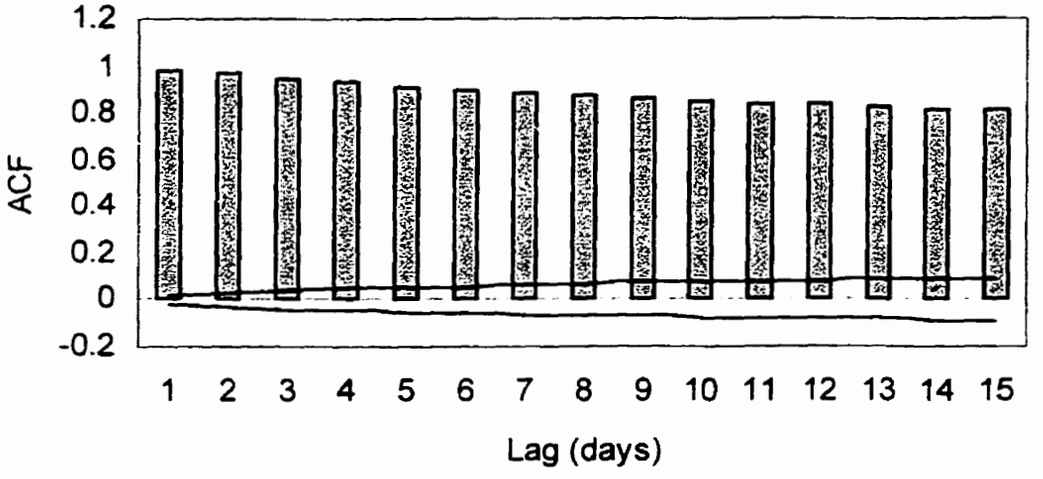

a)




b)




c)




d)

**Fig. 6.16 Autocorrelation function of the noise time series of (a) 2% noise, (b) 4% noise, (c) 10% noise, and (d) 50% noise**

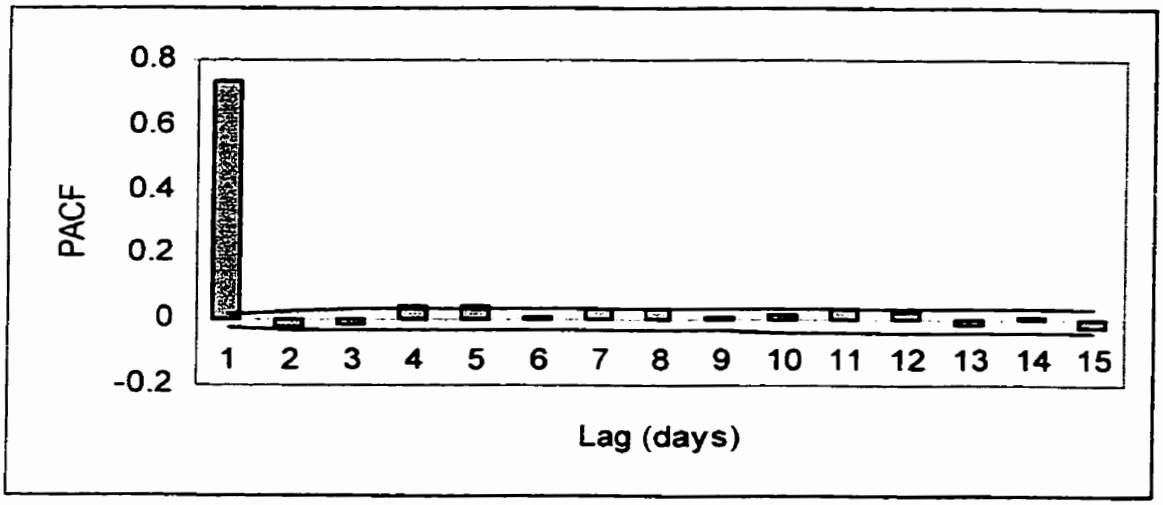



**Fig. 6.17 Partial autocorrelation function of the noise time series of 10% noise level**

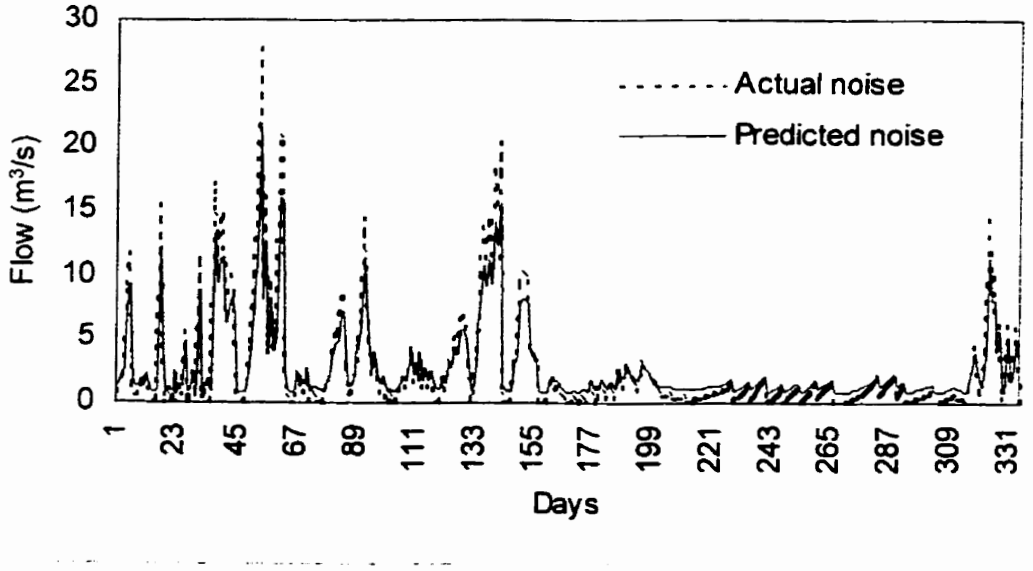



**Fig. 6.18 Actual and predicted values of the noise series using AR(1)**

*d) Preservation of the underlying dynamics*

Sivakumar et al. (1999b) and Jayawardena and Gurung (2000) reported the necessity of verifying the noise reduction method by plotting the phase space and computing the correlation dimension for both noise free and noisy time series. Jayawardena and Gurung (2000), using the visual verification, mentioned that significant changes in the shape of the phase-space diagrams as they became less noisy imply the enhancement in determinism. However, such a graph like Figure 6.19 (after Jayawardena and Gurung, 2000) does not reflect any indication of maintaining the original structure or dynamics of the raw data. In the case of the English River data, the correlation dimension for both the raw data and the 10% noise-reduced data are estimated. Figure 6.20 shows that less variations are observed in the noise-reduced data, however, the embedding dimension (m = 8) can be obtained from both. The fractal dimensions of the noisy attractor and the noise-reduced one are 2.39 and 2.26 respectively. The significance of such a fraction difference is beyond the scope of present discussion. In conclusion, reducing the noise may give better visualization of the correlation dimension plot, nonetheless, the difference is of little importance to the practical use in hydrologic applications. This conclusion is supported by the statement of Sivakumar et al. (1999b) that an enlarged scaling region (for estimating the correlation dimension) could also be observed in the correlation dimension plots even after disturbing the clean underlying deterministic signal (by noise).

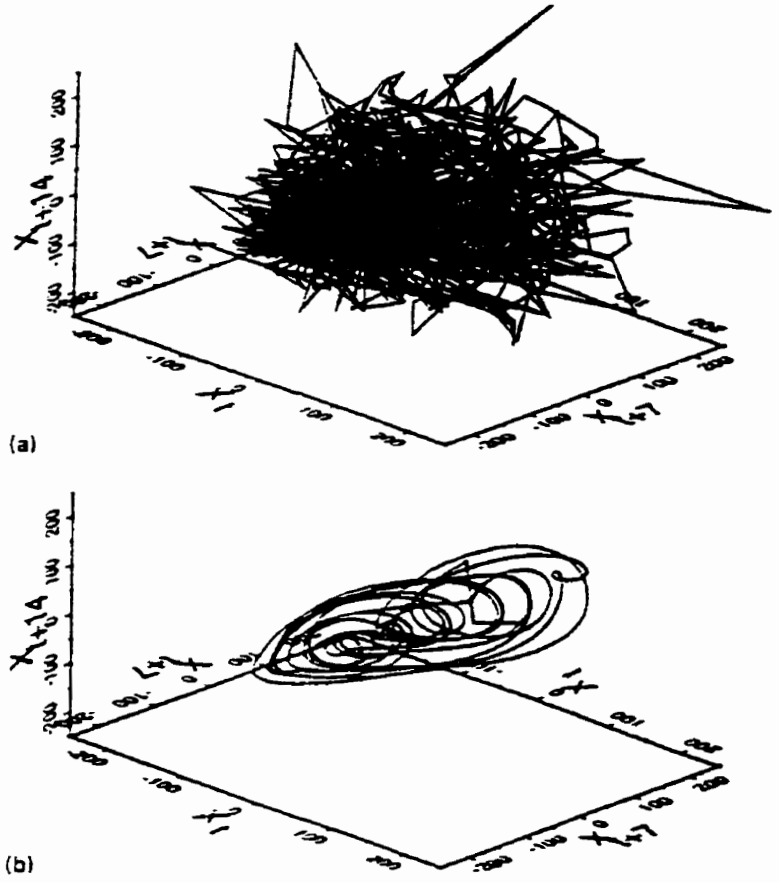



**Fig. 6.19 Phase-space diagram of sea surface temperature (S Index) data. (a) raw data, (b)after noise removal [after Jayawardena and Gurung (2000)].**

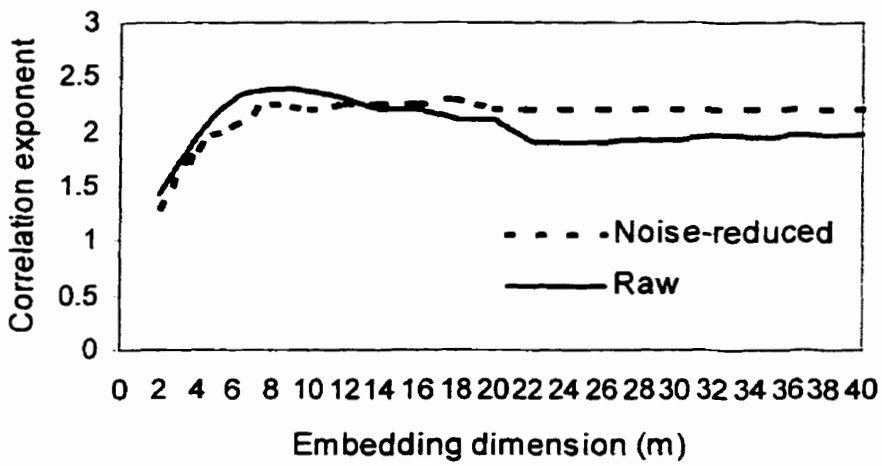
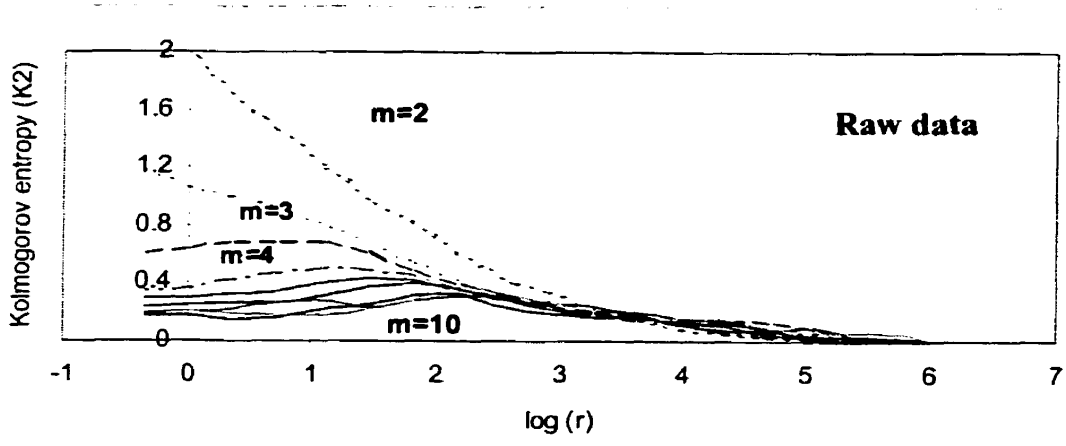


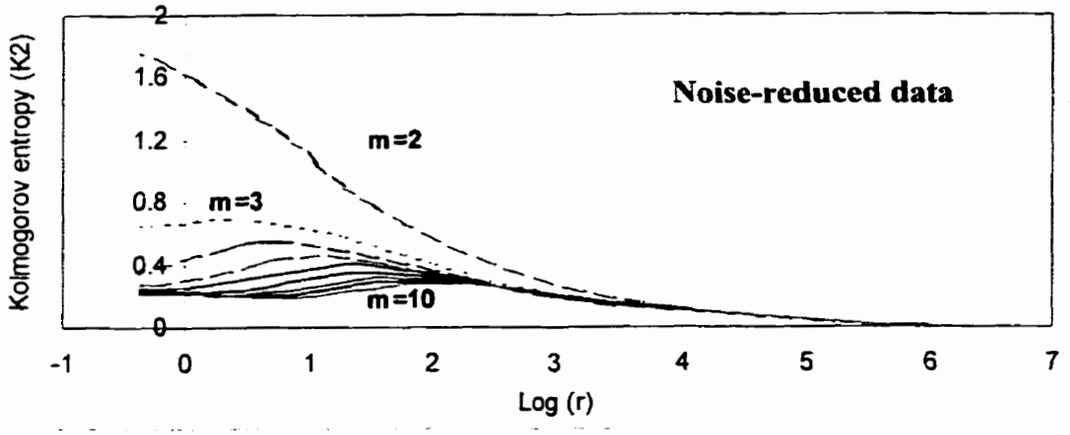

**Fig. 6.20 Correlation exponent (dimension) of the raw and noise-reduced data of the English River.**

In this section, for comparison, the Lyapunov exponents and the Kolmogorov entropy are estimated for both raw and 10% noise-reduced data. The largest Lyapunov exponents of the raw and noise-reduced data are 0.0091 and 0.044 and the sums of all positive exponents ($h$) in both cases are 0.012 and 0.085 respectively. The inverse value of $h$ denotes a mean time scale up to which predictability ($T_p$) may be possible (Fraedrich, 1987). The time $T_p$ is ~ 83 and ~12 days for raw and noise-reduced data respectively. As it has been shown earlier that the prediction accuracy, apparently, improved after removing the assumed noise but the time scale, within which predictability is possible, has been shortened. Two possible reasons can be given for this phenomenon, first, the removed component is strongly autocorrelated, and the removal of it has affected the

length of the $T_p$. Second, artificial chaoticity may have been added to the data, shortening the time within which predictability is possible (in harmony with the definition of chaotic systems making only short-term prediction possible). Figure 6.21 shows the correlation entropy for various embedding dimensions (m). A plateau can be observed in both cases, raw and noise-reduced data. Curves for different m values are merging together, except for m= 2, creating that horizontal plateau, which is an indication of existence of chaos (Kantz and Schreiber, 1997). Probably, that plateau can be detected easier with the noise-reduced data.




a)




b)

**Fig. 6.21 Entropy of the English River data. (a) raw data and (b) 10% noise-reduced data**

*Discussion*

The fact that hydrologic data may be contaminated by the measurement noise is beyond doubt. The process of measuring, averaging, and converting the data into computer file-format contribute to the process of introducing noise to the data. The doubts raised in this thesis, regarding the noise reduction process in hydrologic data, are associated with the way it is conducted. Revisiting some of the literature on noise reduction in the field of physics, one can notice that the major concern is better understanding of the system rather than using the filtered data for further analysis. It has been mentioned that the idea of separation of noise is to understand the system (Schouten et al. 1994) while computation of the correlation dimension, which is affected by noise, was the concern of

Schreiber and Grassberger (1991). Also, it was indicated that less noisy trajectory should be consistent with the dynamics as well as with the measured data (Schreiber, 1993). It implies that getting the less noisy attractor helps understand the system and estimate different invariants rather than replacing the measured data. Only in Kostelich and Schreiber (1993), the possibility of having prediction before and after noise reduction was mentioned.

As suggested in section 5.3 of the previous chapter, the correlation dimension, as one of the chaotic invariants, can be used to configure the model required to predict or infill the missing data. Hence, if noise reduction can help facilitate the computation of the chaotic invariants, and therefore, identify the correct model configuration then it will be of a valuable significance to the modeling process. However, noise reduction should not be used as a way to manipulate the original data and make it obedient to a specific, pre-selected model (Elshorbagy et al. 2001b).

The prediction of 10-days ahead by the LLM and ANN models using both raw and 10% noise-reduced data are shown in Figure 6.22. Also the Mean squared error (MSE) and the mean relative error (MRE) of the prediction averaged over 100 days are given in Table 6.4. The extremely high prediction accuracy obtained using the noise-reduced data are illusive because artificial chaoticity may have been introduced to it due to the filtering process. Also the significant difference between two different models (LLM and ANNs) in the rate of the prediction accuracy improvement after noise reduction raises doubts concerning the perception of noise by different models. It is believed that accepting 5%

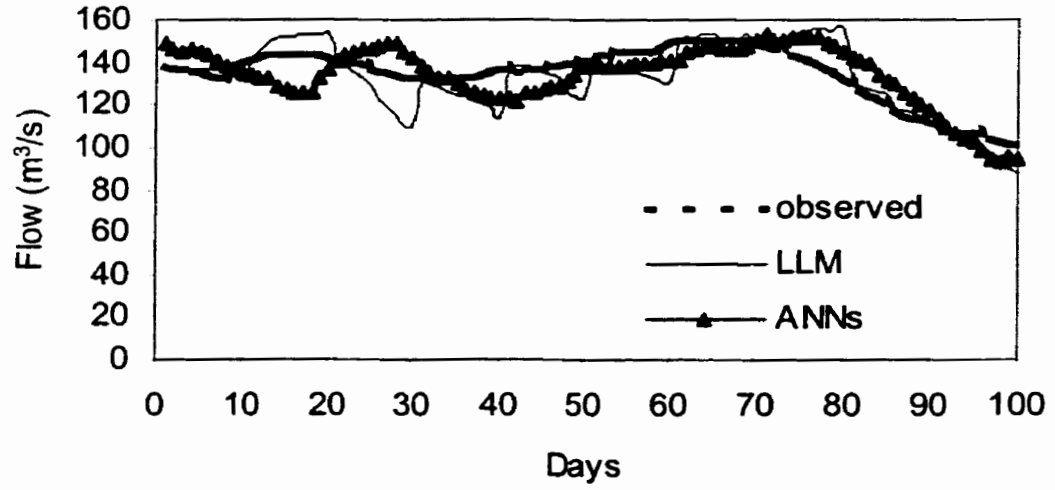
or 6% relative error in the prediction with certainty that the true data are used is better than obtaining 1.3% relative error using highly "uncertain" filtered data. It is also noted that the LLM applied in this thesis performed better than the traditional nearest neighbor model used in section 6.2, which is similar to the method of Farmer and Sidorowich (1987). The promising results of ANNs and LLM encourage to devote efforts towards establishing better models for chaotic data rather than filtering the data to cope with the available models (Elshorbagy et al. 2001b).

**Table 6.4. Prediction accuracy of the raw and noise reduced data of the English River.**

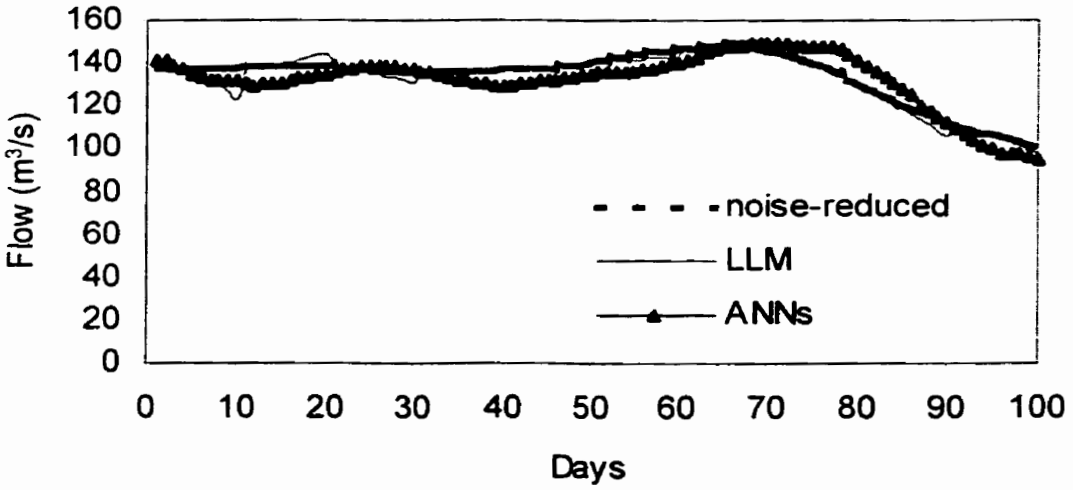
| | Raw data | | 10% noise-reduced | | % |
|---|---|---|---|---|---|
| | MSE | MRE % | MSE | MRE % | improvement |
| LLM | 85 | 5 | 8 | 1.3 | 91 |
| ANN model | 80 | 6 | 33 | 3.8 | 59 |

The issue of noise reduction assumes a high significance whenever a result depends sensitively on the details of the embedding procedure (Kantz and Schreiber, 1997). Further, in hydrologic data, when a record length of years measured in a specific way, it is highly useful to treat it as the real observations for further analysis. In hydrologic analysis, when reservoir operation, design of water structures, prediction, and others are concerned, there is a high chance that future observations will be obtained in the same fashion (noise included). Therefore, it is argued here that filtering the data and conducting analysis based on that artificially produced time series might be of less utility

in hydrologic applications where the noise-free data are not expected to be widely available.




a)




b)

**Fig. 6.22 Predicted data of the English River using LLM and ANNs. (a) predicted versus observed data and (b) predicted versus noise-reduced data.**

*6.4.6 Comparison between two noise reduction algorithms*

The issue of noise reduction in hydrologic data and its reliability is discussed in the previous section based on the results obtained by two noise reduction methods: Schreiber and Grassbeger (1991) and Schreiber (1993). The noise reduction method, reported by Grassberger et al. (1993), is employed in this section to complete the analysis regarding the noise issue and also is compared to the method of Schreiber (1993). The method of Grassberger et al. (1993) is different from other nonlinear noise reduction methods in the way it perceives and removes the noise. This method is briefed below.

*Locally projective nonlinear noise reduction method, Grassberger et al. (1993):*

It is a more sophisticated method that makes use of the hypotheses that the measured data are composed of the output of a low-dimensional dynamical system and of random or high-dimensional noise. The effect of noise is to spread the data off the low-dimensional manifold. The idea of the projective nonlinear noise reduction scheme is to identify the manifold and to project the data onto it (Hegger et al. 1999).

Suppose the dynamical system form a q-dimensional manifold $V$ containing the trajectory. All embedding vectors $\underline{y_i}$ would lie inside another manifold $\tilde{V}$ in the embedding space. For each $\underline{y_i}$ there exists a correction $\Delta\underline{y_i}$, with $\|\Delta\underline{y_i}\|$ small, in such a way that $\underline{y_i} - \Delta\underline{y_i} \in \tilde{V}$ and that $\Delta\underline{y_i}$ are orthogonal on $\tilde{V}$. So, vectors have to be over-embedded in m-dimensional spaces with m > q.

Both of the nonlinear noise reduction schemes are applied through the TISEAN Software (Hegger et al. 1999). As mentioned earlier, TISEAN allows for removing noise using a parameter (v) as a percentage of standard deviation of the raw data or a parameter (r) as a radius of neighborhood. Using the method of Schreiber (1993), the level of removed noise can increase monotonically by increasing the level of (v); whereas in the method of Grassberger et al. (1993), the amount of removed noise increases with the decrease of the projection dimension (q). The maximum level of noise reduction is achieved with $q = 2$, which is the minimum possible value for q. The removed noise is considered to be the ratio of the standard deviation of the noise component to the standard deviation of the cleaned signal. Accordingly, the amounts of noise removed from the raw data of the English River are estimated, after many trials, to be 5.5% and 2.2% using the methods of Schreiber (1993) and Grassberger et al. (1993), respectively.

The mean squared error (MSE) and the mean relative error (MRE) of the estimated data of the English River corresponding to the raw data and cleaned signals using the ANN models are tabulated in Table 6.5. It should be noted that the missing data are supposed to be ten missing consecutive observations at a time, which are estimated as one group.

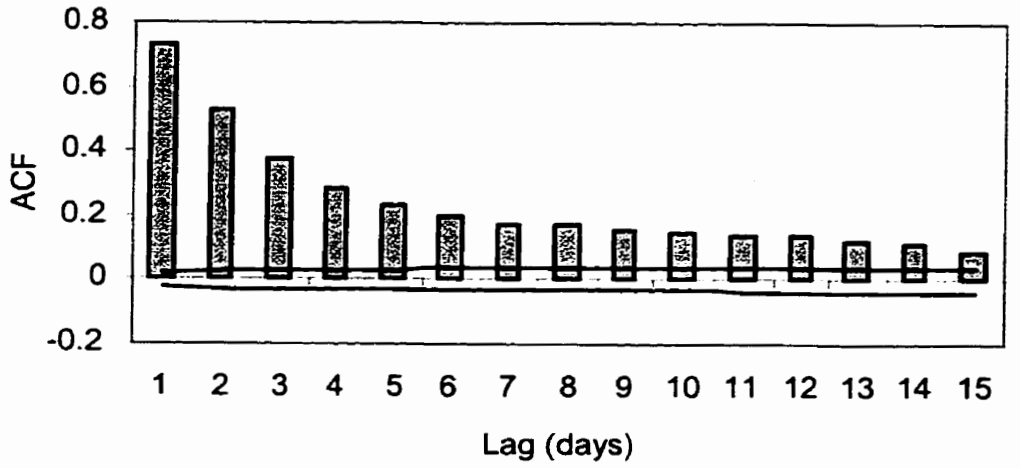
**Table 6.5. Estimation accuracy of the raw and noise-reduced data of the English River**

| Model | Raw data | | Noise-reduced (Schreiber 1993) | | Noise-reduced (Grassberger et al. 1993) | |
|---|---|---|---|---|---|---|
| | MSE | MRE % | MSE | MRE % | MSE | MRE % |
| ANNs | 80 | 6 | 33 | 3.8 | 93 | 4.3 |

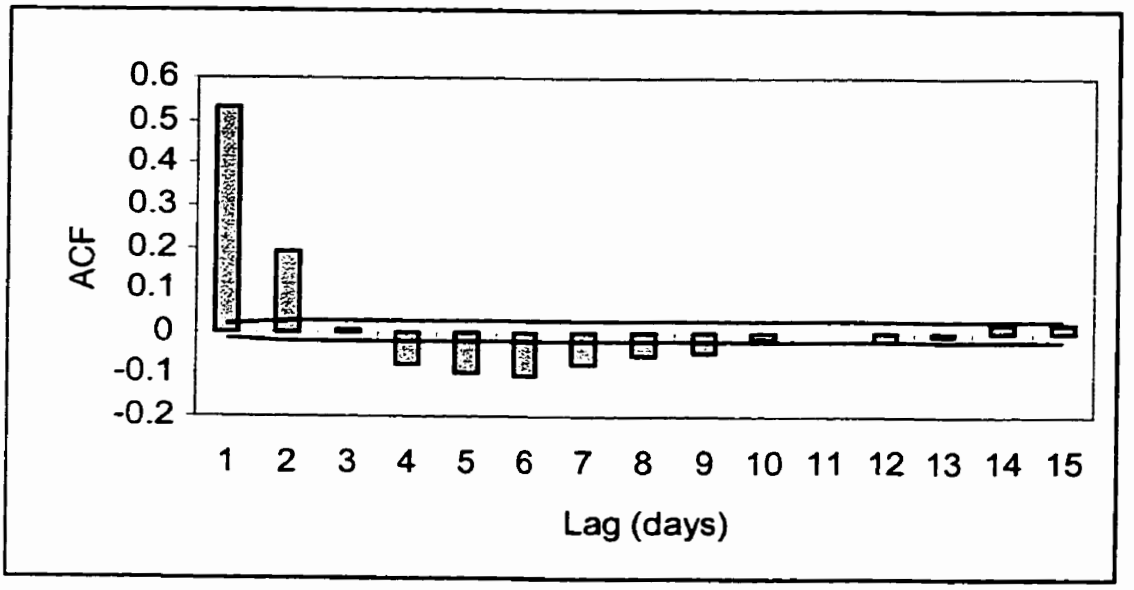
*Discussion*

The results given in Table 6.5 show a significant improvement in the accuracy of estimating the missing data when noise is reduced using the method of Schreiber (1993). Using the method of Grassberger et al. (1993), a little improvement in the MRE is achieved while some deterioration of the MSE occurs. Therefore, one cannot conclude that the method of Grassberger et al. (1993) has led to any improvement in the estimation accuracy of the missing data.

The removed component (noise) is not expected to have any structure that can be modeled. In this section, the removed noise, using both of the nonlinear noise reduction methods, is tested with regard to its autocorrelation structure. Figures 6.23 and 6.24 show the autocorrelation (ACF) and the partial autocorrelation (PACF) functions of the noise component removed by the noise reduction methods under consideration. The slow decay of the ACF in Figure 6.23a indicates the existence of significant autoregressive component that can be modeled by ARMA-type models. The absence of such a slow decay of ACF in Figure 6.23b indicates that the noise removed by the method of Grassberger et al. (1993) abides by the definition of noise to be removed. On the other hand, as shown in Table 6.5, no significant improvement in the accuracy of estimating the missing data is achieved using the noise-reduced data that is filtered by the method of Grassberger et al. (1993). This apparently marginalizes the utility of employing such a method for noise reduction (Elshorbagy et al. 2000d). The method of Schreiber (1993) that has improved the accuracy of estimating the missing data seems to be removing a stochastic component, which can be further analyzed and modeled.

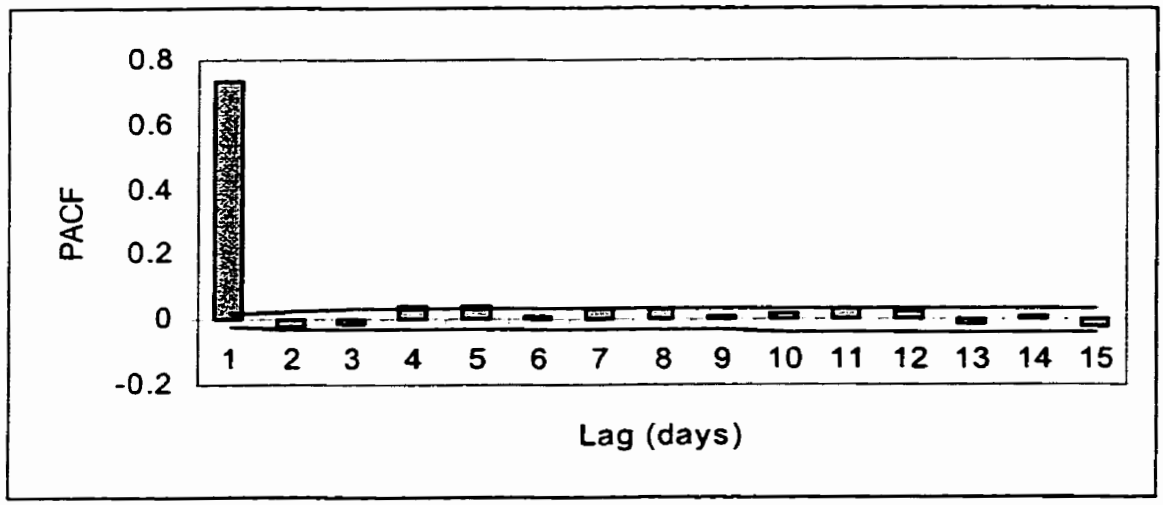



a)

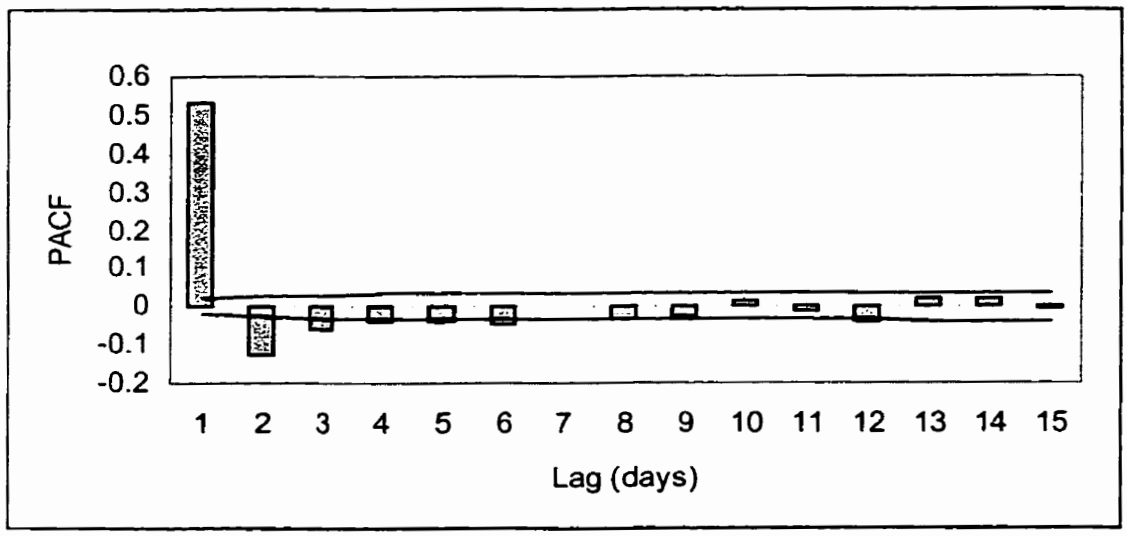



b)

**Fig. 6.23 Autocorrelation function of the removed noise, a) method of Schreiber et al. (1993) and b) method of Grassberger et al. (1993).**




a)




b)

**Fig. 6.24 Partial autocorrelation function of the removed noise, a) method of Schreiber et al. (1993) and b) method of Grassberger et al. (1993).**

*6.4.7 case study 2: two cross-correlated series*

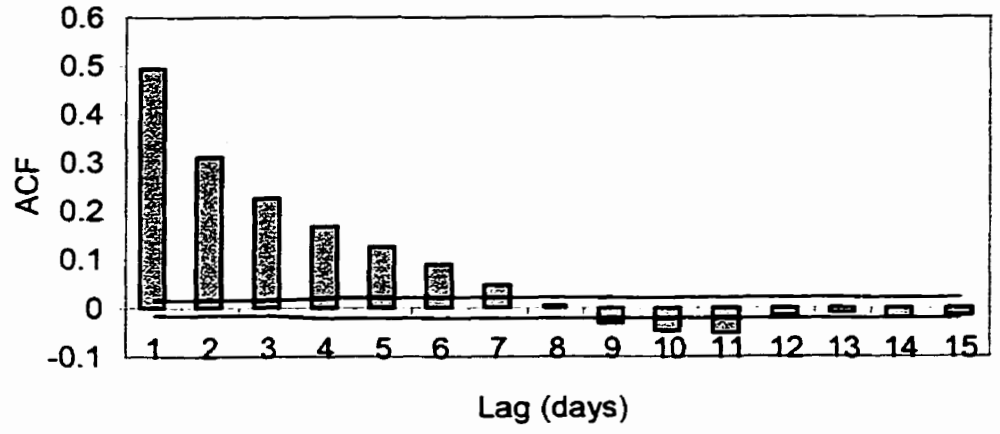
The locally projective nonlinear noise reduction method (Grassberger et al. 1993) is applied in this section to reduce the noise of both *target* and *reference* rivers. Analysis of the raw data has been conducted in section 6.3. Four different values of q are considered and consequently the removed noise level is calculated. For each noise-reduced data set, the cross-correlation between the Little River and the Reed Creek is calculated and summarized in Table 6.6. The stopping criteria for removing noise is a disputable issue in chaos-related literature. Porporato and Ridolfi (1997) have kept on removing noise until the amount of noise removed is not worth the time and effort consumed to do so; whereas Sivakumar et al. (1999b) have used the improvement in the prediction accuracy as an indication of the correctness of the level of the reduced noise. Both approaches have been critically reviewed in the previous section, and their deficiencies have been highlighted.

**Table 6.6. Different levels of reduced noise and the corresponding cross-correlation values.**

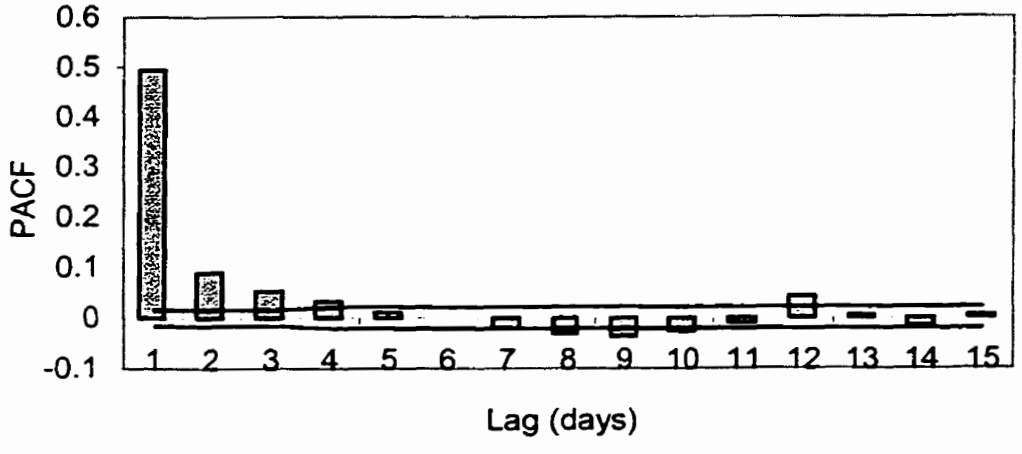
| Projection dimension (q) | Noise level reduced (%) | | Cross-correlation |
|---|---|---|---|
| | Little River | Reed Creek | |
| 19 = m | 0.0 | 0.0 | 0.733 |
| 17 | 2 | 1.9 | 0.733 |
| 15 | 3.7 | 3.7 | 0.732 |
| 10 | 9.1 | 9.5 | 0.720 |
| 5 | 18.7 | 20.6 | 0.709 |

In this section, the removed noise is assessed in two ways. First, the autocorrelation function (ACF) and the partial autocorrelation function (PACF) of the noise component are investigated to ensure the randomness and the absence of internal structure. Second, the cross-correlation between the noise-reduced signals of the two time series is calculated to observe the effect of noise reduction on the cross-correlation structure of the two rivers. At q values of 10 and 5, the cross-correlation starts to deteriorate (Table 6.6). Further, a slow decay of the ACF and sudden cutoff of the PACF are observed, which is an indication of the possible existence of autoregressive-type structure. At q values of 17 and 15, the cross-correlation does not decrease and less significant autocorrelation exists. Therefore, a noise level of 3.7%, which is achieved at q value of 15, is considered in this analysis (Figures 6.25 and 6.26).

The experiments conducted with the raw data in section 6.3 are repeated with 3.7% noise reduced time series. The actual and estimated missing data, using LR and ANN models, of the Reed Creek are shown in Figure 6.27. One can observe that using multiple local models (sub-models) can improve the accuracy of estimating the missing data.

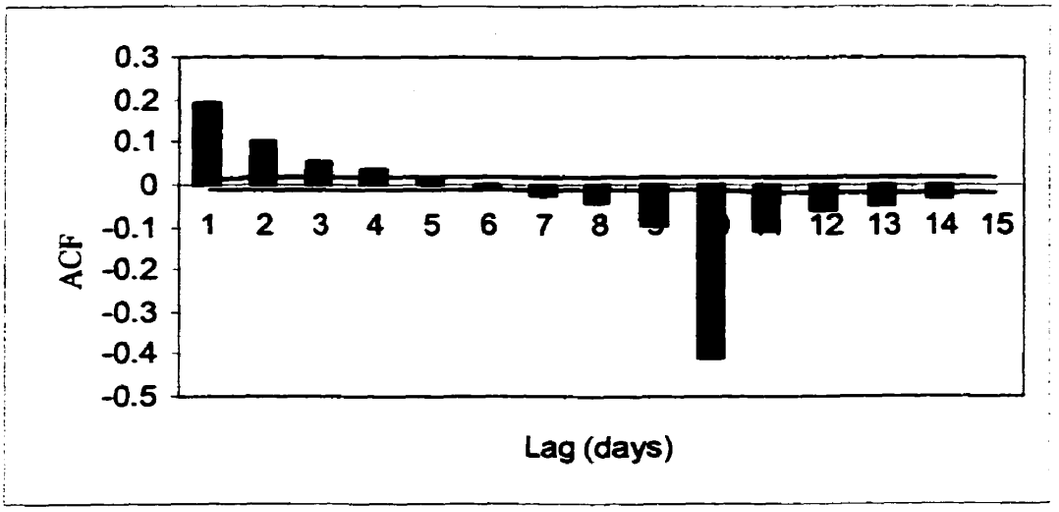



a)

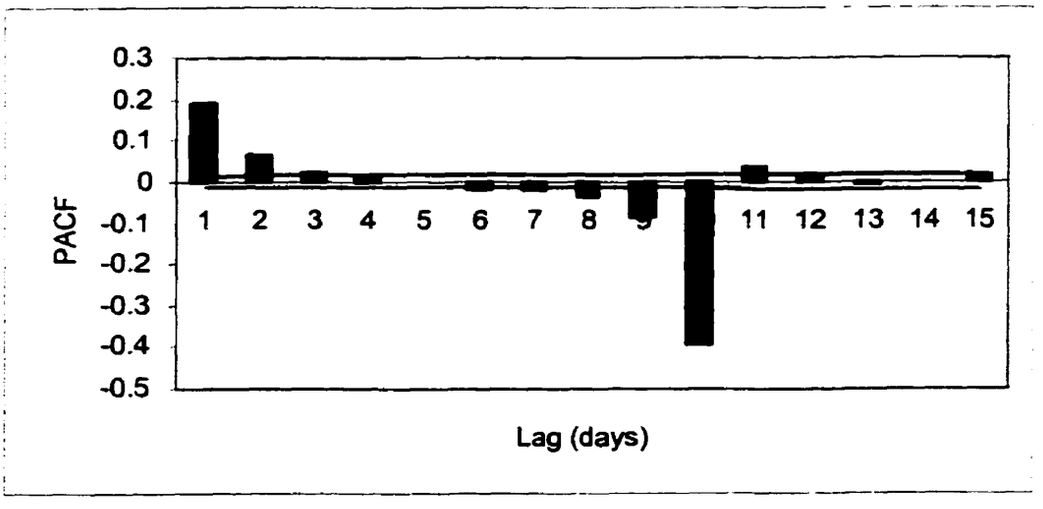



b)

**Fig. 6.25 Correlation structure of the 18.7 % (q=5) noise removed from the Little River data, a) autocorrelation function and b) partial autocorrelation function**

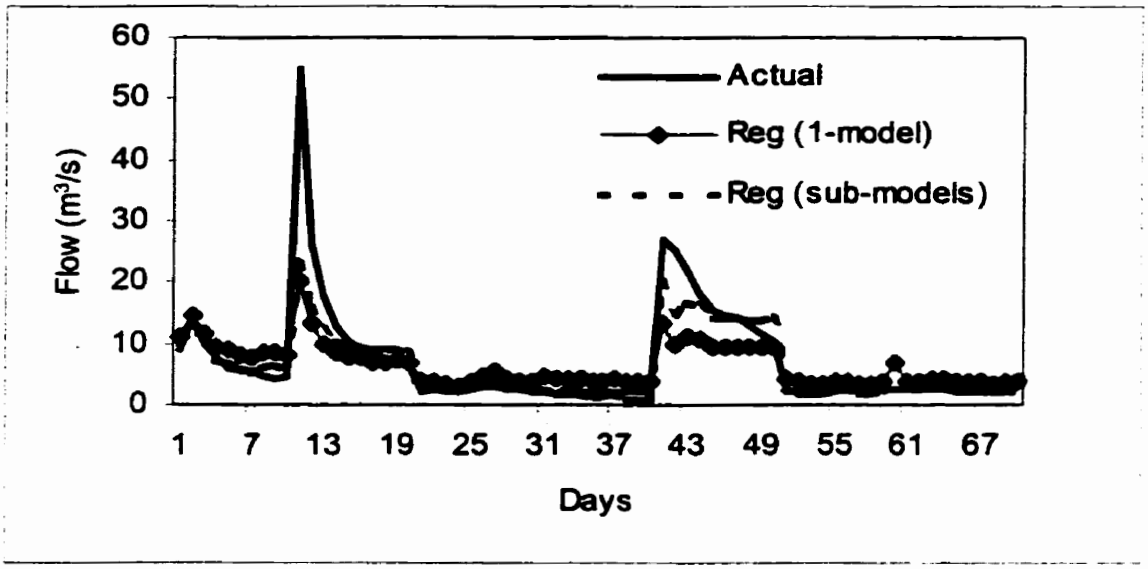



a)

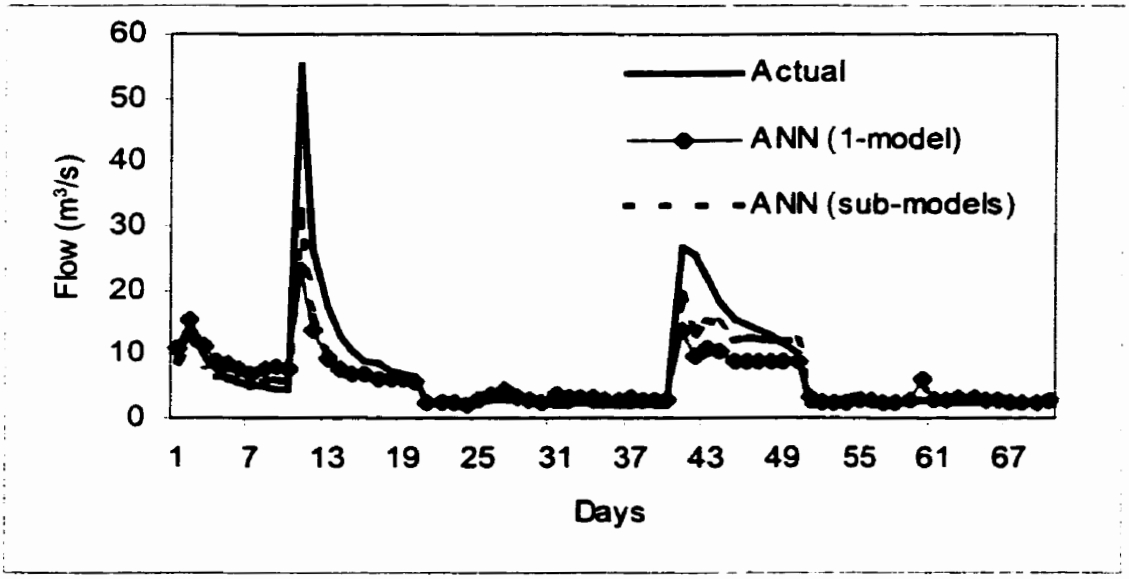



b)

**Fig. 6.26 Correlation structure of the 3.7 % (q=15) noise removed from the Little River data, a) autocorrelation function and b) partial autocorrelation function**




a)




b)

**Fig. 6.27 Actual and estimated data of the Reed Creek noise-reduced data using a) linear regression and b) artificial neural networks**

*Discussion*

The MSE and the MRE of estimating the missing data resulting from different experiments are given in Table 6.7. It is evident that ANN models are superior to the LR models given that the experiment type is fixed (i.e., either one global model or local models are considered).

**Table 6.7. The mean squared and the mean relative errors of the estimated values.**

| Experiment | Raw data | | Noise-reduced data | |
|---|---|---|---|---|
| | MSE | MRE | MSE | MRE |
| **LR (global model)** | 32.30 | 0.55 | 33.49 | 0.54 |
| **LR (local models)** | 19.64 | 0.23 | 20.65 | 0.24 |
| **ANN (global model)** | 26.79 | 0.31 | 29.78 | 0.30 |
| **ANN (local models)** | 15.29 | 0.17 | 15.00 | 0.18 |

The values presented in Table 6.7 indicate also that the process of nonlinear noise reduction has not helped in improving the accuracy of estimating the missing data. Neither the MSE nor the MRE has shown the possibility of improving the correlation between the *reference* and the *target* rivers after reducing the "alleged" noise. This result is harmonious with the fact that the cross correlation values, given in Table 6.6, do not increase with the increase of the removed noise level. The cross-correlation value starts to deteriorate when higher levels of "alleged" noise are decreased, which raises serious

doubts about the utility and validity of the currently used noise reduction concept in chaos-related applications in hydrology.

## 6.5 Summary and concluding remarks

The existence of chaotic behavior (nonlinear deterministic dynamics) in the daily flows of the English River, Ontario, Canada and the Little River and Reed Creek, Virginia, USA, has been investigated. The correlation dimension, the Lyapunov exponent, the Kolmogorov entropy, and the method of surrogate data have been used in the analysis to test the existence of chaos. There are sufficient indications to believe that these streamflows have some chaos and the data could be modeled by the time delay embedding method. Based on the attractor dimension, the minimum and sufficient number of variables, essential to model the dynamics of the daily flows can be approximately identified.

The ANN models show superiority in the accuracy of estimating the missing data, over the K- nearest neighbor and local linear models, which is attributed to the capability of the ANNs to capture the nonlinear dynamics and generalize the structure of the attractor over the whole data set. This result is inferred when the model is applied to the whole data set. More significantly in chaotic time series, local models are preferred over global models in terms of the MSE and MRE criteria, when global models that are applied to the whole data set are compared to local models applied to data sections in the direct vicinity

of the missing data. The last result is inferred from the application of the chaos theory to the case of two cross-correlated time series.

Few different nonlinear noise reduction techniques have been employed in this thesis to investigate the effect of noise reduction on the analysis of chaotic time series. The noise reduction applications in hydrologic literature have been critically reviewed and some doubts are raised in this thesis about the use of the prediction accuracy as a tool for diagnosing the level of noise. The promising results of both models, ANN and LLM, in the case of single time series and the ANN and LR models in the case of bi-series encourage the process of seeking tools to handle the raw chaotic time series for obtaining better results. Noise reduction process can help, sometimes as in the case of the English River, in facilitating the process of estimating the chaotic invariants, and therefore configuring models such as the ANNs. Smoothing the original data, by noise reduction, and using them for further analysis is a questionable approach and therefore should be discouraged. However, seeking tools for perceiving and removing the noise from hydrologic data can be of a significant value to the field of water resources, given that the removed component can be proved, beyond doubt, to be a white noise.

# Chapter 7

# 7. CONCLUSIONS

7.1 Summary of the thesis

7.2 Research contribution

7.3 Possible research extension

7.4 Study limitations

---

*" Disputing with thee concerning the truth after it was made manifest, as if they were being driven to death while they see it."*

*The Glorious Qura'n (8: 6)*

## 7.1 Summary of the thesis

This thesis comprises two parts; the first part, which covers up to chapter 4 relies mainly on concepts of statistical and stochastic hydrology. However, a nonparametric modeling technique such as the artificial neural networks (ANNs) is employed for comparison purposes. The second part, which covers chapters 5 and 6, is based on the principles of chaos theory where the hydrologic time sequences are perceived as chaotic time series.

A different approach from the traditional single-valued approach for estimating missing values has been proposed and presented in this dissertation. This approach, named "group approach", is based on perceiving the hydrologic time series as a sequence of groups of values rather than single-valued observations. The methodology of applying the group approach to periodic hydrologic data differs from that required for non-periodic data. In the first part of this thesis, monthly streamflows have been addressed as representatives of periodic hydrologic data. Pertinent literature has been reviewed, and the approach and methodology adopted to employ the group approach have been outlined. It has been shown that the concept of data grouping is recognized, both implicitly and explicitly, in the hydrologic literature. For simplicity, in some cases, it has been avoided after recognition. Proximity measures, from the pattern recognition domain, have been presented as representatives of the unsupervised learning technique to partition the data into segments and classes. Models, representing four different techniques for modeling the streamflow data as groups, have been briefly described.

The multivariate AR(1) model has been manipulated to treat univariate time series according to the group approach. Accordingly, the autoregressive groups, ARG(1), and periodic autoregressive groups, PARG(1), models have been developed in this thesis. Techniques capable of treating data vectors, such as multivariate partitioning modeling (MPM) and multivariate multiple regression (MMR) have been employed to model the data groups. ANNs, with their flexibility in handling groups and single-valued observations, have been used for comparison because of their previously established utility in water resource applications.

Chapter 4 demonstrates that monthly streamflow data can be clustered into a definite number of classes according to the group approach. The flow values, their variances, and autocorrelation coefficients are reasonable characteristics that can be used for data segmentation and verification of the procedures. The results of the application of different techniques for estimation of missing segments indicate that the autocorrelation among elements of each segment plays a major role in the success of the group approach. In both, single series and bi-series cases, artificial neural network models show superiority over other models, applied in this thesis, in estimating missing segments using the group approach. However, care must be exercised when ANN models are applied to cases of partially-missing segments. ANNs tend to violate the group structure of the data to minimize the overall errors of the estimates.

Other techniques, represented by group time series models (ARG(1) and PARG(1)) and MPM, are more consistent in handling data groups. Their accuracy of estimating the

missing data improves when a part of the missing segment becomes available. The discussion given earlier, in chapter 4, on the utility of the group approach for forecasting applications indicates that the group approach may have potential for other applications in water resources. The success of the group approach is contingent on the segmenting process and the selection of the models that can handle groups. Segmentation procedures and models used in this thesis are not claimed to be exhaustive. Improvement in these two aspects of the group approach may reflect on the final results of the application. Finally, the group approach is not proposed as a replacement for the traditional single-valued approach but rather as a sound alternative when periodic hydrologic data show clustering characteristics and as a strong candidate when consecutive observations are missing.

The second part of this thesis, chapters 5 and 6, focuses on the non-periodic hydrologic data. Daily streamflow data sets are considered as representatives of the non-periodic data. The existence of the chaotic behavior (nonlinear deterministic dynamics) in the data has been investigated. The correlation dimension, the Lyapunov exponent, the Kolmogorov entropy, and the method of surrogate data have been used in the analysis. There are sufficient indications to believe that the streamflows have some chaos and the data could be modeled by the time delay embedding method. Based on the attractor dimension, the minimum number of variables essential to model the dynamics of the daily flows has been identified and also, the number of variables sufficient to embed the attractor of the data set under consideration has been estimated. The dimension of the attractor is perceived, in this thesis, as the width of the window encompassing a group of

individual observation. This perception has been verified by the following procedures: (1) structuring the data set in groups of dimension equivalent to the embedding dimension of the attractor, (2) configuring the model using the identified embedding dimension, and (3) estimating missing data according to the pre-designed data structure and model configuration. The satisfactory output of the proposed modeling methodology is considered as evidence to support the group approach when handling daily hydrologic data.

When a single time series has been considered in the analysis and the modeling tool has been applied to the whole data set, the ANN models have shown superiority in the accuracy of estimating the missing data over the K-nearest neighbor and local linear models. This result is attributed to the capability of the ANNs to capture the nonlinear dynamics and to generalize the structure of the attractor across the entire data set. When two cross-correlated time series were under consideration, local models have been preferred over global models in terms of the MSE and MRE criteria. This result is ascribed to the chaotic behavior of the time series, and occurs when global models that are applied to the whole data set are compared to local models applied to data sections in the direct vicinity of the missing data.

The noise reduction applications in hydrologic literature have been critically reviewed and doubts are raised in this thesis about the use of the prediction accuracy as a tool for diagnosing the level of noise. The promising results of the models adopted in this research encourage the process of seeking tools to handle the raw chaotic time series and

obtain better results. Noise reduction process can help facilitate the process of estimating the chaotic invariants, and therefore helps configure the models. Smoothing the original data, by noise reduction algorithms, and using them for further analysis is a questionable approach and therefore should be discouraged.

In summary, the utility and applicability of data in-filling procedures developed in this thesis lie in their capability of enhancing the information content of hydrologic time series through improved estimation of missing records. Consequently, this improved information would help improve the decision making process with respect to hydrologic problems, such as reservoir operation and hydropower generation.

For periodic streamflow time series, the proposed methodology can be adapted for forecasting purposes as demonstrated in chapter four. This methodology is capable of providing updated forecast without the necessity of re-modeling of the entire time series. Such a privileged advantage offered by this methodology would increase the forecast efficiency.

## 7.2 Research contribution

The group approach is proposed in this thesis to handle the hydrologic data. The case of estimation of consecutively missing hydrologic observations is used as an application to apply the proposed approach and models. Different situations are considered assuming both stochastic and chaotic hydrologic domains. Periodic and non-periodic streamflow

data are investigated and utilized to test the validity and utility of the proposed approach. However, a larger number of diversified real-life applications is recommended to signify the practicality of the group approach along with the proposed techniques and models. The contribution of this thesis to the field of water resources can be categorized under three categories of contribution scale, not regarding their significance but rather their conceptual level: level-1 contribution, level-2 contribution, and level-3 contribution.

### *7.2.1 level-1 contribution*

The group approach, which perceives the hydrologic data as a sequence of groups of observations rather than individual values, is considered to be the major contribution of this thesis at the conceptual level. The approach is applied to periodic streamflows through the case of estimating consecutively missing data. However, it is expected that the group approach can be applied to other types of periodic data and cases of hydrologic applications. The group approach is identified and shown to be conceptually different from the traditional single-valued approach that is currently employed in hydrology.

Although, it looks as if the group approach does not literally apply to the non-periodic data through the use of principles of chaos theory, the analysis of chaotic time series lends another support to the group approach. In this thesis, it is shown that for a segment of missing data, there is a window of observations that can be modeled to bring the best estimate of the missing data. Applications to single chaotic time series and bi-series are given in chapter 6. The association of specific observations together and the utility of

information contained in specific windows are considered to be among the characteristics of the group approach.

*7.2.2 level-2 contribution*

This level of contribution is perceived to be of a nature that is related to the application under consideration. Its significance emerges from the technicalities that are embedded in the adopted techniques or models. However, it also adds to the field of water resources at the conceptual level with less generality than the level-1 contribution.

(1) The way in which the partially missing segment is treated in chapter 4 is considered a beneficial addition to the process of forecasting the hydrologic data. When an observation is measured, a significant improvement in the prediction accuracy of the remaining values, lying in the same group, can be achieved. This improvement cannot be realized without identifying the boundaries of each group, which is accomplished using the group approach. (2) Configuration of the ANN models is not left, in this thesis, completely to trial and error. One of the chaotic invariant (the correlation dimension) is used to determine the number of inputs to the network in order to achieve the desired output. This is considered a beneficial step towards establishing "chaotic models" similar to the existing stochastic models. (3) Given the fact that chaos implies: nonlinear dynamics, it is shown, in this thesis, that the effect of the dynamics and sensitivity to initial conditions could be more significant than the effect of the nonlinearity. This result is inferred from the comparison of results made regarding local versus global models and linear versus nonlinear models in chapter 6. (4) The analysis of two cross-correlated chaotic time series

has not been tackled before in hydrology literature. In this thesis, it is illustrated how the chaoticity can affect the analysis of multiple chaotic time series. (5) The issue of noise that may exist in hydrologic data is discussed in details. Facts and fads about its existence, effect, and removal are clarified in an unprecedented manner.

*7.2.3 level-3 contribution*

(1) The commonly used models in stochastic hydrology, autoregressive (AR) and periodic autoregressive (PAR) models, are modified to handle the group approach in this thesis. Thus, ARG(1) and PARG(1) models are developed to consider arrays of inputs and outputs rather than single-valued observations. Consequently, the group approach can be well perceived within the context of stochastic hydrology. (2) The modification formula given as Equation 4.11 allows for improving the accuracy of estimating a few consecutively missing observations or forecasting a few time steps ahead when one observation becomes available. This procedure saves the effort and time required to re-model the whole process. It also reinforces the concept of association among observations of the same group. (3) The ability of the pooled mean squared error (PMSE), developed in this thesis, to tackle the inadequacy of the MSE and the MRE as measures of the level of model performance is demonstrated. The efficacy of the PMSE and its relative superiority to the MSE and the MRE are presented through hypothetical and real examples.

## 7.3 Possible research extension

Improvements in the approaches, methodologies, and models developed in this thesis are possible at various levels and in different parts of this research. Some of the opportunities for possible future research directions are briefed below.

In the group approach, adopted with periodic data, unsupervised classification technique along with some conceptual constraints is used to segment the time series into groups. Other models and techniques from the pattern recognition domain can be investigated to enhance the segmentation process, and therefore improving the modeling accuracy.

Only the bi-series case is considered in this thesis with the group approach to limit the number of parameters involved in the analysis. However, there exist venues for conducting the group approach with the multivariate case especially if contemporaneous and unidirectional relationships among time series are considered. Considering contemporaneous and unidirectional relationships will allow for eliminating a significant number of parameters. When ARG (1) and PARG(1) are developed, the moving average component is not considered. It might be worth investigating how the moving average component can improve the modeling process with the group approach.

In the analysis of chaotic time series, the issue of the fractal dimension requires more research on the physical meaning of fractal dimension of a time series. Is there a difference between a dimension with small fraction close to an integer (e.g., 1.01) and a

fractal dimension of a bigger fraction (e.g., 1.50)? Does this number mean a higher or lower chaoticity? Will it have any effect on the accuracy of modeling the time series? Addressing these questions can significantly improve the understanding of chaos in hydrology, which is still in its infancy.

The effect of noise on the chaotic time series is extensively discussed in this thesis. Finding a methodology or an algorithm that can remove the noise component (independent time series) without affecting the dynamics of the signal is difficult but not impossible. Devising a method that can filter the signal providing that the removed component is structure-free should be sought.

The case of cross-correlated bi-series is addressed in this thesis as a start of multiple chaotic time series analysis. Multivariate chaotic time series analysis has not been reported in literature and deserves to receive the researchers' attention. This can open the door for further studies on chaotic data generation and records extension, keeping in mind that chaotic behavior and other characteristics of the time series are to be maintained.

## 7.4 Study limitations

Several limitations of the analysis can be noted in this thesis. Applications are limited to streamflow data. However, the group approach applied to periodic data or to non-periodic data can be easily applied to other hydrologic data (e.g., precipitation, temperature, water quality ....etc..). Also, a limited number of rivers are used in this thesis where the

overriding objective of the application section is to illustrate the approach and how it can be applied to real-world problems. Similarly, the cases of having the entire and partial missing segments are considered in this thesis. Other patterns of missing data can be investigated following the same procedures presented in this work.

In the segmentation process, adopted to segment periodic data, only consecutive observations are allowed to form a group. Clustering inconsequent observations in one group is possible using pattern recognition techniques, however, it is believed that it hardly can find justification on physical grounds in hydrology. Also, the visual inspection of the data is used in this research to help segment the data into groups. This approach gives more flexibility to the analyst. Automation of the segmentation process, which is not adopted in this thesis, can be applied by defining definite number of data ranges and to segment the data accordingly. Lengths of different groups can be different from each other as explained in chapter 3 of this thesis. Nonetheless, only groups of equal length are considered to make the application of time series models (ARG and PARG) easier.

Applying the principles of chaos theory to non-periodic data, the length of the missing segment, to be estimated, is restricted to be equal to or less than the time delay $\tau$ to follow the methodology proposed in chapter 5 of this thesis.

---

" *We raise to degrees of wisdom and knowledge whom We please: but over all endued with knowledge is One, The All-Knowing.*"

*The Glorious Qura'n (12: 76)*

# REFERENCES


Alley, W.M. and Burns, A.W., 1983. Mixed-station extension of monthly streamflow records. *J. Hydraul. Eng.*, ASCE, 109(10): 1272-1284.

Anderson, J. A., 1995. An Introduction to Neural Networks. *MIT*, USA.

Angelbeck, D.I. and Minkara, R.Y., 1994. Strange attractors and chaos in wastewater flow. *J. Environ. Eng., ASCE*, 120(1):122-137.

Barnsley, M., 1988. Fractals everywhere. *Academic Press* ,New York, USA.

Bastarache, D., El-Jabi, N., Turkkan, N., and Clair, T. A., 1997. Predicting Conductivity and Acidity for Small Streams Using Neural Networks. *Canad. J. of Civ. Engin.*, 24: 1030-1039.

Beale, E.M.L. and Little, R.J.A., 1975. Missing values in multivariate analysis. *J. R. Statist. Soc. B*, 37(1): 129-145.

Beauchamp, J.J., Downing, D.J. and Railsback, S.F., 1989. Comparison of regression and time series methods for synthesizing missing streamflow records. *Water Resour. Bull.*, 25(5): 961-975.

Bennis, S., Berrada, F. and Kang, N., 1997. Improving single-variable and multivariable techniques for estimating missing hydrological data. *J. Hydrol.*, 191: 87-105.

Ben-Zvi, M. and Kesler, S., 1986. Spatial approach to estimation of missing data. *J. Hydrol.*, 88: 69-78.

Berkowitz, B., Ben-Zvi, M. and Berkowitz, J., 1992. A spatial time-dependent approach to estimation of hydrologic data. *J. Hydrol.*, 135: 133-142.

Booy, C. and Morgan, D.R., 1985. The effect of clustering of flood peaks on a flood risk analysis for the Red River. *Can. J. Civ. Eng.*, 12: 150-165.

Broomhead, D.S. and King, G.P., 1986. Extracting qualitative dynamics from experimental data. *Physica*, 20: 217-236.

Camacho, F., McLeod, A.I. and Hipel, K.W., 1987. The use and abuse of multivariate time series models in hydrology. MacNeill, I.B. and Umphrey, G.J. (eds.). *Stochastic hydrology, D. Reidel Publishing Company*, 27-44.

Carpenter, W. C. and Barthelemy, J. F., 1994. Common Misconceptions About Neural Networks As Approximators. *J. Comp. Civ. Engin.*, 8(3): 345-358.

Casdagli, M., 1991. Chaos and deterministic versus stochastic nonlinear modeling. *J. Royal Stat. Soc. B*, 54(2): 303-328.

Casdagli, M., 1989. Nonlinear prediction of chaotic time series. *Physica D*, 35: 335-356.

Chilardi, P. and Rosso, R., 1990. Comment on "Chaos in rainfall" by I. Rodriguez-Iturbe et al. *Water Resour. Res.*, 26(8): 1837-1839.

Correll, D.L., Jordan, T.E. and Weller, D.E., 1999. Effect of inter-annual variation of precipitation on stream discharge from Rhode River sub-watersheds, *Amer. Water Resour. Assoc.*, 35(1): 73-82.

D'Astous, F. and Hipel, K.W., 1979. Analyzing environmental time series. *J. Env. Eng. Div., ASCE*, 105(EE5):979-992.

Davies, M., 1994. Noise reduction schemes for chaotic time series. *Physica D*, 79: 174-192.

Dax, A., 1985. Completing missing groundwater observations by interpolation. *J. Hydrol.*, 81: 375-399.

Dhar, V. and Stein, R., 1997. Chapter 6. Intelligent Decision Support Methods. *Prentice-Hall, Inc.*, USA.

Doscher, C., 1997. Applying nonlinear dynamics to hydrologic systems: A review. *App. Eng. Agr.*, 13(2): 199-207.

Dubes, R.C., 1987. How many clusters are best ?- An experiment. *Pattern Recognition*, 20(6): 645-663.

Duda, R. and Hart, P., 1973. Pattern classification and scene analysis. *John Wiley& Sons*, New York, USA.

Elshorbagy, A., Panu, U. S. and Simonovic, S. P. 2000a. Group-based estimation of missing hydrological data. I. Approach and general methodology. *Hydrol. Sci. J.*, 45(6): 849-866.

Elshorbagy, A., Panu, U. S. and Simonovic, S. P. 2000b. Group-based estimation of missing hydrological data. II. Application to streamflows. *Hydrol. Sci. J.*, 45(6): 867-880.

Elshorbagy, A., Simonovic, S. P. and Panu, U. S. 2000c. Performance evaluation of artificial neural networks for runoff prediction. *J. Hydrol. Engng, ASCE*, 5(4): 424-427.

Elshorbagy, A., Simonovic, S.P. and Panu, U.S., 2001a. Estimation of missing streamflow data using principles of chaos theory. *J. Hydrol. (accepted for publication).*

Elshorbagy, A., Simonovic, S. P. and Panu, U. S. 2001b. Noise reduction in chaotic

hydrologic time series: Facts and doubts. *J. Hydrol.* (*accepted for publication*).

Elshorbagy, A., Panu, U.S. and Simonovic, S.P. 2000d. Group-based estimation of missing data in chaotic hydrologic time series. *Hydrol. Sci. Tech., AIH*, (under review).

Elshorbagy, A., Panu, U.S. and Simonovic, S.P. 2001c. Analysis of cross-correlated chaotic streamflows. *Hydrol. Sci. J.*, (*accepted for publication*).

Elshorbagy, A., Panu, U.S. ans Simonovic, S.P., 1999. Investigations into group-based data in-filling techniques. *Proc. Annual Conf. Can. Soc. Civ. Eng.*, Regina, Saskatchewan, June 2-5, 337-348.

Embrechts, M., 1994. Basic concepts of nonlinear dynamics and chaos theory. In: G.J. Deboeck (Ed.), Trading on the Edge: Neural, Genetic, and Fuzzy Systems for Chaotic Financial Markets. *NY: John Wiley & Sons, Inc.* pp. 265-279.

Farmer, J.D. and Sidorowich, J.J., 1987. Predicting chaotic time series. *Phys. Rev. Lett.*, 59(8): 845-848.

Feder, J., 1989. Fractals. New York: *Plenum Press*.

Fiering, M.B., 1962. On the use of correlation to augment data. *J. Amer. Statist. Assoc.*, 57(297): 20-32.

Flick, T.E. and Lee K. J., 1986. A Combinatorial Approach for Classification of Patterns with Missing Information and Random Orientation. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 8 (4): 482-490.

Fraedrich, K., 1987. Estimating weather and climate predictability on attractors. *J. Atmos Sci.*, 44(4): 722-728.

Fraedrich, K., 1986. Estimating the dimensions of weather and climate attractors. *J.*

*Atmos Sci.*, 43(5): 419-432.

Fraser, A.M. and Swinney, H.L., 1986. Independent coordinates for strange attractors from mutual information. *Phys. Rev. A,* 33(2): 1134-1140.

Freeman, J. A. and Skapura, D. M., 1991. Neural networks. Algorithms, applications, and programming techniques. *Addison-Wesley Inc.*, USA.

French, M. N., Krajewski, W. F., and Cuykendall, R. R., 1992. Rainfall forecasting in space and time using a neural network. *J. of Hydrol.*, 137: 1-31.

Frison, T., 1994. Nonlinear data analysis techniques. In: G.J. Deboeck (Ed.), Trading on the Edge: Neural, Genetic, and Fuzzy Systems for Chaotic Financial Markets. *NY: John Wiley & Sons, Inc.* pp. 280-296.

Gilroy, E.J., 1971. Reliability of a variance estimate obtained from a sample augmented By multivariate regression. *Water Resour. Res.*, 6(6): 1595-1600.

Gnanadesikan, R., 1977. Methods for statistical data analysis of multivariate observations. *John Wiley & Sons, Inc.*, USA.

Grassberger, P., Hegger, R., Kantz, H. and Schaffrath, C., 1993. On noise reduction methods for chaotic data. *Chaos* 3(2): 127-141.

Grassberger, P. and Procaccia, I., 1983. Measuring the strangeness of strange attractors. *Physica D,* 9: 189-208.

Griffith, D.A., Haining, R.P. and Bennett, R.J., 1985. Estimating missing values in space-time data series. in *Time series analysis: theory and practice 6.* Anderson, O.D., Ord, J.K. and Robinson, E.A., eds., *Elsevier Science Publishers* B.V., North-Holland.

Gupta, A. and Lam, M.S., 1996. Estimating missing values using neural networks. *J. Oper. Res. Soc.*, 47(2): 229-238.

Gyau-Boakye, P. and Schultz, G.A., 1994. Filling gaps in runoff time series in West Africa. *Hydrol. Sci. J.*, 39(6): 621-636.

Hegger, R., Kantz, H. and Schreiber, T., 1999. Practical implementation of nonlinear time series methods: The TISEAN package. *Chaos* 9: 413-440.

Henderson, H.W. and Wells, R., 1988. Obtaining attractor dimensions from meteorological time series. *Adv. Geophys.*, 30: 205-237.

Hense, A., 1987. On the possible existence of a strange attractor for the Southern Oscilation. *Beitr. Phys. Atmos.*, 60(1): 34-47.

Hirsch, R.M., 1982. A comparison of four streamflow record extension techniques. *Water Resour. Res.*, 18(4): 1081-1088.

Hirsch, R.M., 1979. An evaluation of some record reconstruction techniques. *Water Resour. Res.*, 15(6): 1781-1790.

Hsu, K. L., Gupta, H. V., and Sorooshian, S., 1995. Artificial neural network modeling of the rainfall-runoff process. *Water Resour. Res.*, 31(10): 2517-2530.

Hughes, D.A. and Smakhtin, V., 1996. Daily flow time series patching or extension: a spatial interpolation approach based on flow duration curves. *Hydrol. Sci. J.*, 41(6): 851-871.

Hurst, H.E., 1956. Methods of using long-term storage in reservoirs. *Proc. Inst. Civ. Eng.*, 1: 519-543.

Hurst, H.E., 1951. Long-term storage capacity of reservoirs. *Trans. Am. Soc. Civ. Eng.*, 16: 770-808.

Ishibuchi, H., Miyazaki, A., Kwon, K. and Tanaka, H., 1993."Learning from incomplete training data with missing values and medical application." *Proc. Int. Joint Conf. On Neu. Net.*, Nagoya, Japan, 2: 1871-1874.

Islam, S., Bras, R.L. and Rodriguez-Iturbe, I., 1993. A possible explanation for low correlation dimension estimates for the atmosphere. *J. Appl. Meteor.*, 32: 203-208.

Jayawardena, A.W. and Gurung, A.B., 2000. Noise reduction and prediction of hydrometeorological time series: dynamical systems approach vs. stochastic approach. *J. Hydrol.*, 228: 242-264.

Jayawardena, A.W. and Lai, F., 1994. Analysis and prediction of chaos in rainfall and streamflow time series. *J. Hydrol.*, 153: 23-52.

Jeong, G.D. and Rao, A.R., 1996. Chaos characteristics of tree ring series. *J. Hydrol.*, 182: 239-257.

Johnson, R.A. and Wichern, D.W., 1988. Applied multivariate statistical analysis. *Prentice Hall*, NJ, USA.

Kantz, H. and Schreiber, T., 1997. Nonlinear time series analysis. *Cambridge University Press*, UK.

Karunanithi, N., Grenney, W.J., Whitley, D., and Bovee, K., 1994. Neural networks for river flow prediction. *J. Comp. in Civ. Eng., ASCE*, 8(2): 201-220.

Kawamura, A., McKerchar, A.I., Spigel, R.H. and Jinno, K., 1998. Chaotic characteristics of the Southern Oscillation Index time series. *J Hydrol.*, 204: 168-181.

Keppenne, C.L. and Nicolis, C., 1989. Global properties and local structure of the

weather attractor over Western Europe. *J. Atoms. Sci.*, 46(15): 2356-2370.

Khalil, M., Panu, U. and Lennox, W., 1998. Estimation of missing streamflows: a historical perspective. *Proc. Annual Conf. Can. Soc. Civil Eng.*, Halifax, Nova Scotia, June 10-13: 235-246.

Kittler, J., 1978. Classification of Incomplete Pattern Vectors Using Modified Discriminant *Functions. IEEE Transactions on Computers*, C-27: 367-375.

Knotters, M. and Van Walsum, P.E.V., 1997. Estimating fluctuation quantities from time series of water-table depths using models with a stochastic component. *J. Hydrol.*, 197: 25-46.

Kojiri, T., Unny, T.E. and Panu, U.S., 1994. Cluster based pattern recognition and analysis of streamflows. In: Hipel, K.W., McLeod, A.I., Panu, U.S. and Singh, V.P. (Eds.), Time series analysis in hydrology and environmental engineering. *Kluwer Academic Publishers*, The Netherlands.

Kostelich, E.J. and Schreiber, T., 1993. Noise reduction in chaotic time-series data: A survey of common methods. *Phys. Rev. E*, 48(3): 1752-1763.

Kuczera, G., 1987. "On maximum likelihood estimators for the multisite lag-one streamflow model: complete and incomplete data cases." *Water Resour. Res.*, 23(4): 641-645.

Kuligowski, R.J. and Barros, A.P., 1998. Using artificial neural networks to estimate missing rainfall data. *Amer. Water Resour. Assoc.*, 34(6): 1437-1447.

Lall, U., Sangoyomi, T. and Abarbanel, H.D.I., 1996. Nonlinear dynamics of the great salt lake: nonparametric short-term forecasting. *Water Resour. Res.*, 32(4): 975-985.

Lettenmaier, D.P., 1980. Intervention analysis with missing data. *Water Resour. Res.*, 16(1): 159-171.

Little, R.J.A. and Rubin, D.B., 1987. Statistical analysis with missing data. *John Wiley & Sons*, New York.

Liu, Q., Islam, S., Rodriguez-Iturbe, I. and Le, Y., 1998. Phase-space analysis of daily streamflow: characterization and prediction. *Adv. Water Resour.*, 21: 463-475.

Makhuvha, T., Pegram, G., Sparks, R. and Zucchini, W., 1997a. Patching rainfall data using regression methods. 1. Best subset selection, EM and pseudo-EM methods: theory. *J. Hydrol.*, 198: 289-307.

Makhuvha, T., Pegram, G., Sparks, R. and Zucchini, W., 1997b. Patching rainfall data using regression methods. 2. Comparisons of accuracy, bias and efficiency. *J. Hydrol.*, 198: 308-318.

Mandelbrot, B.B., 1983. The fractal geometry of Nature. New York: *W.H. Freeman.*

Maren, A., Harston, C., and Pap, R., 1990. Handbook of neural computing applications. *Academic Press, Inc.*, USA.

Matalas, N.C., 1967. Mathematical assessment of synthetic hydrology. *Water Resour. Res.*, 3(4): 937-945.

Matalas, N.C. and Jacobs, B., 1964. A correlation procedure for augmenting hydrologic data. *U.S. Geol. Surv. Prof. Pap. 434-E.*

Mikosch, T. and Wang, Q., 1995. A Monte-Carlo method for estimating the correlation dimension. *J. Stat. Phys.*, 78: 799-813.

Mohan, T.R.K., Rao, .J.S. and Ramswamy, R., 1989. Dimension analysis of climatic data. *J. Clim.*, 2: 1047-1057.

Moran, M.A., 1974. On estimators obtained from a sample augmented by multiple regression. *Water Resour. Res.*, 10(1): 81-85.

Mott, P., Sammis, T.W. and Southward, G.M., 1994. Climate data estimation using climate information from surrounding climate stations. *Appl. Eng. In Agric.*, 10(1): 41-44.

Mujumdar, P.P. and Kumar, D.N., 1990. Stochastic models of streamflow: some case studies. *Hydro. Sci. J.*, 35(4): 395-410.

Panu, U.S., Khalil, M. and Elshorbagy, A., 2000. Streamflow data infilling techniques based on concepts of groups and neural networks, in: R.S. Govindraju and A.R. Rao(Eds.), *Artificial Neural Networks in Hydrology. Kluwer Academic*, The Netherlands, 235-258.

Panu, U.S., 1992. Application of some entropic measures in hydrologic data infilling procedures. in *Entropy and energy dissipation in water resources.* Singh, V.P. and Fiorentino, M., (Eds.), *Kluwer Academic Publishers*, the Netherlands: 175-192.

Panu, U.S. and Unny, T.E., 1980a. Stochastic synthesis of hydrologic data based on concepts of pattern recognition, I. General methodology of the approach. *J. Hydrol.*, 46: 5-34.

Panu, U.S. and Unny, T.E., 1980b. Stochastic synthesis of hydrologic data based on concepts of pattern recognition, II. Application of natural watersheds. *J. Hydrol.*, 46: 197-217.

Panu, U.S., Unny, T.E. and Ragade, R.K., 1978. A feature prediction model in

synthetic hydrology based on concepts of pattern recognition. *Water Resour. Res.*, 14(2): 335-344.

Pasternack, G.B., 1999. Does the river run wild? Assessing chaos in hydrological systems. *Adv. Water Resour.*, 23(3): 253-260.

Pawlak, M., 1993. Kernel classification rules from missing data. *IEEE Trans. Inform. Theory*, 39(3): 979-988.

Pedreira, C.E. and Parente, E., 1995. Neural networks with missing values attributes. *Proc. IEEE Int. Conf. Neu. Net.*, 6: 3021-3023.

Pegram, G., 1997. Patching rainfall data using regression methods. 3. Grouping, patching and outlier detection. *J. Hydrol.*, 198: 319-334.

Peters, E.E., 1994. Fractal market analysis. New York: *John Wiley & Sons*.

Peters, E.E., 1991. Chaos and order in the capital markets: A new view of cycles, prices, and market volatility. New York: *John Wiley & Sons*.

Porporato, A. and Ridolfi, L., 1997. Nonlinear analysis of river flow time sequences. *Water Resour. Res.*, 33(6): 1353-1367.

Procaccia, I., 1988. Complex or just complicated?. *Nature*, 333(9): 498-499.

Puente, E. and Obregon, N., 1996. A deterministic geometric representation of temporal rainfall: results for a storm in Boston. *Water Resour. Res.*, 32(9): 2825-2839.

Querios, C.E. and Gelsema, E.S., 1988. Incomplete data sets. In: E.S. Gelsema and L.N. Kanal (Eds.), Pattern Recognition and Artificial Intelligence. Amsterdam: *Elsevier*. pp. 237-255.

Raman, H. and Sunilkumar, N., 1995. Multivariate modeling of water resources time series using artificial neural networks. *Hydrol. Sci. J.*, 40 (2): 145-163.

Rodriguez-Iturbe, I., De Power, B.F., Sharifi, M.B. and Georgakakos., 1989. Chaos in rainfall. *Water Resour. Res.*, 25(7): 1667-1675.

Salas, J.D., 1992. Analysis and modeling of hydrologic time series. Maidment, D.R. (ed.). *Handbook of Hydrology, McGraw-Hill, Inc.*, USA.

Salas, J.D., Tabios III, G.Q. and Bartolini, P., 1985. Approaches to multivariate modelling of water resources time series. *Water Resour. Bull.*, 21(4): 683-708.

Salas, J.D., Delleur, J.W., Yevjevich, V. and Lane, W.L., 1980. Applied modelling of hydrologic time series. *Water Resour. Publ.*, Littleton, Colorado, USA.

Sangoyomi, T.B., Lall, U. and Abarbanel, H.D.I., 1996. Nonlinear dynamics of the great salt lake: dimension estimation. *Water Resour. Res.*, 32(2): 149-159.

Sauer, T., 1992. A noise reduction method for signals from nonlinear systems. *Physica D*, 58: 193-202.

Scargle, J.D., 1990. Studies in astronomical time series analysis, IV, Modeling chaotic and random processes with linear filters. *Astrophys. J.*, 359: 469-482.

Schouten, J.C., Takens, F. and Van den Bleek, C.M., 1994. Estimation of the dimension of a noisy attractor. *Phys. Rev. E*, 50(3): 1851-1861.

Schouten, J.C. and van den Bleek, C.M., 1994. RRCHAOS: a menu driven software package for chaotic time series analysis. *Reactor Research Foundation*, Delft, The Netherlands.

Schreiber, T., 1993. Extremely simple nonlinear noise-reduction method. *Phys. Rev. E*, 47(4): 2401-2404.

Schreiber, T. and Grassberger, P., 1991. A simple noise-reduction method for real data. *Phys. Lett. A*, 160(5): 411-418.

Shamseldin, A. Y., 1997. Application of a Neural Network Technique to Rainfall-Runoff Modeling. *J. Hydrol.*, 199: 272-294.

Sharifi, M.B., Georgakakos, K.P. and Rodriguez-Iturbe, I., 1990. Evidence of deterministic chaos in the pulse of storm rainfall. *J. Atmos. Sci.*, 47(7): 888-893.

Shih, S.F. and Cheng, K.S., 1989. Generation of synthetic and missing climatic data for Peurto Rico. *Water Resour. Bull.*, 25(4): 829-836.

Shvarster, L., Shamir, U. and Feldman, M., 1993. Forecasting hourly water demands by pattern recognition approach. *J. Water Resour. Plann. Manag., ASCE*, 119(6): 611-627.

Simonovic,S.P., 1995. Synthesizing missing streamflow records on several Manitoba streams using multiple nonlinear standardized correlation analysis. *Hydrol. Sci. J.*, 40(2): 183-203.

Sivakumar, B., Sorooshian, S., Gupta, H.V. and Gao, X. 2001. A chaotic approach to rainfall disaggregation. *Water Resour. Res.*, 37(1): 61-72.

Sivakumar, B., 2000. Chaos theory in hydrology: important issues and interpretations. *J. Hydrol.*, 227: 1-20.

Sivakumar, B., Phoon, K., Liong, S. and Liaw, C., 1999a., Comment on "Nonlinear analysis of river flow time sequences" by Amilcare Porporato and Luca Ridolfi. *Water Resour. Res.*, 35(3): 895-897.

Sivakumar, B., Phoon, K., Liong, S. and Liaw, C., 1999b. A systematic approach to noise reduction in chaotic hydrological time series. *J. Hydrol.*, 219: 103-135.

Sivakumar, B., Liong, S., Liaw, C. and Phoon, K., 1999c. Singapore rainfall behavior: chaotic?. *J. Hydrol. Eng., ASCE*, 4(1): 38-48.

Sivakumar, B., Liong, S. and Liaw, C., 1998. Evidence of chaotic behaviour in Singapore Rainfall. *J. Am. Water Resour. Assoc.*, 34(2): 301-310.

Sorooshian, S., Duan, Q., and Gupta, V. K., 1993. Calibration of Rainfall-Runoff Models: Application of Global Optimization to the Sacramento Soil Moisture Accounting Model. *Water Resour. Res.*, 29 (4): 1185-1194.

Sorooshian, S., Gupta, V. K., and Fulton, J. L., 1983. Evaluation of Maximum Likelihood Parameter Estimation Techniques for Conceptual Rainfall-Runoff Models: Influence of Calibration Data Variability and Length on Model Credibility., *Water Resour. Res.*, 19 (1): 252-259.

Takens, F., 1981. Detecting strange attractors in turbulence. In: D.A. Rand and L.S. Young (Eds.), Dynamical Systems and Turbulence. *Warwick, Lecture Notes Math.* pp. 366-381.

Terry, W.R., Lee, J.B. and Kumar, A., 1986. Time series analysis in acid rain modeling: evaluation of filling missing values by linear interpolation. *Atmos. Env.*, 20 (10): 1941-1945.

Theiler, J., Eubank, S., Longtin, A., Galdrikian, B. and Farmer, J.D., 1992. Testing for nonlinearity in time series: The method of surrogate data. *Physica D*, 58: 77-94.

Theiler, J., 1986. Spurious dimension from correlation algorithms applied to limited time-series data. *Phys. Rev. A*, 34: 2427-2432.

Titterington, D.M. and J. Sedransk., 1989. "Imputation of Missing Values Using Density Estimation". *Stat. & Prob. Lett.*, 8: 411-418.

Titterington, D.M. and G.M. Mill., 1983. Kernel-Based Density Estimates from Incomplete Data. *J. Roy. Stat., Soc. B*, 45 (2): 258-266.

Tou, J.T. and Gonzalez, R.C., 1974. Pattern recognition principles. *Addison-Wesley*, Massachusetts, USA.

Tsonis, A.A. and Elsner, J.B., 1988. The weather attractor over very short time scales. *Nature*, 333: 545-547.

Valencia, D.R. and Schaake, J.C., 1973. Disaggregation processes in stochastic hydrology. *Water Resour. Res.*, 9(3): 580-585.

Van Geer, F.C. and Zuur, A.F., 1997. An extension of Box-Jenkins transfer/noise models for spatial interpolation of groundwater head series. *J. Hydrol.*, 192: 65-80.

Vogel, R.M. and Stedinger, J.R., 1985. Minimum variance streamflow record augmentation procedures. *Water Resour. Res.*, 21(5): 715-723.

Wang, Q., 1995. Correlation dimension estimates og global and local temperature data. *J. App. Meteo.*, 34: 2556-2564.

Wang, Q. and Gan, T.Y., 1998. Biases of correlation dimension estimates of streamflow data in the Canadian prairies. *Water Resour. Res.*, 34(9): 2329-2339.

Wolf, A., Swift, J.B., Swinney, H.L. and Vastano, J.A., 1985. Determining Lyapunov exponents from a time series. *Physica D*, 16:285-317.

Yevjevich, V., 1987. Stochastic models in hydrology. *Stoch. Hydrol. Hydraul.*, 1: 17-36.

Young, G.K., Orlob, G.T. and Roesner, L.A., 1970. Decision criteria for using stochastic hydrology. *J. Hydrau. Div., ASCE*, 96 (HY4): 911-926.

Zealand, C.M., Burn, D.H., and Simonovic, S.P., 1999. Short term streamflow forecasting using artificial neural networks. *J. of Hydrol.*, 214: 32-48.

Zeng, X., Pielke, R.A. and Eykholt, R., 1992. Estimating the fractal dimension and the predictability of the atmosphere. *J. Atmos. Sci.*, 49(8): 649-659.

Zhang, K., Douglas, B.C. and Leatherman, S.P., 1997. East coast storm surges provide unique climate record, *EOS, Trans. AGU*, 78(37): 389-397.

Zhang, T. and Berndtsson, R., 1991. Analysis of soil water dynamics in time and space by use of pattern recognition. *Water Resour. Res.*, 27(7): 1623-1636.

Zhu, Q., 1990. On the Minimum Probability of Error of Classification with Incomplete Patterns. *Pattern Recognition*, 23 (11): 1281-1290.

## *Appendix A*

*A segment*: is a vector of consecutive observations. For example, in case of monthly streamflows, Jan, Feb, Mar, and Apr can form *segment 1*; May, Jun, Jul, and Aug are elements of *segment 2*; and Sep, Oct, Nov, and Dec constitute *segment 3*. A segment can be referred to as *group, pattern vector, or object*. Each segment may indicate different flow conditions (e.g., wet, dry, semi wet,...etc..).

*A segment type*: is a collection of all segments from the entire record that have similar flow conditions (e.g., wet conditions). Words such as *class, pattern class, season, group type, or cluster* are used in literature of different domains.

*An element*: is the smallest unit of a segment, which is the monthly flow in case of monthly time series record. It may also be referred to as *"attribute"*.

*The dimensionality*: represents the number of dimensions in the problem and is defined according to the number of elements in the segments (e.g., in our example of this section, it is a 4-dimensional problem). Based on these definitions, a 50-year record of monthly flows has (50*12) = 600 observations that are segmented into (600/4) = 150 *segments*. These *4-dimensional* segments are clustered into three *segment types*.

## *Appendix B*

Examples of pre-analysis of the consistency of the Little River and Reed Creek data are given below and illustrated in Figures B.1, B.2, B.3, and B.4.

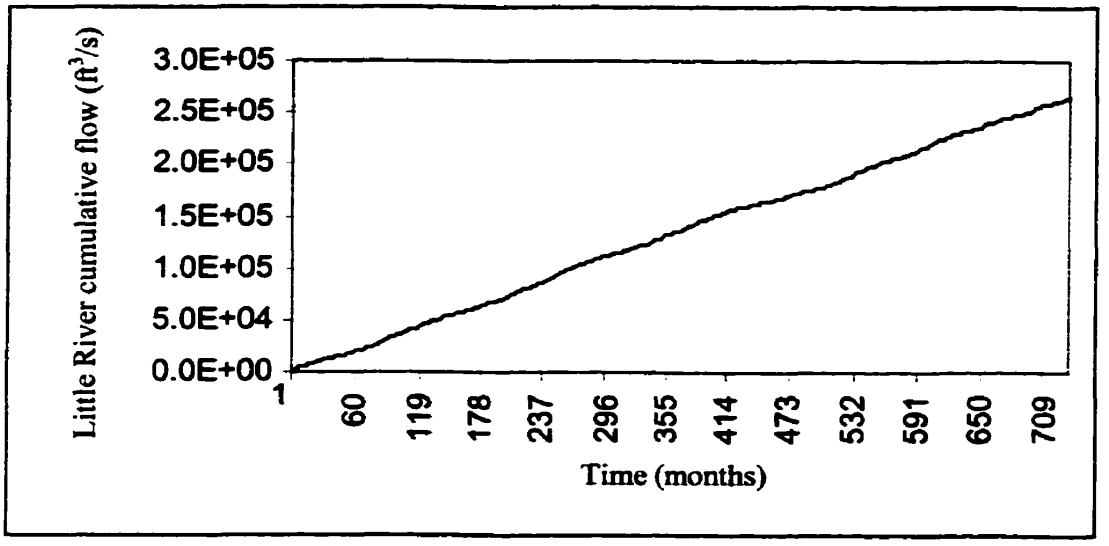



**Fig. B.1. The mass curve of the Little River flow data.**

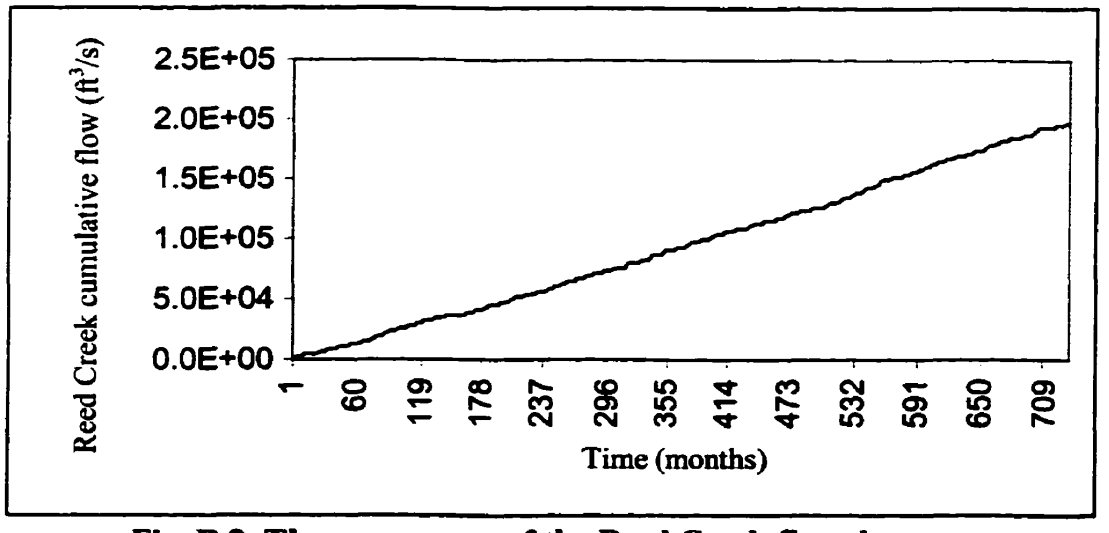



**Fig. B.2. The mass curve of the Reed Creek flow data.**

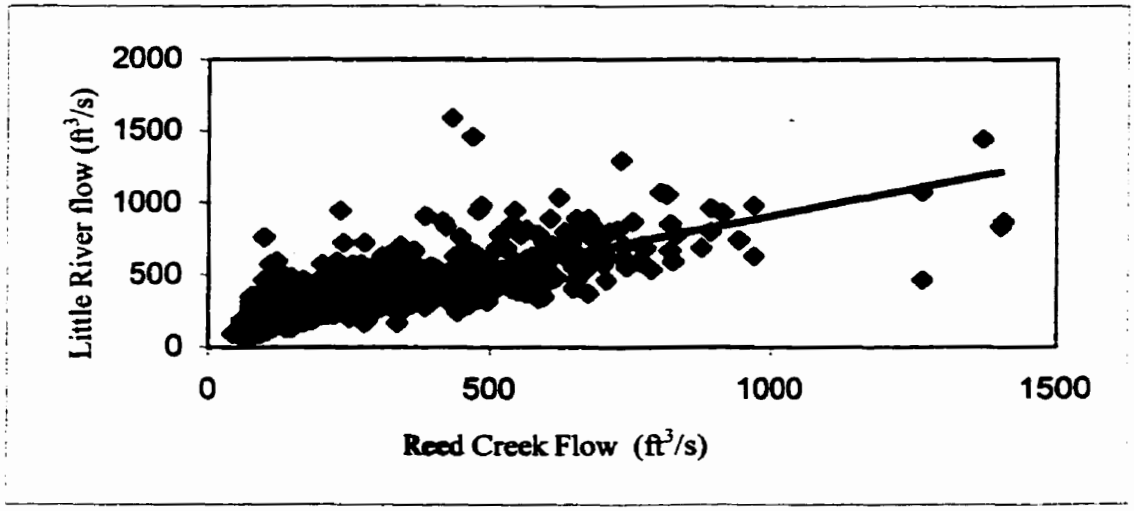



**Fig. B.3. The relationship between the Little River and the Reed Creek flow data.**

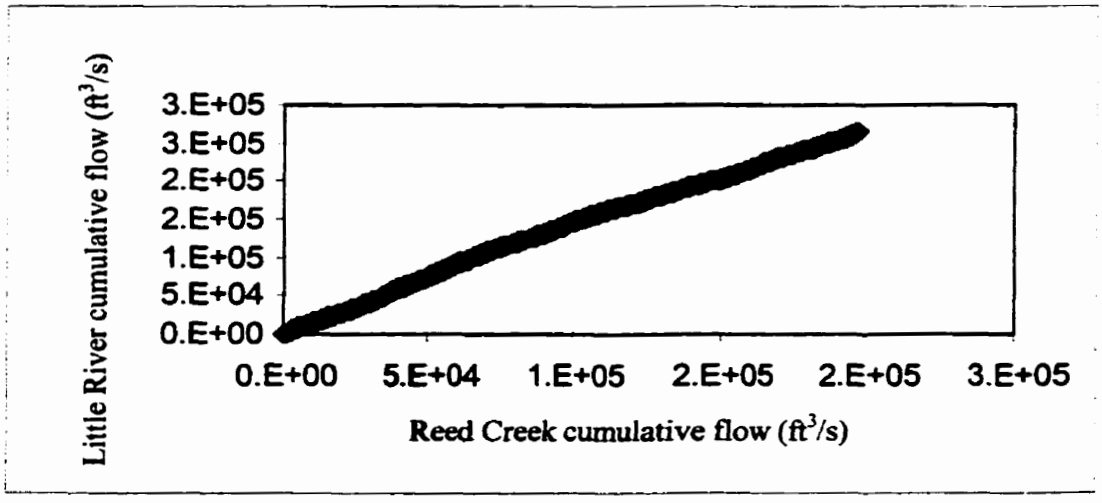



**Fig. B.4. The double mass curve of the Little River and the Reed Creek.**

The plots of the mass curves of the Little River and the Reed Creek flow data (Figures B.1 and B.2) show the absence of any significant change in the slope of the trend line. This indicates that no major change occurred in the course of the rivers regarding their hydraulics or hydrology. Figure B.3 shows the trend of the relationship between the flows of the two rivers. The graph provides a physical bases for the correlation between the two rivers that is calculated using statistical means. It also encourages the use of any of the two rivers as a *reference river* to estimate the missing data of the other one, which is the *target river*. Figure B.4 complements the conclusion drawn from the previous three figures by showing that the nature of the relationship between the two rivers has been always consistent throughout the time horizon of the historical data available. There is no significant change of the slope of the double mass curve, which is calculated by computing the running sum of both rivers.

*Appendix C*

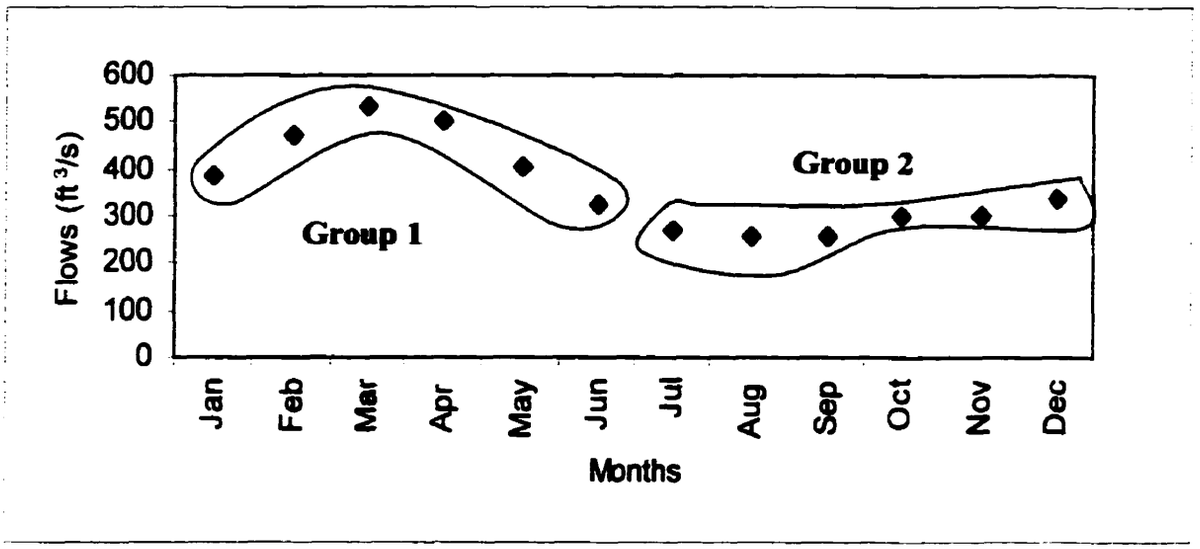



**Fig. C.1. Groups (segments) of the Little River monthly data.**

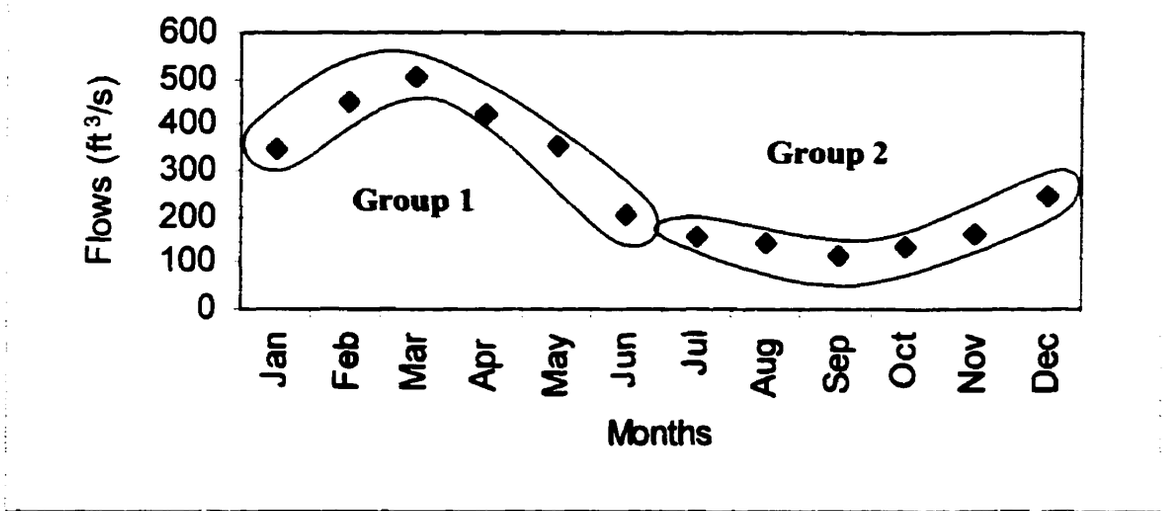



**Fig. C.2. Groups (segments) of the Reed Creek monthly data.**

## *Appendix D*

The monthly data of the English River are arranged in a column as shown in column (3) of Table D.1. The average value of each group is calculated (column (4) of Table D.1). Then the group autocorrelation coefficient can be obtained in the traditional way based on the time sequence of the average values of column (4). In other words, the monthly time sequence (36.2, 35.5, 33.1, 44.4, 106,.....) is replaced by the time sequence (37.3, 64.43,....) only when the group autocorrelation is to be calculated. In this case the single-valued autocorrelation coefficient is 0.80 while the group autocorrelation equals to 0.65.

The same procedures are conducted with the Little River data, which are shown in Table D.2. The monthly time sequence (325, 538, 763, 499, 574, 802, 481,.....) is replaced by the time sequence (583.50, 355.67,....) only when the group autocorrelation and group cross-correlation with the Reed Creek are to be calculated. A table similar to Table D.2 can be established for the Reed Creek. The single-valued lag-1 autocorrelation coefficients for the Little River and the Reed Creek are 0.64 and 0.63 respectively, while the lag-1 group autocorrelation coefficients of the same rivers are 0.46 and 0.45.

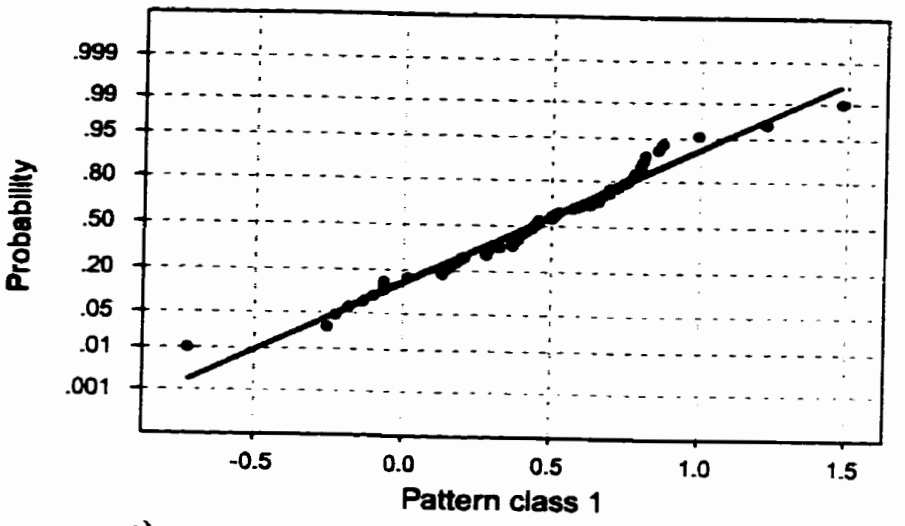
**Table D.1. The grouping of the English River data for the calculation of the group autocorrelation.**

| Group (1) | Month (2) | Monthly value ($m^3/s$) (3) | Group average ($m^3/s$) (4) |
|---|---|---|---|
| G1-year1 | January | 36.20 | 37.30 |
| | February | 35.50 | |
| | March | 33.10 | |
| | April | 44.40 | |
| G2-year1 | May | 106.00 | 64.43 |
| | June | 74.80 | |
| | July | 42.70 | |
| | August | 34.20 | |
| G3-year1 | September | 24.90 | 22.78 |
| | October | 21.50 | |
| | November | 21.30 | |
| | December | 23.40 | |
| G1-year2 | January | 23.90 | 21.53 |
| | February | 21.60 | |
| | March | 20.30 | |
| | April | 20.30 | |
| G2-year2 | May | 79.80 | 70.2 |
| | June | 59.80 | |
| | July | 65.50 | |
| | August | 75.70 | |
| G3-year2 | September | 52.40 | |
| | October | 36.20 | |

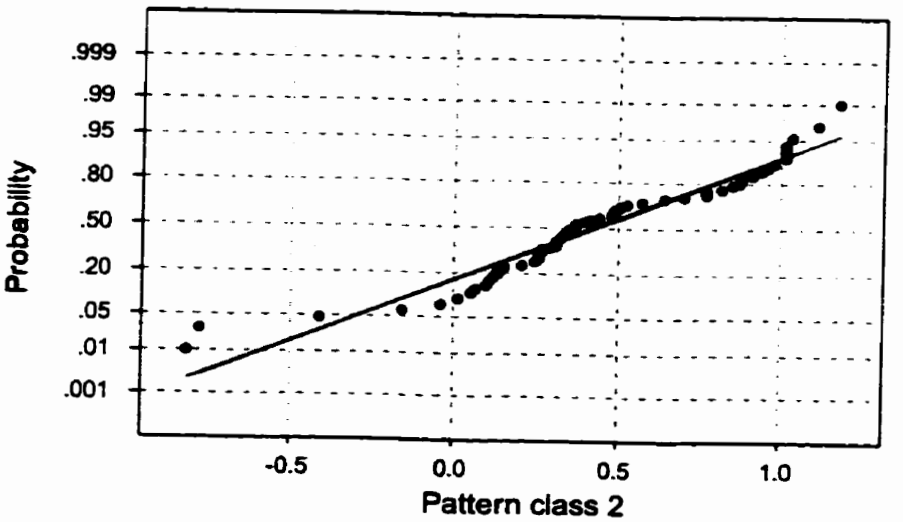
**Table D.2. The grouping of the Little River data for the calculation of the group autocorrelation.**

| Group (1) | Month (2) | Monthly value ($ft^3/s$) (3) | Group average ($ft^3/s$) (4) |
|---|---|---|---|
| G1-year1 | January | 325 | 583.50 |
| | February | 538 | |
| | March | 763 | |
| | April | 499 | |
| | May | 574 | |
| | June | 802 | |
| G2-year1 | July | 481 | 355.67 |
| | August | 298 | |
| | September | 218 | |
| | October | 502 | |
| | November | 336 | |
| | December | 299 | |
| G1-year1 | January | 412 | 333.83 |
| | February | 431 | |
| | March | 383 | |
| | April | 310 | |
| | May | 246 | |
| | June | 221 | |
| G2-year2 | July | 108 | |
| | August | 139 | |
| | September | 140 | |
| | October | 126 | |

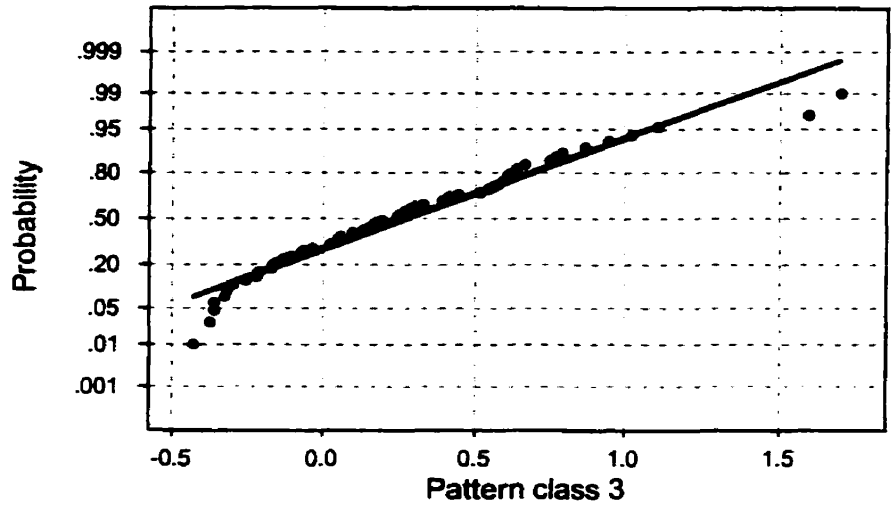
## *Appendix E*




a)

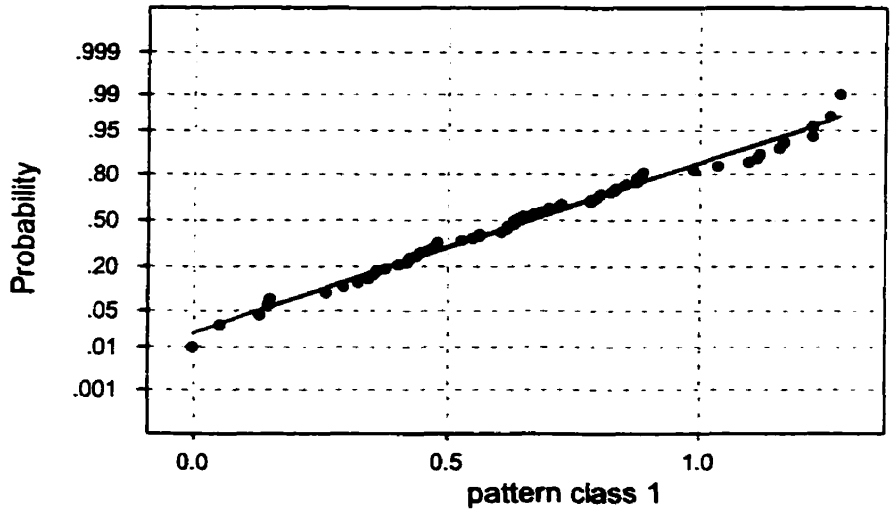



b)




c)

**Fig. E.1. Normal probability plots for the data groups of the English River. a) pattern class 1, b) pattern class 2, and c) pattern class 3.**




a)

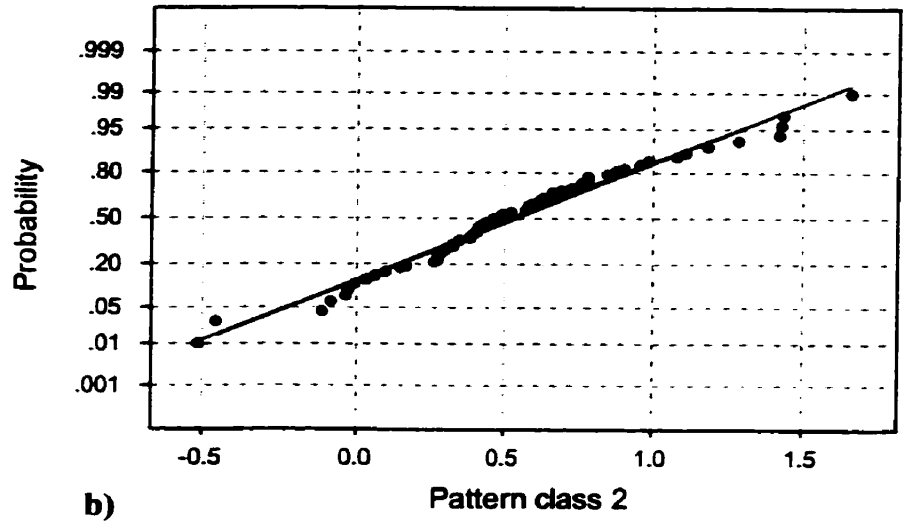
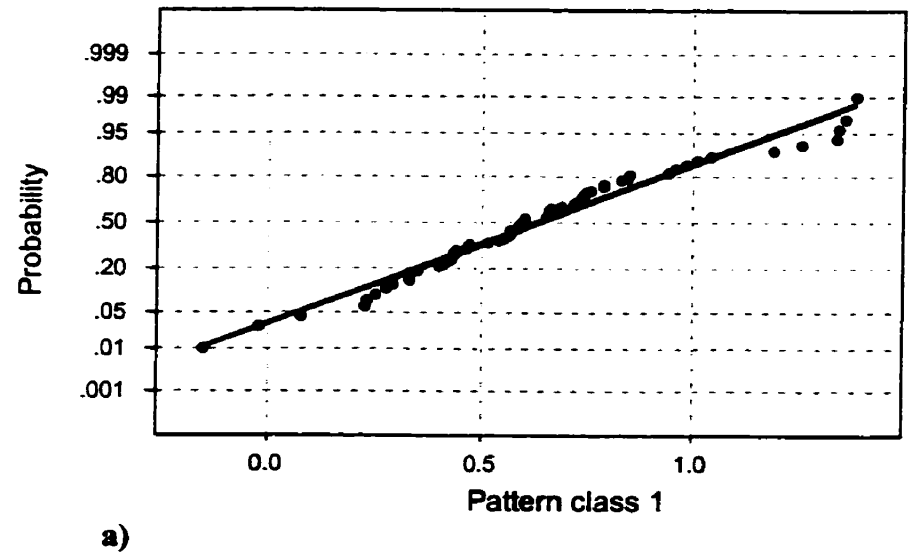



**Fig. E.2. Normal probability plots for the data groups of the Little River. a) pattern class 1 and b) pattern class 2.**



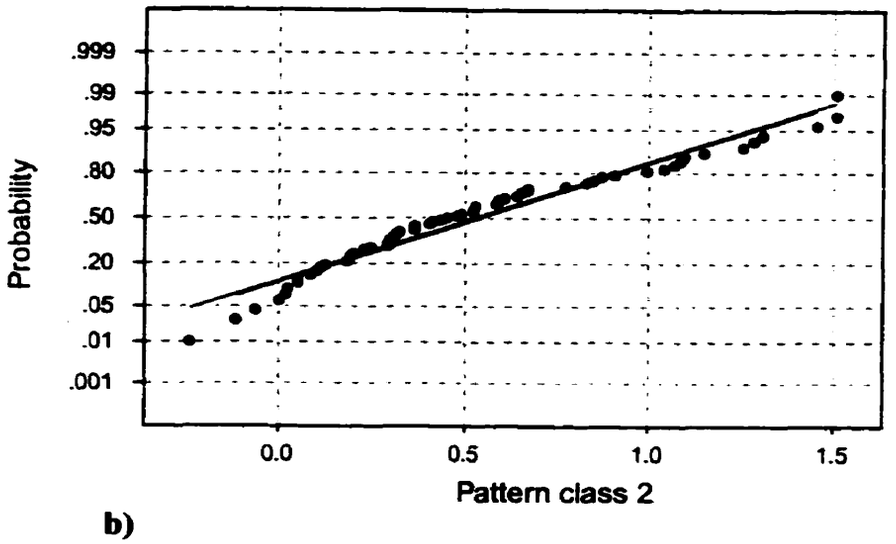



**Fig. E.3. Normal probability plots for the data groups of the Reed Creek. a) pattern class 1 and b) pattern class 2.**

The group multivariate normality test is conducted as follows:

After segmenting the data into groups, which are categorized under definite number of pattern classes (three classes for the English River and two for both of the Little River and the Reed Creek), The center of each class is estimated.

The Mahalanobis distance ($d_M$) between each group and the center of the class is calculated. The values of $d_M$ are treated as the new data for which normality test is to be conducted.

The Log transformation has brought the data close to the normal distribution as shown in Figures E.1, E.2, and E.3.

## *Appendix F*

**Development and parameter estimation of the ARG(1) model**

*English River data (single series case):*

$$\hat{A}_1 = \hat{M}_1 \hat{M}_0^{-1} \tag{F.1}$$

$$\hat{B}\hat{B}^T = \hat{M}_0 - \hat{A}_1 \hat{M}_1^{-1} \tag{F.2}$$

$$\hat{M}_k = E\left[Z_t Z_{t-k}^T\right] = \begin{bmatrix} \rho_k^{11}, \rho_k^{12}, \ldots\ldots\ldots\ldots\ldots \rho_k^{1n} \\ .. \\ .. \\ .. \\ \rho_k^{n1}, \qquad\qquad\qquad \rho_k^{nn} \end{bmatrix} \tag{F.3}$$

Where T denotes the transpose of the matrix, $\rho_k^{ij}$ = corr $[Z_t^{(i)}, Z_{t-k}^{(j)}]$ is the lag-k cross-correlation coefficient between $Z_t^{(i)}$ and $Z_{t-k}^{(j)}$ for $i \neq j$. In order to calculate the correlation coefficient for the ARG(1) model, the monthly data of the English River have to be arranged in a way that facilitates the calculations. The single time series is manipulated to resemble the case of four time series as follows:

| | Year 1 | | | Year 2 | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| **Time series 1:** | Jan | May | Sep | Jan | May | ... | ... | ... | ... |
| **Time series 2:** | Feb | Jun | Oct | Feb | Jun | ... | ... | ... | ... |
| **Time series 3:** | Mar | Jul | Nov | Mar | Jul | ... | ... | ... | ... |
| **Time series 4:** | Apr | Aug | Dec | Apr | Aug | ... | ... | ... | ... |

$$\hat{M}_o = \begin{bmatrix} R_o^{11} & R_o^{12} & R_o^{13} & R_o^{14} \\ R_o^{21} & R_o^{22} & R_o^{23} & R_o^{24} \\ R_o^{31} & R_o^{32} & R_o^{33} & R_o^{34} \\ R_o^{41} & R_o^{42} & R_o^{43} & R_o^{44} \end{bmatrix} \tag{F.4}$$

**Table F.1. Arrangement of the English River data for calculations of $M_o$**

| $R_o^{11}$ | | $R_o^{12}$ | | $R_o^{13}$ | | $R_o^{14}$ | |
|---|---|---|---|---|---|---|---|
| Jan | Jan | Jan | Feb | Jan | Mar | Jan | Apr |
| May | May | May | Jun | May | Jul | May | Aug |
| Sep | Sep | Sep | Oct | Sep | Nov | Sep | Dec |
| Jan | Jan | Jan | Feb | Jan | Mar | Jan | Apr |
| May | May | May | Jun | May | Jul | May | Aug |

$$\hat{M}_o = \begin{bmatrix} 1.00 & 0.835 & 0.720 & 0.527 \\ 0.835 & 1.00 & 0.916 & 0.698 \\ 0.720 & 0.916 & 1.00 & 0.833 \\ 0.527 & 0.698 & 0.833 & 1.00 \end{bmatrix}$$

$$\hat{M}_1 = \begin{bmatrix} R_1^{11} & R_1^{12} & R_1^{13} & R_1^{14} \\ R_1^{21} & R_1^{22} & R_1^{23} & R_1^{24} \\ R_1^{31} & R_1^{32} & R_1^{33} & R_1^{34} \\ R_1^{41} & R_1^{42} & R_1^{43} & R_1^{44} \end{bmatrix} \qquad \text{(F.5)}$$

**Table F.2. Arrangement of the English River data for calculations of $M_1$**

| $R_1^{11}$ | | $R_1^{12}$ | | $R_1^{13}$ | | $R_1^{14}$ | |
|---|---|---|---|---|---|---|---|
| Jan | May | May | Feb | May | Mar | May | Apr |
| May | Sep | Sep | Jun | Sep | Jul | Sep | Aug |
| Sep | Jan | Jan | Oct | Jan | Nov | Jan | Dec |
| Jan | May | May | Feb | May | Mar | May | Apr |
| May | Sep | Sep | Jun | Sep | Jul | Sep | Aug |

$$\hat{M}_1 = \begin{bmatrix} 0.428 & 0.609 & 0.724 & 0.804 \\ 0.420 & 0.554 & 0.632 & 0.609 \\ 0.424 & 0.513 & 0.570 & 0.548 \\ 0.302 & 0.349 & 0.381 & 0.375 \end{bmatrix}$$

$$\hat{A}_1 = \begin{bmatrix} -0.114 & 0.046 & 0.232 & 0.639 \\ -0.026 & -0.022 & 0.455 & 0.259 \\ 0.083 & -0.046 & 0.346 & 0.249 \\ 0.097 & -0.022 & 0.159 & 0.207 \end{bmatrix}$$

$$\hat{B}\hat{B}^T = \begin{bmatrix} 0.339 & 0.322 & 0.263 & 0.218 \\ 0.322 & 0.578 & 0.537 & 0.443 \\ 0.263 & 0.537 & 0.655 & 0.599 \\ 0.218 & 0.443 & 0.599 & 0.840 \end{bmatrix}$$

**Development and parameter estimation of the PARG(1)**

*English River data (single series case):*

$$Z_{\nu,\tau} = A_{1,\tau}\, Z_{\nu,\tau-1} + B_\tau\, \varepsilon_{\nu,\tau}$$

The procedures of parameter estimation are similar to those pursued with ARG(1) model.

The parameters: $\hat{M}_{01}, \hat{M}_{02}, \hat{M}_{03}, \hat{M}_{11}, \hat{M}_{12}, \hat{M}_{13}, \hat{A}_{11}, \hat{A}_{12}, \hat{A}_{13}$ are to be estimated.

$$\hat{M}_{01} = \begin{bmatrix} 1.00 & 0.812 & 0.706 & 0.436 \\ 0.812 & 1.00 & 0.906 & 0.600 \\ 0.706 & 0.906 & 1.00 & 0.758 \\ 0.436 & 0.600 & 0.758 & 1.00 \end{bmatrix}, \qquad \hat{M}_{11} = \begin{bmatrix} 0.399 & 0.490 & 0.543 & 0.593 \\ 0.376 & 0.462 & 0.448 & 0.241 \\ 0.283 & 0.341 & 0.336 & 0.184 \\ 0.200 & 0.229 & 0.217 & 0.114 \end{bmatrix}$$

$$\hat{A}_{11} = \begin{bmatrix} 0.075 & 0.150 & -0.007 & 0.476 \\ 0.006 & 0.220 & 0.384 & -0.185 \\ 0.026 & 0.112 & 0.320 & -0.139 \\ 0.042 & 0.095 & 0.171 & -0.092 \end{bmatrix}$$

$$\hat{M}_{02} = \begin{bmatrix} 1.00 & 0.764 & 0.610 & 0.535 \\ 0.764 & 1.00 & 0.885 & 0.763 \\ 0.610 & 0.885 & 1.00 & 0.925 \\ 0.535 & 0.763 & 0.925 & 1.00 \end{bmatrix},$$

$$\hat{M}_{12} = \begin{bmatrix} 0.221 & 0.502 & 0.691 & 0.839 \\ 0.241 & 0.408 & 0.556 & 0.646 \\ 0.356 & 0.432 & 0.514 & 0.556 \\ 0.346 & 0.427 & 0.506 & 0.538 \end{bmatrix}$$

$$\hat{A}_{12} = \begin{bmatrix} -0.317 & 0.176 & -0.546 & 1.379 \\ -0.091 & -0.068 & -0.133 & 0.870 \\ 0.128 & -0.096 & 0.013 & 0.550 \\ 0.115 & -0.093 & 0.079 & 0.475 \end{bmatrix}$$

$$\hat{M}_{o3} = \begin{bmatrix} 1.00 & 0.929 & 0.844 & 0.610 \\ 0.929 & 1.00 & 0.956 & 0.730 \\ 0.844 & 0.956 & 1.00 & 0.815 \\ 0.610 & 0.730 & 0.815 & 1.00 \end{bmatrix},$$

$$\hat{M}_{13} = \begin{bmatrix} 0.669 & 0.840 & 0.942 & 0.983 \\ 0.647 & 0.801 & 0.898 & 0.945 \\ 0.635 & 0.774 & 0.866 & 0.910 \\ 0.361 & 0.393 & 0.421 & 0.477 \end{bmatrix}$$

$$\hat{A}_{13} = \begin{bmatrix} -0.306 & 0.059 & 0.670 & 0.581 \\ -0.215 & -0.033 & 0.637 & 0.582 \\ -0.139 & -0.104 & 0.625 & 0.562 \\ 0.195 & -0.158 & 0.066 & 0.419 \end{bmatrix}$$

## *Appendix G*

*An example of a chaotic system*

In the literature of the chaos theory, Hénon map (Kantz and Schreiber, 1997) is one of the most commonly used examples of a chaotic system. The map:

$$x_{n+1} = a - x_n^2 + by_n, \qquad y_{n+1} = x_n \tag{G.1}$$

results in irregular solutions for different choices of $a$ and $b$. There are initial conditions for which trajectories remain in a bounded region when $|b| \leq 1$. When $a = 1.4$ and $b = 0.3$, the sequence of $x_n$ will not be periodic but ***chaotic***. Different systems, with 5000 points, are generated using Hénon map and illustrated as Figures G.1 and G.2.

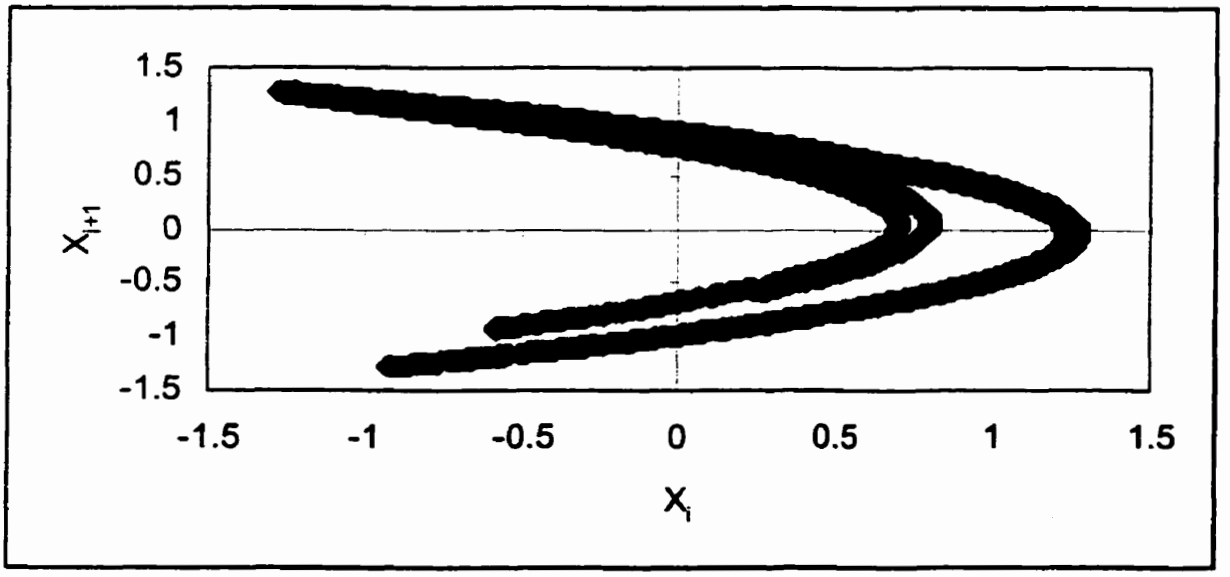
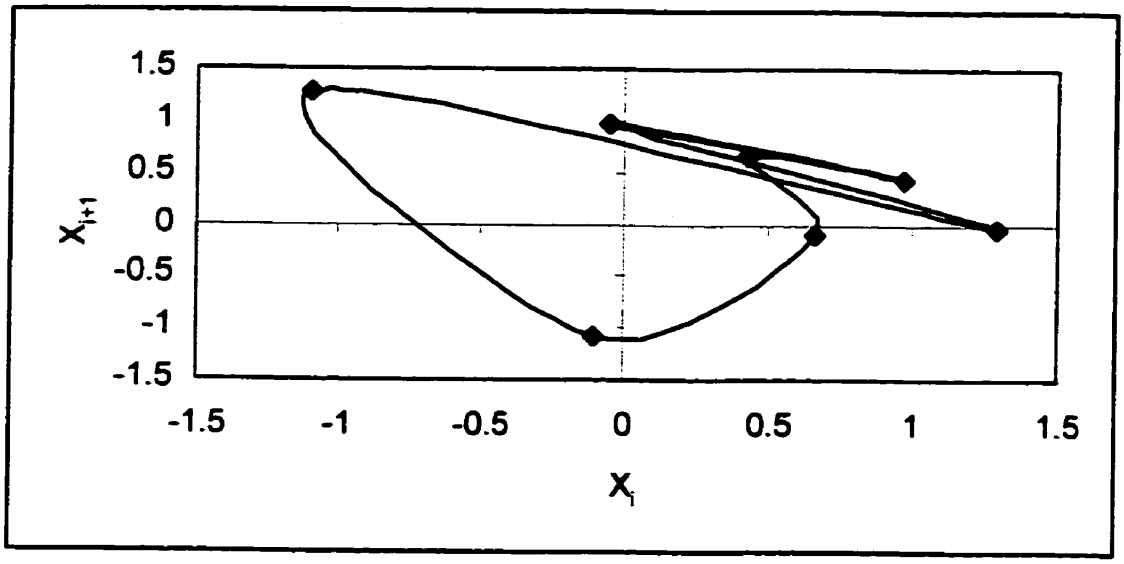
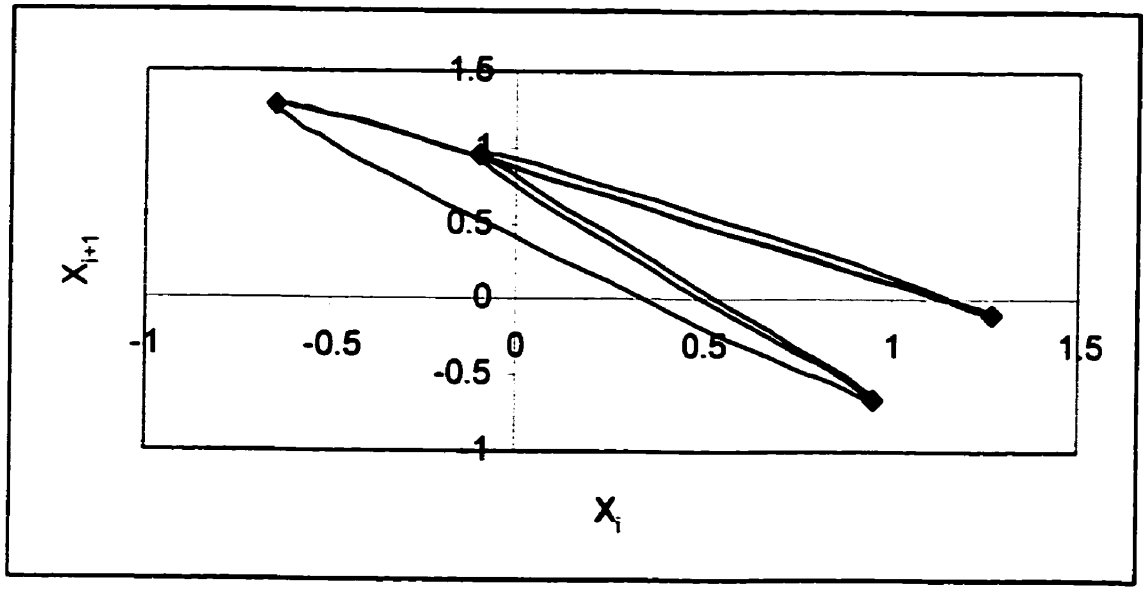


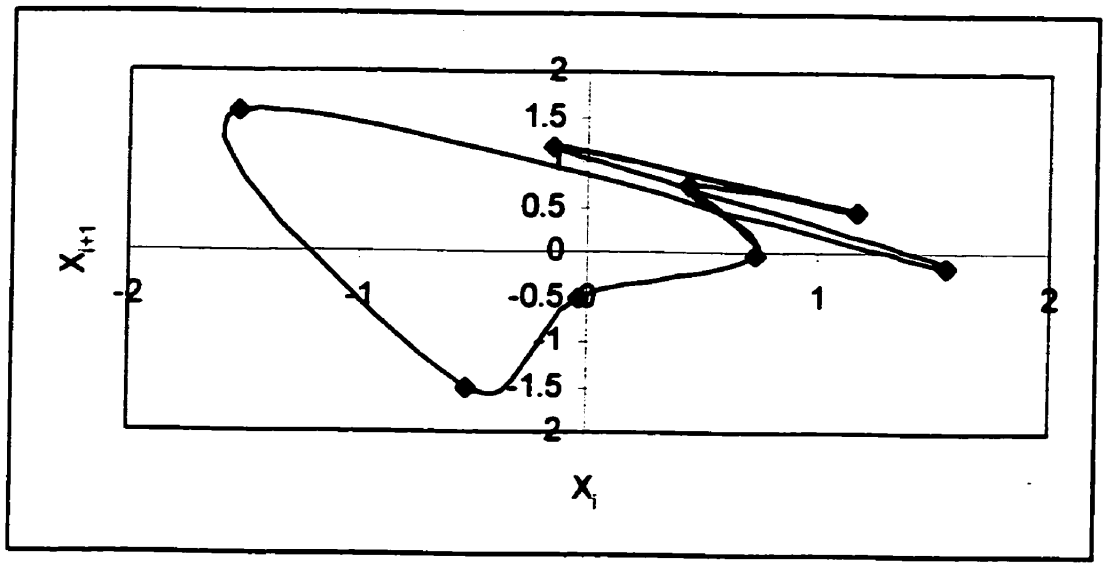

**Fig. G.1. Chaotic data generated using Hénon map (a=1.4 and b=0.3).**


1.5
1
0.5
0
-0.5
-1
-1.5
$X_{i+1}$
-1.5
-1
-0.5
0
0.5
1
1.5
$X_i$

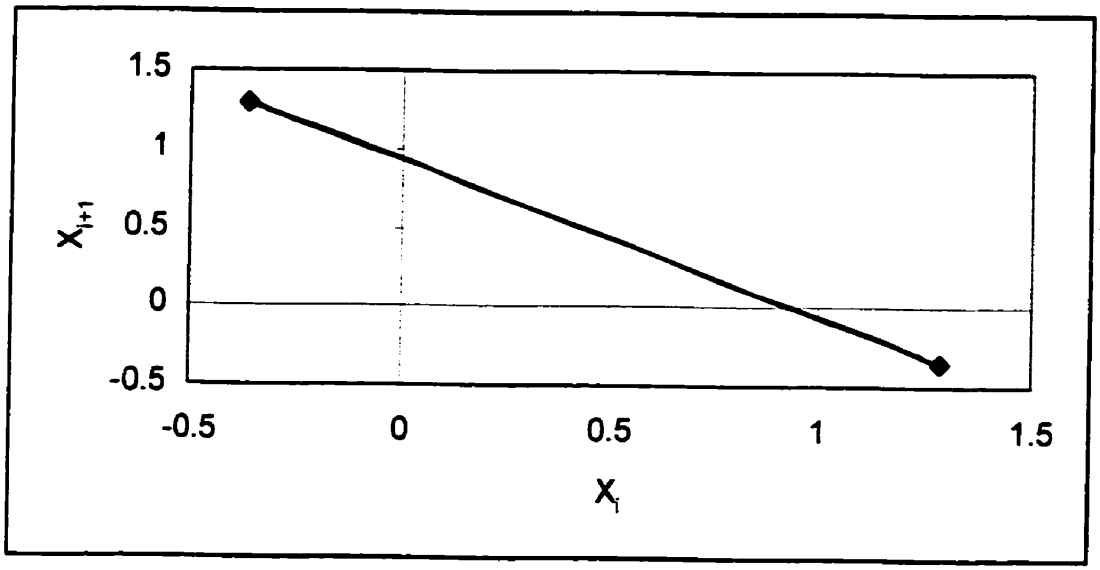
**G.2a**


1.5
1
0.5
0
-0.5
-1
$X_{i+1}$
-1
-0.5
0
0.5
1
1.5
$X_i$

**G.2b**




G.2c




G.2d

**Fig. G.2. Non-chaotic data generated by Hénon map. a) a=1.25 and b=0.3, b) a=1.0 and b=0.3, c) a=1.0 and b=0.5, and d) a=0.75 and b=0.3.**

NQ

62636

UMI

MICROFILMED 2002

# INFORMATION TO USERS

This manuscript has been reproduced from the microfilm master. UMI films the text directly from the original or copy submitted. Thus, some thesis and dissertation copies are in typewriter face, while others may be from any type of computer printer.

**The quality of this reproduction is dependent upon the quality of the copy submitted.** Broken or indistinct print, colored or poor quality illustrations and photographs, print bleedthrough, substandard margins, and improper alignment can adversely affect reproduction.

In the unlikely event that the author did not send UMI a complete manuscript and there are missing pages, these will be noted. Also, if unauthorized copyright material had to be removed, a note will indicate the deletion.

Oversize materials (e.g., maps, drawings, charts) are reproduced by sectioning the original, beginning at the upper left-hand corner and continuing from left to right in equal sections with small overlaps.

Photographs included in the original manuscript have been reproduced xerographically in this copy. Higher quality 6" x 9" black and white photographic prints are available for any photographs or illustrations appearing in this copy for an additional charge. Contact UMI directly to order.

ProQuest Information and Learning
300 North Zeeb Road, Ann Arbor, MI 48106-1346 USA
800-521-0600

# THE INFLUENCE OF ROCK FABRIC ON EXCAVATION DAMAGE IN THE LAC DU BONNET GRANITE.

BY

RICHARD ALLEN EVERITT

A thesis
Submitted to the Faculty of Graduate Studies
In Partial Fulfillment of the Requirements
for the Degree of

DOCTOR OF PHILOSOPHY

Department Of Civil and Geological Engineering
University of Manitoba
Winnipeg, Manitoba

June 06, 2001

National Library of Canada
Acquisitions and Bibliographic Services
395 Wellington Street
Ottawa ON K1A 0N4
Canada

Bibliothèque nationale du Canada
Acquisitions et services bibliographiques
395, rue Wellington
Ottawa ON K1A 0N4
Canada

*Your file Votre référence*

*Our file Notre référence*

The author has granted a non-exclusive licence allowing the National Library of Canada to reproduce, loan, distribute or sell copies of this thesis in microform, paper or electronic formats.

The author retains ownership of the copyright in this thesis. Neither the thesis nor substantial extracts from it may be printed or otherwise reproduced without the author's permission.

L'auteur a accordé une licence non exclusive permettant à la Bibliothèque nationale du Canada de reproduire, prêter, distribuer ou vendre des copies de cette thèse sous la forme de microfiche/film, de reproduction sur papier ou sur format électronique.

L'auteur conserve la propriété du droit d'auteur qui protège cette thèse. Ni la thèse ni des extraits substantiels de celle-ci ne doivent être imprimés ou autrement reproduits sans son autorisation.

0-612-62636-9

Canada

**UNIVERSITY OF MANITOBA**
**FACULTY OF GRADUATE STUDIES**
**FINAL ORAL EXAMINATION OF THE PHD THESIS**

**The undersigned certify that they have read, and recommend to the Faculty of Graduate Studies for acceptance, a PhD thesis entitled:**

**THE INFLUENCE OF ROCK FABRIC ON EXCAVATION DAMAGE IN THE LAC DU BONNET GRANITE.**

**BY**

**RICHARD ALLEN EVERITT**

**In Partial fulfillment of the requirements for the PhD Degree**

.....
**Dr. E. Z. Lajtai, Advisor**

.....
.....
.....
.....

.....
**External Examiner**
**Dr. A. R. Cruden**
**Department of Geology**
**University of Toronto**
**Toronto, Ontario**

**Date of Oral Examination:.....June 6, 2001.....**

**The Student has satisfactorily completed and passed the PhD Oral Examination.**

.....
**Advisor**

.....
**Chair of PhD Oral**

.....
.....
.....
.....

**(The signature of the Chair does not necessarily signify that the Chair has read the thesis.)**

THE UNIVERSITY OF MANITOBA

FACULTY OF GRADUATE STUDIES

*****

COPYRIGHT PERMISSION

# THE INFLUENCE OF ROCK FABRIC ON EXCAVATION DAMAGE IN THE LAC DU BONNET GRANITE.

BY

RICHARD ALLEN EVERITT

A Thesis/Practicum submitted to the Faculty of Graduate Studies of The University of Manitoba in partial fulfillment of the requirement of the degree

of

DOCTOR OF PHILOSOPHY

RICHARD ALLEN EVERITT © 2001

Permission has been granted to the Library of the University of Manitoba to lend or sell copies of this thesis/practicum, to the National Library of Canada to microfilm this thesis and to lend or sell copies of the film, and to University Microfilms Inc. to publish an abstract of this thesis/practicum.

This reproduction or copy of this thesis has been made available by authority of the copyright owner solely for the purpose of private study and research, and may only be reproduced and copied as permitted by copyright laws or with express written authorization from the copyright owner.

# ABSTRACT

The Canadian concept for disposing of nuclear fuel waste involves placing and sealing it in excavations 500 to 1000 m deep in plutonic rocks of the Canadian Shield. Investigations pertaining to the Canadian concept for used nuclear fuel waste disposal in plutonic rocks were conducted at AECL's Underground Research Laboratory (URL). The URL is excavated within the Archean granite of the Lac Du Bonnet batholith (LDBB), approximately 120 km northeast of Winnipeg, Manitoba, at the western edge of the Canadian Shield.

At the URL, AECL has undertaken a wide range of experiments to test the feasibility of burying nuclear fuel waste in granitic rocks. The investigations have included detailed mapping of geological features, stress measurement, and a recording of excavation response.

The purpose of this thesis is to show how the geological conditions affect excavation stability, in particular how the fabric of granite influences excavation damage development including overbreak around tunnel or shaft perimeters.

Most of the design of the excavations has been based on the assumption that the host granite is homogeneous (i.e. texturally and mechanically homogeneous). This thesis demonstrates, through a program of field observations and laboratory testing, that the Lac

du Bonnet, and probably most granites, are highly heterogeneous and anisotropic at both the macro and the micro scale.

Field investigations included detailed mapping. This had four objectives:

1. to define the general site characteristics including: 1) the geology of the site, 2) the fault and fractures domains, and 3) the in-situ stress domains. This was supported by field and laboratory investigations including fracture, foliation and microcrack analyses. This work defined the nature and types of rock mass heterogeneities in the rock mass.
2. to locate and describe evidence where fabric anisotropy has influenced rock mass response. This was not restricted to only present-day excavation response, but also included observations from ancient faults and fractures, whose distribution and orientation provided the first hint of the anisotropic behaviour of the LDBB granite. In this regard, the ancient fractures provide a natural analogue for the long-term / large-scale response of the granite to stresses.
3. to characterize areas where excavation damage, including overbreak, occurred, and
4. to select and characterize areas where samples were collected for rock properties testing.

For this last purpose, three locations were selected at the 420 m depth (420 Level) to provide a representative sample of the range of the compositional and textural heterogeneity present in the LDBB, and its affect on rock strength and rock mass response. These are:

- a fine grained granite dyke with subvertical flow banding,

- a medium grained and weakly to moderately layered granite forming the main mass of the batholith, and
- a generally coarse grained and leucocratic unit that occurs as sills and recrystallized zones in the main phase.

For the sake of convenience, these are referred to as the fine, medium and coarse grained granites, respectively.

The three case studies were sited in newly excavated tunnels at the 420 Level (420 m depth of the URL) where the damage due to stress relief can be observed in the process of developing. The sites are in close proximity to each other. They are distant from faults or major fractures, and are in the same in-situ stress-domain. As such, they are subject to the same stress magnitudes and orientations.

At each of the three locations, the tunnel profile was surveyed, and the geology and excavation damage were then added to the profile. An array of boreholes was then designed for each of the three locations. The arrays were designed to intersect the most prominent foliation / microcrack set at 90°, and then at 20-30° increments until the core axis was parallel to the foliation. Boreholes, twelve in all, provided a total of 171 samples along these profiles.

The core was logged for foliations and microcracks, and then cut to provide oriented samples for uniaxial compression tests. Structure orientations were measured and converted from a relative (borehole) coordinate system to true orientations.

Strength anisotropy and the influence of foliation were then systematically investigated for each of the fine, medium and coarse grained granites using rock mechanics tests that are sensitive to planar anisotropy. This included: 103 Point Load tests to determine preferred microcrack directions and their relative strengths, and *57* Brazilian tests to determine the average tensile strength and any directional anisotropy. In addition to these tests, the borehole arrays provided samples at varying angles to the fabric with which to test for anisotropy in the uniaxial compressive strength of each rock type.

The excavations and boreholes at the URL site provide a unique cross section 1100 m deep through an Archean batholith. The underground mapping and boreholes clearly indicate that what appears on the surface to be a homogeneous Archean granite, is in fact coarsely layered with vertical gradations in lithology, in batholith structure, in the development of both natural and excavation-induced fractures, and in alteration.

Layering, grain size, foliation and microcrack alignment influence excavation damage development, including the depth and shapes of overbreak. The reason for this is that the tensile strength along layering, foliation and microcracks alignment is significantly lower than in direction unaffected by these. Similarly, grain size has a strong influence. Both the tensile and the compressive strength of granite are higher in the finer grained variety of the Lac du Bonnet granite.

# ACKNOWLEDGEMENTS

I would first like to thank my wife, Maria, for her tireless support and encouragement and faith in my pursuit of a doctorate degree, and also my daughter Louisa, for being so patient while I was immersed in my university work.

This thesis represents a culmination of ideas and experience gained from working in a unique research environment. I am indebted to AECL Research for the financial support provided to pursue graduate studies at the University of Manitoba. AECL also funded my studies of the microfractures in the Lac du Bonnet granite, and the uniaxial compression testing of the pegmatite. These were initially components of this thesis, but are now presented separately owing to time limitations. Ontario Power Generation also provided financial support. I would particularly like to thank Mark Jensen and Frank King for their involvement. I would also like to acknowledge the advice and encouragement given to me by Anton Brown, Derek Martin, Cliff Davison, Ed Dzik, Jason Martino, Keith Nuttall, Rod Read, Denver Stone and Gary Simmons. Gig Hiltz, Ed Jacobs and Paul Gann provided invaluable assistance with figure preparation, and Dave Woodcock with the preparation and retrieval of the core samples used in this investigation.

Finally, I would also like to express sincere thanks to my advisor, Dr. Emery Lajtai, and to my examination committee (Drs. Brian Stimpson, Alexander Cruden, Neil Chandler Nash Soonawala, Nabil Bassim, and Vijay Chatoorgoon) for their guidance and recommendations that have helped in the completion of this thesis.

# SYMBOLS AND NOTATION

| | |
|---|---|
| x,y,z | *Rectangular Cartesian coordinates* |
| E | *Young's modulus* |
| G | *Shear modulus* |
| V | *Poisson's ratio* |
| $\sigma_1$ | *Maximum principal stress component* |
| $\sigma_2$ | *Intermediate principle stress component* |
| $\sigma_3$ | *Minimum principle stress component* |
| $\sigma_c$ | *Uniaxial compressive strength* |
| $\sigma_t$ | *Tensile strength* |
| $\sigma_{ci}$ | *Crack initiation stress* |
| $\sigma_{cd}$ | *Crack damage stress (start of unstable crack growth)* |
| Is | *Point load strength index* |
| $\sigma_{tB}$ | *Brazilian tensile strength* |
| k | *ratio of Young's moduli in tension and compression* |
| $\varepsilon_v$ | *volumetric strain* |
| $\varepsilon_a$ | *axial strain* |
| $\varepsilon_c$ | *circumferential strain.* |

# CONTENTS

ABSTRACT ..... ii

ACKNOWLEDGEMENTS ..... vi

SYMBOLS AND NOTATION ..... vii

LIST OF FIGURES ..... xiv

LIST OF TABLES ..... xxxvi

Chapter 1 INTRODUCTION ..... 1

1.1. Problem Statement ..... 3

1.2. Objective, Scope of Work, and Thesis Organization ..... 14

Chapter 2 LITERATURE REVIEW ..... 28

2.1. Introduction ..... 28

2.1.1. Influence of pore fluid on strength ..... 30

2.1.2. Influence of specimen size ..... 31

2.1.3. Influence of anisotropy ..... 34

2.2. The Experimental Determination Of Strength Anisotropies ..... 39

2.2.1. Determination of Tensile Strength ..... 40

2.2.2. Point Load Test ..... 41

2.2.3. Brazilian Tensile Strength Tests ..... 43

2.2.4. Compressive strength ..... 45

2.2.5. Application to this Thesis ..... 47

2.3. Anisotropy in Granite ..... 48

2.3.1. Introduction ..... 48
2.3.2. Field evidence for primary and secondary fabric anisotropy in granite. ..... 53
2.3.3. Application to this thesis ..... 56
2.4. Field And Laboratory Evidence For Mechanical Anisotropy In Granite ..... 57
2.4.1 Large-scale (e.g. decimetre) compositional layering or heterogeneity. ..... 57
2.4.2. Influence of texture ..... 58
2.4.3. Mineral alignment and natural microcracking ..... 59
2.4.4. Foliations and mineralogy ..... 61
2.4.5. Microcracks ..... 62
2.4.6. Conclusion --- Influence of Fabric Anisotropy in Other Granites ..... 65
2.5. Rock Properties Testing in the Lac du Bonnet Batholith. ..... 65
2.6. Conclusions ..... 72

Chapter 3 SITE DESCRIPTION ..... 75
3.1. General Geology of the Lac du Bonnet Batholith ..... 75
3.2. Underground Research Laboratory ..... 76
3.3. Batholith Structure ..... 77
3.4. Outcrop-scale Rock Units ..... 85
3.4.1. Fine Grained Granite ..... 86
3.4.2. Medium Grained Granite ..... 87
3.4.3. Coarse Grained Granite ..... 88

3.5. Rock Fabric ..... 88

3.6. Faults and Fractures ..... 93

3.7. In-situ Stresses ..... 99

3.7.1 Convergence Measurements ..... 101

3.7.2 Overcoring ..... 102

3.7.3 Hydraulic Fracturing ..... 104

3.7.4 Microseismic Monitoring ..... 107

Chapter 4 NATURAL MICROCRACKS ..... 110

4.1. Terminology ..... 113

4.2. Scope of Work ..... 115

4.3. Sampling ..... 116

4.4. Natural vs. Induced Microcracks ..... 117

4.5. Microcrack Distribution ..... 118

4.6. Natural Microcrack Sets ..... 122

4.7. Natural Microcrack Length ..... 124

4.8. Conclusions of Microfracture Study ..... 125

Chapter 5 FIELD OBSERVATIONS ON THE INFLUENCE OF GEOLOGY ON FRACTURE DEVELOPMENT ..... 126

5.1. Evidence of Fabric Control of Natural and Induced Fractures at the URL ..... 126

5.2. Fabric Control of Large Subhorizontal Faults ..... 128

5.3. Fabric Control of Subvertical Natural Joints ..... 131

5.4. Fabric Control of Excavation - Induced Fractures ..... 135

5.5. Fabric Control of Excavation - Induced Fractures in Horizontal Tunnels at the 240 Level ..... 136

5.6. Fabric Control of Excavation - Induced Fractures in Horizontal Tunnels at the 420 Level ..... 140

5.7. Fabric Control of Excavation - Induced Fractures about the Vertical Shaft at the 300 Level ..... 146

5.8. Conclusion of Field Investigations into Fabric Control of Excavation – Induced Fractures ..... 149

Chapter 6 SAMPLE COLLECTION FOR ROCK PROPERTIES TESTING ..... 150

6.1. Coarse Grained Granite ..... 151

6.2. Medium Grained Granite ..... 152

6.3. Fine Grained Granite ..... 153

Chapter 7 POINT LOAD TESTING ..... 161

7.1. Introduction ..... 161

7.2. Specimen Preparation ..... 161

7.3. Methodology ..... 164

7.4. Determination of Orientation ..... 166

7.5. Summary of Point Load Testing Results ..... 171

7.5.1. Presentation ..... 171

7.5.2. General Observations. ..... 173

7.5.3. Influence of Grain Size on Tensile Strength ..... 174
7.5.4. Influence of Fabric on Tensile Strength ..... 174
7.5.5. Coarse Grained Granite, Borehole 418-051-RT 2, 3, 4, 5. ..... 174
7.5.6. Medium Grained Granite, Borehole 417-071-RT 4, 5. ..... 178
7.5.7. Fine Grained Granite, Borehole 421-032-RT1, 2, 3, 4. ..... 180

Chapter 8 BRAZILIAN TESTS ..... 183
8.1. Rock Type and Specimen Preparation ..... 189
8.2. Brazilian Test Results ..... 191
8.3. General Observations. ..... 192
8.4. Fine Grained Granite, Borehole 421-032-RT1, 2, 3, 4. ..... 192
8.5. Medium Grained Granite, Borehole 417-071-RT 4, 5. ..... 199
8.6. Coarse Grained Granite, Borehole 418-051-RT 2, 3, 4, 5. ..... 204
8.7. Summary and Conclusions from the Brazilian Tests ..... 210

Chapter 9 UNIAXIAL COMPRESSION TESTS ..... 212
9.1. Introduction ..... 212
9.2. Rock Type and Specimen Preparation ..... 212
9.3. Test and Apparatus And Procedure ..... 213
9.4. Presentation Of Results ..... 213
9.5. Interpretation of Stress-strain Curves ..... 214
9.6. Summary of Observations ..... 217
9.6.1. Influence of Grain Size ..... 217

9.6.1. Influence of Grain Size ..... 217

9.6.2. Influence of Rock Fabric ..... 217

9.6.3. Summary and Conclusions from Uniaxial Compression Tests ..... 219

Chapter 10 SUMMARY ..... 228

10.1. Outline of Investigation ..... 229

10.2. Identification of Anisotropy in the "Homogeneous Lac du Bonnet Batholith ..... 231

10.3. Natural Microcracks ..... 231

10.4. Influence of Grain Size on Rock Strength ..... 233

10.5. Influence of Rock Fabric on Compressive Strength ..... 233

10.6. Influence of Grain Size on Tensile Strength ..... 234

10.7. Fabric-induced Anisotropy of Tensile Strength ..... 234

10.8. Influence of Natural Microcracks on Tensile Strength ..... 235

10.9. Summary of Variations by Rock Type ..... 236

10.10. Application Of Results ..... 238

REFERENCES ..... 243

Appendix A POINT LOAD TESTS ..... 276

Appendix B BRAZILIAN TESTING ..... 337

Appendix C UNIAXIAL COMPRESSION TESTS ..... 371

# LIST OF FIGURES

Figure 1. Regional setting of the Lac Du Bonnet batholith of Manitoba, the study area at AECL's Underground Research Laboratory (URL), and the excavations comprising the URL. Locations referred to in the text are shown. ..... 2

Figure 2. Schematic representation of excavation damage types observed in the URL. Breakouts (1) occur in all shafts and tunnels below 255 m, whether bored or blasted. Spalling and other excavation-induced fractures (2,3,4) occur at all levels in the URL, but are most intensely developed in the highly stressed rock below 255 m. ..... 8

Figure 3. Spalling developed along a tunnel wall at the 240 Level of the URL during its construction.. The tunnel axis is parallel to $\sigma_1$ (25.8 MPa), $\sigma_2$ (16.8 MPa) is nearly horizontal while $\sigma_3$ (12.8 MPa) is nearly vertical..... 9

Figure 4. Influence of rock fabric on excavation damage - example from the 240 Level. This figure presents a schematic view of the fabric elements present at the 240 Level, and the damage associated with different tunnel orientations. Both tunnels are constructed in the same medium grained, coarsely layered granite, but differ in their orientation. The profile of Room 209 is in the $\sigma_1 - \sigma_3$ plane, while the profile of Room 207 is in the $\sigma_2$ - $\sigma_3$ plane, Mapping of the two tunnel

profiles revealed that the damage about Room 209 is limited to minor surface exfoliation, with a symmetrical distribution of excavation-induced (stress-relief) fractures. In contrast, the damage about Room 207 is more extensive, with a markedly asymmetrical distribution of excavation-induced (stress-relief) fractures. A 1x 1.5 m slab spalled from the tunnel in the area where the perimeter is parallel to the layering..... 10

Figure 5. Influence of rock fabric on excavation damage - example from the 300 Level. In the shaft at this depth, weak or incipient breakouts were occurring in the NE and SW quadrants of the shaft, due to spalling of slabs from the shaft wall. Most spalling occurred in the NE quadrant, and additional loose (partially spalled) slabs at this location was indicated by the hollow sound when hit hard by a scaling bar. The corresponding breakout in the SW quadrant was, in comparison, less developed with less loose rock. Mapping of the station floor at this depth provided a cross section of the damage about the shaft, and an explanation for the difference in breakout development described above. In the NE quadrant, spalling was enhanced by parting along the subvertical foliation in the fine grained granite. In the SW quadrant, the foliation was subhorizontal, and therefore not favourably oriented to promote fracture development under these stress conditions..... 11

Figure 6. Influence of geology on breakout development within a single drill-and-blast tunnel. Room 407 at the 420 Level of the URL provides an example of the influence of subtle differences in geology on breakout development in granite where conventional excavation is used. The tunnel profiles developed in the medium grained and fine grained granites are shown in A and B, respectively. In the medium grained granite, the parting of the rock along planes of biotite facilitates breakout development. In the fine grained granite, a biotite foliation also exists but it is subvertical and normal to the tunnel axis. As such, its presence does not influence breakout development for this excavation orientation and geometry. ..... 12

Figure 7. Influence of geology on breakout development within a single drill-and-blast tunnel. Room 421 at the 420 Level of the URL provides an example of the influence of subtle differences in geology on breakout development in the fine grained granite where conventional excavation is used. In the previous figure it was noted that ... "a biotite foliation also exists in the fine grained granite, but it is subvertical and normal to the tunnel axis. As such, its presence does not influence breakout development for this excavation orientation and geometry". In Room 421 however, the tunnel axis is parallel to the strike of the foliation, and spalling now occurs along this structure..... 13

Figure 8. Distribution of measurements relative to rock structure. The orientation of the arrays relative to the local rock structure is shown. The influence of the subhorizontal layering is tested by the arrays in the coarse and medium grained granite, while the influence of the subvertical layering in the fine grained granite is tested by the third array in the fine grained granite, which occurs as dykes intruding the medium and coarse grained granites. The geometry of the complete sample array at each location is also shown. ..... 25

Figure 9. Approximate extent of the zone of core discing around a tunnel at the 420 Level..... 26

Figure 10a. Simplified geological map of the upper portion of the URL access shaft (continued next page, refer to text for explanation)..... 80

Figure 10b. Simplified geological map of the lower portion of the URL access shaft (refer to text for explanation). ..... 81

Figure 11. Block diagrams summarizing the structures in the Lac du Bonnet Batholith as seen at AECL's URL. A: Large-scale compositional layering and the low-dipping thrust faults, which are concordant with the layering. B: Faults, joints and natural microcracks within the same area..... 82

Figure 12. Rock units and structures recognized at depth in the URL, and selected for rock properties testing in this study. Each photograph is of a vertical tunnel face. The scales are approximately the same, each

photo showing an area of approximately 3.5 m by 2.5 m. The measuring stick is 1m long. The fine and medium grained granites are shown in the upper photograph (A). The coarse grained and (locally xenolithic) granite is shown in the lower photograph (B). ..... 83

Figure 13. Typical appearance in core, of the granite types selected for rock properties testing in this investigation. From top to bottom are the fine-, coarse-, and medium- grained granites, respectively, and the pink (altered) equivalent of the coarse granite. ..... 84

Figure 14. Grain sizes and variation with mineralogy, of the three rock types selected for rock properties testing. ..... 85

Figure 15. A summary of fabric elements present at the 420 Level, listed in order of relative age. The orientations of each structure are summarized on the stereographic projections, and in the accompanying conceptual block models which also summarize the age relationships..... 91,92

Figure 16. Summary of the faults and subvertical fractures as seen in the walls of access shaft (the shaft from surface to 255 m has a rectangular cross section, which changes to a circular cross section below this depth). The shaft below 290 m is unfractured, and for this reason is not shown. ..... 95

Figure 16b. Fracture Zone 2 as seen in the access shaft and its environs. The vertical section shows the offset determined from multiple borehole intersections and shaft mapping. The incremental nature of offset and

the sense of movement is indicated in the inset (a portion of the shaft wall). Fracture orientations and the average fault orientation in the hangingwall, fault interior and footwall are shown in the stereographic projections. (modified from Everitt and Brown, 1996)..... 96

Figure 16c. Faulting in the Late Archean -Early Proterozoic. The low dip thrust faults and associated splays are believed to be associated with the craton-scale faulting shown (modified from Osmani 1991)..... 97

Figure 17. In-situ stress domains in the URL in relation to the fracture domains (stress data from Martin 1989) For explanation refer to text..... 98

Figure 18. Examples of microcrack varieties as seen in a coarse grained phase of the Lac Du Bonnet Batholith, exposed on a tunnel wall at the 240 Level of the URL. The whitish linear groove running the length of the photograph is the remnant of a percussion-drilled perimeter blasthole as seen on the tunnel wall. The rock immediately adjacent to the hole appears whiter than the rock surrounding it due to abundant blast-induced microfractures. A much larger but related (20 cm long) excavation induced "microfracture" runs parallel to the blasthole. These are easily distinguished from natural microcracks by the presence of mineral infilling or alteration, as in the prominent hematite -stained microcrack in the large microcline crystal near the centre of the photograph. ..... 112

Figure 19. Photomicrograph of a single hematite-filled natural microcrack in quartz. Photograph by C. McGregor..... 117

Figure 20. Map of natural microcracks (hematite or fluid inclusion-filled planes) within a sample from the 420 Level..... 120

Figure 21. Summary stereographic projections of poles to fracture and microcracks sets from the 130, 240 and 420 Levels. Infilling types: He –hematite, Chl – chlorite, Go – goethite, Ca – carbonate. Principal stress orientations are indicated (see Table 2). . ..... 121

Figure 22. Schematic illustration of the sets of natural microcracks in quartz at the URL..... 123

Figure 23. Variation with depth of the average length of natural microfractures in quartz..... 124

Figure 24. Influence of batholith contacts and internal structure on fault orientation. Structure contours showing the depth (in km) to the base of the Lac du Bonnet Batholith are shown in (A). Dekametre-scale layering within the pluton varies locally in orientation, but is broadly parallel to the lower contact. The low-dipping fault zones (e.g. Fracture Zone 2) are controlled in location and orientation by the layering. Structure contours on the lower contact of Fracture Zone 2 are shown in B..... 129

Figure 25. Fracture Zone 2 as seen in the URL shaft. An example of fabric control of fracture orientation and location on the dekametre-scale. This photomosaic presents an unrolled 360° view of the walls of the access shaft over a vertical distance of about 8 m. The squares of the

metal grid visible at the bottom of the mosaic are each 40 by 40 cm. The apparent sinusoidal trace of the fault is an artifact formed by the intersection of a plane (the fault) with an unrolled cylinder (the shaft wall). ..... 130

Figure 26. The influence of rock fabric on natural fractures in the Lac du Bonnet Granite. Histogram depicting the frequency of several classes of natural fractures (faults and joints) as a function of the dekametre-scale fabric of the Lac du Bonnet Batholith. Variation in fracture frequency with granite fabric type, based on data from a 240-m scan-line down the URL upper vent raise. ..... 132

Figure 27. The influence of rock fabric on natural fractures in the Lac du Bonnet Granite - An example of fabric control of fracture orientation and location on the grain scale - natural microcracks developed in a zone of coarse microcline crystals ..... 133

Figure 28. Control of fracture location and extent due to variations in granite at the 240 Level. Fractures are most frequent in the coarse grained granite sills, and in the fine grained granite dyke, while the main phase medium grained granite is largely unfractured. Note: these are all phases of the Lac du Bonnet Batholith, and not separate unrelated rock units. ..... 134

Figure 29. Influence of rock fabric on excavation damage - example from the 240 Level. This figure depicts the damage associated with different tunnel orientations in the same rock type – a medium grained granite. While

apparently massive in hand specimen, more careful inspection reveals that the granite possesses both a weak mineral alignment (foliation) and a decimetre-scale layering. These apparently weak structures nevertheless impart a mechanical anisotropy to the rockmass, which is evident in the differences in excavation damage between these two profiles. The tunnel profile of Room 207 is asymmetric, with induced fractures paralleling the gneissosity being not only more frequent, but also larger than those in other directions, with spalling developing in the upper right corner (also shown in Figure 30). In Room 209, the layering is not favourably oriented to serve as exfoliation planes, and - the damage profile is symmetrical. ..... 137

Figure 30. Influence of rock fabric on excavation damage - example from the 240 Level. This photograph shows the surface of the large spall formed by separation along the plane of layering and parallel biotite alignment in the granite in Room 207 (shown in Figure 29). The inset shows the field of view and the relationship between the tunnel profile, layering and the spall. This type of damage is not present in the tunnels, which are normal, or near normal to the strike of the layering. Spalling along the layering is arrested wherever the layering is absent or recrystallized - as in this case – by intrusion of fine grained granite dykes on the right and left sides of the field of view. ..... 139

Figure 31. Influence of rock fabric on excavation damage - example from the 420 Level ..... 142

Figure 32. Influence of geology on breakout development within a single line-drilled tunnel. Room 415 at the 420 Level of the URL provides an example of the influence of subtle differences in geology on breakout development in granite. The geology, notch depth and construction are shown. ..... 143

Figure 33. Influence of geology on breakout development within a single drill-and-blast tunnel. Room 407 at the 420 Level of the URL provides an example of the influence of subtle differences in geology on breakout development in granite where conventional excavation is used. The tunnel profiles developed in the fine grained and medium grained granites are shown in A and B, respectively. ..... 144

Figure 34. Influence of geology on breakout development within a single drill-and-blast tunnel. Room 421 at the 420 Level of the URL provides an example of the influence of subtle differences in geology on breakout development in the fine grained granite where conventional excavation is used. In the previous figure it was noted that ... "a biotite foliation also exists in the fine grained granite, but it is subvertical and normal to the tunnel axis. As such, its presence does not influence breakout development for this excavation orientation and geometry". In Room 421 however, the tunnel axis is parallel to the strike of the foliation, and spalling now occurs along this structure. A similar occurrence occurred in the shaft at the 300

Level, with spalling developed along this foliation where it paralleled the shaft wall (see Figure 35 and section 5.7 following). ..... 145

Figure 35. Simplified geological map showing the excavation damage at the 300 Level station. ..... 148

Figure 36. General location of the 420 Level, and the names of rooms referred to in this text. ..... 154

Figure 37. 420 Level geology, upper ramp level (depth 354 m). ..... 155

Figure 38. 420 Level geology, main level (depth 415 m). ..... 156

Figure 39. 420 Level geology, lower level, (depth 437 m). ..... 157

Figure 40. The geology, tunnel profile, excavation damage, and sampling array at the coarse grained granite location in Room 418. ..... 158

Figure 41. The geology, tunnel profile, excavation damage, and sampling array at the medium grained granite location in Room 417. ..... 159

Figure 42. The geology, tunnel profile, excavation damage, and sampling array at the medium grained granite sample site, Room 421. ..... 160

Figure 43. Point Load testing, rotation of loading direction relative to foliation. Rock cylinders were placed vertically in a loading frame and subjected to axial point loading as shown. The orientation of the test-induced fracture, and the fabrics present were recorded, and converted to give the true orientation. ..... 167

Figure 44. Method for recording orientation of test-induced fracture. ..... 168

Figure 45. Conversion of relative orientations to true orientations. In each case, the "R" and red arrow represent the reference line painted on the up-dip side of the core, and oriented to show the trend and plunge directions. A: A series of samples with test-induced fractures. In this view, the core is aligned along the trend of the borehole, but "held" in a vertical orientation so that it plunges 90°. The strike of each fracture is recorded relative to the reference line. B : The data from A plotted on a lower hemisphere equal area projection, with the reference line and the poles to the induced fractures indicated. C: As A, but with the core rotated to the correct plunge for this borehole. D: Lower hemisphere equal area projection of the data in C. The reference line "R" rotates from its position in B, to the location shown here. The poles to each fracture rotate in the same direction and orientation.. ..... 169

Figure 46. Comparison of point load strength Indices by rock type (relative average grain size). ..... 172

Figure 47. Point Load test results from the coarse grained granite. For each borehole, a pair of histograms shows the strike of test-induced fractures and the strike of the foliations in the same samples..... 176

Figure 48. Schematic representation of the variation in tensile strength in the coarse grained granite and its correlation with the fabrics present. In shallow dipping boreholes (< 45 °), axial point loading of the core produces fractures following the strongest fabric (referred to for

convenience as the large scale layering, but in fact a mixture of metre to millimetre scale compositional layering, mineral alignment, schlieren (in some cases showing evidence of slip) and natural microfractures in quartz). In steeply dipping boreholes, the layering is not favourably oriented with respect to axial loading, and the plane of failure then "jumps" to other, steeper dipping foliations (this transition is reflected in the values of the average tensile strength). These relationships are shown schematically above. ..... 177

Figure 49. Point Load test results from the medium grained granite. This pair of histograms shows the strike of test-induced fractures (upper) and the strike of the foliations (lower) from the same samples. ..... 179

Figure 50. Point Load test results from the fine grained granite. The histograms show the strike of the test-induced fractures (upper), and the strike of the foliations (lower) from the same samples. This granite hosts a well developed subvertical flow banding and a low-dipping natural microcrack set. Fractures induced by axial point loading preferentially follow the (low-dipping) plane of the natural microcracks (which strike ~90° in this 2-D view), the subvertical foliation being ignored. This relationship between the natural microcrack set and the point-load induced fractures is displayed to better effect in Figure 51 following. ..... 181

Figure 51. Lower hemisphere equal area plot of poles to fractures induced by axial point loading, in fine grained granite from boreholes 421-032-RT1, 2, 3 and 4. ..... 182

Figure 52. Lower hemisphere equal area plot of poles to natural fractures (fluid inclusion or alteration-filled planes in quartz) obtained by universal stage mapping of thin section from the granodiorite. The induced fractures in Figure 51 share the same general orientation as the natural fractures shown here. ..... 182

Figure 53. Configuration of loading for the Brazilian test as seen on a single specimen. ..... 186

Figure 54. Brazilian tests, rotation of loading direction relative to foliation. ..... 187

Figure 55. Conversion of relative orientation to true orientation. ..... 188

Figure 56. Summary statistics for the Brazilian tests from the entire borehole array within the Fine Grained Granite. The array includes boreholes 421-032-RT1, 2, 3, 4. These were arranged in a horizontal plane, and oriented to intersect the subvertical flow banding at angles of ~60, ~45, ~ 10, and ~00, respectively. The data from each borehole includes loading orientations ranging from 0° to 90° to the foliation in each borehole. This was done to test for systematic variation within the horizontal plane, or put in another way - to examine the affect of borehole orientation when sampling for rock properties testing. The

results indicate that there is significant direction -dependent anisotropy..... 195

Figure 57. Brazilian test results for the Fine Grained Granite from boreholes 421-032-RT1, 2, 3, and 4, showing the variation in the estimated tensile strengths of each sample plotted against the corresponding sample density. If the damage zone around the excavation affected the samples, it would be expected that such damage would lower both the rock density and the tensile strength, and that the specimens affected would be closest to the excavation. Instead, there is no correlation between density and location (as given by the sample numbers). Tensile strength actually decreases with increasing sample density, which is the opposite of what would be expected if damage was present. ..... 196

Figure 58. Brazilian test results for the Fine Grained Granite from boreholes 421-032-RT1, 2, 3, and 4. Variation in the estimated tensile strengths of each sample against the inclination of the loading axis relative to the flow banding. The three insets show the orientation of the loading axis (arrows) relative to the flow banding as seen in a cross section of the core. The best-fit linear regression curve, the equation for the curve, and the $R^2$ correlation coefficient are included. In this example, the tensile strength is somewhat lower when loading is done parallel to the flow banding, but the trend is dispersed. ..... 197

Figure 59. Brazilian test results for the Fine Grained Granite from boreholes 421-032-RT1, 2, 3, and 4. Variation in the estimated tensile strengths of each sample against the inclination of the loading axis relative to the preferred orientation of natural microcracks. The three insets show the orientation of the loading axis (arrows) relative to the microcracks as seen in a cross section of the core. The best-fit linear regression curve, the equation for the curve, and the IN-SITU correlation coefficient are included. In this example, the tensile strength parallel to the natural microcracks is ~ 65% of that perpendicular to the microcracks. ..... 198

Figure 60. Summary statistics for the Brazilian tests from the entire borehole array within the Medium Grained Granite. The array includes boreholes 417-071-RT4 and 5. These were arranged in a vertical plane, and oriented to intersect the subvertical flow banding at angles of ~45, and ~00, respectively. The data from each borehole includes loading orientations ranging from 0° to 90° to the foliation in each borehole. This was done to test for systematic variation within the horizontal plane, or put in another way - to examine the affect of borehole orientation when sampling for rock properties testing. The results indicate that there is significant direction -dependent anisotropy..... 201

Figure 61. Brazilian test results for the Medium Grained Granite from boreholes 417-071-RT4 and 5. Variation in the estimated tensile strengths of

each sample against the corresponding sample density. There is no correlation between density and location (as given by the sample numbers). Tensile strength decreases with increasing sample density, which is the opposite of what would be expected if damage was present. ..... 202

Figure 62. Brazilian test results for the Medium Grained Granite from boreholes 417-071-RT4, and 5. Variation in the estimated tensile strengths of each sample against the inclination of the loading axis relative to the compositional layering. The three insets show the relative orientation of the loading axis (arrows) relative to the flow banding as seen in a cross section of the core. The best-fit linear regression curve, the equation for the curve, and the $R^2$ correlation coefficient are included..... 203

Figure 63. Summary statistics for the Brazilian tests from the entire borehole array within the Coarse Grained Granite. The array includes boreholes 418-051-RT1, 2, 3, 4. These were arranged in a vertical plane, and oriented so as to intersect the compositional layering at angles of ~90°, ~30°, ~15°, ~ 0°, and ~10°, respectively as shown by the schematic illustrations - (the tunnel profile is represented by the ellipse, a xenolithic zone by the grey shading, the layering by the light grey dashed lines, and the boreholes by the solid lines Note: the near-vertical borehole RT1 was not useable due to discing). The data from each borehole includes loading orientations ranging from 0° to 90° to the foliation in each borehole. This was done to test for

systematic variation within the vertical plane, or put in another way - to examine the affect of borehole orientation when sampling for rock properties testing. The results indicate that there is significant direction -dependent anisotropy in tensile strength. ..... 207

Figure 64. Brazilian test results for the Coarse Grained Granite from boreholes 418-051-RT2, 3, 4, and 5. Variation in the estimated tensile strengths of each sample against the corresponding sample density. This suite differs from the fine and medium grained granites in that tensile strength increases with increasing density. This does not however, suggest sample damage, as there is no correlation between strength and the depth of the sample from the borehole collars. Instead, this relationship appears to be due to a combination of factors, including the orientation of the borehole relative to the layering, and to differences in rock composition and degree of fabric development as indicated in the boxes above. ..... 208

Figure 65. Brazilian test results for the Coarse Grained Granite from boreholes 417-071-RT2, 3, 4, and 5. Variation in the estimated tensile strengths of each sample against the inclination of the loading axis relative to the low-dipping compositional layering. The three insets show the relative orientation of the loading axis (arrows) relative to the flow banding as seen in a cross section of the core. The best-fit linear regression curve, the equation for the curve, and the $R^2$ correlation coefficient are included. ..... 209

Figure 66. Idealized stress strain diagram for a brittle rock with the stages of brittle deformation as represented by the axial, volumetric and lateral stress-strain curves. ..... 216

Figure 67. X-Y-Z comparison chart of the minimum uniaxial compression and (Brazilian) tensile strengths of the 3 rock types. The stresses corresponding with crack closure, the onset of dilatancy and the peak load failure were determined from the stress / strain curves in Appendix C. The data for the very coarse granite is from Everitt (1995). The figure clearly indicates that the various measures of strength decrease as the grain size and heterogeneity of the granite increase. Compare with Figures 68 and 69 showing the mean and maximum uniaxial compression and (Brazilian) tensile strengths, respectively. ..... 222

Figure 68. X-Y-Z comparison chart of the mean uniaxial compression and (Brazilian) tensile strengths of the 3 rock types. The stresses corresponding with crack closure, the onset of dilatancy and the peak load failure were determined from the stress / strain curves in Appendix C. The data for the very coarse granite is from Everitt (1995). The figure clearly indicates that the various measures of strength decrease as the grain size and heterogeneity of the granite increase. Compare with Figures 67 and 69 showing the minimum and maximum uniaxial compression and (Brazilian) tensile strengths, respectively. ..... 223

Figure 69. X-Y-Z comparison chart of the maximum uniaxial compression and (Brazilian) tensile strengths of the 3 rock types. The stresses corresponding with crack closure, the onset of dilatancy and the peak load failure were determined from the stress / strain curves in Appendix C. The data for the very coarse granite is from Everitt (1995). As was shown in Figure 67 (minimum values), the various measures of strength decrease as the grain size and heterogeneity of the granite increase. This trend is obscured by the specimen-scale heterogeneity in the coarser granites, where the mean and maximum strengths are more variable (Figure 68 and 69). In other words, depending on the location and orientation of the borehole, the coarsest granite can be either weaker than, or stronger than, the medium grained granite..... 224

Figure 70. Anisotropy of compressive strength related to rock fabric, is most strikingly displayed by the fine grained granite. This figure compares the variation in the various strength estimates with orientation. Orientation refers to the angle between the strike of the dyke and the sampling boreholes, which ranges from 10°, 35°, 50°, and 75° in boreholes –RT4, - RT3, - RT2 and - RT1, respectively. The values for the peak load failure, onset of dilatancy, crack initiation, and microcrack closure are derived from the uniaxial compression data (stress /strain curves) in Appendix C. The tensile strength is from the Brazilian tests. In this rock, each of the various measures of rock

strength decrease together in a systematic manner as the orientation of the sampling borehole rotates towards parallelism with the strike of the rock unit contact. The weakest direction correlates with the strike of the dykes and of rare subvertical fractures found at this level. ..... 225

Figure 71. Influence of low-dipping fabric on rock strength for the coarse grained granite. This figure compares the variation in the various strength estimates with orientation. The values for the peak load failure, onset of dilatancy, crack initiation, and microcrack closure are derived from the uniaxial compression data (stress /strain curves) in Appendix C. The tensile strength is from the Brazilian tests. The data presented in this figure clearly indicate that directional anisotropy is also evident in the coarse grained granite. In the coarse grained granite, samples are weakest in compression when the centimetre- to decimetre –scale layering intersects the sample axis at an angle between 20° and 40°. The rock is stronger (at the specimen scale) when the sample is within the plane of layering, or sharply inclined to it. ..... 226

Figure 72. Influence of low-dipping fabric on rock strength. X-Y-Z comparison chart of the variation, with orientation, of the estimated average strengths for the medium grained granite. The values for the peak load failure, onset of dilatancy, crack initiation, and microcrack closure are derived from the uniaxial compression data (stress /strain curves) in Appendix C. The tensile strength is from the Brazilian tests.

Samples are weakest in compression when the centimetre- to decimetre –scale layering intersects the sample axis at 30°. The rock is stronger (at the specimen scale) when the sample is within the plane of layering, or sharply inclined to it. ..... 227

# LIST OF TABLES


Table 1. In-situ stresses measured at the 240 Level ..... 6

Table 2. Scope of work ..... 16

Table 3. Summary of Site Characterization by the Author. ..... 23

Table 4. Rationale for Selection of the Sample Sites. ..... 24

Table 5. Sample and Testing Summary ..... 24

Table 6. Sample Properties ..... 27

Table 7. Plutons Suites in the Superior Province (Thurston et al., (eds.) 1991, 1992 a,b). ..... 54

Table 8. Comparison of Petrological Units from surface and subsurface mapping of the Lac du Bonnet Batholith. ..... 79

Table 9. In-situ Stresses at the URL. ..... 109

Table 10. Microfracture types. ..... 111

Table 11. Sample and testing summary ..... 163

Table 12. Sample Properties ..... 163

Table 13. Average Point Load Strength Index Rock Type and Borehole. ..... 170

Table 14. Average Point Load Strength Index Rock Type and Borehole. ..... 170

Table 15. Comparison of Calculated Brazilian Tensile Strengths. ..... 185

Table 16. Sample and testing summary ..... 190

Table 17. Summary Properties. ..... 190

Table 18. Brazilian test results Fine Grained Granite from boreholes 421-032-RT1,-2,-3,4. ..... 200

Table 19. Brazilian Test Data Medium Grained Granite (Borehole 417-071-RT4,5) ..... 206

Table 20. Brazilian Test Data Coarse Grained Granite (Borehole 418-051-RT2,3,4,5) ..... 206

Table 21. Summary tables of peak load failure and other strength measures as determined by manual inspection of the stress-strain curves. (tensile strength data from Brazilian tests) ..... 220

Table 22. Uniaxial Compression Testing Summary by Rock Type and Borehole(tensile strength data from Brazilian tests) ..... 221

# Chapter 1

# INTRODUCTION

The understanding of the mechanism of fracture development around underground openings is important for the design of a safe and stable openings in mining and civil engineering applications. In the design of nuclear waste disposal vaults, a special concern is the potential for significant propagation of existing fractures, or the creation of new fractures (excavation damage) through the construction of such a facility. Interconnected fractures could reduce the effectiveness of the rock as a barrier to contaminant movement by providing pathways for groundwater to transport contaminants released from a disposal vault.

The Canadian concept for disposing of nuclear fuel waste involves placing and sealing it in excavations 500 to 1000 m deep in plutonic rocks of the Canadian Shield. Investigations pertaining to the Canadian concept for used nuclear fuel waste disposal in plutonic rocks are being conducted at AECL's Underground Research Laboratory (URL). The URL is excavated within the Archean granite of the Lac Du Bonnet batholith (LDBB), approximately 120 km northeast of Winnipeg, Manitoba, at the western edge of the Canadian Shield (Figure 1). The main purpose of AECL's investigations at the URL is to develop geologic data and experience that can then be used in locating a permanent nuclear waste facility.

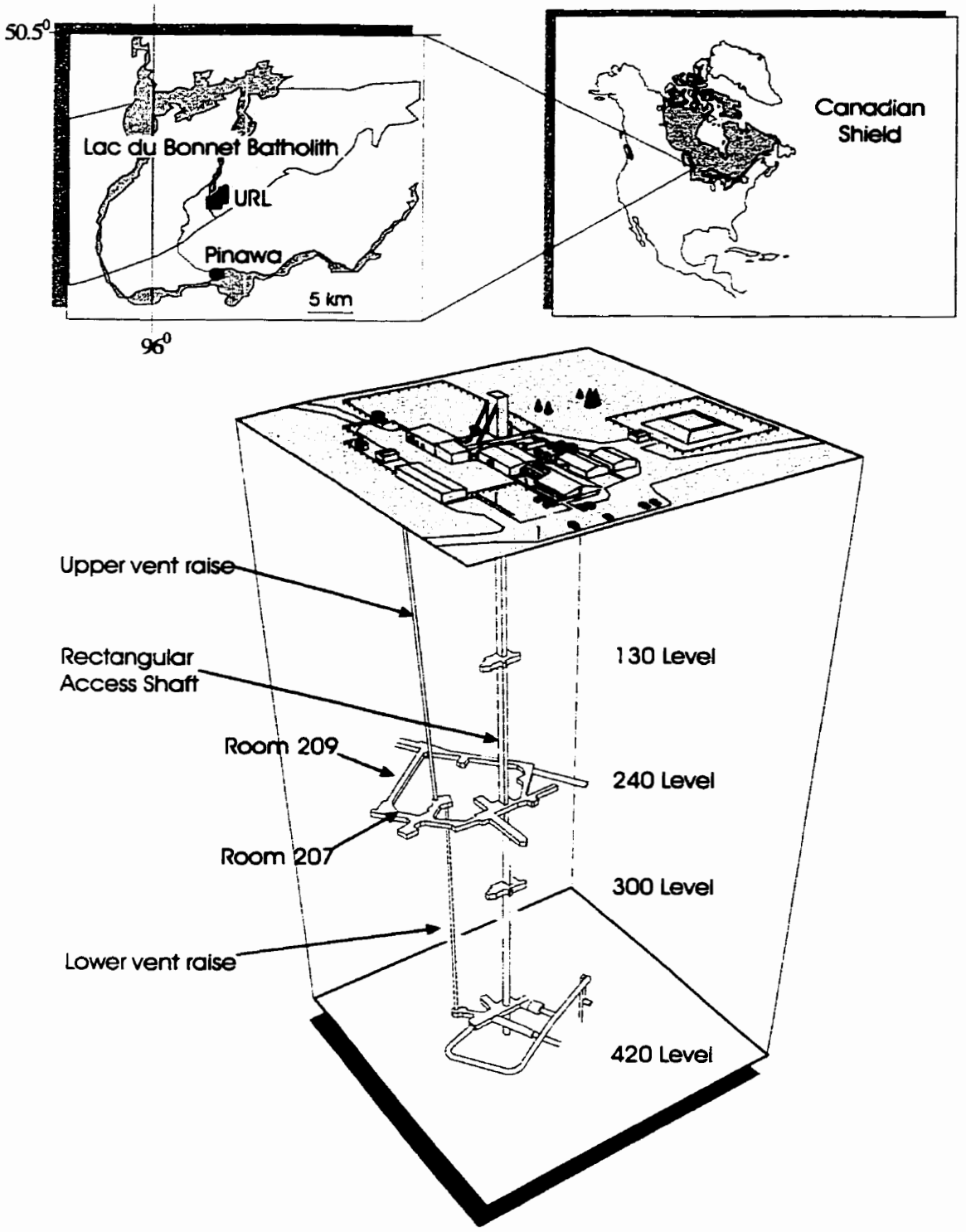



Figure 1. Regional setting of the Lac Du Bonnet batholith of Manitoba, the study area at AECL's Underground Research Laboratory (URL), and the excavations comprising the URL. Locations referred to in the text are shown.

## 1.1. Problem Statement

Mapping of the shafts and the various tunnels at the URL by the author, conducted during its construction, has indicated that excavation-induced fractures were far more numerous than natural fractures except within fault zones. In the shaft below Fracture Zone 2, in most of the 240 Level, and in all of the 420 Level, excavation-induced fractures are the only fractures present. Excavation-induced fractures are distinguished from natural fractures by the absence of mineral infilling or wall rock alteration, by their position and geometry relative to blasthole remnants, and by their position and geometry relative to the tunnel profiles or tunnel face. In general, they form an onion-skin pattern of surfaces, similar to the exfoliation fractures seen in surface outcrops, but in the subsurface their orientation and size is dependent on their location relative to the tunnel perimeter, and to the location of tunnel faces during excavation.

Prior to the excavations of the URL, there have been a number of studies to predict the response of the local granitic rocks to the new stress conditions imposed through the wide-ranging excavations, consisting of a central shaft and connecting tunnels and raises (Figure 1). Not one of these studies anticipated the massive fracture development (excavation damage of AECL terminology) that was later observed at compressive stress concentrations. There is one thing common to these and even to most of the later studies. They have all assumed that the granite is homogeneous, isotropic and relatively unfractured. From this flowed the next assumption that the in-situ strength of the granite should not be that much different from the intact rock strength determined in the laboratory using small drill core samples.

Although the in-situ stresses at the URL site have been found to be much higher than at comparable depths elsewhere in the Canadian Shield, the stresses are still not high enough to cause even minor fracture development in the intact rock samples used in the laboratory investigations. A one-half reduction of in-situ strength relative to the unconfined compressive strength due to environmental effects (heat, humidity and long-duration static loading) have been predicated by others (e.g. Lajtai et al., 1987) and some of the same factors have certainly been active at the URL site. This would explain some excavation damage, especially in the highest stress environment at 420 m depth. They do not however explain:

1. the fact that failure in the form of slabbing also exists under the much lower stresses present at the 240 m level, and
2. that documented differences in excavation damage within a single tunnel or shaft subject to the same in-situ stresses and excavated by the same means and schedule.

Examples of the excavation damage include:

1. slabbing and spalling of rock about tunnel perimeters at the 240 m depth (Figure 2; also Figure 3 and Figure 4)
2. slabbing and spalling of rock about the shaft perimeter at 300 m depth (Figure 2; also Figure 5),
3. breakout notches around the tunnels, bored vent raises, and the conventionally excavated (i.e. drill-and-blast) shafts (Figure 2), and
4. blast- or stress-relief fractures radiating away from the blasthole remnants preserved on tunnel faces (Figure 2).

None of the above can be explained while assuming homogeneous and isotropic conditions. Evidence that anisotropy of rock structure (i.e. average grain size, variation in

grain size, number and type of foliations, microcrack alignments and their orientations) was contributing to excavation damage development was first obtained during routine mapping of the shaft and tunnels during the most early phases of construction by the author, in his role as the URL site geologist. Damage development has been observed to vary in response to textural and structural variations in the granite in both the vertical (Figure 4) and horizontal planes (Figure 5).

The tunnels at the 240 Level demonstrate the influence of subtle fabric on excavation damage development in tunnels (Figure 4). The axis of tunnel (Room) 207 is parallel to the maximum principal stress and to the strike of the gneissosity, while Room 209 is oriented perpendicular to them. There was hardly any damage around the tunnel oriented parallel to the intermediate principal stress, despite that fact that this orientation would be the most critical. One would expect the least damage in Room 207 which is oriented parallel to the maximum principal stress. Nevertheless, a 1 x 1.5 m slab spalled from the area where the perimeter is parallel to the layering in the medium grained granite.

The tunnel profile of Room 207 is asymmetric, with induced fractures paralleling the gneissosity being not only more frequent, but also larger than those in other directions. Damage here also includes a 1 x 1.5 m slab spalled from the area where the perimeter is parallel to the layering in the medium grained granite. In Room 209, the layering is not favourably oriented to serve as exfoliation (i.e. failure) planes, and as a result, the damage profile is symmetrical. The orientation and magnitudes of the in-situ stress at this location are listed in Table 1. In this example, it is also noteworthy that the greatest

damage developed not in the tunnel subject to the higher $\sigma_1 - \sigma_3$ stress difference (25.5 – 12.8), but in the tunnel subject to the lower $\sigma_2 - \sigma_3$ stress difference (16.8 – 12.8). In fact, the perimeter stress calculated from the Kirsch stress equation should vary between 20 – 40 MPa only. No fracture would be expected for a rock that has uniaxial compressive strength of about 200 MPa. Clearly, while the damage is initiated by stress release about the tunnel profile, the geometry and extent of this damage is very dependent upon the presence of the low-dipping layering.

Table 1.

In-situ stresses measured at the 240 Level.

(Martin and Kozak, 1992)

| | Magnitude | trend | plunge |
|---|---|---|---|
| $\sigma_1$ | 25.5 MPa | 228° | 8° |
| $\sigma_2$ | 16.8 MPa | 135° | 23° |
| $\sigma_3$ | 12.8 MPa | 335° | 65° |

The cross section of the shaft excavation damage zone (provided by the floor of the 300 Level shaft station) further demonstrates the influence of subtle fabric on excavation damage development (Figure 5). Incipient breakout notches are developing in the NE and SW quadrants of the tunnel perimeter, in response to the stress orientations shown. Damage development varies with rock type, being best developed in the fine grained granite, and is weak or absent in the medium grained granite. This is the opposite of what would be expected based on the published strength data on these two rock types

(e.g. Martin, 1993), and from what would be expected based on the excavation response observations from the 420 Level (Read and Martin, 1990; 1991; 1992; and 1996). Mapping by the author, of the damage zone and geology at the 300 Level has revealed that the "anomalous damage" in the fine grained granite is due to the separation of slabs along a vertical foliation (planes of aligned biotite). Again, the conclusion is that while damage is initiated by stress release about the tunnel profile, the geometry and extent of this damage is very dependent upon the presence of a suitably oriented fabric element. The purpose of this thesis is to show how the geological conditions affect excavation stability, in particular how the fabric of granite influences excavation damage development including overbreak around tunnel or shaft perimeters. Fabric refers to natural planes of weakness in the form of pre-excavation microcrack alignments, compositional layering, foliation and changes in average grain size.

This space intentionally left blank for effective pagination.

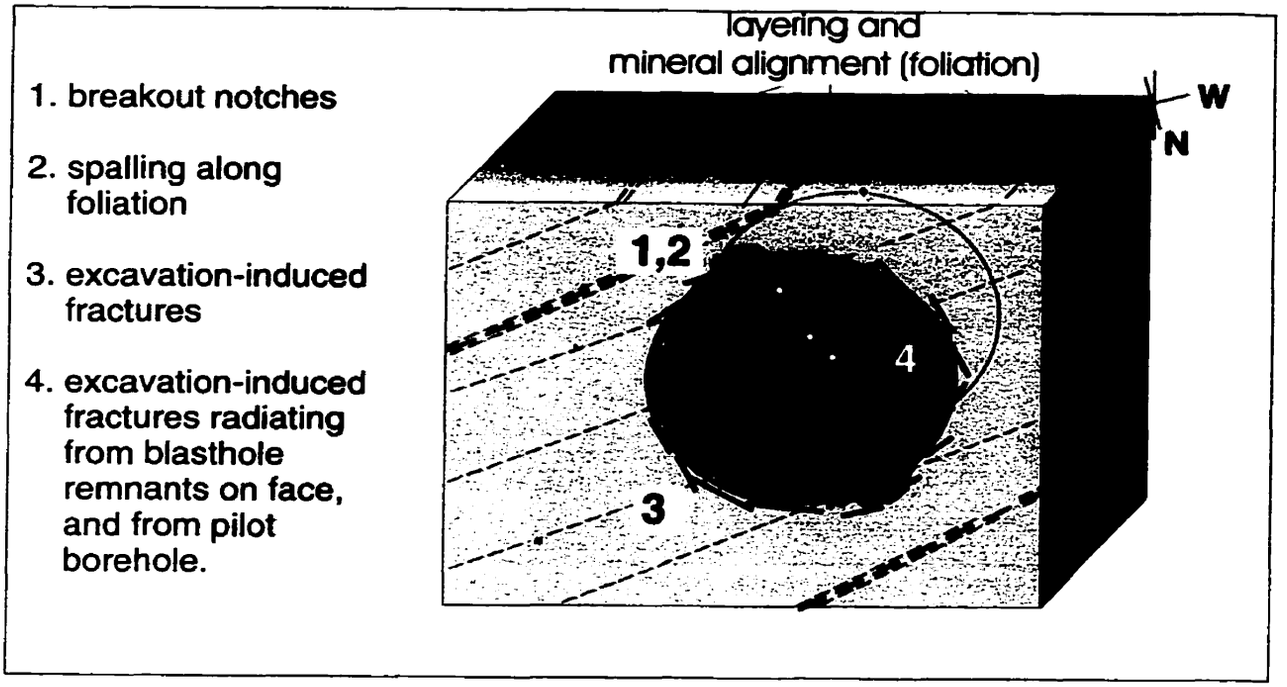



Figure 2. Schematic representation of excavation damage types observed in the URL. Breakouts (1) occur in all shafts and tunnels below 300 m, whether bored or blasted. Spalling and other excavation-induced fractures (2,3,4) occur at all levels in the URL, but are most intensely developed in the highly stressed rock below 300 m.

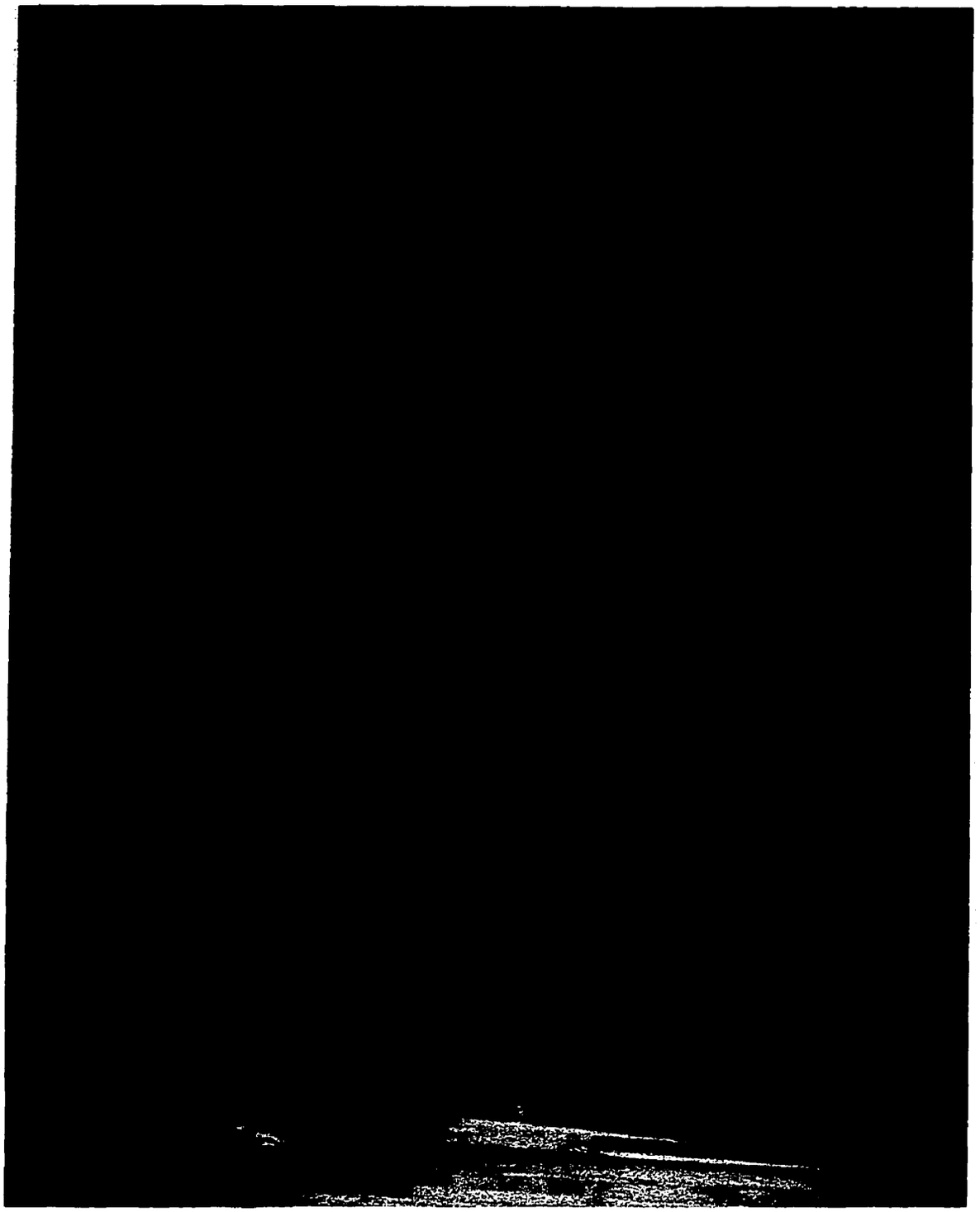


Figure 3. Spalling developed along a tunnel wall at the 240 Level of the URL during its construction. The tunnel axis is parallel to $\sigma_1$ (25.8 MPa), $\sigma_2$ (16.8 MPa) is nearly horizontal while $\sigma_3$ (12.8 MPa) is nearly vertical.

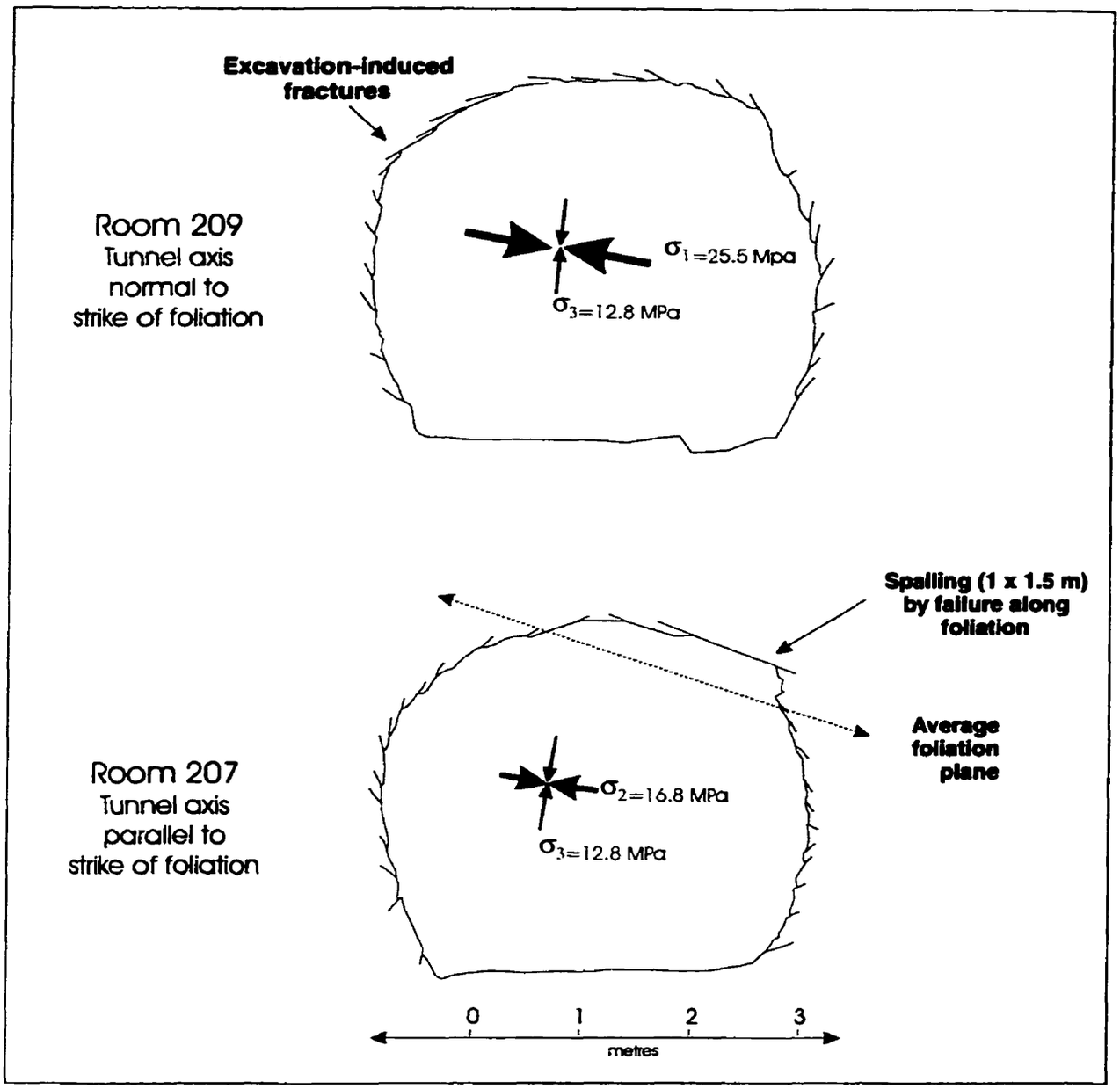



Figure 4. Influence of rock fabric on excavation damage - example from the 240 Level. This figure presents a schematic view of the fabric elements present at the 240 Level, and the damage associated with different tunnel orientations. Both tunnels are constructed in the same medium grained, coarsely layered granite, but differ in their orientation. The profile of Room 209 is in the $\sigma_1$ / $\sigma_3$ plane, while the profile of Room 207 is in the $\sigma_2$ / $\sigma_3$ plane. Mapping of the two tunnel profiles revealed that the damage about Room 209 is limited to minor surface exfoliation, with a symmetrical distribution of excavation-induced (stress-relief) fractures. In contrast, the damage about Room 207 is more extensive, with a markedly asymmetrical distribution of excavation-induced (stress-relief) fractures. A 1x 1.5 m slab spalled from the tunnel in the area where the perimeter is parallel to the layering.

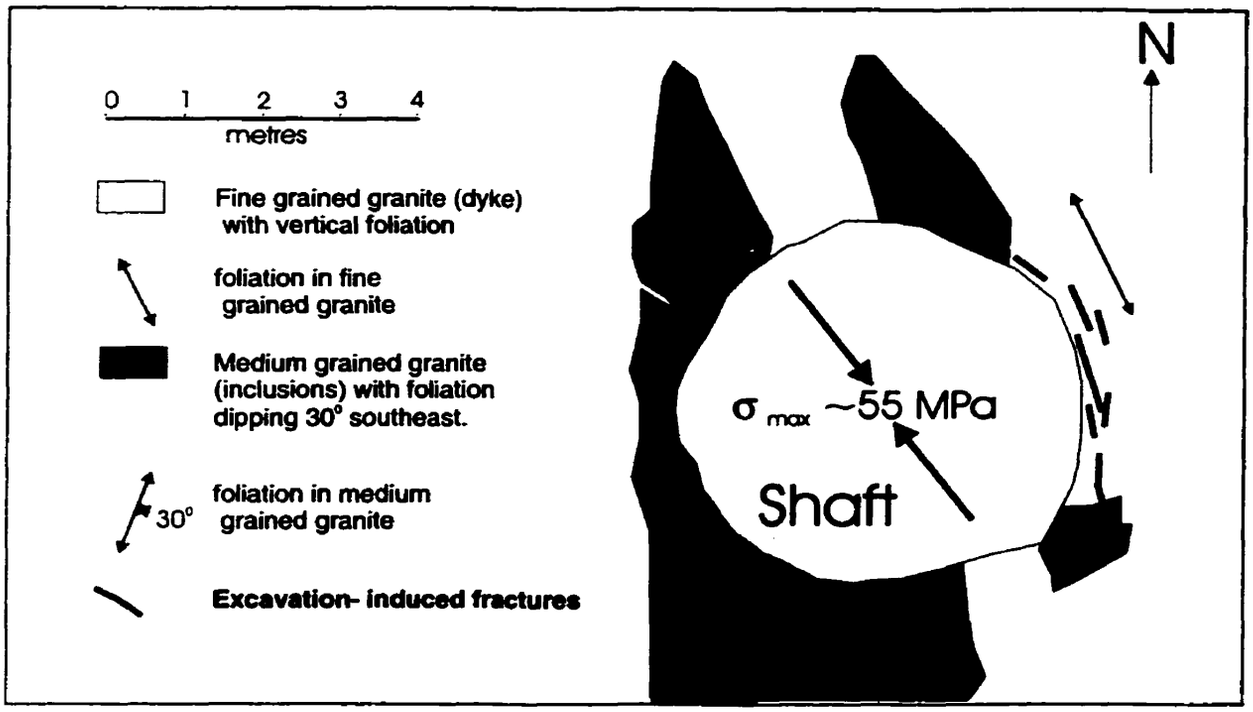



Figure 5. Influence of rock fabric on excavation damage - example from the 300 Level. At this depth, weak or incipient breakouts were occurring in the NE and SW quadrants of the shaft, due to spalling of slabs from the shaft wall. Most spalling occurred in the NE quadrant, and the presence of additional "loose" (incipient or partially spalled slabs) was indicated by a hollow sound when hit hard with a scaling bar. The corresponding breakout in the SW quadrant was, in comparison, insignificant and intact (no loose). Mapping of the station floor at this depth provided a cross section of the damage about the shaft, and an explanation for the difference in breakout development described above. In the NE quadrant, spalling was enhanced by parting along the subvertical foliation in the fine grained granite. In the SW quadrant, the foliation was subhorizontal, and therefore not favourably oriented to promote fracture development under these stress conditions.

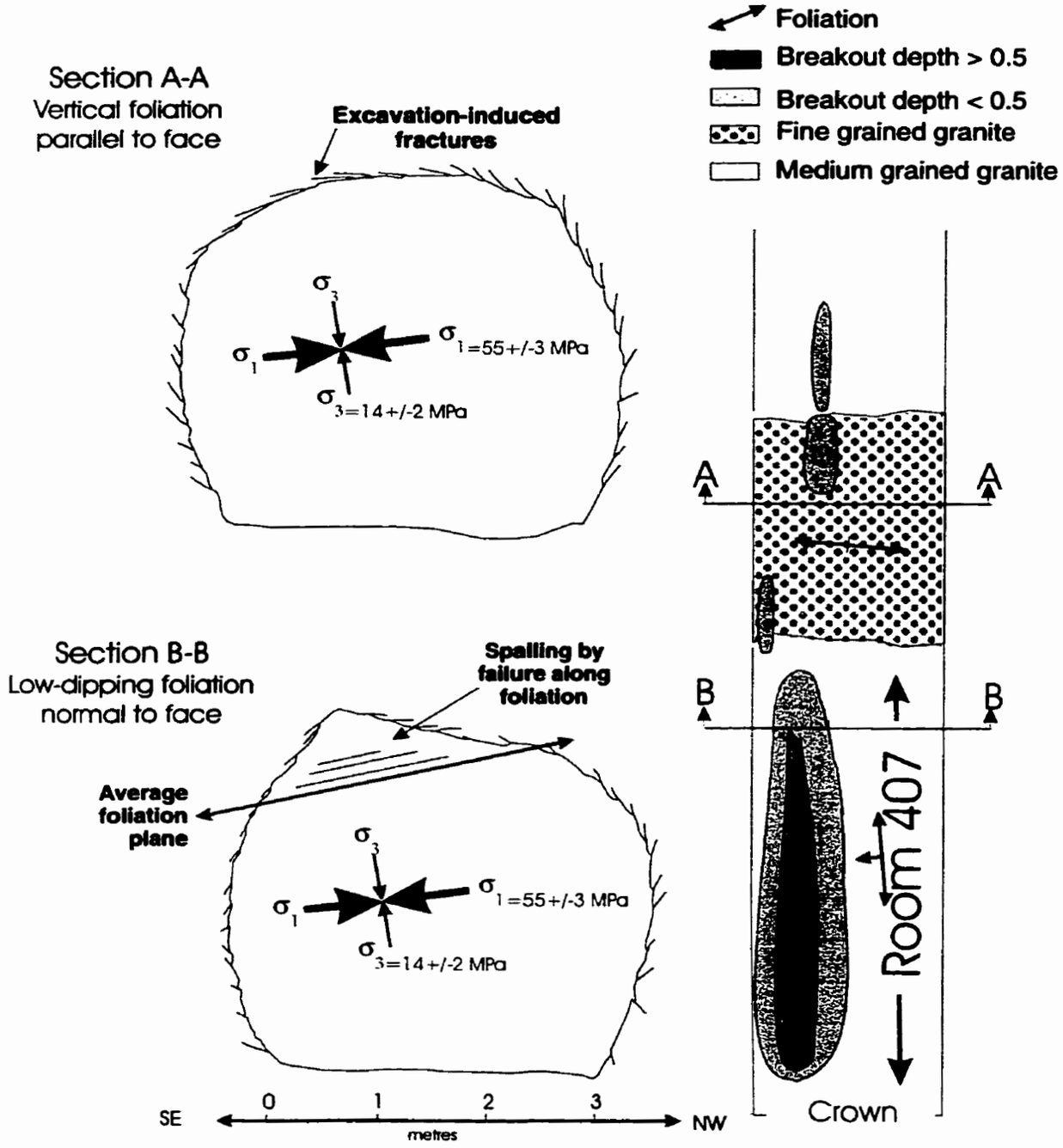



Figure 6. Influence of geology on breakout development within a single drill-and-blast tunnel. Room 407 at the 420 Level of the URL provides an example of the influence of subtle differences in geology on breakout development in granite where conventional excavation is used. The tunnel profiles developed in the medium grained and fine grained granites are shown in A and B, respectively. In the medium grained granite, the parting of the rock along planes of biotite facilitates breakout development. In the fine grained granite, a biotite foliation also exists but it is subvertical and normal to the tunnel axis. As such, its presence does not influence breakout development for this excavation orientation and geometry.

Figure 7. Influence of geology on breakout development within a single drill-and-blast tunnel. Room 421 at the 420 Level of the URL provides an example of the influence of subtle differences in geology on breakout development in the fine grained granite where conventional excavation is used. In the previous figure it was noted that ... "a biotite foliation also exists in the fine grained granite, but it is subvertical and normal to the tunnel axis. As such, its presence does not influence breakout development for this excavation orientation and geometry". In Room 421 however, the tunnel axis is parallel to the strike of the foliation, and spalling now occurs along this structure.

## 1.2. Objective, Scope of Work, and Thesis Organization

The objective of this research is to investigate and document the contribution of rock fabric on excavation damage development. This thesis demonstrates, through a program of field observations and laboratory testing, that foliation and other seemingly minor textural variations in the rock are major contributors to the observed variations in excavation damage, and the lower than expected strength.

The rock properties testing conducted in this study was intended to define the influence of foliation and preferred microcrack orientations in creating strength anisotropies, which in turn influence excavation damage development. Techniques sensitive to planar anisotropy in rock were identified from the literature. It is recognized that fabric is just one of the factors contributing to lowering of strength from what is observed under laboratory conditions. Other factors include e.g. saturated versus unsaturated conditions, the duration of loading, the rate of loading, and whether loading is cyclic. Treatment of these factors is outside the scope of this investigation.

This investigation is the first of its kind in both scope and emphasis. The essential contribution of this work is to take the data from AECL's Underground Research Laboratory (URL) and document the contribution of rock fabric in granite to the observed excavation damage. This damage includes overbreak due to stress relief, and to visible excavation-induced fractures. This work presents a unique and unprecedented data set on:

1. the considerable fabric variations present within what was previously thought to be a type example of a homogeneous granite, and
2. the relationships between excavation damage and geology,

This thesis will summarize:

1. what was learned from the URL that is useful in future site characterization,
2. how geological features in granite influence excavation damage (i.e. the link between geology and engineering), and
3. the role of rock fabric on rock mass response (including rock properties testing and in-situ stress measurement).

An outline of the scope of work and the organization of this thesis is provided in Table 2, with details provided in the following sections.

Table 2.

Scope of Work

Literature Review. ..... (Chapter 2)

Rock Mass Description.

- General Geology, Fault and Fractures, In-situ stress domains ..... (Chapter 3)
- Fracture, foliation and microcrack mapping ..... (Chapter 4)

Field Evidence of Fabric Anisotropy on Rock Mass Response.

- Ancient Faults and Fractures ..... (Chapter 5)
- Excavation Induced Fractures ..... (Chapter 5)

Selection of detailed study areas for sampling ..... (Chapter 6)

- Excavation damage mapping ..... *3 linescan samples and tunnel profiles.*
- Drilling ..... *3 arrays, 12 boreholes.*
- Sampling ..... *171 samples.*
- Characterization of sample structure and texture. ..... *171 sample surface maps.*
- Classification of Fabric elements ... *foliations, lineations, and microcrack sets.*
- Measurement of fabric orientations ..... *~500 orientations.*

Rock Properties Testing ..... (Chapter 7)

- for preferred microcrack directions and their relative strengths. ..... *103 Point Load Tests.*. (Chapter 8)
- for tensile strength and anisotropy. ..... *57 Brazilian tests.*. (Chapter 9)
- for rock strength. ..... *11 uniaxial compression tests.* (Chapter 10)

Summary and Conclusions ..... (Chapter 11)

The site description (Chapter 3) provides a summary of the general geology, with emphasis on the foliations, natural microcracks, and other structures present in the granite. Also included is a summary of the in-situ stresses. The range of site characterization work previously completed by the author is summarized in Table 3. For the most part the work listed in Table 3 defines the general geology, excavation damage, petrography, natural microcracks, and the influence of rock structure on rock mass response.

The site characterization work and the determination of the ancient or natural microcracks (Chapter 4) were major elements of the scope of work, providing new data on these important structures and their variation with depth in a granite pluton. The results of the analysis of natural microcracks is provided in Chapter 4. In Chapter 5, field evidence for anisotropic behaviour of the rock mass is presented. The case is made that the rock mass behaved as an anisotropic medium when first subject to brittle deformation in the Precambrian, and that the present-day excavation response is simply a continuation of this behaviour.

Locations (sample areas) selected to examine the interaction between rock fabric, rock mass anisotropy, and excavation damage are described in Chapter 6. Three locations were chosen for mapping of excavation damage and for collection of a suite of samples for rock properties testing. These provide a representative sample of the range of the

compositional and textural heterogeneity present in the Lac du Bonnet granite, and their affect on rock strength and rock mass response. These are:

1. fine grained granite dykes with subvertical flow banding,
2. medium grained and weakly to moderately layered granite which forms the main mass of the batholith, and
3. generally coarse grained and leucocratic unit that occurs as sills and recrystallized zones in the main phase.

For the sake of convenience, these are referred to as the fine, medium and coarse grained granites, respectively.

The three study areas were sited in newly excavated tunnels at the 420 Level (420 m depth of the URL) where the damage due to stress relief can be observed in the process of developing. The sites are in close proximity to each other. They are distant from faults or major fractures, and are in the same in-situ stress-domain. As such, they are subject to the same stress magnitudes and orientations.

Excavation damage mapping was conducted at each of the study areas. "As-built" and "final" tunnel cross-sections were obtained to show progressive damage development, and to indicate the overbreak due to spalling. A linescan survey was then done for each study area, along the line of the surveyed tunnel profiles. Excavation induced fractures encountered along this profile were mapped and their orientation and surface characteristics noted. The rock types and the local structures were also mapped. Examples of excavation-induced fractures formed by parting along compositional

layering or other fabrics present at each case study site were noted. This mapping revealed those visible differences in excavation response (damage development) correlated with variations in the average grain size, the structure type and structure orientation. For example: the degree of excavation damage, including breakout development, follows the progression shown in Table 4. Within a given rock unit, the damage in the form of breakout notches is most pronounced in tunnels constructed along the strike of the layering.

At each study area, oriented core was obtained from an array of boreholes as listed in Table 5. The arrays were designed to intersect the most prominent foliation / microcrack set at 90°, and then at 20-30° increments until the core axis was parallel to the foliation. All samples are segments of HQ- sized core (46 mm diameter), the choice of specimen diameter being largely dictated by budget constraints. Specimens were prepared in accordance with the procedure described by Gyenge and Ladanyi (1977a,b), Gyenge (1980), and the method suggested by the International Society for Rock Mechanics (Brown, 1981). The end surfaces of each of the specimens were ground flat to 0.015 mm to ensure that they were parallel to each other and perpendicular to the axis of the specimen. The average dimensions and bulk densities of samples from each of the three rock types are listed in Table 6. The bulk density ranged from 2.54 to 2.72 gm/cm$^3$, these corresponding to the leucocratic and melanocratic layers in the coarse grained granite. Additional information on sample properties is presented in the individual chapters on rock properties testing.

Sample locations within the boreholes were chosen to minimize the possibility of damage caused by proximity to the excavation. Such damage is represented by discing of the core (which is relatively easy to see), and by the opening of microcracks along grain boundaries or other directions (which is not easily detected). The term "discing" refers to the tendency for core to split spontaneously into thin discs by the formation of closely spaced fractures perpendicular to the core axis. In general, discing is the result of tensile failure, which occurs wherever stresses radial to the borehole are high relative to: 1) the stresses along the borehole axis, and 2) the unconfined compressive strength of the rock (Engelder, 1993). According to the author, core discing occurs only in a high stress environment, which he defines as greater than 1 km depth on average, but which can be found at any depth in the URL. The extent of discing in any given area is a function of the room geometry (stress concentrations around the tunnel perimeter), and of the rock type, the fine grained granite being less susceptible to discing than the coarser grained types.

To avoid core damage, either from discing or from the opening of grain boundaries, samples were taken from the ends of each borehole as shown in Figure 9, a schematic representation of the general area affected by discing in core as seen in arrays surrounding the tunnels of the 420 Level (D.R. Woodcock, unpublished data).
To test for non-visible damage, rock properties tests results (such as tensile strength from the Brazilian Tests), were plotted against the specific gravity of the sample on scatterplots. If reduced tensile strengths were associated with lowered specific gravity, and these in turn correlated with distance from the excavation, it would be assumed that

damage, in the form of microcracking, was present. The results (presented in Chapters 6 and 7), showed no correlation with distance from the tunnel wall, and as such, are assumed to be free of damage due to proximity of the excavation.

The core was logged for foliations and microcracks, and then cut to provide oriented samples for rock properties testing. Logging procedures, particularly those relevant to distinguishing natural and induced fractures, followed recommendations by Kulander et al. (1980). Each of the 171 samples were also mapped individually for: foliations, microcracks, alteration products along grain boundaries, nature of grain boundaries including microcracking along boundaries, mineralogy variations, and mineral orientations. The sample maps for each specimen are presented in Appendices A and B.

The rock properties testing program is described in Chapters 7 to 9. Evidence for natural weakness planes (mechanical anisotropy) was defined by the preferred orientations of fractures induced by 103 axial Point Load tests (Chapter 7). Directional anisotropy in tensile strength was determined by the 57 Brazilian tests (Chapter 8) and in compressive strength by 11 uniaxial compression tests (Chapter 9). In each of these chapters, correlations between the rock properties tests and measured foliations and natural microcrack orientations are presented. To better examine the complex geometries between these data and the various structural elements in each rock type, all the above were plotted in their proper three-dimensional context using stereographic projections.

A summary of results is provided in Chapter 10. A relative ranking of rock strength correlating with average grain size is presented. The results are then compared with field observations of excavation damage development, and with published rock strengths. Directional anisotropies in rock strength are compared to the orientation of known fabric elements such as natural microcracks directions and some foliations. The results are consistent with tunnel mapping observations regarding preferential development of excavation damage in some rock types and in certain directions.

This page intentionally left blank for effective pagination.

## Table 3. Summary of Site Characterization by the Author

General Geology

Everitt, R., McMurry, J., Brown, A. and Davison, C. 1996. Geology of the Lac du Bonnet Batholith, Inside and Out: AECL's Underground Research Laboratory, Southeastern Manitoba. Field Excursion B-5: Guidebook, Geological Association of Canada - Mineralogical Association of Canada Joint Annual Meeting 1996, Winnipeg, Manitoba, Canada. 1996 May 30.

Everitt, R.A., Gann, P., Brown, A., Woodcock, D.R., Sikorsky, R.I., and R. B. Ejeckam. 1998. Subhorizontal Layering in the "Homogeneous" Lac du Bonnet Batholith at the Underground Research Laboratory, South Eastern Manitoba. In: Special Symposium on Extraction, Transport and Emplacement of Granitic Magmas, Ottawa '97 GAC-MAC Annual Meeting. Journal of Structural Geology. Volume 20, No. 9/10, p. 1291-1304.

Everitt, R.A. and Brown A. 1996. Geological Mapping of AECL Research's Underground Research Laboratory. A Cross Section of Thrust faults and Associated Fractures in the Roof Zone of an Archean batholith. Proceedings of Fractured and Jointed rock masses, a regional conference of the International Society for Rock Mechanics. June 1-6th 1992. Granlibakken, California. Vol. 1, p. 1-11.

Everitt, R.A., Gann, P., and Boychuk, D.M. 1993. Mine-by Experiment: Part 7 - Geological Setting and General Geology. AECL Technical Record RC-1080.1.

Excavation Damage - shaft

Everitt, R.A. and Passmore W. 1993. Shaft Extension Data Report - Wall Topography and Geology Map.

Everitt, R.A., Chernis, P.J., Good, D.H., and Grogan, A. 1989. Mapping of the Excavation Damage Zone around the Circular Access Shaft at Atomic Energy of Canada Limited's Underground Research Laboratory. In: Excavation Response in Geological Repositories for Radioactive Waste. Proceedings of an NEA Workshop, April 1988, Winnipeg, Canada, p. 271-282.

Excavation Damage - 240 Level

Everitt, R.A. and Read, R.S. 1989. Geology of the 240 Level of the Underground Research Laboratory, Volume I, General Geology. AECL Technical Record TR 491- 1, URL EXP 006 R001.

Everitt, R.A., Chernis, P.J. and Good, D.H. 1989. Geology of the Room 209 Floor. Contributions to the Excavation Damage Assessment Experiment. 5 p., 4 figures, 4 tables and 2 maps.

Lang, P. A., R. A. Everitt, E. T. Kozak and C. C. Davison. 1988. Underground Research Laboratory Room 209 Instrument Array: Pre - Excavation Information for Modellers. 1988 December, AECL 9566-1.

Lang, P.A., Babulic, P.J., Bilinsky D.M., Everitt, R.A., Spinney, M.H., Kozak, E.T. and Davison, C.C. 1990. Underground Research laboratory Room 209 Instrument Array: Measured Response to Excavation, Volumes 1 and 2, Atomic Energy of Canada Limited Report AECL 9566-3.

Excavation Damage - 420 Level

Everitt, R.A. and Woodcock, D.R. 1994. Contribution to the HFT Characterization Summary Report. Mapping Results for the 600 mm Heated Failure Test Borehole. Unpublished AECL Internal Memorandum. URL-01,84.41, AGB-94-190.

Everitt, R.A., Brown, D., A. Boychuk, D.M., and Gann, P. 1994. Excavation damage and organic growth in a 1.2m diameter borehole. Poster presentation at the 5th Annual International High-Level Radioactive Waste Management Conference & Exposition. May 22-26th 1994.

Everitt, R.A., Gann, P., and Boychuk, D.M. 1993. Mine-by Experiment: Part 7 - Geological Setting and General Geology. AECL Technical Record RC-1080.1

Petrography and Petrofabrics

Everitt, R.A. 1993b. Petrography of Granitic Rock Samples from the 420m Level of the Underground Research Laboratory, Pinawa, Manitoba. Prepared by Dana Kelly, David S. Peck and Richard S. James, Department of Geology, Laurentian University, Sudbury, Ontario, P3E 2C6. URL Library Identifier 01.73.03.

Correlations between Mineralogy and Crack Initiation

Everitt, R. A. 1995. Unpublished internal memorandum. URL-01.84.41, R-3002-07-00, Status of Pegmatite Testing Study. (A draft report on the testing results). Fracture Initiation and Propagation by Compressive Loading and Unloading in Component Mineral Phases of the Lac du Bonnet Batholith.

Natural Microcracks

Everitt, R.A. 2000. Synopsis of natural microcrack work at the Underground Research Laboratory. Unpublished report.

Table 4. Rationale for Selection of the Sample Sites

Variations in Excavation Damage with Average Grain Size and Foliation.

1. Coarse Grained Granite (Room 418)
   - Strong low dipping foliation and compositional layering
   - Variable but significant damage developing within hours

2. Medium grained Granite (Room 417)
   - Weak low dipping foliation
   - Significant damage developing within hours

3. Fine Grained Granite (Room 421)
   - Strong subvertical foliation and subhorizontal natural microcracks
   - Least damage, developing over a period of days or weeks.

Table 5. Sample and Testing Summary

| Granite Type | Borehole | Orientation | | test type and number | | |
|---|---|---|---|---|---|---|
| | | Trend | Plunge | Uniaxial compression | Brazilian | Point Load |
| 1. Coarse Grained | 418-051-RT1 | 324.90° | 77.10° | | *discing** | |
| | 418-051-RT2 | 143.61° | 54.76° | 1 | 8 | 15 |
| | 418-051-RT3 | 143.40° | 35.40° | 1 | 8 | 6 |
| | 418-051-RT4 | 143.56° | 19.70° | 1 | 7 | 6 |
| | 418-051-RT5 | 143.38° | 0.40° | 1 | 4 | 12 |
| 2. Medium Grained | 403-014-MB2 | 225° | 00° | 1 | | |
| | 417-071-RT1 | Not drilled. | | | | |
| | 417-071-RT2 | 190.43° | 59.44° | | *discing* | |
| | 417-071-RT3 | 190.20° | 39.48° | | *discing* | |
| | 417-071-RT4 | 189.10° | 19.50° | 1 | 6 | 6 |
| | 417-071-RT5 | 189.70° | -0.20° | 1 | 4 | 8 |
| 3. Fine Grained | 421-032-RT1 | 133.99° | 0.23° | 1 | 4 | 10 |
| | 421-032-RT2 | 157.61° | 0.46° | 1 | 6 | 7 |
| | 421-032-RT3 | 175.81° | 0.80° | 1 | 5 | 10 |
| | 421-034-RT4 | 199.05° | 0.00° | 1 | 6 | 13 |

* The term "discing" refers to the tendency for core to split spontaneously into thin discs by the formation of closely spaced fractures perpendicular to the core axis. For further information on discing see page 20.

Coarse grained Granite

**strong low-dipping fabric**

*4 uniaxial compression tests*

*27 Brazilian Tests*

*40 Point Load Tests*

Medium grained Granite

**weak low-dipping fabric**

*3 uniaxial compression tests*

*9 Brazilian tests*

*20 Point Load Tests*

Fine grained Granite

**subvertical fabric**

*4 uniaxial compression tests*

*21 Brazilian tests*

*43 Point Load Tests*

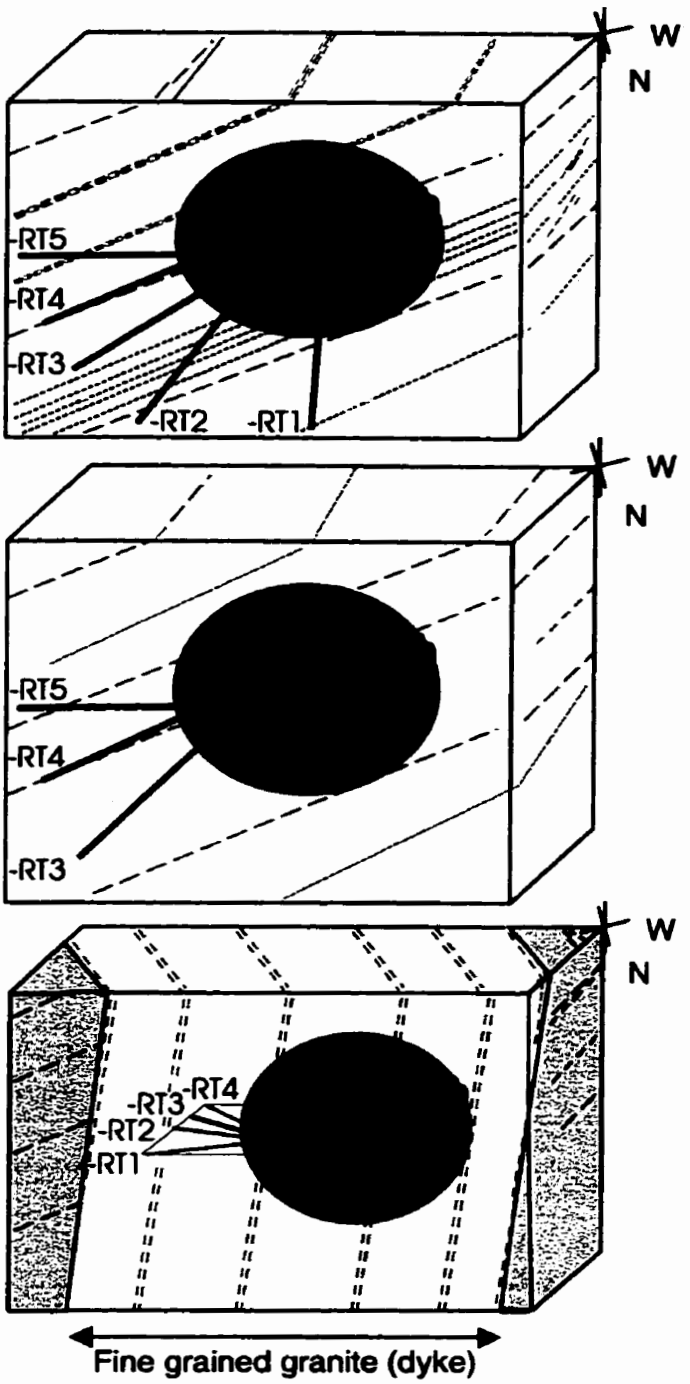



Figure 8. Distribution of measurements relative to rock structure. The orientation of the arrays relative to the local rock structure is shown (layering / foliation indicated by the line symbols). The influence of the subhorizontal layering is tested by the arrays in the coarse and medium grained granite, while the influence of the subvertical layering in the fine grained granite is tested by the third array in the fine grained granite, which occurs as dykes intruding the medium and coarse grained granites. The geometry of the complete sample array at each location is also shown.

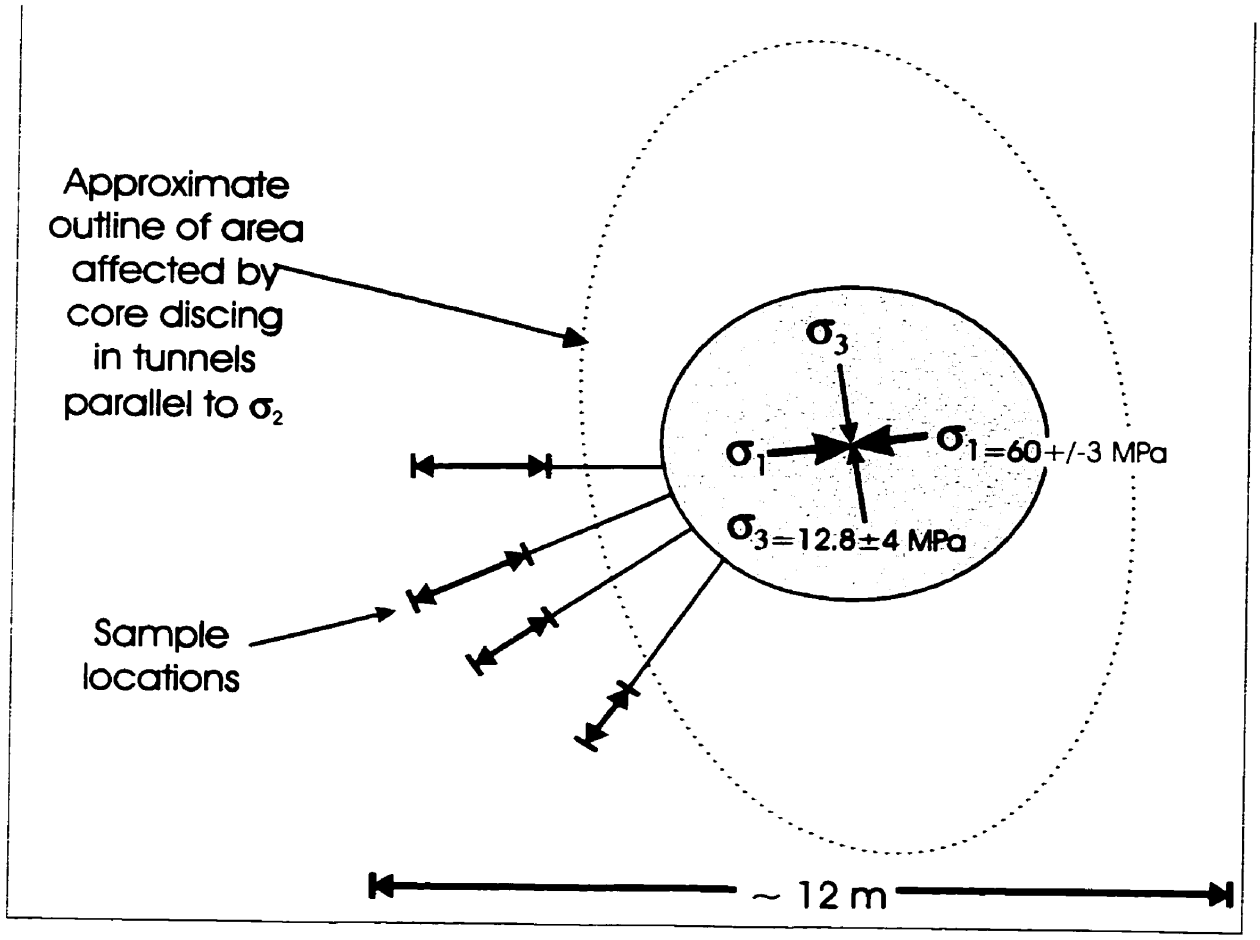



Figure 9. Approximate extent of the zone of core discing around a tunnel at the 420 Level.

Table 6. Sample Properties

| Granite type | Number of samples | Diameter (mm) | Thickness (mm) | T/D | Density gm/cm$^3$ |
|---|---|---|---|---|---|
| | | max | max | | max |
| | | mean | mean | (average) | mean |
| | | min | min | | min |
| Fine grained granite | 21 | 44.91 | 28.36 | | 2.67 |
| | | | 24.93 | 0.56 | 2.64 |
| | | 44.86 | 22.24 | | 2.61 |
| | | 44.82 | | | |
| Medium grained granite | 10 | 44.80 | 25.86 | | 2.63 |
| | | 44.77 | 22.56 | 0.58 | 2.62 |
| | | 44.73 | 28.06 | | 2.60 |
| Coarse grained granite | 27 | 44.81 | 27.55 | | 2.72 |
| | | 44.60 | 25.05 | 0.56 | 2.65 |
| | | 44.43 | 22.03 | | 2.54 |

# Chapter 2

# LITERATURE REVIEW

## 2.1. Introduction

Understanding the mechanism of fracture development around underground openings is important for the design of a safe and stable opening. This understanding applies to mining and civil engineering works as well as nuclear waste disposal vaults. The Canadian concept for disposing of nuclear fuel waste involves placing and sealing it in excavations 500 to 1000 m deep in plutonic rocks of the Canadian Shield. Of concern is the potential for significant propagation of existing fractures or the creation of new fractures (excavation damage) during construction. Interconnected fractures could reduce the effectiveness of the rock as a barrier to contaminant movement by providing pathways for groundwater to transport contaminants released from a disposal vault.

The literature was reviewed for studies that considered the role of rock mass anisotropy in excavation-induced displacements around tunnels with particular emphasis on cylindrical tunnels in highly stressed brittle rock masses. More specifically, the objectives of this literature review were to identify information relevant to the objectives of this thesis regarding:

1. the factors affecting rock strength,
2. the influence of foliation and other fabrics on strength and excavation response in rocks where these structures are well developed and first tested (slates, schists, etc.), and for granites in general,
3. the laboratory methods designed to test for and quantify anisotropy in rock mass response at a variety of scales and in different rock types,
4. the recognition of primary and secondary fabrics within plutons, and
5. previous related work within the Lac du Bonnet Batholith.

A recent compilation by the United States National Research Council (1993) reviewed the state of the art regarding the stability problems that occur in sediments and metamorphic rocks, with less emphasis on granite or other crystalline intrusives or extrusives. The editors concluded that field observations on circular openings of varying scales have highlighted the complex interrelationship between:

1. the stress field (either static or transient),
2. the geometry of the opening (radius and time dependence), and
3. the in-situ material characteristics (strength, discontinuities, and rockmass anisotropy).

Factors inducing strength reduction in rock include:

1. time-dependence,
2. influence of pore fluid on strength,
3. influence of specimen size, and
4. influence of anisotropy.

The first three factors, although important in predicting rock mass response to excavation, are of secondary importance to this thesis, affecting only the collection of samples and the rock properties testing. Each of these is examined in the sections 2.2.1 to 2.2.4 following. The influence of anisotropy, the subject of this thesis, is reviewed in the most detail.

### 2.1.1. Influence of pore fluid on strength

The presence of water in the pores of a rock causes a reduction in its uniaxial compressive strength. Hoek and Brown (1985, p. 155) showed that the presence of water caused the strength of samples of shale and sandstone to reduce by a factor of 2 from oven dried to saturated specimens. The following ratios between the uniaxial compressive strengths of dry and saturated specimens were reproduced by Hoek and Brown (1985), from Broch (1974), and Colback and Wild (1965).

| | |
|---|---|
| Quartz diorite | 1.5:1 |
| Gabbro | 1.7:1 |
| Gneiss (normal to foliation) | 2.1:1 |
| Gneiss (parallel to foliation) | 1.6:1 |
| Sandstone | 2:1 |
| Shale | 2:1 |

Lajtai et al. (1987), on the other hand, found that a reduction in strength for Lac du Bonnet granite due to moisture was strongly time-dependent. In a fast uniaxial

compression test, the change in strength was negligible. However the strength reduction in long-term loading tests was considerable, corresponding to a ratio of 2:1.

Application to this thesis: As a precaution, samples for this thesis were stored and tested under the same conditions (i.e. dry), and tested at the same time. Accordingly, the differences observed between rock types are considered to be intrinsic differences between the rock types and different sample orientations, rather than to differences in moisture content.

### 2.1.2. Influence of specimen size

One of the most common explanations given for the difference between the laboratory strength of rock, and the rock strength found in-situ, is the effect of specimen scale (Martin, 1993). However, the rock mechanics literature contains a number of conflicting observations on the effect of specimen size upon the strength of rock. As noted by Hoek and Brown (1985, p.157) for the general case involving all rock types:

1. some authors report no change in rock strength with change in specimen size (e.g. Hodgson and Cook, 1970; and Obert et al., 1946), while
2. others note significant strength reductions with increasing specimen size (e.g. Mogi 1962, Bieniawski, 1972; Pratt et al., 1972; Protodiakonov and Koifman, 1963; and Hoskins and Horino, 1969).

The case for scale-dependence specific to the Lac du Bonnet granite, was examined in a comprehensive study conducted by Jackson and Lau (1990), Lau and Gorski (1991 a,b,c) and Martin (1993).

Uniaxial compressive strength and scale-dependence

The study by Jackson and Lau (1990) was conducted specifically to study the effect of specimen size on the ambient-temperature uniaxial mechanical properties of the Lac du Bonnet grey granite. According to the authors:

1. Specimens up to 63 mm, while showing some variation, indicated no trends with increasing specimen size.
2. However, as specimen diameters increased from 63 mm to 294 mm, the uniaxial compressive strength and tangent modulus of elasticity decreased.

Jackson and Lau (1990) concluded that " ..This decrease is well explained by the increased probability of the existence of a critical flaw or flaws in larger specimens (Weibull's "weakest link" theory) and is adequately described by any of the Hoek and Brown, Gyenge and Herget or Weibull models". Martin (1993) concluded, " the effect of scale on the compressive strength of Lac du Bonnet granite is not very significant" for the normal core diameter range (i.e. below 63 mm).

Field evidence collected by the author for this thesis supports the interpretation by Jackson and Lau 1990. The samples for the study by Jackson and Lau (1990) and also by

Lau and Gorski (1991 a, b, c) were taken from what was described as "massive grey granite". In reality, the granite from which their samples were taken is layered, with the layering varying between 100 and 300 mm in thickness, and the layers separated by thin planes of aligned biotite (Everitt and Read, 1989, also Figures 10, 12 and 13 of this thesis). It is suggested by the author that the decrease in strength in the larger samples obtained by Jackson and Lau (1990), reflects the increasing likelihood that these samples contain more than one layer and layer boundary (i.e. their critical flaw or flaws interpretation).

For this thesis, the diameter (43 mm) for uniaxial compression test samples is in the mid-range where no size effect was noted by Jackson and Lau (1990) or Martin (1993), and was small enough to keep the samples as texturally and structurally homogeneous as possible. This last point was a concern as samples chosen for this thesis were specifically chosen to avoid having multiple orientations of foliations.

Tensile strength and scale-dependence.

Hodgson and Cook (1970), and Martin (1993) showed that under uniform stresses, tensile strength would be independent of size. However, in situations with high stress gradients, such as are created by Brazilian testing, strength will be a function of sample size, with Brazilian tensile strength showing an increase at the smaller scales. This was based on a comparison of test results from samples with diameters of 25 mm, ~30 mm, ~35 mm, 43 mm, ~ 50 mm, ~55 mm, 63 mm and ~105 mm. Nevertheless, it has been shown that tensile strength is independent of sample size for the most common core diameters (35 –

63 mm). For the LDBB, Martin (1993) concluded that samples larger than 63 mm diameter were beyond any scale effect and provided an accurate measure of tensile strength for granite from the Lac du Bonnet granite. Similar results were obtained by Wijk et al. (1978) for the Bohus granite.

In this authors view, the caution about sample size relative to the scale of textural / structural domains in the rock also applies here. That is, the decrease in strength in the larger samples may reflect, at least in part, the increasing likelihood that these samples contain more than one layer and layer boundary (i.e. critical flaw). Given this caveat, the sample size (43 mm) used for this study is regarded as acceptable, especially as the objective of the testing is to test for directional anisotropy related to fabric within a given suite of samples. The absolute tensile strength of the LDB granite is of secondary importance, and as this study demonstrates – tensile strength is rock unit and orientation specific anyway.

### 2.1.3. Influence of anisotropy

A rock is mechanically anisotropic when the mechanical behaviour changes for different stress orientations. Foliations or lineations formed by the parallel alignment of micaceous minerals, or of rod-like minerals like hornblende, impart a strong textural and mechanical anisotropy. Thus, metamorphic rocks like schist and slate are often markedly directional in their behaviour (Davis, 1984; Goodman, 1989; and Behrestaghi, 1998). Anisotropy also occurs in layered mixtures of different components, as in banded gneisses, or in alternating sequences of sandstone, shales, cherts, and limestones (United

States National Research Council 1993, Hoek and Brown). In all such rocks, strength varies continuously with direction and demonstrates pronounced minima when the planes of symmetry of the rock structure are oblique to the major principal stress.

According to Goodman (1989), strength anisotropy can be evaluated best by systematic laboratory testing of specimens drilled in different directions from an oriented block sample.

Experimental deformation studies of the Martinsburg slate by Donath (1961), and a oil shale (intercalated marlstone and kerogen). McLamore (1966) demonstrated the mechanical effect of well developed mechanical anisotropy, by rotating the loading axis relative to the fabric in increments from 0° to 90°, the angle between the two being referred to as $\psi$. The results indicate that a strong mechanical anisotropy will control the orientation of the failure plane over a broad range of loading geometries, and that this control is reflected by the variation in breaking strengths. Donath (1961) concluded:

$\psi < 45°$ failure is preferentially developed along the plane of the slaty cleavage, with breaking strength at its lowest when the fabric is at 30° to the loading axis.

$\psi = 45°$ to 75° the failure planes reach a maximum of about 60° to the loading axis, and the breaking strength continued to increase.

P = 90° the plane of failure returns to about 30° (considered the ideal angle as would exist in a homogeneous rock), and the breaking strength reaches its maximum at this geometry.

The ratio of minimum to maximum unconfined compressive strength in the slate was approximately 25% (Donath, 1961), and approximately 17% in the oil shale (McLamore, 1966). The minimum strength occurred at approximately 30° to the fabric plane in both examples. Similar results in sandstones were noted by Hoek and Brown (1980), and in metamorphic rocks – with the lowest axial strength (at various confining pressures) at 30° to loading axis.

Examples of structural or textural anisotropies affecting rock mass response in sedimentary or metamorphic rocks are common in the literature (Hoek and Brown, 1980; Goodman et al., 1993). The causative structures include: 1) spaced singular discontinuities at the tunnel-scale, such as bedding planes and joints, and 2) fabrics such as foliation, slaty cleavage, or lineations which are pervasive at the tunnel and hand - specimen scales. The latter should be considered as an intrinsic characteristic of the intact portion of jointed rock (Behrestaghi et al., 1996).

The anisotropic strength behaviour of four varieties of schists was examined by Behrestaghi et al. (1996), through the testing of specimens with varying orientation of schistosity with respect to the major principal stress under uniaxial and triaxial conditions. The results of this study were anisotropic deformation properties and natural planes of fracture defined by foliations. This structural anisotropy influences rock mass

response through either the development of unstable block geometries, and by anisotropic response of the intact rock including the development of fractures in orientations related to the fabric (also USNRC, 1993; Goodman et al., 1993).

The role of bedding in controlling breakouts in excavations has been described by Komar and Frohne (1973) and by Haimson and Edl (1972). Oriented Brazilian compression tests (Ewy et al., 1988b) determined that the failure location was controlled by rock strength anisotropy, with failure occurring where the tangential stress around the hole was parallel to the bedding plane. These examples are analogous to the situations at the URL described in Chapter 1 of this thesis (Figures 4, 5 and 6), with the foliation and compositional layering in the LDBB granite corresponding to the bedding and schistosity, respectively.

The conclusions by the United States National Research Council (1993) regarding the role of rock mass anisotropy, and their significance to this thesis were as follows:

1. "... The role of anisotropy , discontinuities and imperfections will undoubtedly play an even stronger role in engineering situations outside of a controlled laboratory environment ...",

2. many factors affect the displacements and ultimate stability of both shaft and tunnel wall rock. "... Foremost among these are the magnitude and ratio between the horizontal principal stresses, the extent of rockmass anisotropy and geological

discontinuity, the time lag between excavations and support installation, and the influence of intersections into the primary opening, such as shaft stations and loading pockets".

These factors were considered in the selection of sample locations to assess the role of rockmass anisotropy for this thesis. All samples are from intact rock and are confined to a single in-situ stress domain. All samples are from the portions of tunnels remote from intersections with adjoining excavations, from the ends of tunnel, and from the ends of blast rounds.

3. "...that there are no strict lithology rules..." .. their synopsis emphasizing the importance of the in-situ stress field in combination with the materials strength characteristics, and highlighting the risk of arbitrarily correlating excavation stability with lithology. This may be re-phrased as the risk of assuming that there is a typical behaviour from any given rock type.

This last statement is especially noteworthy, as will be demonstrated in this thesis, where statements such as the following from the literature are shown to be oversimplifications.

1. The LDBB is a type example of a homogeneous granite (Hoek et al., 1995) where the authors noted that there was "no influence of geology" to be seen in the damage zones about the URL excavations.

2. The fine grained phase of the LDBB granite is stronger and less susceptible to damage than the medium grained granite (Martin, 1990; Martin, 1993; Lau et al., 1991 a, b, c).

As was noted earlier in Chapter 1 of this thesis, and as will be examined in the chapters following, the extent of excavation damage within a particular variety of the LDBB granite depends upon the orientation of its fabric relative to the orientations of the excavation and the in-situ stresses.

## 2.2. The Experimental Determination Of Strength Anisotropies

This section identifies methodologies from the general literature which were used, in circumstances similar to those of this thesis, to define mechanical anisotropy and its relationship to fabric anisotropy. Specific objectives were:

1. to determine if there are naturally occurring directions of preferred fracture in otherwise intact rocks,
2. to correlate these with variations in the tensile strength, and
3. to determine the compressive strength and its variations with orientation.

### 2.2.1. Determination of Tensile Strength

Tensile fracture is a ubiquitous feature of rock failure ranging in size from micro-fractures to the large-scale explosive failure accompanying rock bursts. Hence the resistance of rock to failure in tension, the tensile strength, is one of the fundamental parameters of rock strength.

Although the measurement of tensile strength is performed in many projects in a routine manner, little attention has been paid to possible variation in tensile strength with the direction of loading (Lajtai, 1980; Duncan, 1987).

Methods to determine the tensile strength, and its variation with direction, include the point loading and line-loading (Brazilian) tests.

In the point-loading method, core discs are loaded axially between a pair of indentors while in line-loading method; the disc is placed on edge and loaded in the direction of its diameter between two flat platens.

In the line-load or Brazilian method, the diametral fracture plane is constrained by the position of the loading indentors, while in the point-load test, the diametral fracture plane is free to take any orientation. In practice, failure tends to be controlled by heterogeneities in the rock, such as foliations or microcracks, which are close to or parallel to the plane of failure under this loading geometry. Heterogeneities well outside this plane, such as incipient core discing, do not affect the test.

According to Lajtai (1980) "... The Brazilian test seems to yield a more accurate definition of both the tensile strength and its variation with direction ... (and is the more commonly applied tool for this application in rock mechanics).... The point load test is more suitable for the determination of the minimum value of tensile strength, and ...(in determining if there are preferred directions of fracture)...".

### 2.2.2. Point Load Test

This method has been used to determine if there are preferred orientations of failure under conditions of tensile fracture, in what may otherwise appear to be intact rock. Testing under controlled conditions has shown that, in artificially produced isotropic materials, such as a concrete aggregate composed of cement and fine grained sand, tensile fractures induced by Point Load testing develop in random orientations (Paulman, 1966). However, the same test method applied to a variety of rock types is consistent in producing one or more preferred orientations of test-induced tensile fractures, whose orientations are, upon inspection, found to be controlled by natural microcracks or others structures which: 1) reflect the rock's deformation history, and 2) which are pervasive at the specimen scale. Examples include:

1. Test-induced fractures correlate with the microscopic fractures in sandstones (Cardium sandstone - Reik and Currie, 1974), (Mispec, Perry and McCoy sandstones - Lajtai and Alison, 1979), and (various Devonian to Carboniferous sandstones from New Brunswick and Prince Edward Island – Lajtai, 1980; Lajtai and Stringer, 1991).

2. Test-induced fractures correlate with the microscopic fractures in granite (Shield Quarry, Manitoba - Duncan, 1987, and the Charcoal granite - Friedman and Bur, 1974). In the latter example, a preferred orientation of residual strain and framework of feldspar grains with approximately the same attitude as the microscopic fractures was considered as contributing to the fracture anisotropy).

More commonly in rock engineering, the Point Load test is used as a relatively simple method to determine a strength index, the Point Load strength index, which is related to the uniaxial compressive strength of rock (Broch and Franklin, 1972; Bieniawski, 1975; Broch, 1983). The Point Load Strength Index" is given by the following (Goodman, 1989):

$$\text{Point Load strength index (MPa)}, I_s = P/D^2 ,$$

where P = the load to break the specimen, and

D = the distance between the platen contacts.

According to Goodman (1993), tests are normally done on pieces of drill core at least 1.4 times as long as the diameter. In practice, there is a size effect so a correction must be made to reduce results to a common size. Point Load strength is found to fall by a factor of 2 to 3 as one proceeds from cores with a diameter of 10 mm to diameters of 70 mm; therefore, size standardization is required. The Point Load index is reported as the Point Load strength of a 50 mm core (size correction charts are given by Broch and Franklin).

### 2.2.3. Brazilian Tensile Strength Tests

The Brazilian test is a recognized as an indirect but convenient method for estimating the tensile strength (Goodman, 1989). It has been found that a rock core will split along the diameter and parallel to the cylinder axis when line loaded on its side in a compression machine. The reason for this can be demonstrated by examining the stress inside a disk loaded at opposite sides of a diametral plane. In such a configuration the horizontal stresses perpendicular to the loaded diameter are uniform and tensile. At failure their magnitude determines the tensile strength:

$$Brazilian Tensile strength, \sigma_{t_B} = \frac{2P}{\pi D t}$$

where P = load at failure,

D = diameter of the specimen, and

t = thickness of the specimen.

The "Brazilian tensile strength - $\sigma_{tB}$" (in MPa) is estimated from the test result from the applied stress corresponding to the peak compression load. The actual cause of failure may in fact be more complex, reflecting the interaction of small-scale flaws, compression parallel to the eventual rupture plane, and the horizontal tension induced by this loading.

The Brazilian test has been found to give a tensile strength higher than that of the direct tension test, probably owing to the effect of grain-scale (and larger) discontinuities. However, the Brazilian test method is preferred, as it is much easier to perform this test

than to arrange the precise alignment and end preparation required for a direct tensile test (Goodman, 1989).

Jaeger and Cook (1979) compared tensile strengths results as determined by various test procedures. For three different rock types it was found that Brazilian tensile strengths are generally similar in range to direct uniaxial tensile strengths but only half the value of three-point bending tensile rupture strengths. Lajtai (1982) also found that Brazilian tensile strengths of LDBB granite to be approximately half the three-point bending tensile rupture strength. He attributed strength variations to volume differences as specified by Weibull's theory. Based on Jaeger and Cook's (1979) discussion it may be concluded that a higher tensile strength of intact Shield granite determined using the modulus of rupture tests is likely the function of the mechanics of the physical testing procedure: the Brazilian tensile test sets up a uniform tensile stress normal to the diametral plane of compression of a rock disc being loaded uniaxially, whereas the modulus of rupture test is a three-point bending procedure that sets up non-uniform stresses through the sample in the direct plane of loading. This produces a zone of compression in the upper portion of the block, and a zone of tension in the lower portion of the block. The resulting stress gradient leads to higher tensile strengths.

Duncan (1987) noted that the rupture tensile strength data did not corroborate the same tensile strength anisotropy indicated by the Brazilian tensile strength data, which shows a good correspondence with the microscopic fracture anisotropy. He concluded that ..." the Brazilian tensile strength test is very sensitive to the microscopic fracture properties

that exist...", and that the modulus of rupture test method is not as effective a test procedure for determining tenuous differences in tensile strength in granite.

### 2.2.4. Compressive strength

Methods to investigate the effect of fabric anisotropy on compressive strength and other deformation properties have been described in several references in recent literature. The testing methods are similar, regardless of rock type, and follow the general procedure laid out by the earliest investigations into anisotropy in rocks, such as the by Donath (1964), and for other, less obviously anisotropic rock types as:

- Shale — Mesri and Gibala (1972),
- Metamorphic rocks — Behrestaghi et al. (1996),
- Granite — Feves and Simmons (1976); Duncan (1987); and Kusumi et al. (1997),
- LDBB granite — Duevel and Haimson (1997); Lajtai (1981); Lau and Gorski (1990; 1991a,b,c); and Lau et al. (1994).

The anisotropy or anisotropies in the rockmass are identified prior to testing, using one or more of the following techniques (from Sano et al., 1992; Hoek and Brown, 1980; Goodman, 1993): sample surface mapping, petrographic examination, sonic velocities, variations in tensile strength, or thermal conductivity.

Oriented samples are then taken from the rockmass, with the structural anisotropy (usually slaty cleavage, foliation bedding, or microcracks) inclined at angles from 0° to 90°, usually through increments of 20°-30° to the intended loading axis.

Alternatively, comparatively homogeneous rocks, granites in particular, have been cut to provide cylindrical samples parallel to each of the following: the rift, the grain, and the hardway as defined from the quarrying operations where the samples were collected. In each study, these structures are correlated with sets of natural (ancient) microcrack development. The sample set is therefore intended to investigate the influence of the microcrack sets on compressive strength and other deformational properties (Feves and Simmons, 1976; Duncan, 1987).

Samples are then subject to one or more of the following test procedures: uniaxial or triaxial compression tests, single or cyclic, conducted under saturated or unsaturated conditions (e.g. shale - Mesri and Gibala, 1972; granite - Feves and Simmons, 1976; Duncan, 1987; LDBB granite - Duevel and Haimson, 1997; metamorphic rocks – Behrestaghi et al., 1996).

The data from these tests is then plotted as series of conventional stress-strain diagrams for each sample orientation, or as rose diagrams or x-y plots illustrating the variation in the measured parameter with the (two-dimensional) orientation of one of the structures present. In spite of the fact that most rocks contain more than one fabric element at various angles of intersection, the use of three-dimensional or stereographic projections

to portray such data are rare in the literature, even when the effect of intersecting fabric planes is the topic of discussion (e.g. Figure 77, p. 163 of Hoek and Brown (1980), which shows the variation in rock strength for a rock with multiple discontinuities, by means of a rose diagram of axial strength against discontinuity inclination ).

### 2.2.5. Application to this Thesis

The methods used in this thesis were those outlined in the preceding sections. Point Load tests were employed to determine if there were preferred directions of fracture in the massive rocks of the LDBB, and Brazilian tests to determine if there are variations in the tensile strength with the direction of loading. The Point Load index was calculated for each of the samples tested in this thesis to show strength differences between rock types and from different orientations.

Directional variation in uniaxial compressive strength related to layering and other structures was investigated using a series of samples from an array of boreholes. Oriented samples were then taken from the rockmass, with the structural anisotropy (layering or microcracks) inclined at angles from 00° to 90°, usually through increments of 20°-30°, to the intended loading axis. The data from these tests are plotted as series of conventional stress-strain diagrams for each sample orientation, and then summarized for each rock type using x-y-z diagrams depicting various measures of compressive and tensile strength versus orientation (Chapter 9).

## 2.3. Anisotropy in Granite

### 2.3.1. Introduction

Although rock fabric is a well recognized factor influencing rock strength and response in metamorphic and sedimentary rocks, plutonic rocks (whether mafic to felsic) are still generally regarded as homogeneous and isotropic, with the LDBB being described as a massive (Martin, 1993), and a type example of an isotropic granite in a recently published general reference text (Hoek et al., 1995). This assumption of granite isotropy has been maintained in spite of:

1. nearly a century of literature on the structural and textural variations within plutons, such as presented in the original work by Balk (1937), Cloos (1946), Mayo (1941),

2. repetition of the conclusions in later compilations such as in Dennis (1972), and Davis (1984), and Bouchez (1997),

3. recent case studies appearing in the published literature and major conference proceedings on the complexities of pluton structure by Bouchez (1997), Cruden (1990), Cruden and Lanueau (1994), and

4. for the LDBB in particular, detailed mapping of pluton structure by McCrank (1985), Brown et al. (1989a), Cerny et al. (1987), Stone et al. (1989), Everitt and Brown (1986), Everitt et al. (1990, 1994, 1996 and 1998).

The inappropriateness of the assumption of granite isotropy as it affects geotechnical investigations, including work for a used fuel disposal facility, has been noted by Goodman et al. (1993). They concluded that in addition to joints, granites also contain a variety of foliation planes and schistose zones. The authors also noted that insufficient attention has been paid:

1. to mapping the structural complexity of granite in site investigation programs, and
2. to incorporating this complexity in the design of sampling programs to define rock properties.

It is suggested here, from the literature reviewed for this thesis, that this tendency to refer to granitic plutons as homogeneous and isotropic may be due to a combination of the following four factors:

1. Inadequate site investigations or imperfect evaluation of geological information, as emphasized by Goodman (1993). Davis (1984, p.255) noted that: "... foliations and lineations in granitic rocks are not always conspicuous. Many granitic rocks display very subtle foliations and/or lineations whose physical reason for existence is not at all obvious. Microscopic study is often required to determine the petrographic basis for the foliations and lineations. Foliation and lineation may reflect any number of textural configurations, including aligned feldspars, elongate or discoidal quartz, and closed-spaced fractures or cleavages ".

2. Scale of Observation: It is inherently difficult to recognize fabrics in coarse grained rocks in general, and especially in core samples whose limited dimensions may obscure structures otherwise readily apparent in outcrop. The latter point is especially problematic when samples are routinely shipped from a remote site to another location for lab analyses, without mention of the core orientation relative to any local fabrics.

3. Limitations to site investigation. Most geotechnical work pertaining to plutonic rock originates with the various international nuclear waste programs. Apart from the Canadian program, which is unique in having large volumes of intact or sparsely fractured rock (Davison and Simmons, 1985; Davison et al., 1989, 1993) the granites investigated by other countries with a similar disposal plan, are all sited in naturally fractured blocky rock (Sweden (Stanfors, 1996; Backblom and Olsson, 1996), Finland (Vira, 1996; Tolppanen et al., 1998), France (Raynal, 1996), Spain (Olmo, 1996), Belarus (Kudelsky, 1996), Ukraine (Khrushchov and Starodumov, 1996).

The high fracture frequency in the waste disposal programs of other countries overshadows any flow and mechanical anisotropy that may be induced by foliation and other fabrics. Accordingly, site characterization and design in these programs has concentrated on modelling rock mass response for a discontinuous medium (e.g. Christiansson, 1989; Christiansson and Hamberger, 1991; Kornfalt et al., 1991; Olsson and Conterra AB, 1991; Ouchterlony et al., 1991; Rawlence, 1990). The

rockmass within the fracture-bound blocks tends to be ignored or is considered isotropic.

4. Sampling bias. The almost universal assumption of granite homogeneity that appears in the literature is perhaps the result of a fundamental sampling bias, inherited from the dimension stone industry (DSI) where homogeneity is the prime basis for selecting a quarry, and for its continued development. Dimension stone is defined as: a natural rock material that has been selected, trimmed, or cut to specified or indicated shapes or sizes with or without one or more mechanically dressed surface" (Taylor, 1992).

In a recent compilation of the DSI in Ontario, Marmont (1993) notes that the geology of a prospective dimensional granite quarry should be as simple as possible. In terms of lithology, this means that the quarry should ideally consist of uniform granite/granite gneiss free from crosscutting veins, dikes and inclusions of other rock types.

The dimension stone industry literature therefore shows a strong bias towards the selection and development of the most homogeneous plutons or parts of plutons (Bergeron, 1991a,b,c; Gerow and Bellinger, 1990; Hill, 1989a,b; Hill, 1990a,b; Hinz et al., 1994; Kennedy and Shelock, 1988; Lacey, 1989; Marmont, 1993; Martinsen et al., 1991; Ontario Ministry of Northern Development and Mines, 1989, 1990, 1992;

Papertzian and Farrow, 1995; Verschuren et al., 1986; Verschuren et al., 1989; Verschuren et al., 1985; and Vos et al., 1981).

This bias has likely suppressed the provision of more representative sampling of granite. Such a bias is especially significant when one considers that the published laboratory work identified in this literature review is largely based on samples provided by quarries. Examples include:

- Westerly Granite - e.g. Kronenberg et al. (1990); Feves and Simmons (1976),
- Oshima Granite from the Yamanashi Quarry, Barre Granite from the Rock of Ages Quarry, and Chelmsford Granites from the Fletcher Quarry - Sano et al.(1992).
- Lac du Bonnet Batholith from Cold Spring Quarry - e.g. Svab and Lajtai (1981); Duevel and Haimson (1997); and Martino J.B. (1994).
- Samples of an unnamed granite pluton from the "Shield Quarries of Canada" near Whitemouth, Manitoba. - Duncan (1987).

In spite of these oversights and sampling biases, review of the recent geotechnical literature indicates progress has been made in a number of areas relating to the recognition of the existence of fabric anisotropy in granites, and the effect of this anisotropy on measured rock properties. (e.g. Kronenberg et al., 1990; Goodman et al., 1993; Goodman ,1993; USNRC, 1993). The next sections review progress made in:

1. the recognition of original and deformation fabrics within plutons,
2. and evidence from the dimension stone industry (DSI) and the general geotechnical literature for the influence of this fabric anisotropy in producing mechanical anisotropy in granite.

### 2.3.2. Field evidence for primary and secondary fabric anisotropy in granite.

Evidence for significant fabric anisotropy in granites is increasingly apparent in the geological literature. Relatively recent compilation maps of Ontario (e.g. Stockwell, 1970) did not differentiate between the various types and ages of granite within the vast areas between the greenstone belts, nor did the mapping efforts include much work on the internal structure of these plutons. However, the granites within each geologic province of the Canadian Shield have since been subdivided and re-classified on the basis of composition, internal structure and age of intrusion (Thurston et al., 1991; 1992 a,b; and Card and Poulsen, 1998). Some types most usually predate regional deformation, while other types generally postdate it; however, these relationships vary locally. The timing of intrusion relative to orogeny is important as it defines to some extent:

- the style and extent of fractures found within the batholiths,
- the complexity of the pluton's geometry and internal structure,
- the types of faults, fractures, deformation structures and recrystallization present, and from these
- the in-situ stresses and rock properties.

Table 7 provides a highly abbreviated summary of the plutonic rocks of the Superior Province. The "vast pink areas" shown in the geological maps of Canada from the 1970's (Stockwell, 1970) are now recognized as including 5 separate suites of plutons, each with its own range of compositional variation and structures.

Table 7:
Plutons Suites in the Superior Province (Thurston et al., (eds.) 1991, 1992 a,b).

| | |
|---|---|
| Suites 1,2: | Gneissic Tonalite Suite & Foliated Tonalite Suites: This suite is largely pre-orogenic and as such, is recrystallized and moderately to strongly deformed. It occurs as large complex batholiths to sill-like remnants, which incorporate several compositional varieties. The internal structure of these suites tends to be very complex and heterogeneous with abundant, often kilometre-scale, mafic inclusions and strong compositional layering, and several generations of granitic and mafic dykes. Secondary and relict primary foliations and lineations are widespread. |
| Suite 3: | Granite-Granodiorite Suite: This suite includes the Lac Du Bonnet Batholith, and is often the most abundant pluton type in any given subprovince. In form, many are generally curved sheet-like crescents that may outcrop over a considerable area but be relatively thin in section (less than 2 km thick). Others may reach considerable depths (10-15 km) and may outcrop over an area of 2000 $km^2$ (e.g. the Lac du Bonnet Batholith). This suite is generally regarded as syn-tectonic but, in some areas, it includes multi-phase pulses, some syn-, some late- and some post-tectonic. Their intrusion near the end of the Kenoran Orogeny, their size and abundance, and a prolonged cooling history are believed to have contributed to the large areas of sparsely fractured rock reported in some members of this suite (Stone et al., 1989; Beakhouse, 1991; Breaks, 1991; M. Gerow, Ontario Ministry of Northern Development and Mines, pers. comm., 1993; B. Schmidtke, Manitoba Energy and Mines, pers. comm., 1993; F. Beard, Ontario Ministry of Northern Development and Mines, pers. comm., 1994).<br><br>This suite is homogeneous when compared to the other suites, but mapping of the LDBB (by the author and others) has determined that much of this homogeneity is the result of a sampling bias. |
| Suite 4: | Muscovite-Bearing Suite: This suite includes syn- or post-tectonic batholiths and stocks of regional extent, as well as numerous pegmatitic dykes. Individual plutons are generally less than a few tens of kilometres in area. The granites are commonly very xenolithic and have irregular interfingering contacts with the surrounding host rock (Everitt et. al., 1998). |
| Suite 5: | Diorite-monzonite-granodiorite suite: This suite may appear under a variety of names depending on location and the rationale for classification. As defined in Thurston et al., (eds.1991a), it includes a compositionally diverse group of syn- to post-tectonic batholiths. Although widespread throughout the Shield, they are neither abundant nor individually large. |

Within any single pluton, recent structural analysis has indicated a wide range of textural and structural variations related to the mode of intrusion, timing of intrusion, and the effects of later deformation. Examples, with the structures listed in their relative order of formation, include:

- sheet-like injections of granite between xenolithic zones are common (e.g. Clarke et al., 1998; Cruden et al., 1997; Cruden, 1998), syn-magmatic folding and a subvertical foliation (Roman-Berdiel et al., 1995),
- the existence of both primary and secondary (deformation) foliations within a single pluton (Benn et al., 1998; Gleizes et al., 1998; Djouadi et al., 1997; Riller et al., 1996),
- Textural variations indicative of the progressive ductile deformation of a pluton (Schulmann and Mlcoch, 1996; Riller et al., 1996),
- several sets of fractures including early brittle structures related to the crystallization and cooling of the pluton (Rousell and Everitt, 1981; Guermani and Pennacchioni, 1998),
- partial erasure of such joints by subsequent localized ductile deformation (Guermani and Pennacchioni, 1998; Rousell and Everitt, 1977; 1981), and finally,
- one or more episodes of brittle deformation and joint formation.

The orientation and density of these fractures is often controlled by compositional and textural anisotropy. Examples of fabric control of joint orientation in granite are

described in Reitan and Murphy (1983), Rousell and Everitt (1981), Guermani and Pennacchioni (1997).

### 2.3.3. Application to this thesis

As such, there is abundant field evidence supporting the view that homogeneity of granites is more the exception rather than the rule, and that granites are internally complex with structures that influence their mechanical properties.

The need to recognize this heterogeneity when sampling for geotechnical investigations was emphasized by Bouchez and Gleizes (1995). The authors made special mention of the need to create "....a sampling pattern consistent with mapped rock types..." so that measured rock properties are representative.

This precaution has been overlooked in geotechnical studies, such as in dam foundations as noted earlier by Goodman (1993), and in the sampling for previous rock properties testing done for the Canadian Nuclear Fuel Waste Management Program, as discussed in section 2.5.

The next section examines correlations between rock fabric and measured rock properties, as noted by the dimension stone industry (DSI), and in select references from the literature. The former provides field-scale evidence of fabric/ rock properties with some lab analyses, and is useful in indicating how common granite anisotropy actually is.

The role of layering, foliations, texture, and microcracks in rock response is examined through selected investigations from general geotechnical literature.

## 2.4. Field And Laboratory Evidence For Mechanical Anisotropy In Granite.

This section summarizes the influence of fabric anisotropy, in producing anisotropy in rock properties as determined from select case studies in the literature.

The influence of large-scale layering, grain size, mineral alignment and natural microcracking are presented in turn. However, it is recognized that the effect of each of these fabric anisotropies are interrelated, and in many cases it is not possible to separate out the individual effect caused by (e.g.) the alignment of grain boundaries (foliation) and grain boundary cracks.

### 2.4.1. Large-scale (e.g. decimetre) compositional layering or heterogeneity

Various texture/rock mass response phenomena have been identified in granite by the DSI and in the general engineering literature. Marmont (1993) and others have made the following observations in recent compilations of the DSI.

The DSI follows the American Society for the Testing of Materials (ASTM) specifications for the standards which building stones should meet, and the test methods to establish them (e.g. ASTM, 1989). The more important tests measure water absorption, modulus of rupture, compressive strength, abrasion resistance and flexural strength. Other tests measure the thermal coefficient of expansion, weather resistance, stain resistance, and indentation under concentrated load as a measure of serviceability

(Marmont, 1993). Significantly, it is noted that "....slabs sawn in different orientations from the same rock may have markedly different strengths, and care must be taken to determine the rock's properties in each direction...". Marmont also notes that "...potential planes of weakness may lie parallel to the layering in some granites...". This statement applies equally to the primary layering observed in true granites, as well as to the more obvious metamorphic layering in gneisses.

### 2.4.2. Influence of texture

For this discussion, texture is limited to grain size. Efforts to define the influence of foliation and lineation on mechanical properties are described in following section.

Zarifa (1998) acknowledged that: 1) granitic rocks show a variety of engineering properties that may affect quarrying operations, tunneling, mining, slope stability and the use of rock as a construction material, and 2) that the physical and mechanical properties of granite are a function of the mineralogical and textural characteristics of the rock. He applied correlation analysis to investigate the relationships between petrographical and engineering properties of a variety of granitic rocks from different parts of Turkey. These were subjected to petrographic analysis, the samples were then tested to determine specific gravity, dry and saturated unit weight, water absorption, effective and total porosity, sonic velocity, Schmidt hardness, point load strength index, uniaxial compressive strength, tensile strength and modulus of elasticity. The relationships between these properties and the petrographical characteristics were then described by regression analyses. The study revealed that the influence of the textural characteristics

on the engineering properties appears to be more important than the mineralogy. It also determined that the types of contacts, grain shape and size significantly influence the engineering properties of the granitic rocks. These results are consistent with those obtained from similar tests on the LDBB (Svab and Lajtai, 1981).

Marmont (1993) and Stecich et al. (1992) identified potential structural problems in dimension stone quarrying associated with coarse or very coarse grained granite. According to these authors, very coarse-grained granites are not as widely used as finer-grained varieties, and equigranular varieties are preferred over porphyritic. The coarser or more inequigranular rocks are recognized as being susceptible to time-dependent degradation by stress relief cracking along grain boundaries or along the cleavage planes in feldspars. An example of such degradation is especially well developed in a highly stressed intact outcrop exposed at a quarry near Whitemouth, owned by Canital Granite of Manitoba (Schmidtke, 1986; also Everitt et al., 1994). Specimens, cut and polished to serve as floor tiles, will disintegrate after a storage period of as little as a month due to the propagation of grain boundary microcracks.

### 2.4.3. Mineral alignment and natural microcracking

In addition to the aforementioned responses related to grain size, it has also long been recognized that even the most homogeneous and uniform granites usually possess a microscopic structure which predisposes them to break more easily in one direction than another (Duncan, 1987; Marmont, 1993). This is usually caused by weak alignment of elongate mineral grains, and /or by sets of pervasive natural microcracks, either

intergranular, or transgranular (Richter and Simmons, 1977). The DSI terminology for this mechanical anisotropy is as follows:

rift: the direction along which granite splits most easily is the called rift,

grain: the next easiest direction of splitting,

hard way: the direction along which the rock is hardest to split (also known as the "head").

The rift is commonly horizontal or vertical, the grain is usually perpendicular to the rift, and the head perpendicular to rift and grain. These directions commonly correlate with batholith structure. Stecich et al. (1992) describes the rift in dome-shaped granite intrusions being parallel to the granite's contact with the enclosing rocks. The rift and grain are used to plan the quarrying and processing of stone, with stone cut parallel to the rift being stronger than stone cut across it. Understanding of this is considered critical in large architectural panels (Marmont, 1993).

Many of the granites included in the dimension stone literature are technically gneisses, having undergone one or more episodes of deformation and recrystallization (e.g. Marmont, 1993). Marmont notes that good strength characteristics are associated with the more massive granites with "..complex interlocking grain shapes ...". In gneissic granites a well-developed rift usually accompanies the layering. Many granites also possess a linear structure formed by alignment of mineral grains in the plane of the layering, which frequently determines the orientation of the grain. The hard way cuts

across both layering and lineation. Marmont According to the author "...care must be taken in the retrieval of dimension stone to establish that gneissic layering is not a plane of weakness...".

### 2.4.4. Foliations and mineralogy

Sano et al. (1992) conducted an experimental determination of the elastic constants for the Oshima, Barre and Chelmsford granites. This study is unique in its examination of the effect of mineral anisotropy and its contribution to anisotropic elasticity in rocks. According to the authors, the major rock forming minerals have anisotropic elasticity, with the crystals of quartz, orthoclase and plagioclase being trigonal, monoclinic, and triclinic, respectively. A preferred alignment or segregation of these minerals within a rockmass should therefore result in an anisotropy of the linear elasticity and the coefficients of thermal expansion for the rockmass. To test this hypothesis, their granite samples were cut into polyhedra. These were then used to measure the longitudinal and shear wave velocities in various directions of propagation and polarization, which they then related to the fabric of the samples. The results showed:

1. that oriented microcracks are mainly responsible for the orthorhombic elasticity of both granites under lower confining pressures up to 120 MPa (6-8 km depth).
2. However, a slight anisotropy, oblique to the microcracks, was found even at 180 MPa. This correlated with the preferred alignment of biotite and chlorite.

### 2.4.5. Microcracks:

In addition to structures such as bedding or foliation, oriented pores and cracks are also a significant cause of anisotropy in rocks. The role of natural microcracks, specifically their influence on rock mass properties, was summarized in a review by Sano et al. (1992), and by others as follows:

1. In granitic rocks, oriented microcracks are very typical, and these cracks may be associated with either residual or external stresses (Dale ,1923; Nur and Simmons, 1970; Sprunt and Brace, 1974; Plumb et al., 1984; Kudo et al., 1987; Homand-Etienne and Sebaibi, 1996).
2. Natural microcracks include the following types: 1) submagmatic, quartz – mica filled intergranular fractures (Bouchez et al., 1992), 2) planes of aligned fluid inclusions representing re-sealed ancient microcracks (Jang et al., 1989; Rehrig and Heidrick, 1972), 3) oxide-filled ancient microcracks, and 4) open to partially sealed microcracks formed along grain boundary and mineral cleavages (Ren et al., 1989; Richter and Simmons, 1977).
3. The fabric of microcracks in granite rocks shows a correlation with the anisotropy of physical properties: for example, with sound velocity (Simmons et al., 1975), Young's modulus (Douglas and Voight, 1969), compressibility (Tod et al., 1973), uniaxial compressive strength (Douglas and Voight, 1969), tensile strength (Peng and Johnson, 1972), dilatancy anisotropy under compression (Scholz and Koczynski, 1979).

4. Duncan (1987) in a study of a Precambrian granite, found that the presence of microscopic fractures in the intact rock influences the Brazilian tensile strength of intact rock, such that reduced tensile strengths occur across those planes characterized by a higher degrees of subparallel microscopic fracture development.

Peng and Johnson (1972) found similar variations in Brazilian tensile strength for several orientations of Chelmsford granite. They found that the tensile strength was:

- lowest in the direction normal to the plane of the preferred orientation of microcracks.
- second highest normal to a plane characterized as having a preferred orientation of mica cleavages, and
- generally higher along other orientations for which there were no corresponding fabric elements.

Baek (1997) examined the effect of oriented microcracks on acoustic wave velocity and axial Point Load indices. Specimens were cored from blocks of medium grained leucocratic granite, bearing strong anisotropy parallel to the rift plane. The circumferential acoustic wave velocity was measured to determine the direction of the minimum velocity, presumably perpendicular to the dominant microcrack direction. The direction of fractures induced by the axial point loading was found to be closely related to the direction of natural microcracks in the specimen, interpreted from the circumferential velocity anisotropy, and supported by microscopy.

Chen et al. (1997) examined the correlation between microfracture type and physical properties of two granites. The Inada granite is anisotropic in "...dynamic property or permeability" and splitting planes, while the Kurihashi granodiorite is isotropic with no clear splitting planes. Three types of microfractures were distinguished in both specimens: intracrystalline cracks, intercrystalline cracks and grain boundary cracks. The orientations of these were measured and a total length per unit area calculated. It was concluded that:

1. the Inada granite was anisotropic due to a preferred orientation of intercrystalline and intracrystalline cracks, while

2. the Kurihashi granodiorite was isotropic as it contained no preferred orientation of intercrystalline cracks, but there is a strong preferred orientation of grain boundary cracks and intracrystalline cracks, which the authors correlate with indistinct splitting planes.

Similar results linking mechanical anisotropy and preferred orientations of natural microcracks were obtained by Kusumi et al. (1997) (Effect of structural anisotropy on deformation properties of granite under cyclic loading), and Park (1996 ) (The anisotropy of the granite revealed by tensile tests).

### 2.4.6. Conclusion — Influence of Fabric Anisotropy in Other Granites

Strong evidence from quarrying operations and from rigorous rock properties testing in granite using a variety of sonic and rock properties testing, has indicated that granites are texturally and mechanically anisotropic. Most often this anisotropy can be linked to preferred orientations of natural microcracks. These tend to be enhanced by sampling and stress relief, with the effect markedly less noticeable at high confining pressures typical of depths of 6-8 km. Mineral alignment or mineral segregations also induce mechanical anisotropy to a somewhat lesser degree, but their influence is not affected by confining pressure.

## 2.5. Rock Properties Testing in the Lac Du Bonnet Batholith.

This section reviews relevant case studies from the published literature, which examined the deformational properties of the LDB granite and, where data was available, their variations with direction. Comments on the results of these studies and their significance to this investigation are provided.

Svab and Lajtai (1981) examined the effects of variable mineralogy and mineral habits on tensile fracture paths through LDBB granite. They noted that the most obvious indicator of anisotropy within the LDBB granite was a (i.e. single) foliation, marked by subvertical quartz-biotite lenses and mineral alignment striking 075°. To test for strength anisotropy related to either foliation or to microstructural control of fracture, they used: 1) biaxial tension tests to define the orientation of preferred fracture, and then 2) Brazilian tests to evaluate the strength of the anisotropy. The first technique is sensitive to weakness

directions in the rock, while the second provides more accurate values of the tensile strength associated with these weakness directions.

The samples used by Svab and Lajtai (1981) were from the Cold Spring Quarry near Lac Du Bonnet, Manitoba. The background information on the geology for this study was based on initial mapping results of the Lac du Bonnet Batholith published by Tammemagi et al. (1979; 1980; 1982). At that time, the major phase of the batholith was described as: "...a medium to coarse grained (0.5-20 mm), pink, and relatively homogeneous granite. Foliation is limited to several kilometres along the contact, trending approximately 075° with a steep northwesterly dip at the quarry."

The testing (by Svab and Lajtai, 1981) of vertical core revealed that: 1) the subvertical 075° foliation was a plane of preferred fracture, 2) fractures also clustered around 005° and 155°, but there were no recognized foliations for these orientations, and 3) although there was a foliation / fracture correlation at 075°, tensile strength in this plane was only slightly lower than in any other orientation and "... not statistically different to justify the assumption of a multimodal population..." within the horizontal plane.

Subsequent investigations by the author and others (most notably, Brown et al., 1989; and Everitt et al., 1999) have shown that: 1) the sample location used by Svab and Lajtai (1981) is more homogeneous than most other locations in the LDBB and at the URL, and the foliations and microstructures less well developed, 2) the pink colour of the granite is due to hematite from alteration and recrystallization which partially cements grain

boundaries and natural microcracks, 3) that foliation is not limited to the margins of the batholith, 4) that some foliations and natural microcrack sets present in the LDBB have been partially sealed by mineral infilling and alteration (including the 075° foliation), and 5) that there are several foliation and natural microcrack sets aligned with the 005° and 155° preferred fracture directions noted by Svab and Lajtai (1981) (see Figure 16 of Brown et al., 1989). Accordingly, it is not surprising that a greater degree of mechanical anisotropy was not detected. The two preferred directions of fracture in the horizontal core – 029°/65° NW and 029°/75° SE, may correlate with any number of foliation or natural microcrack sets found in the URL and elsewhere in the LDBB.

Duevel and Haimson (1997) conducted a mechanical characterization of pink Lac du Bonnet granite to test for evidence of nonlinearity and anisotropy in samples of the pink granite. It was taken that the pink granite may better represent the in-situ properties of damaged grey granite recovered from highly stressed rock. Samples of grey granite, obtained from depth, exhibit mechanical anisotropy due to sample damage induced by stress-relief (Martin, 1990). This sample damage includes discing in core, and the opening of grain–scale microcracks along mineral cleavages and grain boundaries due to stress relief.

The samples used by Duevel and Haimson (1997) were from the Cold Spring Quarry near Lac Du Bonnet, Manitoba. The background information on the geology for this study was based on original mapping and core logging results for the Lac du Bonnet Batholith and the URL, published by Brown et al. (1989). Additional comments on the

mineralogy, joint orientation and microcracking were based on Martin (1990), and Martin and Stimpson (1994). The batholith was again described as "... medium to coarse grained (0.5-20 mm), and relatively uniform in composition and texture, but varying in colour from grey to pink". The pink granite was described as "... shows no visible microcracking (Martin, 1990), though joints are more prevalent at the surface, with the most prominent set striking at 020° and dipping vertically, with a spacing of 1–2 m".

To test for mechanical anisotropy, a suite of specimens was taken from core oriented in three orthogonal orientations: east-west (E-W), north-south (NS), and vertical (V). These were then subjected to: sonic velocity tests, uniaxial and triaxial compression tests, and Brazilian and Uniaxial tension tests. These results were then correlated with known fracture orientations, but no microstructure of fabric analysis of the specimens was done.

Sonic velocity tests yielded minor directional differences on the order of 10% - with the velocity high in the V (vertical) direction which the authors attributed to the predominant vertical joint set, and to microcracks. The lowest velocity was the E-W direction, approximately normal to the subvertical 020°-striking joint set. Although no microcracks were observed in the test specimens, the authors concluded that the velocity results suggest that they do in fact exist, and are predominantly subparallel to the outcrop-scale surface joint set.

The uniaxial and triaxial compression testing results point to compressive strength isotropy. For the uniaxial compression, the maximum difference between any two

directions was 6 MPa (less than 3%), while the variation within each direction was less than 5%. The overall average uniaxial compressive strength is 219 (±9) MPa. Duevel and Haimson concluded that the preferred microcrack orientation implied from the sonic velocity results does not affect failure in compression.

Unlike the compression tests, the tension tests showed evidence of strength anisotropy. Brazilian and uniaxial tensile strengths were both found to vary up to 25% between orientations. For both measures of tensile strength, the E-W direction was the weakest while the V direction was the strongest. The tensile modulus displayed more anisotropy than any other property measured in this study. The results of the sonic velocity and tensile strength tests suggested the existence of microcracks aligned with the predominant surface joint set. The authors again correlated the directions of the least and largest tensile strengths with the smallest and largest velocities, respectively, and these in turn with the preferred orientation of vertical microcracks.

The elastic parameters in both compression and tension varied continuously with the stress magnitude, indicating nonlinear stress-strain behaviour. The tensile modulus of deformation was substantially lower than that in compression. The authors concluded "...Our work brings into question the use of isotropic models in excavation design, in which compressive material constants determined at one level of loading are employed regardless of the local state of stress... ".

Duevel and Haimson (1997) concluded: 1) that the pink Lac du Bonnet granite is practically isotropic with respect to uniaxial and triaxial compressive strength, compressive elastic parameters, and sonic velocities, but 2) the tensile strength shows a difference of 25% between directions, and the tensile modulus of deformation showed an even stronger anisotropy, with differences of up to 80% between directions, 3) that this anisotropy is due to natural microcracks with the same orientation as outcrop-scale joints, and 4) ".... As the gray Lac du Bonnet granite differs from the pink variety only by possessing stress-induced microcracks created during sampling, it can be deduced that the preceding results describe the in-situ gray granite as well....".

The criticisms made earlier regarding the study by Svab and Lajtai (1981) are also applicable to the study by Duevel and Haimson (1997). These are:

1. That the Cold Spring quarry is not representative of the LDBB, especially as seen in the URL.
2. That Duevel and Haimson may have combined results from dissimilar phases of the LDBB granite,
3. That their study been done without identifying the fabric elements actually present at the site from where the samples were taken, and
4. That the orientations chosen for sampling are oblique to each of the various subvertical and subhorizontal structures known to exist at the batholith-scale. There is also no indication how these orientations relate to any local structures at the quarry.

Detailed geological mapping (Brown et al., 1989; and Everitt et al., 1999) has shown that: 1) the Cold Spring quarry from which the samples were taken is more homogeneous than most other locations in the LDBB and at the URL, and the foliations and microstructures are less well developed. 2) that colour is not a valid basis on which to categorize the mechanical behaviour of the LDBB granite. As is described in Chapter 3, there are several genetically distinct phases of the LDBB, which differ from each other in their average grain size, range of grain sizes, degree of recrystallization, and in the number, type and orientation of foliations and natural microcracks. Pinking is the result of secondary alteration superimposed on these diverse rock types wherever fractures have opened the batholith to groundwater movement.

Svab and Lajtai (1981) identified a subvertical striking - 075° foliation in their samples from the Cold Spring quarry, which biaxial tension tests revealed to be one of several planes of preferred fracture. The study by Duevel and Haimson (1997) makes no mention of any foliation in their samples, or of the results of the earlier work by others. As such it is difficult to determine: 1) what textural / compositional rock type was actually being tested, 2) what structures were present, and 3) how the test results are affected by the orientation of the samples. Accordingly, it is not surprising that a greater degree of mechanical anisotropy was not detected. Also, it is uncertain to what extent the results can be applied to the various types of grey granite at the 420 Level of the URL, which hosts the various geomechancial experiments.

An alternative approach to defining the in-situ properties of highly stressed rock at depth, would be to identify the granite varieties present at the intended test location and depth, and then to obtain samples of these same varieties where they are exposed at or near the surface. In adopting this approach, it would be necessary to determine if the granite at surface and at depth differ substantially in the characteristics of their natural microcracks.

## 2.6. Conclusions

From the literature - general observations on fabric anisotropy, methodology and earlier site work are as follows:

1. Rock fabric is a recognized factor influencing rock strength and response in metamorphic and sedimentary rocks, but plutonic rocks (mafic to felsic) are generally regarded as homogeneous and isotropic.
2. Most work pertaining to plutonic rock comes from the various international nuclear waste programs. The majority of this is from the Canadian program as Canada is unique in having large volumes of intact or sparsely fractured rock – in other words – the granites of other countries are anisotropic due to their pervasive fracturing.
3. Significant fabric anisotropy has been recognized in other granites.
4. Recent investigations provide evidence that, in other granites, these fabrics, usually foliations or natural microcracks, influence deformational properties,
5. The dimension stone industry (DSI) has long recognized the existence of mechanical anisotropy related to fabric anisotropy in granites, and has utilized these phenomena to facilitate exploration and excavation. In contrast, the inherent anisotropy of granitic rocks has been overlooked by the various radioactive waste disposal

programs, but especially in the Canadian program which is unique in that it is being conducted in large volumes of unfractured and intact rock. Programs in other countries are being conducted in granites which are inherently anisotropic due to abundant joints and other discontinuities.

6. The Lac du Bonnet Batholith is incorrectly referred to as a homogeneous granite in literature as recent as 1998, while Hoek and Brown (1995) incorrectly refer to the Mine-by Experiment at the 420 Level as demonstrating that there is no influence of geology.

The methodologies used in this thesis are based on the findings and recommendations of earlier researchers. An array of oriented samples was obtained. Strength anisotropy and the influence of foliation was then systematically investigated using rock mechanics tests that are sensitive to planar anisotropy (Point Load and Brazilian tests), the former defining planes of naturally preferred fracture. Variations in the compressive strength of the granite related to grain size and fabric was examined through uniaxial compression testing of an array of samples providing a range of angles between the loading axis and the fabric.

Although the material properties of granites (and the LDBB) have been investigated in the literature, these investigations differ from this investigation in the following areas:

1. other studies (e.g. Duevel and Haimson, 1997) have been done without identifying the fabric elements present in the granite tested, or have combined results from

dissimilar phases of the LDBB granite, or have incorrectly identified the orientation of structures present

2. other studies have not identified and classified the foliations and other anisotropies present in granite at a variety of scales,
3. nor have they combined this information with the influence these anisotropies have had on the development of both ancient and excavation-induced fractures.

In these aspects, the present study is unique and original in its approach, scope and application.

# Chapter 3

# SITE DESCRIPTION

## 3.1. General Geology of the Lac du Bonnet Batholith

The study location for this investigation into the influence of foliation on rock mass properties is the Lac du Bonnet Batholith (LDBB). Its location and regional setting are shown in Figure 1. The Lac du Bonnet Batholith is one of many late-tectonic granites emplaced in the western Superior Province of the Canadian Shield towards the end of the 2760-2670 Ma Kenoran event (McCrank et al., 1981; Davis et al., 1986; Stone et al., 1989). The presence of very large sparsely fractured volumes within these plutons is characteristic of this suite, and is attributed at least in part to:

1. their intrusion at or near the end of regional deformation (Brown et al., 1989),
2. to prolonged cooling, which delayed the onset of brittle deformation, due to the size of the bodies and their proximity to each other, and
3. to the distance separating the intrusions regionally from later tectonic events in the Superior Province (Everitt et al., 1990).

To the north, the batholith is in sharp contact with a narrow belt of metavolcanic rocks in the Bird River Greenstone Belt, and to the south it is in gradational contact with gneisses and migmatites of the Winnipeg River Subprovince (Beakhouse, 1977). From a tapered eastern margin, surface exposures of the batholith extend westward 85 km. West of the Winnipeg River, the batholith is largely concealed by glacio-lacustrine sediments and by

Paleozoic carbonate rocks; however, geophysical data indicate that the granitic rocks extend westward at least another 20 km (McRitchie, 1971).

## 3.2. Underground Research Laboratory

The Lac du Bonnet Batholith was selected for investigations by Atomic Energy of Canada Limited (AECL) as part of a program to assess the concept of nuclear waste disposal in plutonic rocks of the Canadian Shield (Whitaker, 1987). The geology of the batholith was defined through surface and airborne studies and by 130 boreholes, some reaching depths of 1100 m. The subsurface structure has been further characterized in detail by the author through geologic mapping of the shafts and galleries of AECL's Underground Research Laboratory (URL).

The URL (Figure 1) includes a vertical access shaft that extends to a depth of 443 metres and several hundred metres of tunnels at depths of 240 and 420 metres (240 and 420 Levels, respectively). Additional exposure is provided by raise-bored ventilation shafts, a timber-framed raise at the 420 Level, and small shaft stations in the main access shaft at depths of 130 m and 300 m. Boreholes from the base of the shaft and lowest level provide an additional 600 m of vertical exposure. These excavations and the boreholes provide an 1100 m cross section of the roof zone of the Lac du Bonnet Batholith, two low-dipping thrust faults and associated splays, and the subvertical intrablock fractures that flank the major faults. The roof zone is marked by deuteric alteration and by shallow-dipping compositional layering. Low-dipping fractures, including thrust faults, parallel the large-scale compositional layering, and are generally confined to the contacts

between massive leucocratic and foliated xenolithic litho-structural domains. The geology of the URL is representative of the batholith.

### 3.3. Batholith Structure

The many late to post-tectonic granites in the western Superior Province, and the LDBB in particular, are noted for their structural and compositional homogeneity over extensive surface areas (Beakhouse, 1991; Cerny et al., 1987; MacLeod, 1980; McCrank, 1985), and have been described as type examples of "massive homogeneous plutons" in the literature. However, mapping in the 443 m deep access shaft at the Underground Research Laboratory (URL), borehole logs, and detailed surface mapping clearly demonstrate that dekametre-scale layering and smaller scale structures are widespread in the LDBB. Simplified geological maps of the upper and lower portions of the URL access shaft are shown in Figure 10 a and b, respectively. The upper shaft (0 to 257 m depth) has a rectangular cross section. From left to right are shown the north, east, south, and west shaft walls. The main unit of the granite is shown in white, while the metavolcanic xenoliths and the three generations of slightly younger dykes, sills and masses are shown in other colours. The shaft between 257 m and 442 m has a circular cross-section. The numerals on the left side of each shaft segment are the lithostructural domain reference numbers, and the depth (in metres) from the shaft collar. To the right side of each segment are short descriptions of the domains or of other distinctive features. Block diagrams summarizing the structures in the Lac du Bonnet Batholith as seen at AECL's URL are shown in Figure 11.

Individual layers or litho-structural domains are distinguished by the type and abundance of xenoliths and late residual or metasomatic segregations, by auto-intrusive contacts, and by variations in style and orientation of mesoscopic foliations (Note: the term auto-intrusive refers to late residual segregations, coming from the same magma as that comprising the bulk of the batholith, and it specifically excludes other dykes or veins related to younger plutons and tectonic events). The litho-structural domains are inconspicuous at surface due to their low dip and the flat topography; however, detailed mapping in the central part of the URL has identified their surface expression as a series of broad and open antiforms and synforms. The textural and compositional layering is clearly delineated in the URL shaft (Everitt and Brown, 1986), where it has been shown to exert a strong control on the localization of low-dip thrust faulting, and on the frequency and properties of the subvertical fractures.

The litho-structural domains host three systems of auto-intrusive dykes, sills, and recrystallized zones whose abundance and mode of occurrence change with depth. Late magmatic granodioritic dykes are the predominant rock type below 300 m depth at the URL, but the swarm narrows rapidly up-dip, and at surface is represented only by narrow zones of alteration and ductile deformation. Late pegmatite-aplite dykes are pervasive across all rock types at surface, but are limited to the larger fine grained granite dykes at 420 m depth at the URL. The distribution of inclusions, alteration and fractures in the LDBB suggest the present topographic surface is close to the original roof zone of the batholith.

Table 8.
Comparison of Petrological Units from surface and subsurface mapping of the Lac du Bonnet Batholith.

| Occurrence | Unit # | Reference Cerny et al.( 1987) | McCrank (1985), Stone et al. (1984) | Everitt et al. (1998) |
|---|---|---|---|---|
| xenoliths | 1,2 | Predominantly tonalites and amphibolites but also some granitoids | Same | Same |
| main unit | 3 | Biotite granite | Biotite granite, and textural variants. | Same * |
| auto-intrusive dykes | 4 | | Irregular pegmatitc masses and porphyroblastic schlieren | Granite to pegmatitc sills, and lesser masses, dykes and porphyroblastic schlieren * |
| auto-intrusive dykes | 5 | Late-tectonic biotite granodiorite dykes | | Granite-granodiorite dykes* |
| auto-intrusive dykes | 6 | Pegmatite dykes | Pegmatite - aplite dykes, and quartz veins | Same |

* Note: For the sake of convenience, units 3,4 and 5 are referred to in this report as the medium, fine and coarse grained granites, respectively.

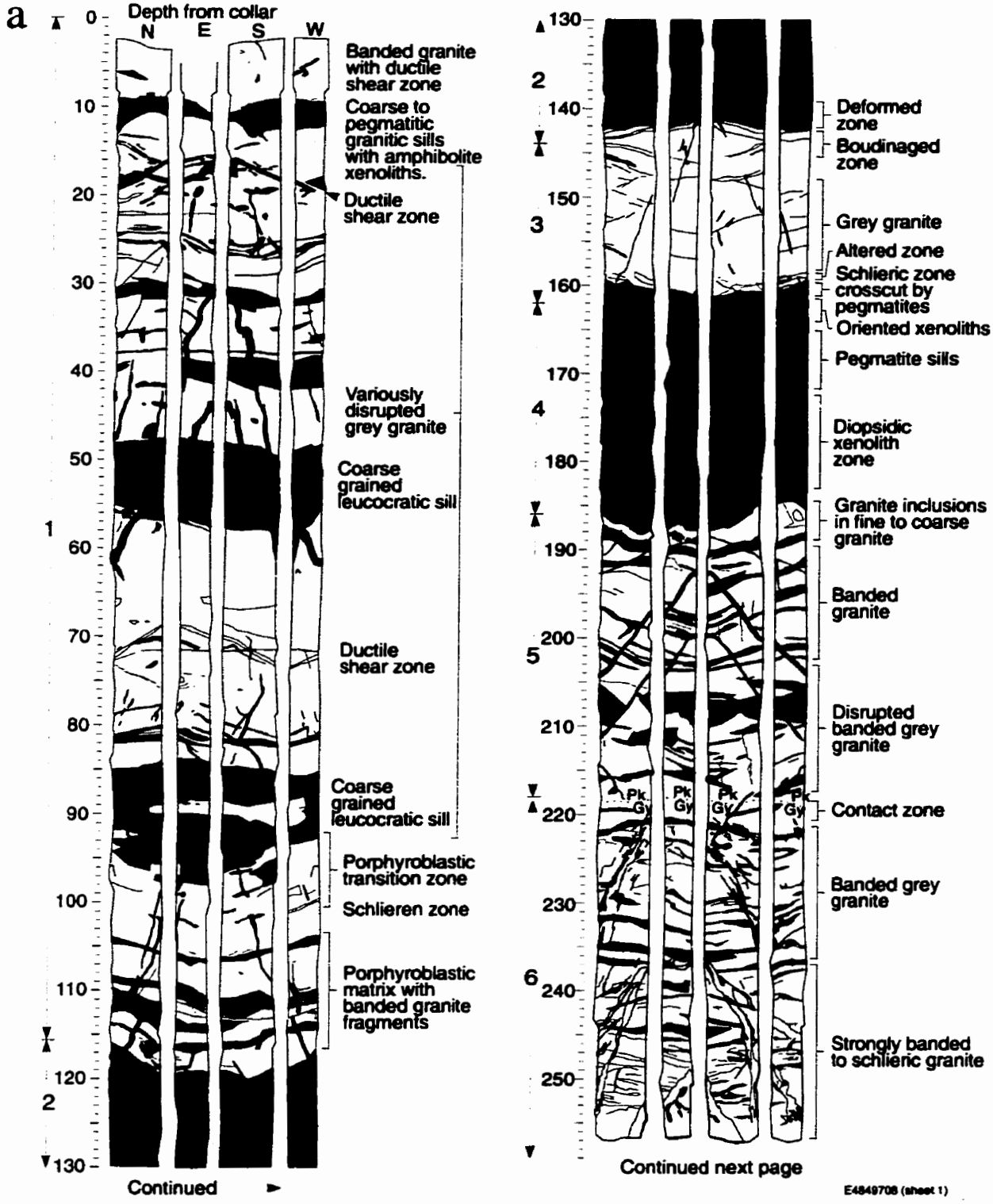



Figure 10a. Simplified geological map of the upper portion of the URL access shaft (continued next page, refer to text for explanation).

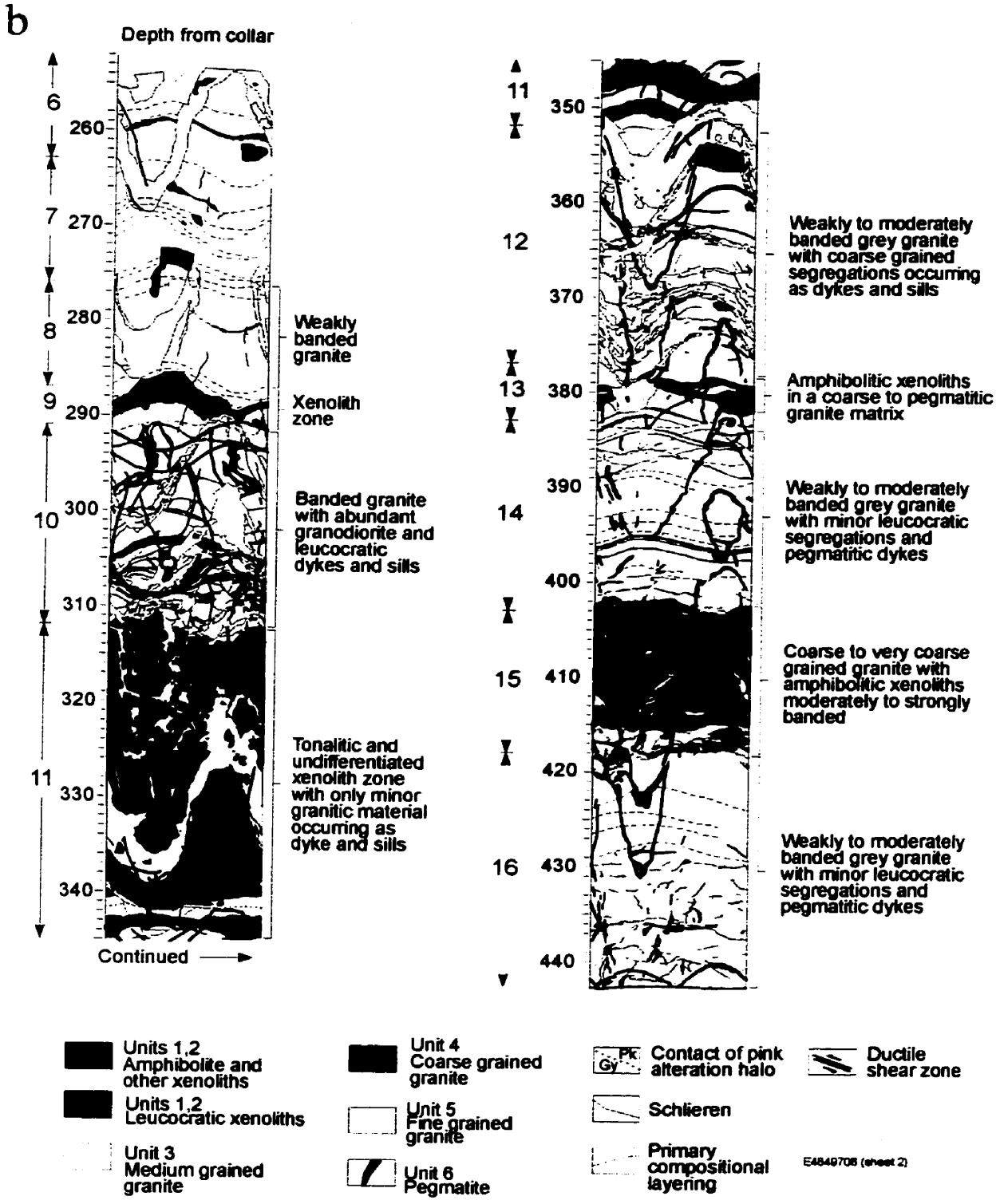



Figure 10b. Simplified geological map of the lower portion of the URL access shaft (refer to text for explanation).

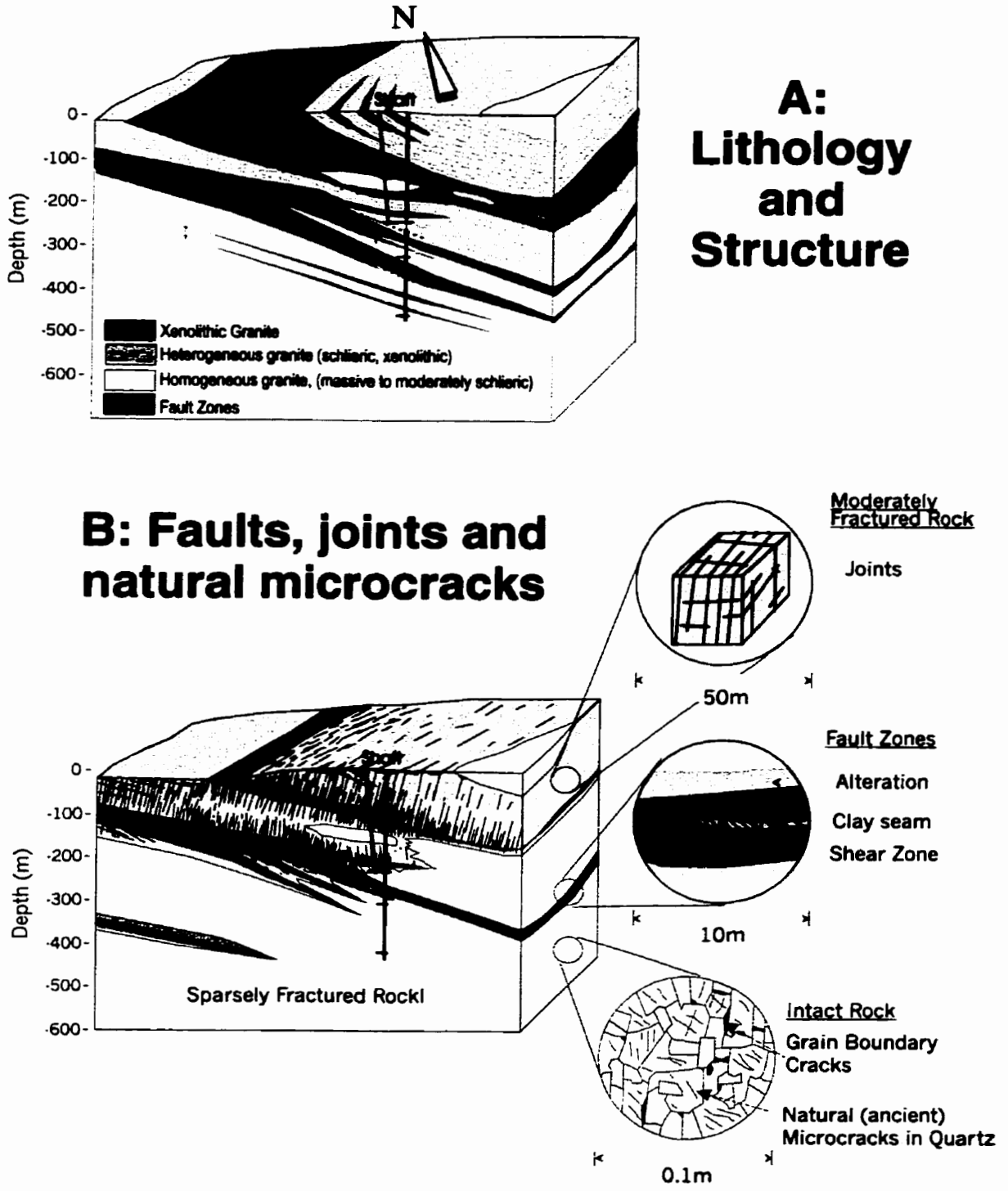



Figure 11. Block diagrams summarizing the structures in the Lac du Bonnet Batholith as seen at AECL's URL. A: Large-scale compositional layering and the low-dipping thrust faults, which are concordant with the layering. B: Faults, joints and natural microcracks within the same area.

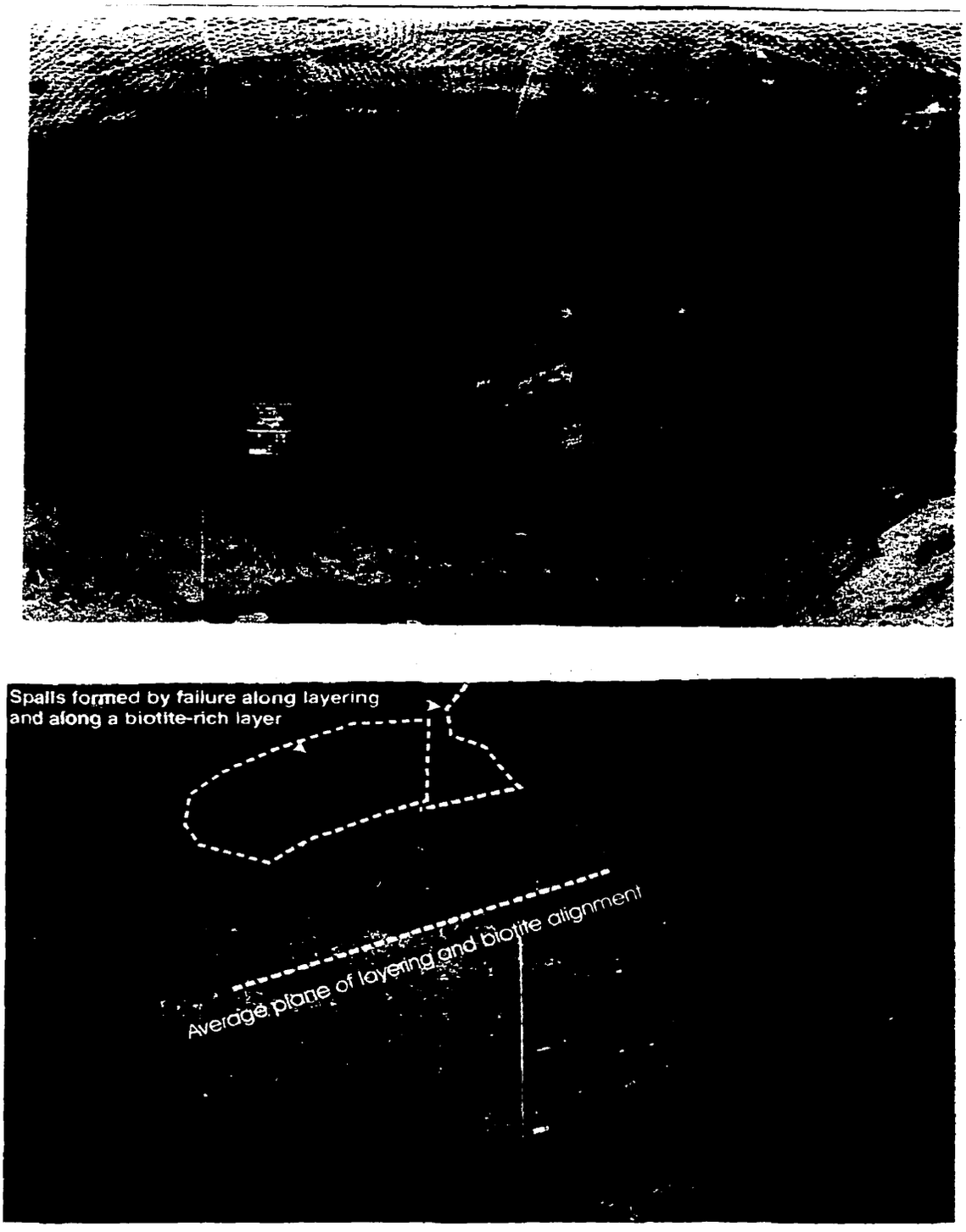



Figure 12. Rock units and structures recognized at depth in the URL, and selected for rock properties testing in this study. Each photograph is of a vertical tunnel face. The scales are approximately the same, each photo showing an area of approximately 3.5 m by 2.5 m. The measuring stick is 1m long. The fine and medium grained granites are shown in the upper photograph (A). The coarse grained and (locally xenolithic) granite is shown in the lower photograph (B).

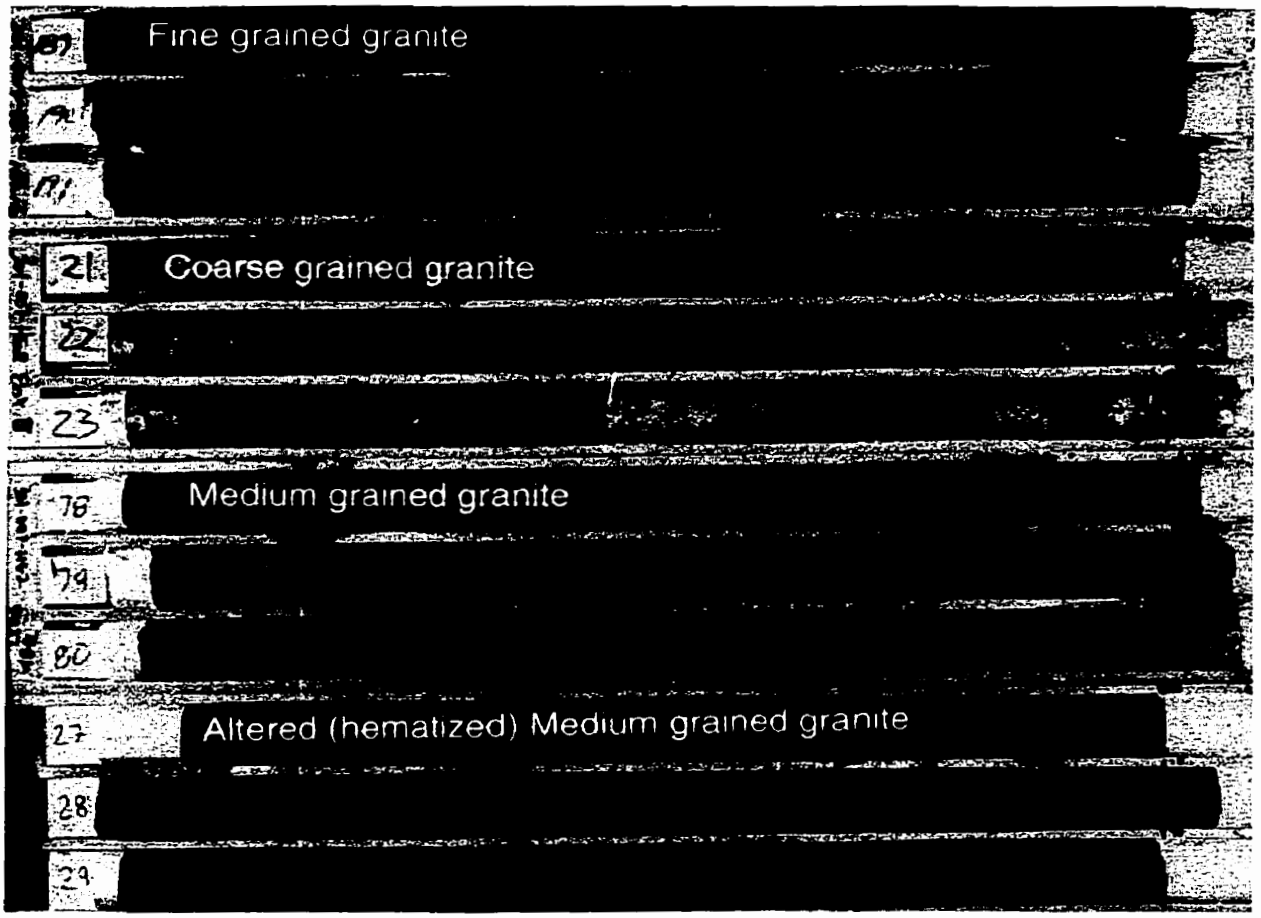



Figure 13. Typical appearance in core, of the granite types selected for rock properties testing in this investigation. From top to bottom are the fine-, coarse-, and medium-grained granites, respectively, and the pink (altered) equivalent of the coarse granite.

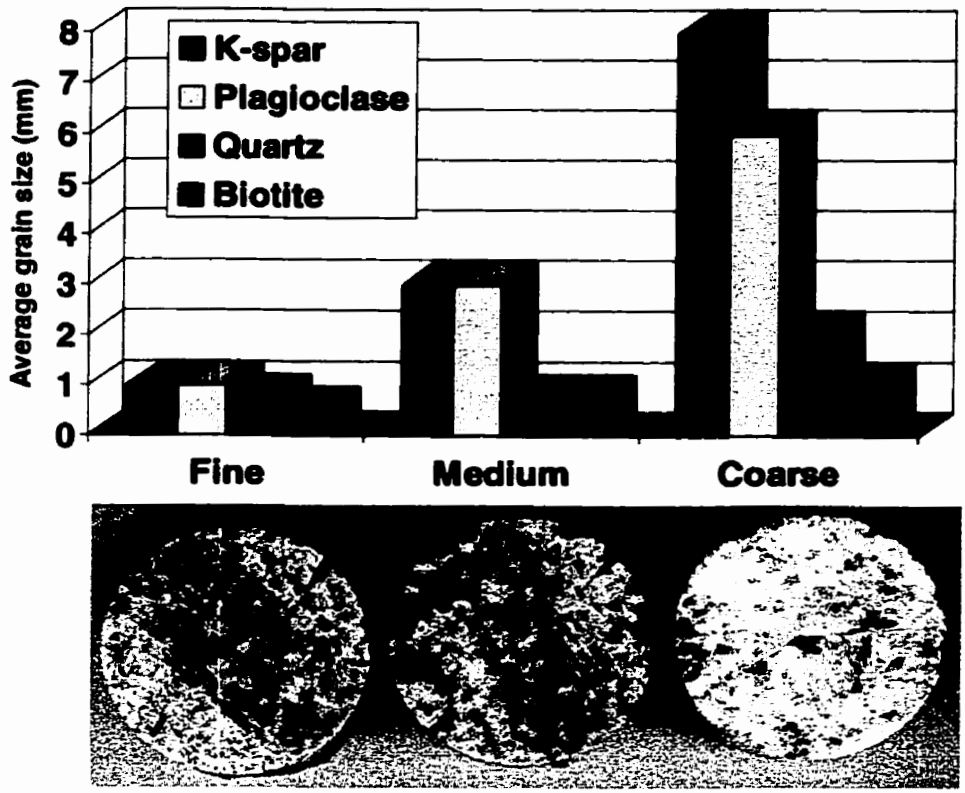



Figure 14. Grain sizes and variation with mineralogy, of the three rock types selected for rock properties testing.

## 3.4. Outcrop-scale Rock Units

On the basis of mineralogy, texture, intrusive relationships, and geochemistry seen from surface exposures, various authors have identified up to six units, with additional textural variants (Table 8). These may be broadly grouped as follows:

- meta-volcanic xenoliths and earlier granites (units 1 and 2 in Table 8),
- the main unit of the granite (unit 3 in Table 8).
- several sets of auto-intrusive dykes (units 4,5 and 6 in Table 8), which range from fine- to very coarse-grained.

Photographs of these units as they appear at the tunnel- and borehole-scales are shown in and Figure 12 and Figure 13, respectively. Colour is not a distinguishing characteristic of the rock units nor is it a guide to compare rock mass response, although it has been used frequently as such in the literature (e.g. Lau and Gorski, 1991a,b). Each of the rock units listed ranges from a uniform grey, to a red or pink where altered. Most reddening is associated with fractures, where the rock has been subject to long-term exposure to groundwater, and cuts across rock unit boundaries (Brown et al., 1989; Stone et al., 1989). Low temperature alteration overprints the higher temperature forms in the area of the major low-dipping fracture zones. Illite replaces epidote and biotite and Fe-oxides are removed, leaving the rock bleached in appearance.

Three distinct textural varieties of the Lac Du Bonnet granite, differing in type and orientation of foliation were identified and selected for testing. These are: 1) fine grained granite dykes with subvertical flow banding, 2) medium grained and weakly to moderately layered granite forming the main mass of the batholith, and 3) a generally coarse grained and leucocratic unit that occurs as sills and recrystallized zones in the main unit. For the sake of convenience, these are referred to as the fine, medium and coarse grained granites, respectively. The grain sizes for each of these units are illustrated in Figure 14.

### 3.4.1. Fine Grained Granite

This rock unit (unit 5 in Table 8) occurs as a dyke swarm, which increases in both size and complexity with depth. A typical exposure of the contact between a fine grained

granite dyke as seen at the 420 Level, and the medium grained granite is shown in Figure 12A. The location selected for sampling was similar to this exposure.

Initial petrographic analyses of samples from the 240 Level confirmed Cerny's (1987) granodiorite classification for this unit, but subsequent analysis of additional samples from the 420 Level indicate this unit straddles the granite-granodiorite boundary of the IUGS classification, with substantial overlap into the (biotite) granite field (IUGS, 1973). At the 240 Level, these dykes comprise up to 10% of the rockmass, and are up to 5 m thick with well-developed chilled margins and flow banding. Flow banding is defined by a preferred alignment of biotite crystals, biotitic aggregates and inclusion trains. At the 300 and 420 Levels of the URL, granodiorite occurs as a stockwork of large dykes and minor interconnecting sills and is the most abundant rock type. The largest dykes at and below the 420 Level are medium grained and, apart from the orientation of the foliation, are indistinguishable in hand specimen from the biotite granite they intrude.

### 3.4.2. Medium Grained Granite

This rock type comprises the main unit of the granite (unit 3 in Table 8, the unshaded area in the shaft geology maps (Figures 10a and b, and Figure 11). Its description matches that of Cerny et al. (1987) and McCrank (1985) from surface mapping, but in the subsurface, its fabric ranges from massive to schlieric, its texture from equigranular to porphyritic. The area selected for sampling of this rock type was similar to the exposure but was free of inclusions and schlieren, and the layering was weakly developed.

### 3.4.3. Coarse Grained Granite

This rock type (unit 4 in Table 8 and shown in Figures 10a and b, and Figure 11) comprises a heterogeneous assemblage of residual leucocratic segregations and porphyroblastic zones. Textures range from aplitic to pegmatitic, and porphyritic to equigranular. The contacts of this unit range from sharp to gradational, and are generally parallel to the layering of the main unit granite. The largest zones (up to 30 m thick) occur in association with xenolithic horizons, and have complex margins and internal fabrics. The photograph shows the actual area where the samples for this rock type were taken. Note the well developed layering. The dark area on the left hand side of the crown is the scar left by the spalling of a slab along a plane of biotite-rich band.

## 3.5. Rock Fabric

A summary of fabric elements present at the (420 Level) level of the case studies is given in Figure 15. The structures are listed in order of relative age. The orientations of each structure are summarized on the stereographic projections, and in the accompanying conceptual block models which also summarize the age relationships. Poles to the gneissosity and larger scale layering cluster about a pole with a strike (dip direction) dip of 032°(122°)/22°. The basic geometries of the fine grained granite and pegmatite dyke systems are similar but are differently oriented. For each system:

- poles plot along great circles but several sets are recognized;
- the subvertical set is dominant;
- the subsidiary sets fall into two groups, those inclined about 45° to 60° from the subvertical set and those whose orientations indicate local control by older fabric

elements (e.g. the southeast dipping fine grained granite sills follow the gneissosity and the subvertical pegmatites follow some of the fine grained granites).

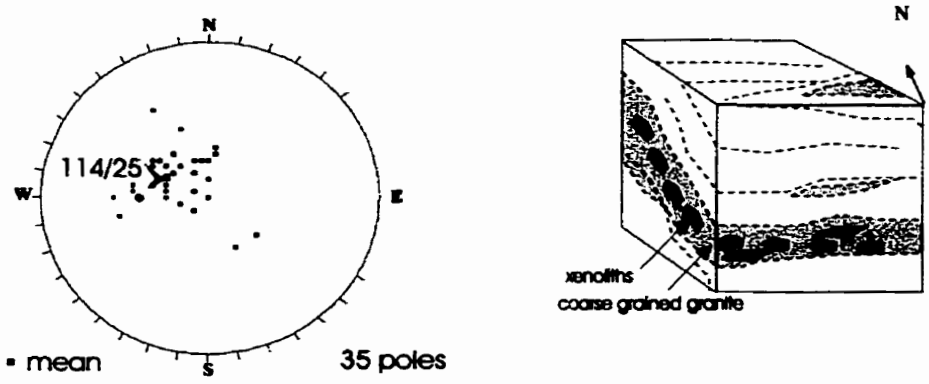
Points to note from the block models include the normal offset of the layering by the subvertical fine grained granite, the orientation of the subsidiary fine grained granite dykes relative to the layering, and the location and pattern of the pegmatites and the minor joints relative to fine grained granite. The pegmatites and joints at the 420 Level are largely confined to the fine grained granite.

The litho-structural domain boundaries and the small scale gneissic layering are thought to represent primary compositional layering in the batholith. The fine grained granite dykes are considered to be auto-intrusive (Everitt and Brown, 1986). Their orientation and general pattern of offset suggest at least local extension of the batholith roof. This may have been initiated by subsidence of the roof zone of the batholith, or by variations in the relative rate of intrusion. The subsidiary, low-dipping sets are thought to represent secondary shear fractures developed to accommodate differential movement within the dyke swarm.

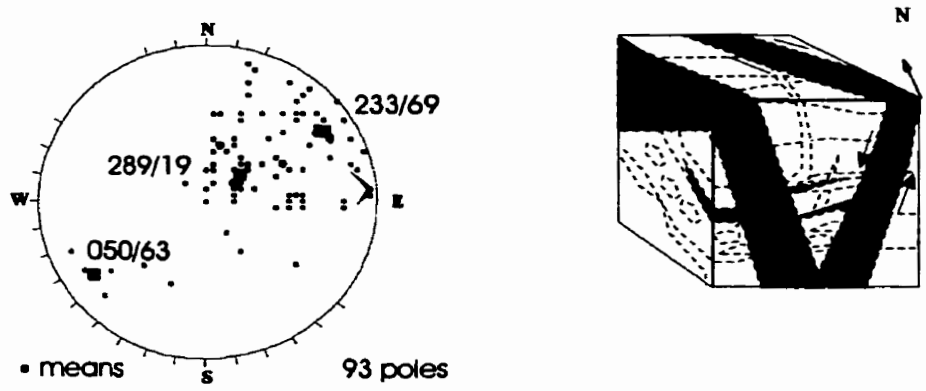
The pegmatite dykes, the quartz veins and the minor joints intersected in some boreholes are largely confined to the fine grained granite dykes, and are interpreted as successive generations of cooling fractures. For the joints, this interpretation is based on their parallelism to the fine grained granite dyke swarm in which they occur, the infilling assemblage of lineated quartz-chlorite-sericite-sulphide, and on the similarity of these fractures to other late- to post-magmatic fractures found elsewhere within the URL. The

**unoxidized sulphides on the fracture surface and the absence of wall rock alteration and detectable seepage support the conclusion that these fractures are closed.**

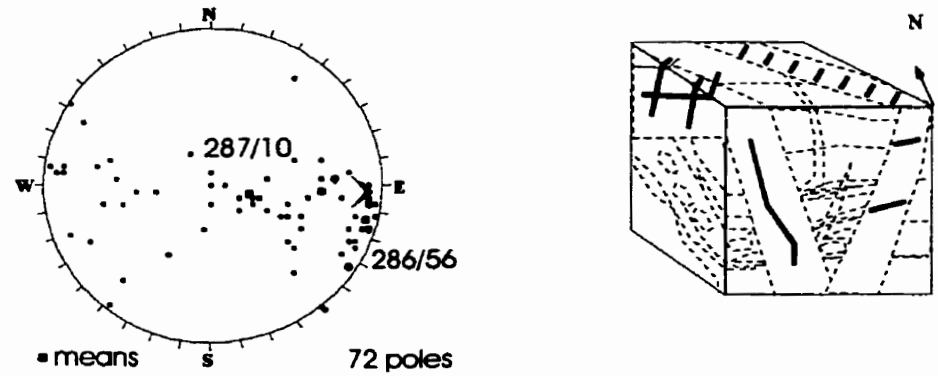
This page intentionally left blank for effective pagination.




A. Compositional and Textural (Gneissic) Layering

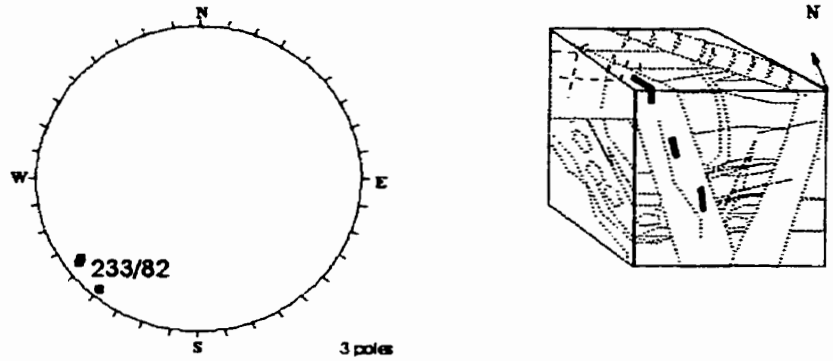



B. Fine Grained Granite Dykes

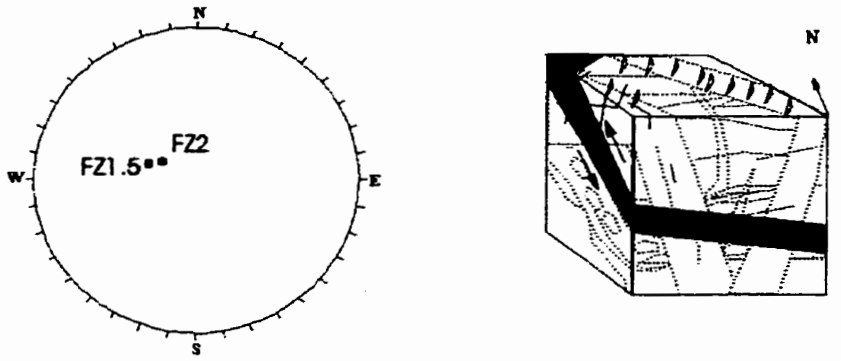



C. Pegmatite Dykes and Quartz Veins

Figure 15. A summary of fabric elements present at the 420 Level, listed in order of relative age. The orientations of each structure are summarized on the stereographic projections, and in the accompanying conceptual block models which also summarize the age relationships.




D. Quartz- Chlorite-filled fractures in the fine grained granite dykes




E. Low-dipping thrust faults (Fracture Zones 1.5 and 2)

Figure 15 (continued). A summary of fabric elements present at the 420 Level, listed in order of relative age. The orientations of each structure are summarized on the stereographic projections, and in the accompanying conceptual block models which also summarize the age relationships.

## 3.6. Faults and Fractures

The URL access shafts provide a cross section of the roof zone of the batholith, and of two low dipping thrust faults and associated splays. In terms of scale, fractures at the URL include:

1) kilometre-scale low-dipping thrust faults (Figures 16a and b),
2) subvertical joints, (Figure 16), and
3) natural (i.e. ancient) microcracks.

The natural microcracks are described separately in Chapter 4.

Shaft mapping has shown that the thrust faults are low-dipping, variously permeable horizons that contain both low-dipping and steeply dipping fractures. The thrust faults have chloritic slip surfaces which grade into complex cataclasite zones where fault movement has been significant (on the order of metres). The cataclasites consist of recrystallized fault rubble cemented by a fine grained chlorite-carbonate matrix, and they are crosscut by chloritic slip surfaces, minor fractures, and seams of soft clay - goethite gouge. This assemblage is in varying stages of groundwater-induced decomposition.

The thrust faults and splays follow the large scale compositional layering, and divide the rockmass into a number of tabular to wedge-shaped blocks. These blocks are crosscut by one or more sets of subvertical joints, the pattern and frequency of which varies from one block (fracture domain) to the next (Figure 16). The factors influencing the pattern of intrablock joints include:

1) the overall distance from the surface,
2) the proximity to the bounding faults, and

3) the local rock type.

With increasing depth, the subvertical joints become less frequent, less continuous, and with fewer preferred orientations. They also become increasingly confined to the immediate margins of the fault zones or to lithological heterogeneities such as dykes. These trends may reflect a general change from three-dimensional, to predominantly two dimensional, strain in the lowermost thrust sheets at the base of the sequence. The subvertical joints were interpreted as extensional intrablock fracturing generated during and after faulting (Everitt and Brown, 1986). Flexing of the fault blocks during thrusting is the most likely mechanism for their initiation. Reactivation may have occurred as a result of regional extension associated with subsequent cycles of uplift and erosion, and likely resulted in continued flexing of the thrust plates, reactivation of existing fractures, and the formation of new fractures.

Samples for this study were taken from intact rock below the lowest thrust fault, at the 420 Level (420 m depth from surface). Drilling and excavation at the 420 Level have exposed six quartz-chlorite filled fractures. All were dry and tight. They appear to be smaller-scale (and lower temperature) equivalents of the pegmatite-aplite dykes, formed by parting along the flow banding during cooling and contraction of the dykes. None were observed in the areas of the 420 Level selected for sampling.

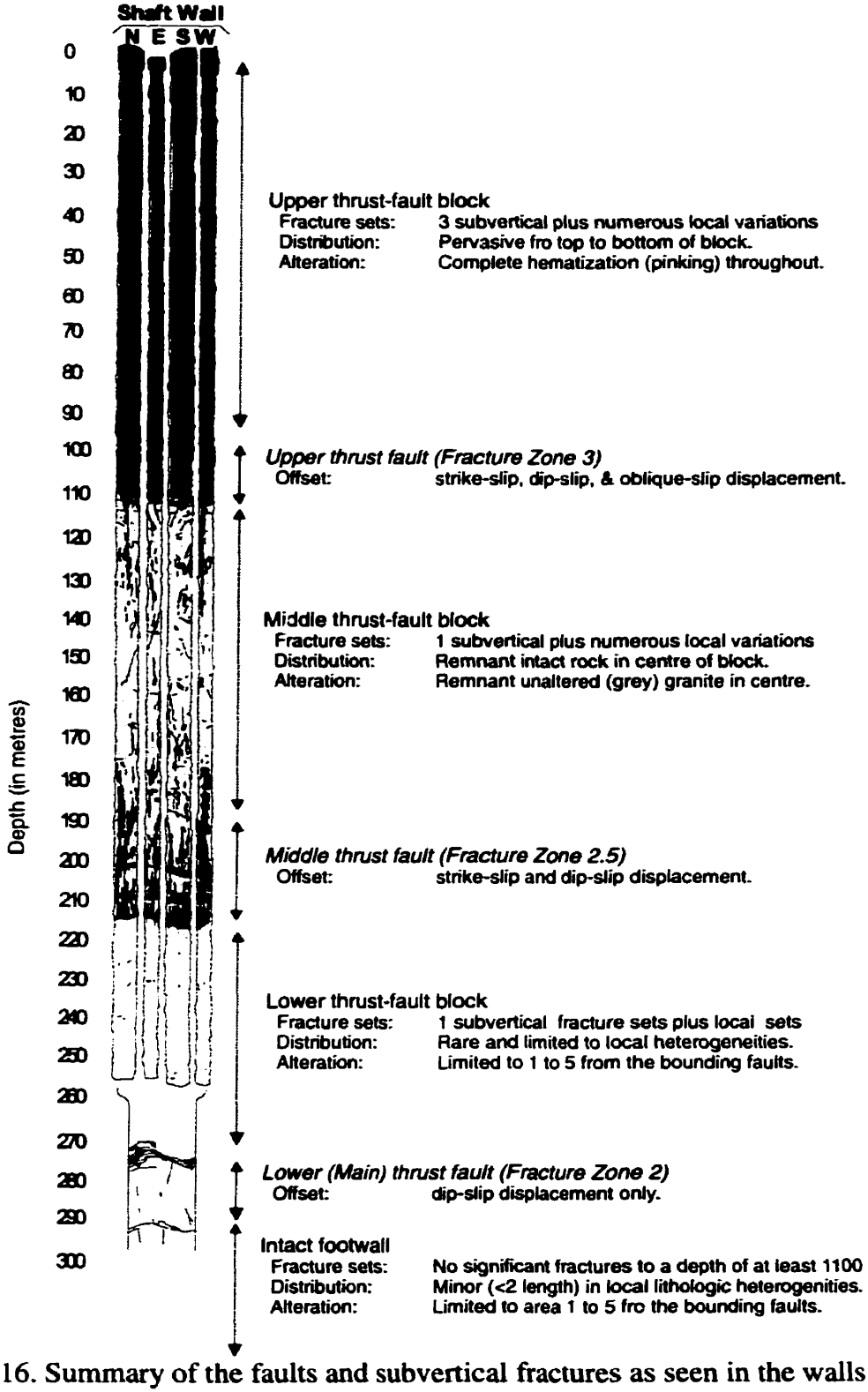



Figure 16. Summary of the faults and subvertical fractures as seen in the walls of the access shaft (the shaft from surface to 255m has a rectangular cross section, which changes to a circular cross section below this depth). The shaft below 290m is unfractured, and for this reason is not shown. Fracture Zone 2 is shown in greater detail in Figure 16 b following.

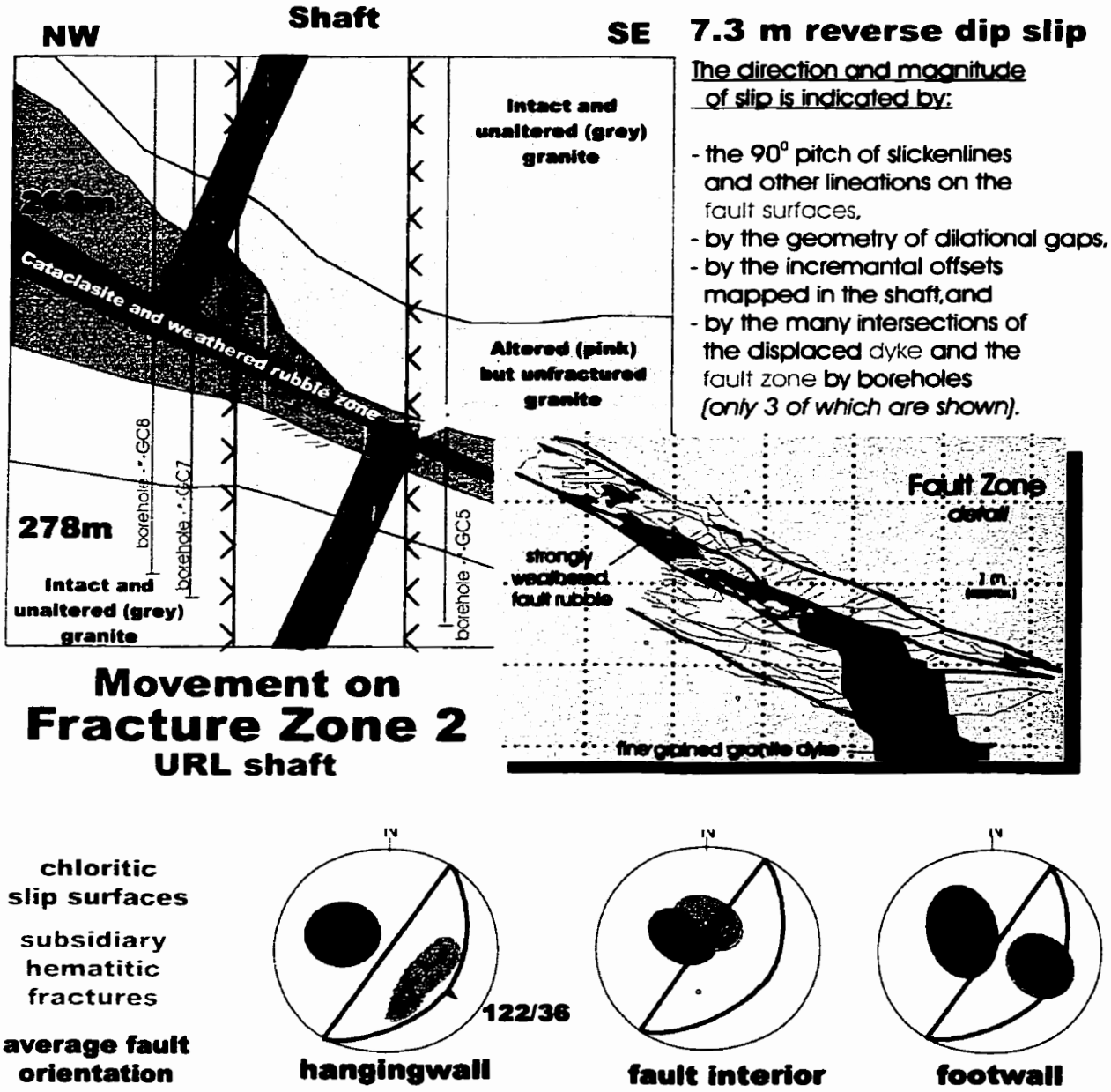



Figure 16b. Fracture Zone 2 as seen in the access shaft and its environs. The vertical section shows the offset determined from multiple borehole intersections and shaft mapping. The incremental nature of offset and the sense of movement is indicated in the inset (a portion of the shaft wall). Fracture orientations and the average fault orientation in the hangingwall, fault interior and footwall are shown in the stereographic projections. (modified from Everitt and Brown, 1996).

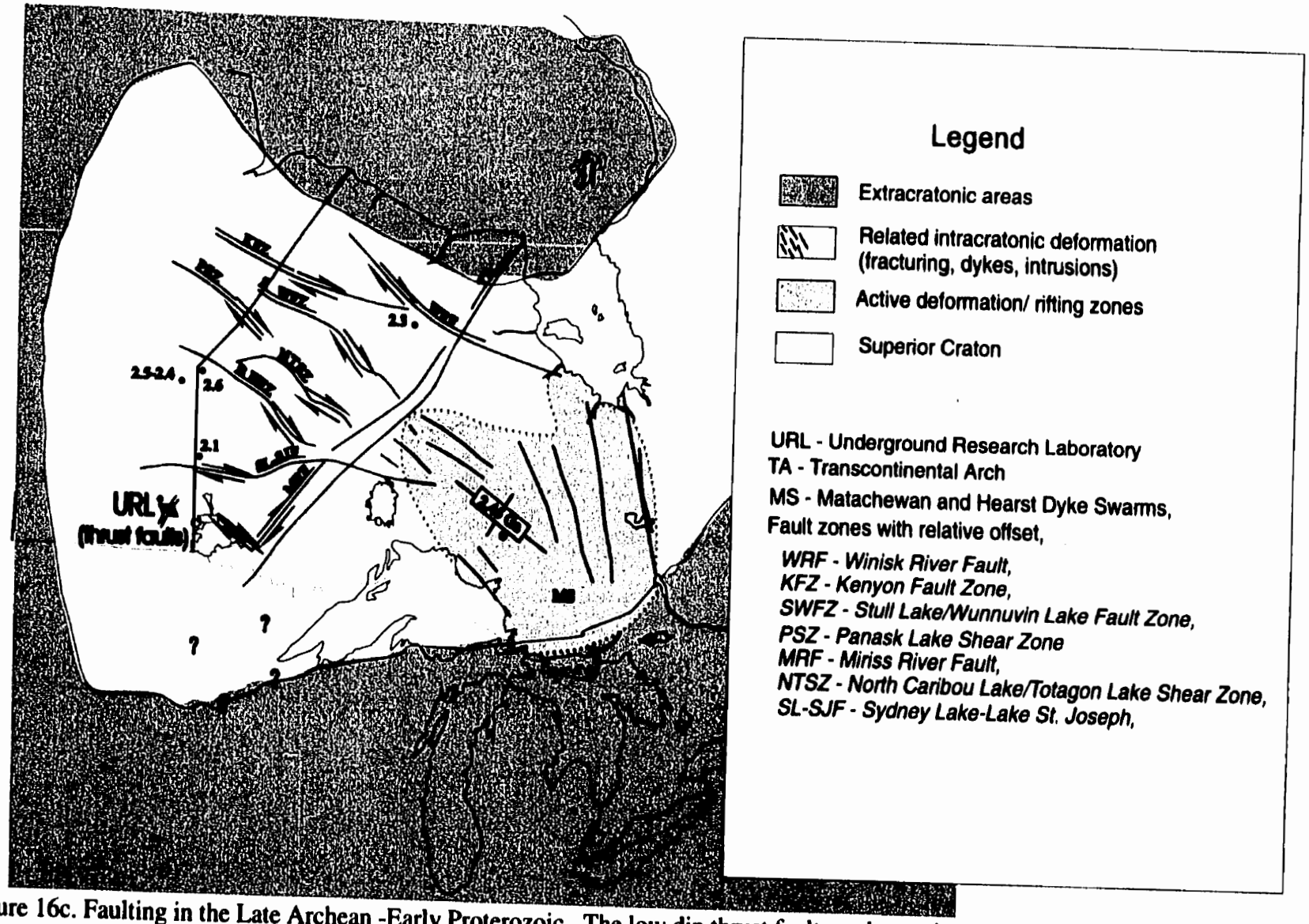



Figure 16c. Faulting in the Late Archean -Early Proterozoic. The low dip thrust faults and associated splays are believed to be associated with the craton-scale faulting shown (modified from Osmani 1991).

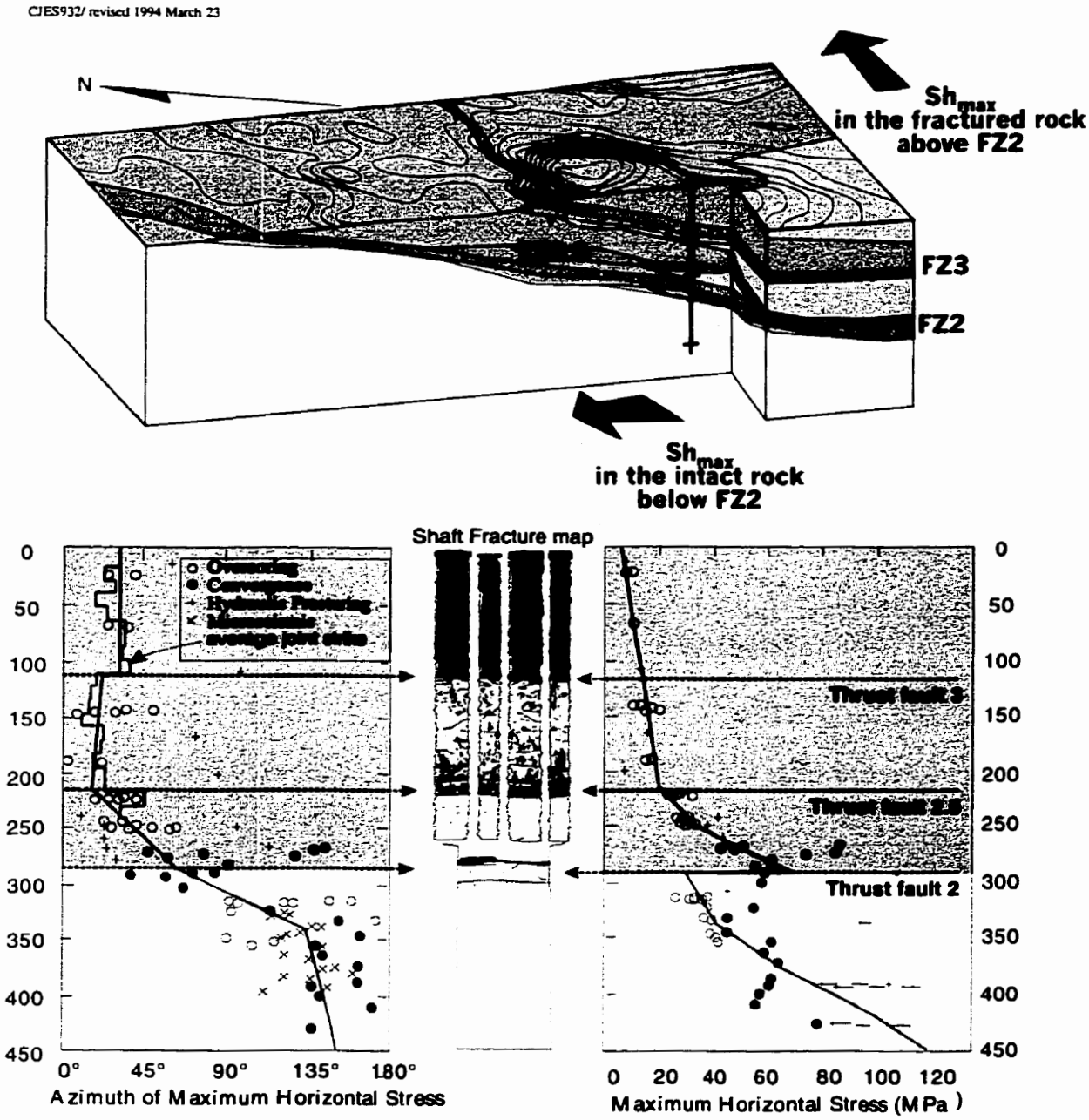



Figure 17. In-situ stress domains in the URL in relation to the fracture domains (stress data from Martin, 1989). The variations in average stress orientation and magnitude with depth are indicated by the dark green lines. For explanation refer to text.

## 3.7. In-situ Stresses

Figure 17 and Table 9 summarize the general orientation of principal stresses at the URL and their geometric relationships with the rock structure. The figure includes the following components:

A. a block model of the URL showing the thrust faults, the jointed rock between them and the intact rock below the main thrust fault (FZ2).

B. two charts depicting the variation in stress orientation and stress magnitude with depth, and

C. the shaft fracture map with the locations of the thrust faults, jointed rock, and intact rock highlighted.

The thrust faults, Fracture Zone 2 (FZ2) in particular, act as stress domain boundaries (Everitt et al., 1990; Martin, 1989b; 1990; 1993). Intact rock below the fault preserves high in-situ stresses, but these have been relieved above the fault by movement and fracture formation. The large arrows on the block model represent the average orientation of the maximum horizontal stress above and below the main thrust fault (FZ2). In the sparsely fractured rock below Fracture Zone 2, the maximum horizontal stress is oriented NW-SE, perpendicular to the strike of the thrust fault. In the fault block above Fracture Zone 2, the maximum principal stress is considerably lower and is oriented NE-SW, parallel to the dominant fracture set and the strike of Fracture Zone 2. This presumably represents a reduction and re-orientation of stresses through the formation of fractures in the hangingwall block. The footwall, including the area of the 420 Level has not been affected by the reorientation and reduction of stresses associated with the formation of the subvertical joints above Fracture Zone 2.

The thrust faults are believed to have formed during the late Archean - early Proterozoic in response to tectonism in the orogenic belt to the south (Figure 16c, also Everitt et al., 1990). Based on the geometry of the fault system, the intermediate and maximum principal stresses were subhorizontal and directed northeast-southwest and northwest-southeast, respectively. The present in-situ stresses at this level in the URL (i.e. in the intact rock beneath FZ2) retain this geometry, which supports the suggestion by Haimson (1990) for a domain within the mid-continent region in which the maximum horizontal principal stress is directed northwest-southeast. Whether this is a remnant stress field or a fortuitous alignment of recent and ancient stresses is unknown.

The in-situ stress measurements shown in this figure are a compilation of results obtained over the duration of the facility (Martin, 1989b; 1990; 1993). The measurement techniques have included:

1. convergence measurements
2. overcoring
3. hydraulic fracturing, and
4. microseismic monitoring

A brief explanation of these techniques, based on Chandler et al. (1996) and other unpublished material provided by Paul Thompson and Jason Martino of AECL follows.

### 3.7.1. Convergence Measurements

Convergence measurements are traditionally taken to obtain information related to the performance of underground support systems, especially in rock masses where the convergence is expected to be several centimetres or more. Convergence measurements in stiff rock masses are more difficult to accurately determine however, with the appropriate instrumentation and an understanding of the time dependent nature of the deformation, it is possible to measure convergence in unsupported rock mass, and to back calculate the horizontal in-situ stresses from the measured convergence. Martin et al. (1996) presents the analytical solution used to calculate the in-situ stresses from convergence measurements and also illustrates the role convergence measurements played in establishing the in-situ stress state at the URL when traditional testing methods, such as overcoring, were not fully successful.

The convergence of the circular shaft was measured at 26 locations between 267 and 430 m depths (Martino, 1989). Convergence pins were installed within 300 mm of the shaft floor at eight locations around the circumference of the shaft, facilitating convergence measurements across four diameters. The Kern Distometer used for taking convergence measurements had a repeatability of $\pm$ 0.1 mm, sufficient for the back analysis of in-situ stresses. The total convergence measured in the shaft typically varied between 1 and 3 mm, and was generally attained once the excavation face had advanced two or three shaft diameters (10 to 15 m) below the elevation of the convergence array.

The circular shaft geometry facilitated the calculation of in-situ stresses from the measured convergence values using a formulation proposed by Goodman (1989) based on the Kirsch equations for stresses and displacements around a circular opening (this was explained in detail in Martin et al., 1996). In-situ measurements of the displacements ahead of an advancing face suggest that 50-60% of the total displacement had already occurred ahead of the advancing excavation. These observations and numerical modelling supported a 60% correction factor for convergence measured in the shaft. Horizontal in-situ stresses calculated from shaft convergence measurements are shown in Figure 17.

### 3.7.2. Overcoring

Overcoring involves installing a strain-measuring instrument either at the flattened bottom of a borehole, or in a small-diameter pilot borehole drilled concentrically at the base of a larger hole then extending the original borehole past the instrument location (i.e. overcoring the instrument) usually by using a special thin-walled bit. For typical overcore tests at the URL, the large hole size is HQ (96-mm-diameter) and the pilot hole, if used, is EWG (38.7 mm diameter). Overcoring relieves the stresses acting on the cylinder of rock containing the instrument, thus inducing strains. The induced strains can then be used to back-calculate the stresses acting on the cylinder prior to overcoring if the stress-strain behaviour and properties of the rock are known. Depending on the strain measuring instrument, either the stresses in the two-dimensional plane orthogonal to the borehole, or the complete three-dimensional stress tensor, can be determined.

The overcore technique of stress measurement using borehole deformation or USBM (United States Bureau of Mines) gauges (Hooker and Bickel, 1974), triaxial strain cells (Leeman, 1967) or doorstopper gauges (Leeman, 1971) is well documented and thoroughly tested. Overcoring is used to determine the stresses in either the two-dimensional plane perpendicular to the borehole using USBM and doorstopper gauges, or the complete three-dimensional stress tensor using CSIR (Council for Scientific and Industrial Research) and SSPB (Swedish State Power Board) cells. All of these instruments have been used at the URL in the determination of in-situ stresses. The CSIR triaxial strain cell, in particular, provides consistent three dimensional stress information. In total, approximately 1000 overcore tests have been conducted at the URL, 350 of which are far-field triaxial stress measurements.

The greatest limitation to conducting overcore stress measurements at the URL is the high horizontal stress below Fracture Zone 2 which frequently induces discing of the core, making elastic interpretation of the results inappropriate.

The doorstopper (Leeman, 1971) is the name given to strain gauge rosettes that can be attached to the flattened end of a borehole and subsequently overcored. Doorstopper gauges have been used on the 420 Level of the URL when core discing makes it impossible to conduct USBM and triaxial cell tests. In theory, doorstopper strain measurements are unaffected by discing of the core. At the URL, the method of Corthesy et al., (1993) was used to determine non-linear and anisotropic elastic constants from doorstopper biaxial tests, and of applying these constants to the results from

overcoring in order to determine the in-situ stresses. This method has been applied to 13 doorstopper tests conducted in four boreholes on the 420 Level.

### 3.7.3. Hydraulic Fracturing

Hydraulic fracturing (hydrofracing) involves isolating a section of a borehole with a straddle packer system and creating fractures by injecting fluid into the zone at a pressure high enough to fracture the rock (Wijesinghe, 1987). The pressure required to fracture the rock (i.e., the breakdown pressure) and the "shut in" pressure or pressure required to re-open the fracture are used to calculate the rock stresses in the plane orthogonal to the borehole axis. The initial breakdown pressure is theoretically a function of the tensile strength, and the near-field tangential stress acting around the borehole.

In hydraulic fracturing, the straddle packer is lowered to the testing location and the packers inflated. Packer pressures are maintained slightly above the pressure in the zone isolated between the tow packers. The zone pressure is increased until the measured pressure drops suddenly indicating that the fluid is lost into a newly created pressure ($P_b$). A break in the plot of pressure versus time is often referred to as the "shut in pressure" ($P_{si}$). The shut-in pressure is often determined using a variety of graphical techniques and is believed to correspond to closure of the fracture. More recently, a slow repressurization test has been used to determine at what pressure the fracture reopens. In most cases, the reopening pressure is very nearly the same as the shut-in pressure but is generally more accurately determined.

The equations for maximum and minimum stress in the plane perpendicular to the borehole ($\sigma_H$ and $\sigma_h$) are as follows:

$$\sigma_H = 3\,\sigma_h + T - P_b - P_0$$

and

$$\sigma_h = P_{si}$$

where T is the tensile strength of the rock and $P_0$ is the ambient pore pressure. The rock pore pressure is often taken to be the water pressure in the fluid filled borehole at the test location, since the borehole provides drainage to the rock where the fracture will initiate.

It is important to note that these equations assume:

1. that one of the principal stresses acts parallel with the axis of the borehole,
2. that the plane of the induced fracture is parallel with the axis of the borehole and propagates in the direction of the maximum principal stress,
3. That the rock mass tested is mechanically isotropic.

As is shown in this thesis, the last assumption does not correlate with either the mapping results done during excavation (this chapter), or with the results of the rock properties testing program presented in the following chapters.

Six series of hydrofrac test have been conducted at the URL in twelve different boreholes. Hydraulic fracturing was conducted in 1981 and 1982 prior to shaft sinking and was the first stress determination used at the URL. Nine tests were conducted in the

borehole URL-1 from 50 to 600 m depth (Haimson, 1982) and a second series of twelve tests was conducted in borehole URL-6 (Bawden, 1982). A third series of fifteen tests was done from the 240 Level in a subvertical probe hole (206-020-PH1) drilled down form the 240 Level (Doe, 1987). A fourth series of tests was conducted in three horizontal boreholes drilled form the 240 Level and two horizontal boreholes drilled from the 420 Level (Doe, 1989). A fifth series of fourteen test was done in borehole 401-009-HF1 drilled downwards form the 420 Level of the URL (Haimson, 1991) and a sixth series was conducted in three holes oriented in three different direction on the 420 Level (Haimson, 1992).

The tests produced consistent results at depths less than 300m, the depth corresponding to the predominant fracture zone (FZ2). Although tests were conducted both above and below FZ2, the induced hydraulic fractures below the fracture zone were either subhorizontal or could not be induced within the limits of the testing equipment. Evidence from the shut-in pressures was sufficient to conclude that the fracture zone acted as a stress domain boundary (i.e. a sharp transition from moderate to high in-situ stress magnitudes). The horizontal stress magnitudes from the hydraulic fracture tests conducted above FZ2 were later confirmed by overcoring stress measurements. The two stress determination methods provided very similar results for stress magnitudes (Figure 17). However, there is a 30° to 60° discrepancy in the azimuth of the major horizontal stress between the two methods. This difference has been attributed to either the natural variation of preferred joint directions in the rock leading to variation in the direction of induced hydraulic fractures, or to the existence of magnetite in the granite at

the URL which may affect the compass directions recorded on impression packers. The data and observations presented in this thesis support the first interpretation – that is, control of fracture orientation by the rock fabric. Evidence of this was first obtained in the first application of hydrofracing at the URL in 1983. Hydrofracing was conducted from a pilot hole centered on the axis of the shaft prior to excavation. The hydrofrac was stained by a dye. Careful mapping of the shaft floor and walls was then done, as the excavation of the shaft provided a rare opportunity to observe the detailed geometry of a hydrofracture (unpublished mapping by the author and A. Brown, 1984). The exposures provided clear and conclusive evidence that the rockmass was not isotropic and that the orientation of the hydrofracture was not determined solely by the in-situ stress orientations. Instead, the hydrofracture followed narrow (<10 cm) pegmatite veins, alternating between vertical and horizontal whenever the veins changed orientation. It should be added, that these veins were tightly sealed and closed surfaces before the hydrofracing.

The presence of these veins and their structural control was not realized at the time of the hydrofrac tests. The evidence for fabric control invalidates several assumptions under which the test is based – namely – the fracture orientation is controlled by the in-situ stress, and that the rock mass is isotropic.

### 3.7.4. Microseismic Monitoring

Microseismic monitoring requires the use of arrays of triaxial accelerometers and / or geophones to record the low energy seismic waves from small scale cracking and microcracking around tunnels. The system is used to identify damage as it occurs during

excavation. This damage is related to stress redistribution around the excavations (Read and Martino, 1996). AE/MS systems have been used at the URL during shaft extension and Mine-by Experiment.

Accelerometers were installed in boreholes collared from the 300 Level and inclined downwards 45° around the shaft extension Young and Talebi (1989). The system was used to monitor events during shaft excavation between 324 and 443 m. In a horizontal plane, most of the microseismic activity was located within 1 to 2 m of the shaft walls and in some cases showed a preferred orientation similar to the zones of microcracking mapped at the 300 Level (Everitt et al., 1988). The events are concentrated in the zone of maximum compression, and indicate a northwest-southeast maximum horizontal stress orientation. This orientation agrees with observations of shaft wall notching, core discing and with results from convergence measurements.

Another array of sixteen accelerometers was installed around the Mine-by tunnel (Talebi and Young, 1990) also showed a concentration of events (Figure 46) in the orientation perpendicular to the maximum stress direction Talebi and Young (1992). The AE/MS system has been shown to provide valuable information regarding the orientation of in-situ stresses in a moderate to high stress environment.

In summary, the convergence measurements provide information on the deformation resulting from excavation, from which the orientation and approximate magnitude of the in situ stresses can be inferred. The microseismic monitoring during excavation identifies regions of high compressive stress concentrations which, in turn, can be used to

infer in situ stress directions. The hydrofracing method provides both the stress orientations and magnitudes, but is susceptible to spurious results resulting from fabric control of hydrofrac development. The overcoring methods have provided the best estimates of in-situ stress orientation and magnitude, however only in intermediate to low stress rock not affected by core discing.

Table 9

In-situ stresses measured at the 240 Level.

(Martin and Kozak, 1992)

| | Magnitude | trend | plunge |
|---|---|---|---|
| $\sigma_1$ | 25.5 MPa | 228° | 8° |
| $\sigma_2$ | 16.8 MPa | 135° | 23° |
| $\sigma_3$ | 12.8 MPa | 335° | 65° |

In-situ stresses measured at the 420 Level.

(Martin, 1993)

| | Magnitude | trend | plunge |
|---|---|---|---|
| $\sigma_1$ | 55± 3 MPa | 135° | 14° |
| $\sigma_2$ | 48± 3 MPa | 47° | 8° |
| $\sigma_3$ | 14± 2 MPa | 73° | 74° |

# Chapter 4

# NATURAL MICROCRACKS

Natural microfractures in quartz and other minerals are visible on close inspection of the coarser grained phases of the Lac du Bonnet granite (Figure 18), and are common in thin section. This chapter provides a compilation of the author's work on natural microfractures in quartz at the URL. The objectives of this work were as follows:

1. Are there preferred orientations or patterns of development for the natural microfractures?
2. Does their frequency and development vary with depth and in a manner consistent with the domains of large-scale fractures and the in-situ stresses?
3. Can microfracture orientations be used to infer the orientations of the larger scale fractures in areas of sparsely fractured rock? This application is considered to be most valuable during the initial stages of site investigation, as it would provide information on which to base the planning of exploration drilling and in the placement of in-situ stress measurements.
4. Are the natural microcracks responsible for the certain phenomena at the tunnel scale, such as the preferred development of excavation-induced fractures in certain orientations?

5. Can natural microcracks be responsible for strength anisotropies identified by rock properties testing (Point Load, Brazilian and uniaxial compression tests)?

This chapter presents the variations in microfracture orientation with depth at the URL, and their spatial relationship to the domains of macroscopic fractures and in-situ stresses. This study was undertaken in three stages: 1) an initial reconnaissance survey, 2) a more detailed reconnaissance survey, and 3) a study of samples from the 420 Level only to correlate with the samples taken for geomechanical testing. The data is presented in Everitt (2000).

Table 10

Microfracture types

| | |
|---|---|
| 1) | Natural microfractures which are dependent on some pre-existing fabric anisotropy, which includes: |
| | a) microfractures developed along grain boundaries, or |
| | b) microfractures formed by parting along crystallographic directions such as the cleavages in feldspars or biotite, and |
| 2) | Natural microfractures that are independent of some pre-existing fabric anisotropy, the most significant of which are those developed in isotropic minerals such as quartz. |

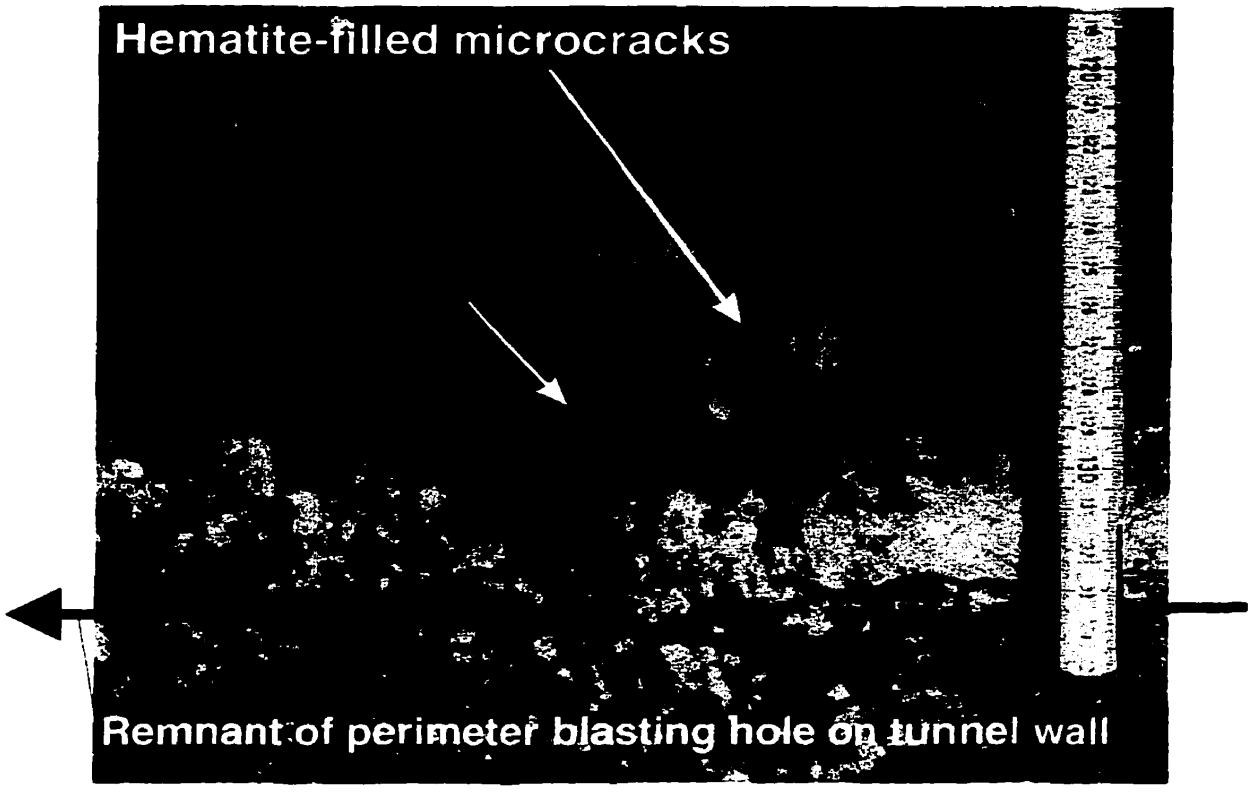



Figure 18. Examples of microcrack varieties as seen in a coarse grained phase of the Lac Du Bonnet Batholith, exposed on a tunnel wall at the 240 Level of the URL. The whitish linear groove running the length of the photograph is the remnant of a percussion-drilled perimeter blasthole as seen on the tunnel wall. The rock immediately adjacent to the hole appears whiter than the rock surrounding it due to abundant blast-induced microfractures. A much larger but related (20 cm long) excavation induced "microfracture" runs parallel to the blasthole. These are easily distinguished from natural microcracks by the presence of mineral infilling or alteration, as in the prominent hematite -stained microcrack in the large microcline crystal near the centre of the photograph.

## 4.1. Terminology

For the purposes of this report, microfractures are defined as grain scale discontinuities. Their importance lies in their influence on rock mass response, whether in the lab during testing, or in the excavations themselves. Preferred directions of natural microcracking have long been recognized and utilized in the quarrying industry. Two general types of natural microfractures can be distinguished, based on their mode of occurrence. These are listed in Table 10.

It is not uncommon for each of these microfracture types to exist together in any given sample, and each or all may be the result of both ancient and recent processes (Figure 18). Stress relief and blast damage in the proximity of an excavation, drilling, and sample preparation tends to accentuate the fabric-related fractures, but can also lead to the formation of new transgranular or intragranular microfractures (including both types 1 and 2). These induced microfractures, along with any changes to the type 1 microfractures, are considered sample damage.

Ancient or natural microfractures are related to geologic events, including the development of the larger scale faults and fractures, and tend to be penetrative (pervasive or statistically homogeneous) through large portions of the rock volume. As such, they are associated with consistent patterns of rock mass response, such as the rift, grain and hardway described in the dimension stone industry.

In a report on exploration targets for the dimension stone industry in Ontario, Marmont (1993) noted that even the most homogeneous and uniform granites usually possess a

microscopic structure that predisposes them to break more easily in one direction than another. This is usually caused by weak alignment of elongate mineral grains, and/or by sets of pervasive natural microcracks, either intergranular, or transgranular (Richter and Simmons, 1977). The direction along which a granite splits most easily is the rift, the next easiest in the grain, while the direction along which the rock is hardest to split is the hard way. The rift is commonly horizontal or vertical, the grain is usually perpendicular to the rift, and the hard way perpendicular to rift and grain (Marmont, 1993). These directions commonly correlate with batholith structure, with the rift in dome-shaped granite intrusions being parallel to the granite's contact with the enclosing rocks.

Experienced stone workers can see or feel the rift and grain, and plan their quarrying and processing accordingly. Stone cut parallel to the rift will be stronger than stone cut across it. Determining the existence and orientation of these natural anisotropies is considered critical in large architectural panels (Marmont, 1993; Stecich et al., 1992).

Of the natural microfractures types recognized, those developed in quartz are the most widely recognized and studied (Richter and Simmons, 1977). They are thought to have originated during cooling (or heating) in the presence of fluids and a deviatoric stress field (Jang et al., 1989). Their development in an isotropic mineral, their pervasiveness on the hand- specimen- scale, and their tendency to occur as sets related geometrically to the causative stresses, make microfractures in quartz useful fabric elements with which to interpret the stress and fracture history of their host. Their use in determining regional

stress orientations related to specific orogenic events have been described by Jang et al. (1989), and Ren et al. (1989).

The natural microfractures in quartz are also the type most easily distinguished from sample damage. In this study (and in general), natural microfractures were distinguished from induced microfractures by the presence of mineral filling or alteration, or by trains of fluid and mineral inclusions. The predominant filling mineral in the URL samples appears to be hematite, but in most instances the filling and fluid inclusion compositions were not optically determinable. Open (i.e. induced microcracks) were abundant in the URL samples. They were usually curviplanar to conchoidal, transgranular, and had no obvious pattern. The filled microfractures in contrast, were usually intragranular, planar and generally displayed some type of repeating pattern on the thin section scale.

## 4.2. Scope of Work

This study of natural microfractures in the Lac Du Bonnet granite at the URL was done in three stages to accommodate funding and construction schedules, and included:

1. an initial reconnaissance survey;
2. a more detailed reconnaissance survey,
3. a study of samples from the 420 Level only, to correlate with the samples taken for geomechanical testing.

In stages 1 and 2, the purpose of sampling was to identify the variation in microfracture pattern, with depth and location relative to the major faults and the large-scale fracture domains. Samples were obtained from each of the following structurally-defined blocks

(as shown previously in Figures 18 and 19, Chapter 3), each of which is bound by a thrust fault, and is represented by considerable differences in the number, size and spatial distribution of outcrop-scale fractures).

- Fracture Domain A: (the pervasively fractured rock between the ground surface and extending to the base of the uppermost thrust fault - Fracture Zone 3).
- Fracture Domains B and C: (the moderately to sparsely fractured rock between Fracture Zones 3, 2.5 and 2), and
- Fracture Domain D: the sparsely fractured rock below FZ2.

The results of the three stages of microcrack study are presented in Everitt (2000). The methodology used and an overview of the results are presented in the following sections.

## 4.3. Sampling

All samples were taken from oriented core, in areas of homogeneous granite away from the low-dipping thrust faults and their alteration halo, and any subvertical fracture zones. Three mutually perpendicular, polished thick sections (100 μm) were made from each specimen, one perpendicular and two parallel to the core axis.

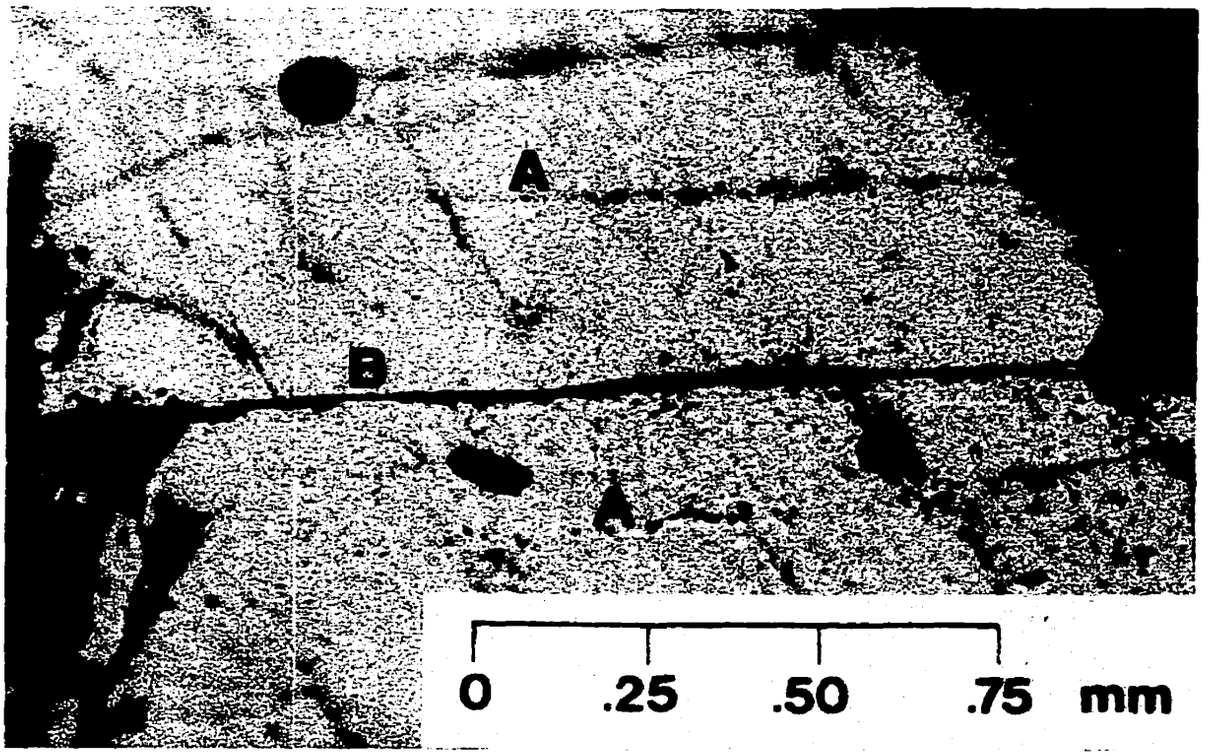



Figure 19. Photomicrograph of a single hematite-filled natural microcrack in quartz. Photograph by C. McGregor. A – Sealed microcrack with fluid and mineral inclusions. B – open microcrack with hematite infilling.

## 4.4. Natural vs. induced microcracks

Natural microcracks were distinguished from induced microcracks by the presence of mineral filling or alteration, or by trains of fluid and mineral inclusions. A single, hematite filled natural microcrack in quartz is shown in Figure 19. The predominant filling mineral appeared to be hematite, but in most instances the filling and fluid inclusion compositions were not optically determinable.

Open cracks with no visible wall rock alteration or filling were abundant, but were not measured to avoid grouping natural microcracks related to the stress relief history of the

rock, with others probably induced by excavation, drilling or sample preparation. Open (i.e. induced microcracks) were usually curviplanar to conchoidal, transgranular, and had no obvious pattern. The filled microcracks in contrast, were usually planar and generally displayed some type of repeating pattern on the thin section scale.

## 4.5. Microcrack Distribution

Microfracture homogeneity on the sample-scale was investigated by systematic mapping of three mutually perpendicular thin sections (Figure 20). Photomicrographs, taken at low magnification were used as a base on which were plotted the grain boundaries, and the location and identifying letter of each measured microcrack.

The mapping was undertaken to determine if there was any indication that the pattern of natural microcracks in quartz was spatially and geometrically related to any textural heterogeneity at the specimen-scale (such as a radial pattern around feldspar phenocrysts or xenoliths). No such evidence was found, and it is concluded that the natural microcracks are pervasive, reflecting the site-scale paleo-strain, and not local-scale textural anomalies.

Quartz is typically interstitial and allotriomorphic in all thin sections. Most of the quartz shows undulose extinction (i.e., the crystal lattice of the quartz is strained). If this strain is due to one or more episodes of cooling and/ or tectonic stresses as described by West (1991), it has not been relieved by recrystallization. This undoubtedly contributes to the

strains measured in rock properties testing at the URL, but the relative magnitude of this influence is not known.

Microfractures were present in both strained and unstrained grains at all levels, and crossed the partially recrystallized margins (sub-grain fabric) of some strongly undulose quartz. This superficially resembled transgranular cracking, but the microfractures are limited by the original grain boundaries.

This spe... ... ...

Figure 20. Map of natural microcracks (hematite or fluid inclusion-filled planes) within a sample from the 420 Level.

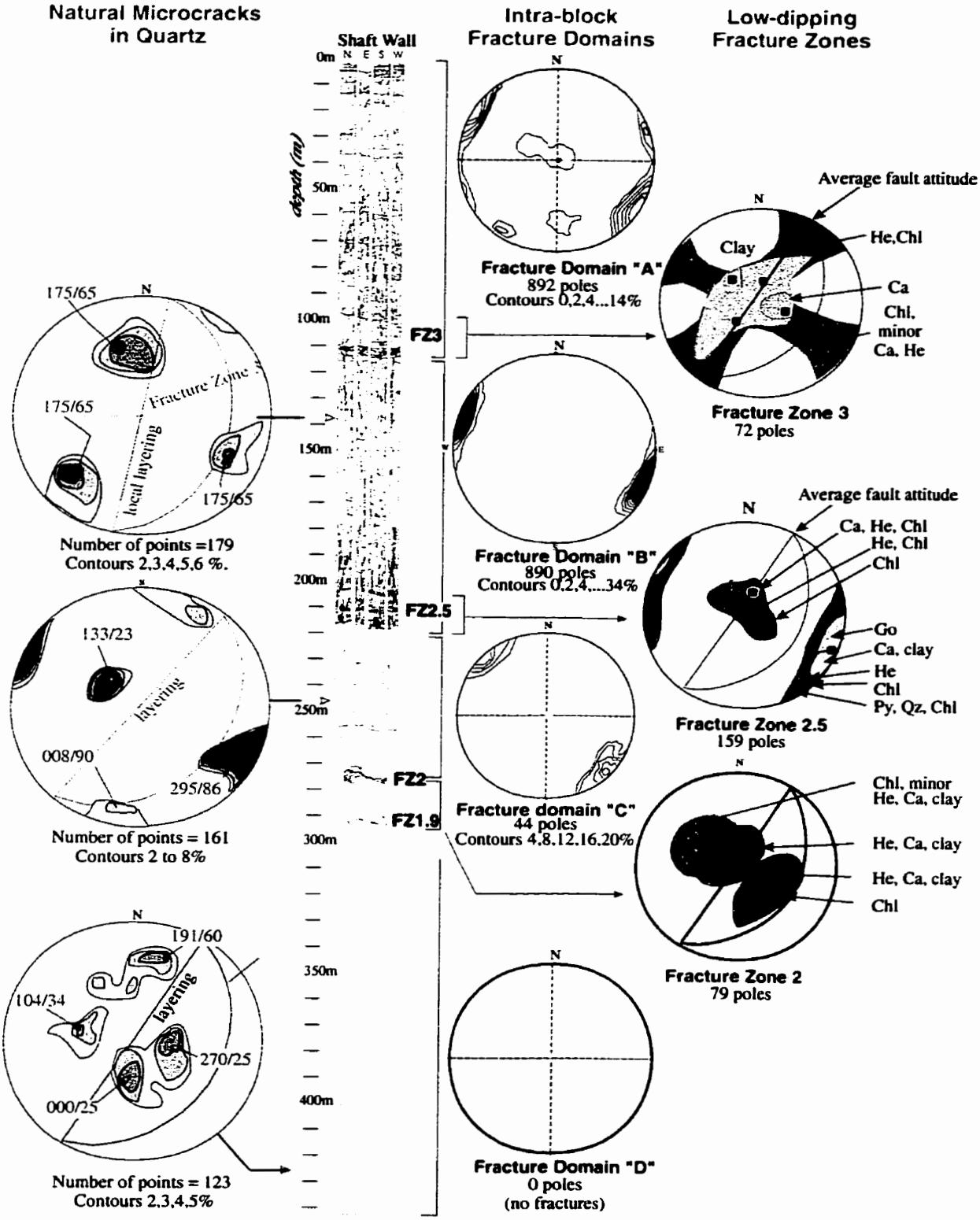



Figure 21. Summary stereographic projections of poles to fracture and microcracks sets from the 130, 240 and 420 Levels. • - average pole to local fracture set in low-dipping fracture zones. Infilling types: He –hematite, Chl – chlorite, Go – goethite, Ca – carbonate.

## 4.6. Natural Microcrack Sets

The orientations of the natural microfractures were then determined along with the orientations of quartz c-axes using a 5 - axis Universal stage, mounted on a petrographic microscope, at a magnification of 10x as described by Jang et al.(1989), and Rousell (1981). The relative orientations measured in each section were converted to true orientations, and these were plotted on summary stereographic projections for each sample.

The results of this study indicate that the pattern and domains of microcracking are consistent with the patterns and domains previously defined for the macro-scale fractures at this site (Figure 21). The microcrack sets at each level correlate well with the sets of macro-scale fractures. This includes a well developed microcrack pattern above Fracture Zone 2, and a weak and predominantly low-dipping pattern below (Figure 22).

Microcracks in the samples from the 420 Level are weakly developed. They are predominantly low dipping, and very similar in orientation to the major low-dipping fault zones. Microfractures are more abundant above Fracture Zone 2 (e.g. from the 240 and 130 Levels), with a well defined low dipping set and two subvertical sets.

The low-dipping set in the sample from the 240 Level parallels the average orientation for Fracture Zone 2, as well as the compositional layering, which this fault follows. The low dipping set in the sample from the 130 Level parallels the orientation for Fracture Zone 3 (based on an average taken over a 500 $m^2$ area), but not the local compositional layering, which deviates here from the regional norm. This is consistent with other URL

surface work, which showed that the low-dipping fracture zones are fabric-controlled where the layering is near the theoretical failure plane, but not where the layering varies substantially in orientation (Brown et al., 1989a).

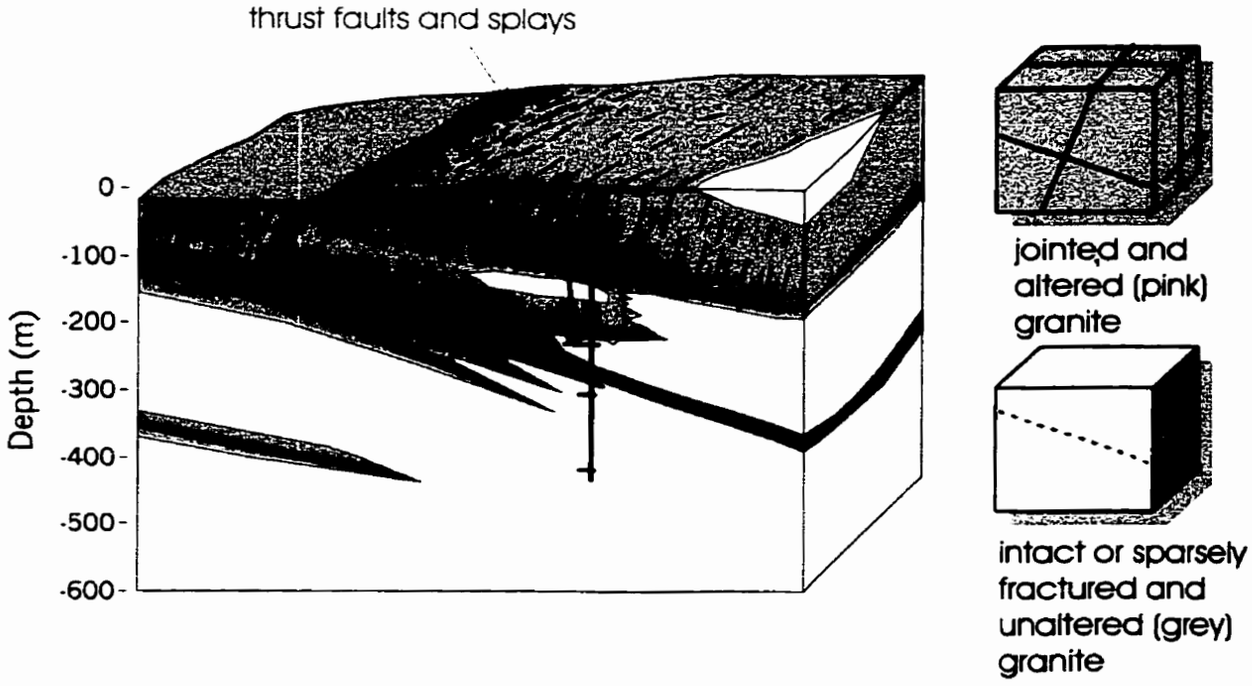



Figure 22. Schematic illustration of the sets of natural microcracks in quartz at the URL.

## 4.7. Natural Microcrack Length

Average microcrack lengths are dependent on depth from surface and average grain size. A summary of natural microcrack lengths for samples ranging from surface to the 420 Level is shown in Figure 23. Mean crack length decreases noticeably with increasing depth from surface. In the moderately fractured rock above Fracture Zone 2, microcracks equal or exceed the grain size, while in the sparsely fractured rock at and below the 240 Level, the microcracks are usually less than the grain size. Microcrack apertures are also dependent on depth and sample disturbance (stress relief) (Chernis, 1981; 1984).

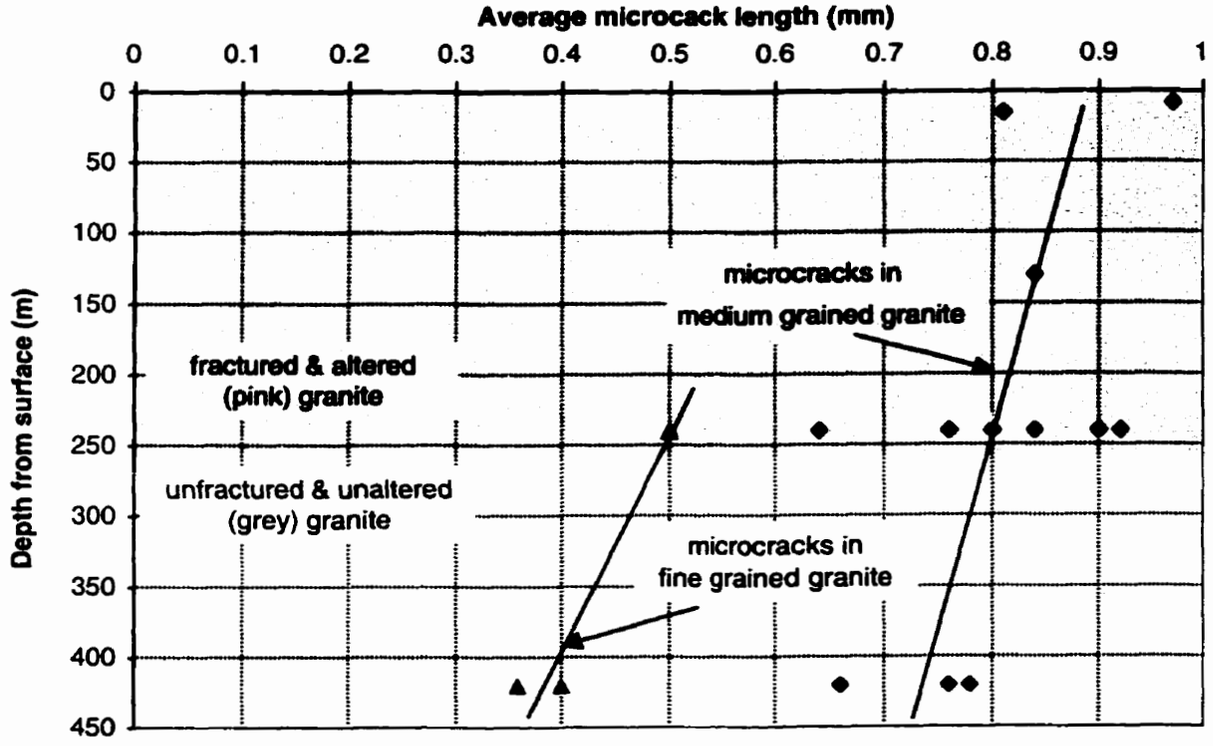



Figure 23. Variation with depth of the average length of natural microfractures in quartz.

## 4.8. Conclusions of Microfracture Study

There are preferred orientations for the ancient natural microfractures at the URL, which are consistent with the patterns and domains previously defined for the macroscopic-scale fractures at this site. This suggests that the microfracture orientations could have been used to provide estimates of the likely orientations of larger scale fractures, not intersected by boreholes, in the sparsely fractured rock at this site. Changes in the overall geometry of microcracks, such as exists on opposite sides of Fracture Zone 2, indicate the existence of different fracture domains.

Natural microfractures in quartz are pervasive at the sample scale, even in areas of sparsely fractured rock, and as such must be considered in assessing sample response to testing. Correlations between microcrack orientation, sample response to testing, and excavation damage at the tunnel-scale exist, and are described in the following chapters.

# Chapter 5

# FIELD OBSERVATIONS ON THE INFLUENCE OF GEOLOGY ON FRACTURE DEVELOPMENT

## 5.1. Evidence of Fabric Control of Natural and Induced Fractures at the URL

The present-day response of an intact rockmass to stress can be gauged by its history of deformation. Fabric control of fractures, for example, is an indication that the rockmass has reacted as an anisotropic medium when previously subjected to brittle deformation. Fabric control refers to the preferential development of fractures, in certain locations and (or) directions, due to the presence of mechanical anisotropy in the rockmass. This anisotropy may be induced by the presence of compositional and textural layering, and by the development of secondary structures such as foliations and lineations. This phenomenon is widely recognized and anticipated in sediments and metamorphic rocks with obvious structure (e.g. Reitan and Murphy, 1983), but is often overlooked in rocks such as granite where the fabric is far more subtle.

The rock mass response in "apparently homogeneous" rocks can be defined thorough a (costly) program of drilling, mechanical testing, and modelling. However, before this is undertaken, physical evidence of a mechanical anisotropy in a rockmass and of its

geometry, can be obtained from an examination of the distribution of the ancient, natural fractures and natural microfractures.

Detailed mapping of the URL excavations, the shaft in particular, and borehole data have clearly indicated that textural and compositional heterogeneities control:

1) the location of large-scale fault zones,
2) the extent, distribution and infilling characteristics of lesser-scale fractures between the fault zones, and
3) the extent and density of excavation-induced fractures (Everitt and Brown, 1986; Everitt et al., 1996).

It has also been demonstrated, in a qualitative manner, that crack initiation and propagation stresses are influenced by rock fabric and grain size. Breakout development for example has been observed to vary in response to textural and structural variations in the granite present at the tunnel- and site-scales (Everitt et al., 1996, and in the sections following). In particular, at compressive stress concentrations the breakout is far more extensive in the coarser grained variants, especially where the foliation coincides with the direction of expected slabbing.

The following sections describe examples where even subtle fabric in the "homogeneous" granite of the LDBB has influenced both natural and excavation-induced fractures at a variety of scales and excavation orientations.

## 5.2. Fabric Control of Large Subhorizontal Faults

The URL access shafts provide a cross section of the roof zone of the Lac Du Bonnet Batholith, and of a system of low dipping thrust faults and associated splays, each extending more than a kilometre along strike, and spaced in the order of 50 to 200 m apart. Detailed mapping by the author has revealed that:

- these faults follow the large scale compositional layering in the roof zone of the batholith, which in turn is defined by the batholith shape. The similarity between the orientation of the fault, and of the floor of the batholith, is obvious (Figure 24).
- that each of the three faults and lesser splays are confined to the more heterogeneous layers (e.g. as shown in Figure 25, Fracture Zone 2 (a low-dipping thrust fault), has developed within a one to two metre thick "plane" of amphibolitic inclusions and biotitic schlieren. The surrounding granite is more weakly layered or massive, and is not fractured. Site drilling has determined that this fault continues to follow this layer over a kilometre from the shaft, along both the dip and strike directions.

This page intentionally left blank for double pagination

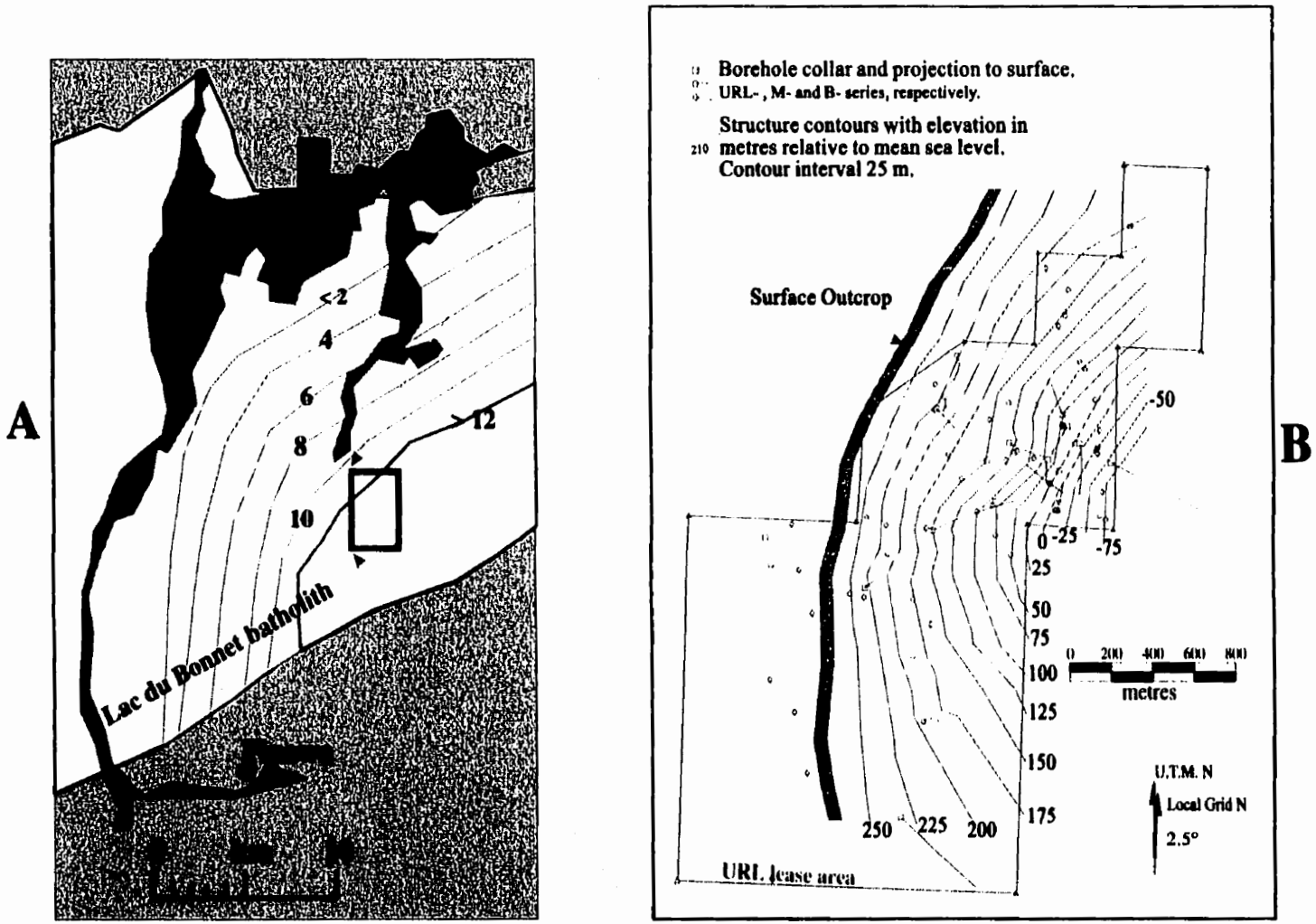



Figure 24. Influence of batholith contacts and internal structure on fault orientation. Structure contours showing the depth (in km) to the base of the Lac du Bonnet Batholith are shown in (A) (as interpreted from profiles appearing in Tomsons et al. (1995) based in turn on gravity and magnetic data). Dekametre-scale layering within the pluton varies locally in orientation, but is broadly parallel to the lower contact. The low-dipping fault zones (e.g. Fracture Zone 2) are controlled in location and orientation by the layering. Structure contours on the lower contact of Fracture Zone 2 are shown in B.

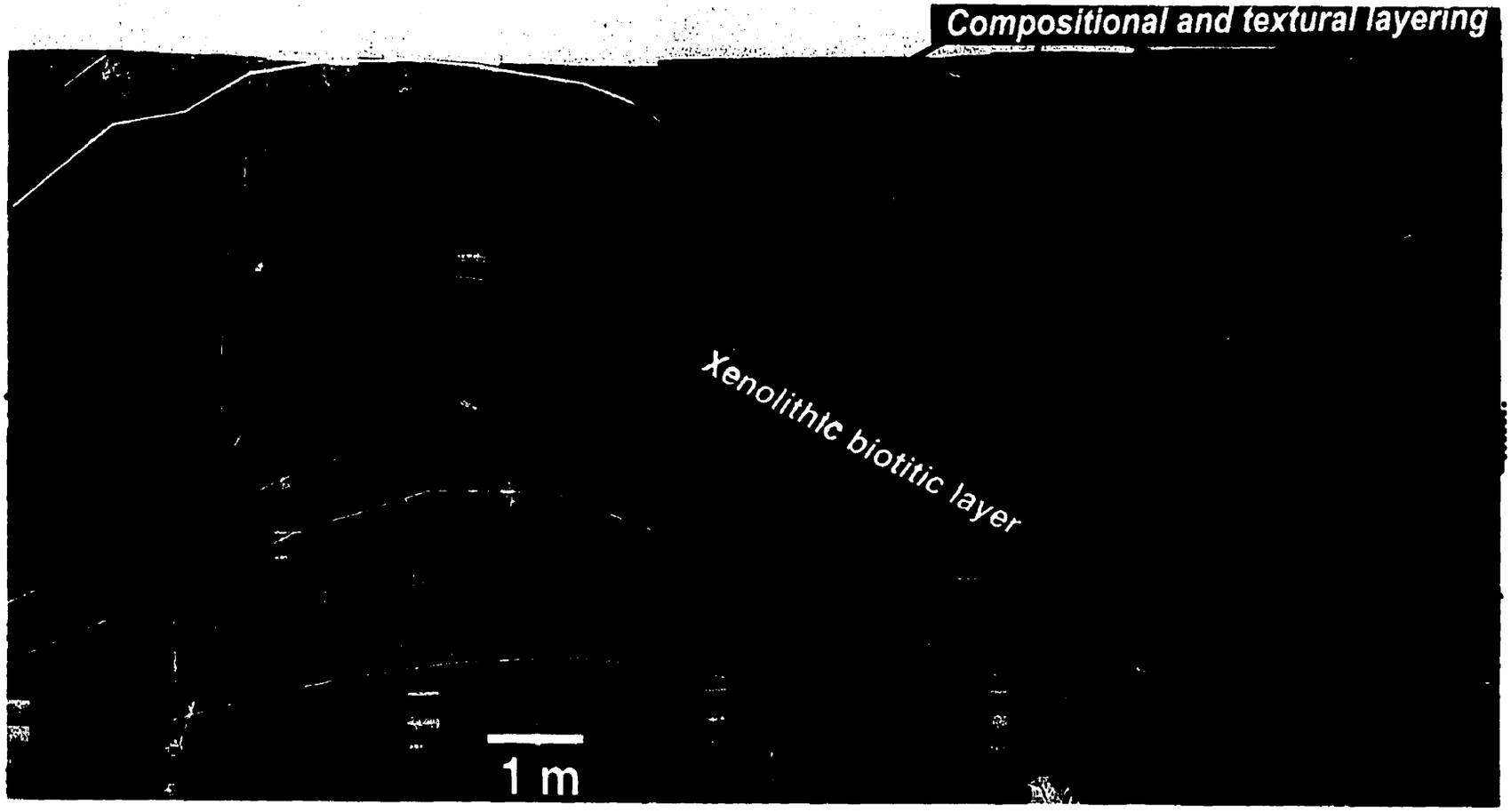



Figure 25. Fracture Zone 2 as seen in the URL shaft. An example of fabric control of fracture orientation and location on the dekametre-scale. This photomosaic presents an unrolled 360° view of the walls of the access shaft over a vertical distance of about 8 m. The squares of the metal grid visible at the bottom of the mosaic are each 40 by 40 cm. The apparent sinusoidal trace of the fault is an artifact formed by the intersection of a plane (the fault) with an unrolled cylinder (the shaft wall).

## 5.3. Fabric Control of Subvertical Natural Joints

The rockmass between the fault zones is host to one or more sets of subvertical fractures, the pattern and frequency of which is affected by variations in the room-scale fabric. This influence is summarized in Figure 26, with the horizontal axis representing the full range of fabric variations, and the vertical axis representing the frequency of several classes of natural fractures (faults and joints) for each of the ten fabric types. The frequency curves are based on a continuous 240-m long scan-line down the URL upper vent raise. It is apparent that fractures, both subvertical and subhorizontal, are most frequent where the granite is mostly strongly layered. This situation is analogous to the situation in sedimentary strata, where frequency for a given lithology varies in a direct linear relationship with bed thickness (Price and Cosgrove, 1994).

Fracture frequency and extent are also affected by local heterogeneities.
As shown in Figure 27, natural hematite -filled microfractures have been preferentially developed within a zone of very coarse microcline crystals. At any given depth and distance from faults, the subvertical fractures at the URL are generally more frequent in structural / textural variants of the host granite, such as the aplitic, granitic and pegmatitic dykes or sills (e.g. Figure 28). Apart from grain size and homogeneity of grain size, these rock units differ from the surrounding host granite in that they contain a variety of internal contacts, schlieren, vein-filled fractures, and foliations. As will be discussed in this work, these structures contribute to strength anisotropy, and promote the formation of both natural and excavation-induced fractures.

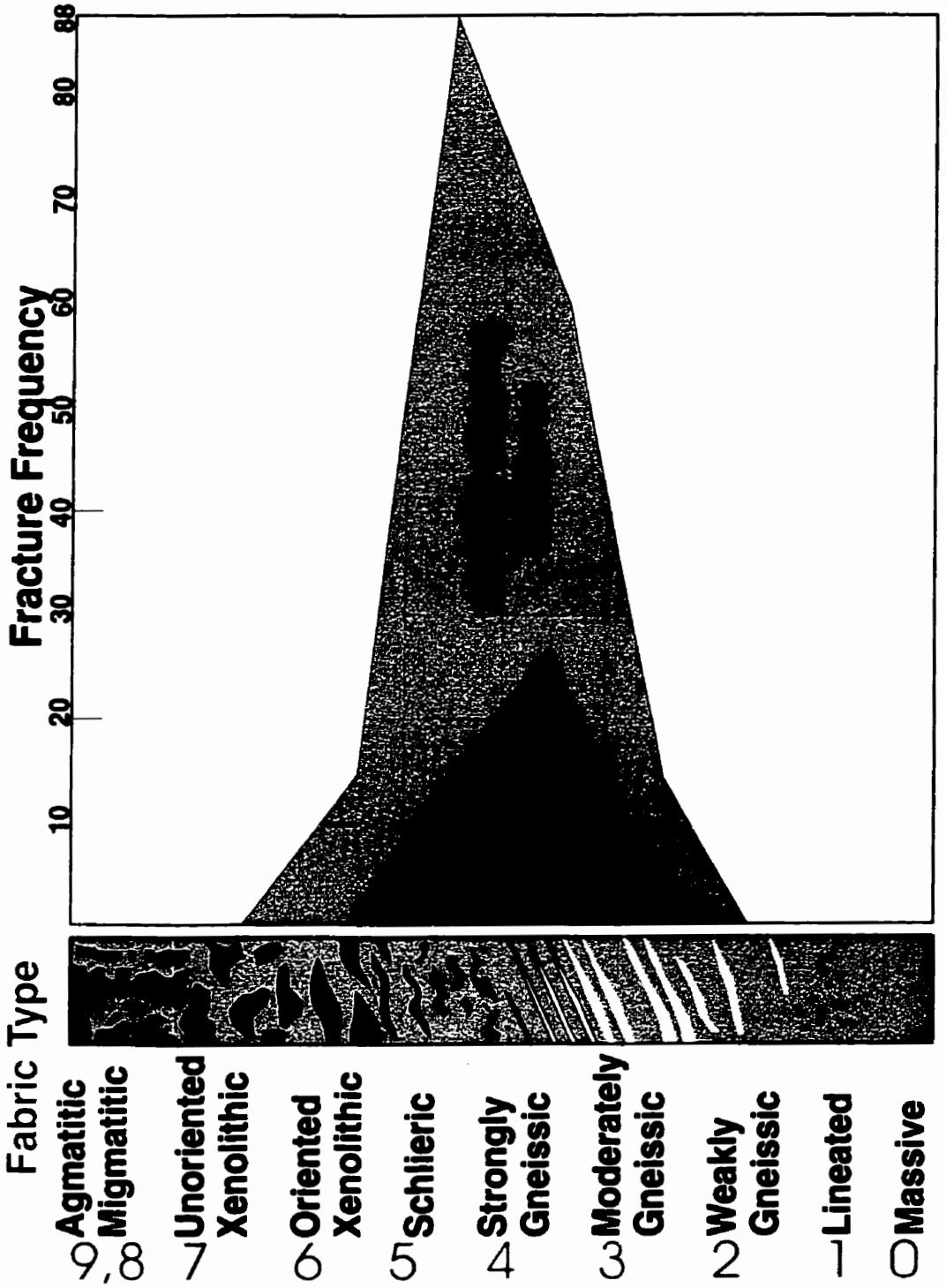



Figure 26. The Influence of Rock Fabric on Natural Fractures in the Lac du Bonnet Granite. Histogram depicting the frequency of several classes of natural fractures (faults and joints) as a function of the dekametre-scale fabric of the Lac du Bonnet Batholith. Variation in fracture frequency with granite fabric type, based on data from a 240-m scan-line down the URL upper vent raise.

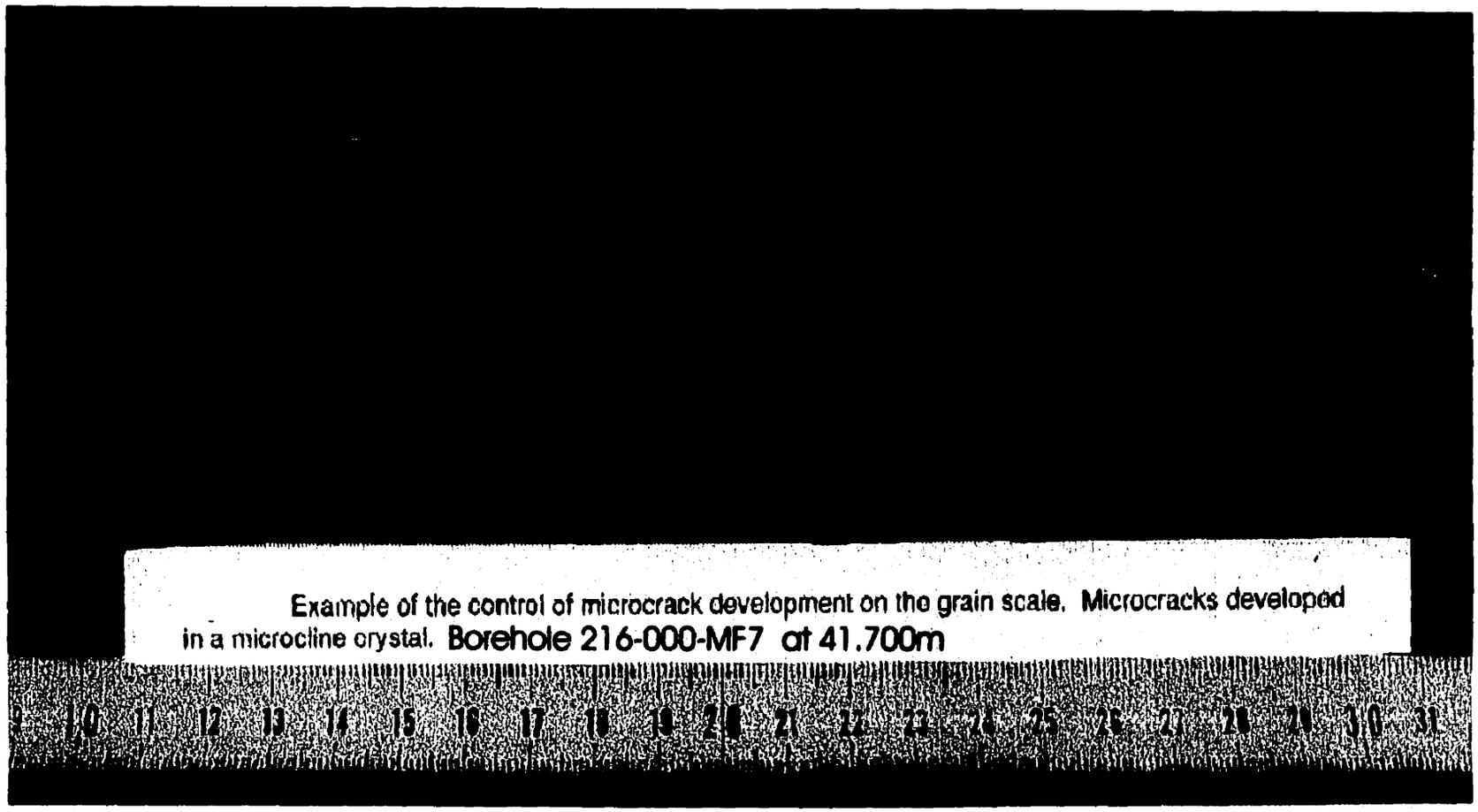



**Figure 27. The Influence of Rock Fabric on Natural Fractures in the Lac du Bonnet Granite - An example of fabric control of fracture orientation and location on the grain scale - natural microcracks developed in a zone of coarse microcline crystals.**

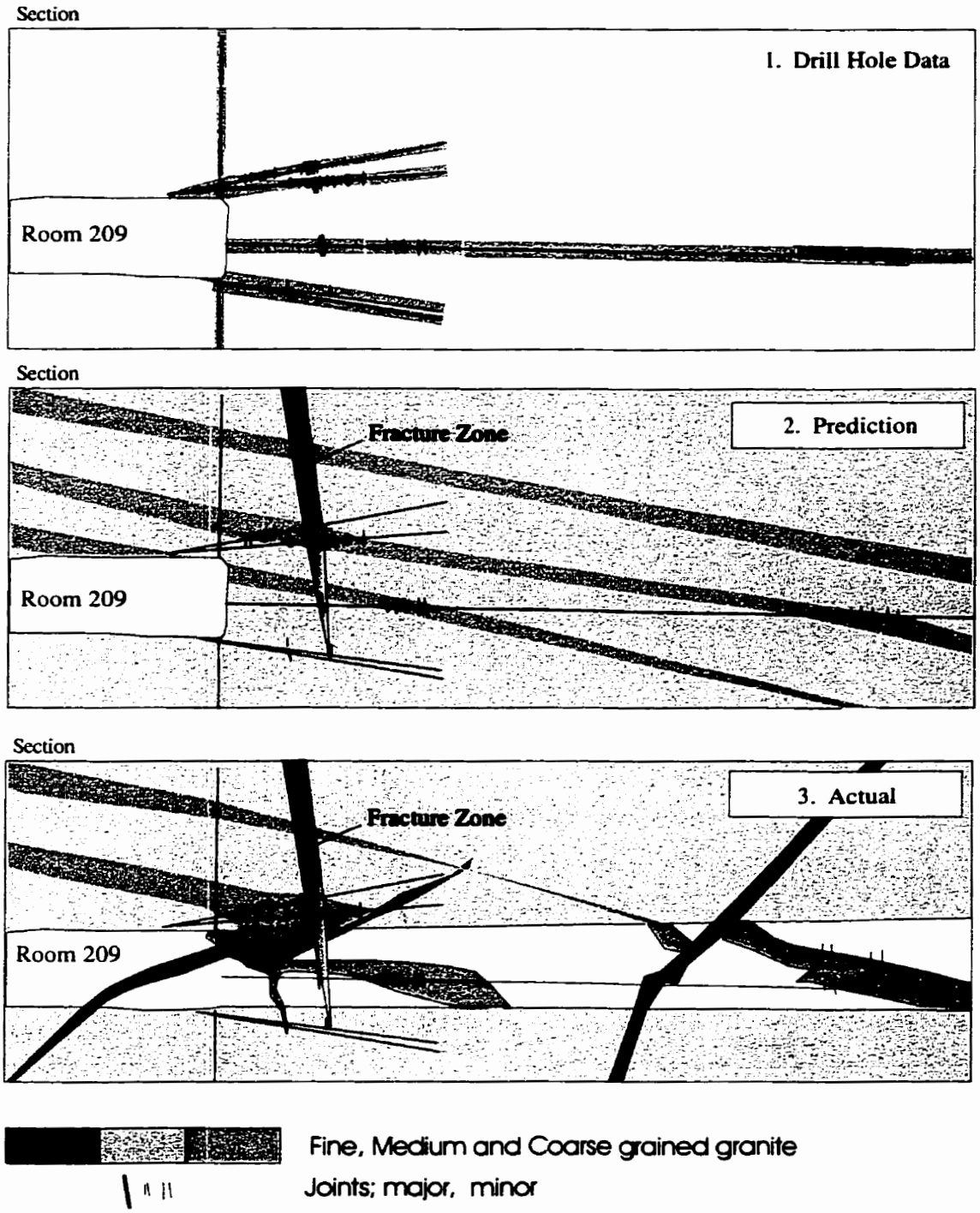



Figure 28. Control of fracture location and extent due to variations in granite at the 240 Level. Fractures are most frequent in the coarse grained granite sills, and in the fine grained granite dyke, while the main phase medium grained granite is largely unfractured. Note: these are all phases of the Lac du Bonnet Batholith, and not separate unrelated rock units.

## 5.4. Fabric Control of Excavation - Induced Fractures

Mapping of the shafts and the various tunnels at the 240 Level indicated that excavation-induced fractures were far more numerous than natural fractures except within the fault zones. In the shaft below Fracture Zone 2, in most of the 240 Level, and in all of the 420 Level, natural fractures (e.g. shear zones and joints, whether open or sealed) are absent, and the only fractures present are those induced by excavation. This result, based on mapping, is supported by the detailed logging of several kilometres of boreholes. An 1100 m long steeply-dipping borehole collared at the 420 Level intersected only one fracture (Everitt and Woodcock, 1999).

Excavation-induced fractures are distinguished from natural fractures by the absence of mineral infilling or wall rock alteration, by their position and geometry relative to blasthole remnants, and by their position and geometry relative to the tunnel profiles or tunnel face. In general, they form an onion-skin pattern of surfaces, similar to the exfoliation fractures seen in surface outcrops, but in the subsurface their orientation and size is dependent on their location relative to the tunnel perimeter, and to the location of tunnel faces during excavation.

Evidence that rock structure was contributing to excavation damage development was first obtained during routine mapping of tunnels during the construction of the 240 Level (Everitt and Read, 1989), and was confirmed by subsequent observations during: 1) the extension of the URL shaft (Everitt and Passmore, 1993; Everitt et al., 1993), and 2) during level development at the 420 Level (Everitt et al., 1993). Damage development has been observed to vary in response to textural and structural variations in the granite in

both the vertical and horizontal planes (Everitt et al., 1996). In particular, at compressive stress concentrations damage is far more extensive where the foliation coincides with the direction of expected slabbing.

In areas where the rock fabric is truly massive (i.e., homogeneous, as in the strict sense of this term), or where the fabric is normal to the excavation axis, the geometry of the excavation damage tends to be symmetrical about the tunnel or shaft profile. This symmetry tends to be lost however; in areas where the excavation axis runs along the "grain" in the rock. This is demonstrated in the following examples from the tunnels at the 240 and 420 Levels, and from the shaft at the 300 Level.

## 5.5. Fabric Control of Excavation - Induced Fractures in Horizontal Tunnels at the 240 Level

The tunnels at the 240 Level demonstrate the influence of subtle fabric on excavation damage development (Figure 29). The axis of tunnel (Room) 207 is parallel to the maximum principal stress and to the strike of the gneissosity, while Room 209 is oriented perpendicular to them. The tunnel profile of Room 207 is asymmetric, with induced fractures paralleling the gneissosity being not only more frequent, but also larger than those in other directions. In Room 209, the layering is not favourably oriented to serve as exfoliation planes, and - the damage profile is symmetrical. In this example, it is also noteworthy that the greatest damage developed not in the tunnel subject to the higher stresses, but in the tunnel subject to the lower stresses. Clearly, while the damage is initiated by stress release about the tunnel profile, the geometry and extent of this damage is very dependent upon the presence of the low-dipping layering.

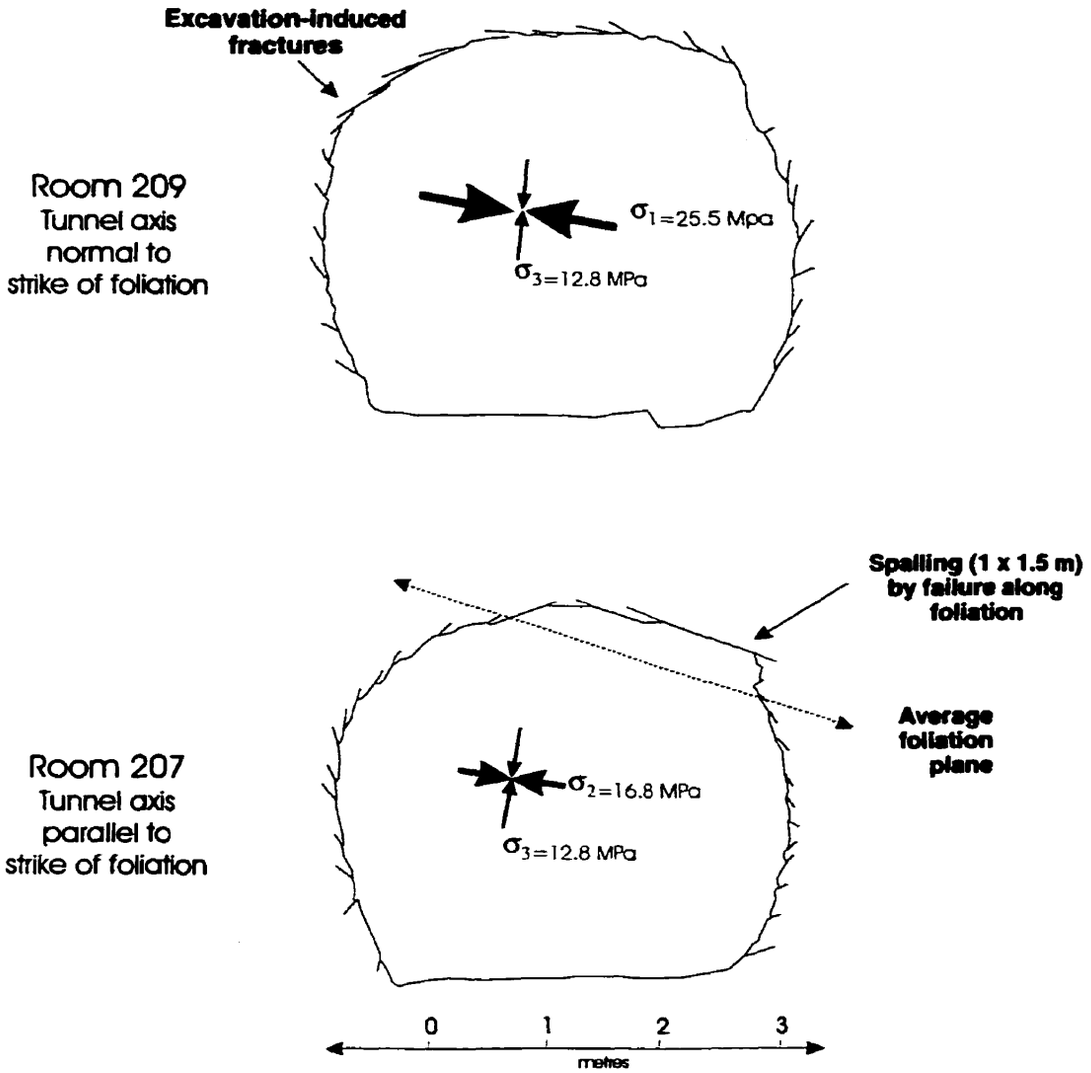



Figure 29. Influence of rock fabric on excavation damage - example from the 240 Level. This figure depicts the damage associated with different tunnel orientations in the same rock type – a medium grained granite. While apparently massive in hand specimen, more careful inspection reveals that the granite possesses both a weak mineral alignment (foliation) and a decimetre-scale layering. These apparently weak structures nevertheless impart a mechanical anisotropy to the rockmass, which is evident in the differences in excavation damage between these two profiles. The tunnel profile of Room 207 is asymmetric, with induced fractures paralleling the gneissosity being not only more frequent, but also larger than those in other directions, with spalling developing in the upper right corner (also shown in Figure 30). In Room 209, the layering is not favourably oriented to serve as exfoliation planes, and - the damage profile is symmetrical.

The excavation damage is also influenced by other fabric elements present. In addition to the layering and foliation shown in the medium grained granite, it is also host to steeply dipping fine-grained granite dykes. Where present, these inhibit the spalling developed in the tunnels parallel to the foliation (Room 207 of Figure 29), as is shown in Figure 30. As will be shown in the section 5.6 following, similar rock type-dependent damage relationships exist at the 420 Level (e.g. cf. Figure 30 with Figure 32) where the development of a breakout notch in the medium grained granite, is arrested wherever the fine grained granite occurs.

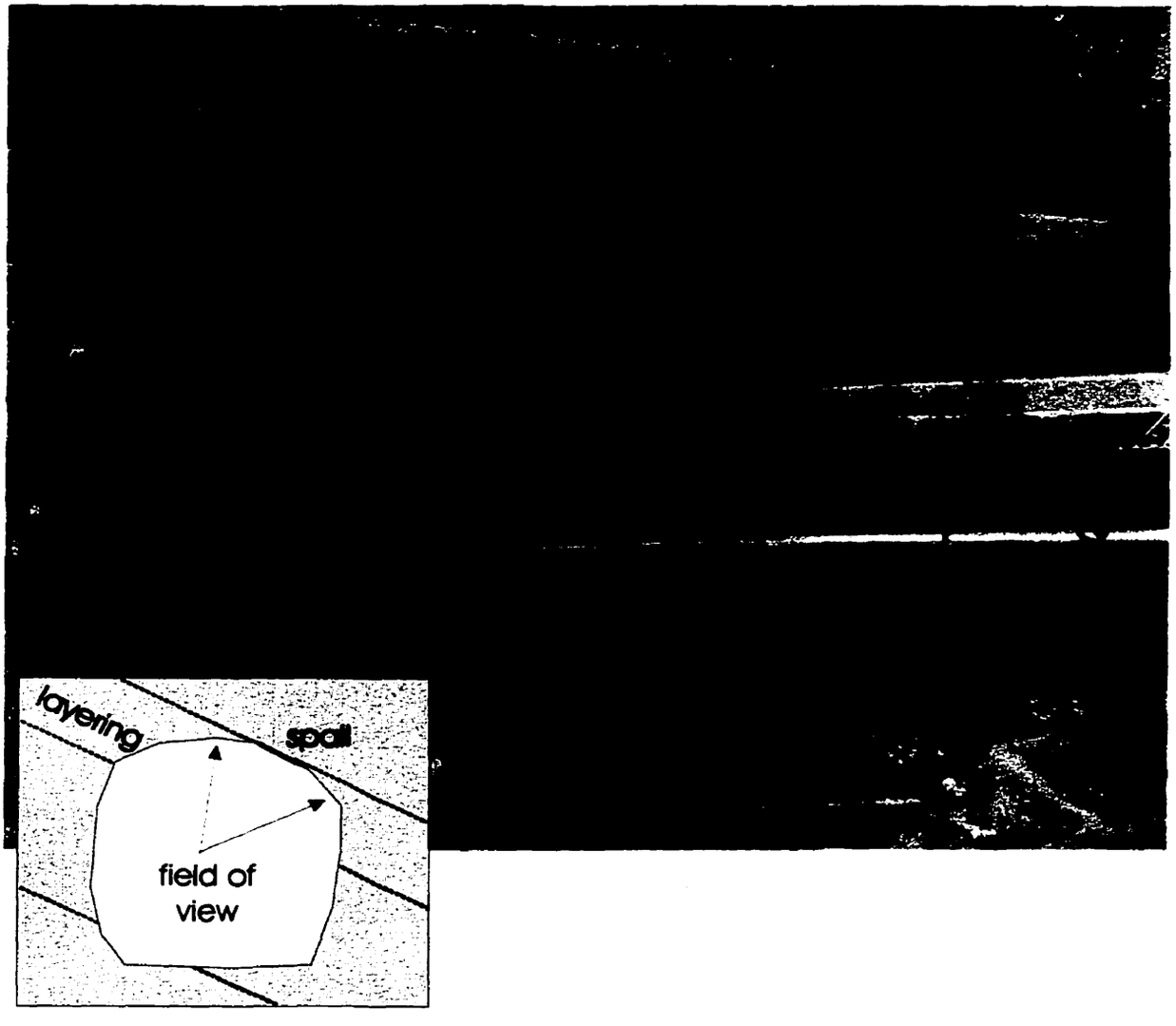



Figure 30. Influence of Rock Fabric on Excavation Damage - example from the 240 Level. This photograph shows the surface of the large spall formed by separation along the plane of layering and parallel biotite alignment in the granite in Room 207 (shown in Figure 29). The inset shows the field of view and the relationship between the tunnel profile, layering and the spall. This type of damage is not present in the tunnels, which are normal, or near normal to the strike of the layering. Spalling along the layering is arrested wherever the layering is absent or recrystallized - as in this case – by intrusion of fine grained granite dykes on the right and left sides of the field of view.

## 5.6. Fabric Control of Excavation - Induced Fractures in Horizontal Tunnels at the 420 Level

Excavation damage characteristic of the 420 Level is shown in Figure 31, and includes:

1. Stress-induced subhorizontal fractures which have propagated from pilot boreholes, along the plane of gneissosity, prior to excavation,
2. Excavation induced fractures of uncertain origin (i.e. blasting and/ or stress redistribution) on the tunnel surface, and
3. Breakout notches in the crown and floor of the tunnel. These notches comprise; (1) a broad zone where the reaming hole half barrels have spalled off, and (2), a central "V shaped" crush zone where the rock disintegrates into slabs and chips by a process of grain-scale fracturing.

The stress-induced subhorizontal fractures that have propagated from pilot boreholes or bootleg remnants (damage types 1 and 2 in Figure 31) can be large and pervasive at the scale of a tunnel face. In the medium and coarse grained granites, they are clearly developed along the plane of gneissosity, but they also occur in the fine grained granite, where they are discordant to the steeply-dipping flow banding, but parallel to the preferred orientation of natural microcracks at this level (as noted in the preceding chapter, and in Chapter 7, Figure 52). The subhorizontal induced fractures are interpreted as propagating from stress concentrations such as boreholes, their orientation controlled by the ambient stress field (i.e. normal to $\sigma_3$), with their development aided by the presence by of favourably oriented foliations and natural microcracks.

The rate of development, location and extent of the breakout notches (damage type 4 in Figure 31) appears to result from a combination of factors, including the sequence of excavation, the angularities in the room geometry, and the variations in local geology. In

excavation, the angularities in the room geometry, and the variations in local geology. In particular, at compressive stress concentrations the breakout is far more extensive where the foliation coincides with the direction of expected slabbing. Subsequent extensional failure (buckling) and spalling along these surfaces led to the development of breakout notches in the crown and floor. Where the layering was reoriented or absent (due to the presence of either fine grained dykes or xenoliths), notch development was much slower to develop (time-dependent, with time varying by rock type), and reduced in size.

The influence of geology on excavation response is illustrated using two examples from the 420 Level. In the first example, (Figure 32) excavation was by carefully controlled line-drilling and mechanical excavation – blasting was avoided so the purely stress—related effects of excavation could be assessed. Damage was allowed to progress to a stable state by careful scaling and limited (temporary) screening. In the second example (Figure 33), excavation was by conventional drill-and-blast methods. The tunnels are parallel to each other, approximately 50 m apart, with the same geology and stress environment. The tunnel axes parallel $\sigma_2$ at this level, with their transverse sections in the $\sigma_1$ $\sigma_3$ plane. In both examples, the tunnels were designed with circular cross-sections.

In both examples, spalling and notch development began almost immediately in areas of medium to coarse grained granite, and progressed rapidly (within a week) to their final form. The rate of notch development in the fine grained granite was much slower, regardless of excavation method, and in some areas never developed at all. For the area within the medium grained granite (in Figure 32), the ceiling and floor notches kept pace

development fell one or more rounds behind the advancing face and was much reduced in depth.

It should be noted however, that all the previous examples were conducted in areas where the tunnel axis was normal to the subvertical foliation in the fine grained granite. In this orientation, the presence of the foliation does not influence breakout development. This situation changes in those few tunnels where the tunnel axis is parallel to the strike of the foliation. As shown in Figure 35, spalling is now preferentially develop along this structure. As will be described in the following section, a similar damage / fabric relationships was observed in the 300 Level shaft and station, with damage being most prevalent in the "generally stronger" fine grained granite.

1. breakout notches
2. spalling along foliation
3. excavation-induced fractures
4. excavation-induced fractures radiating from blasthole remnants on face, and from pilot borehole.

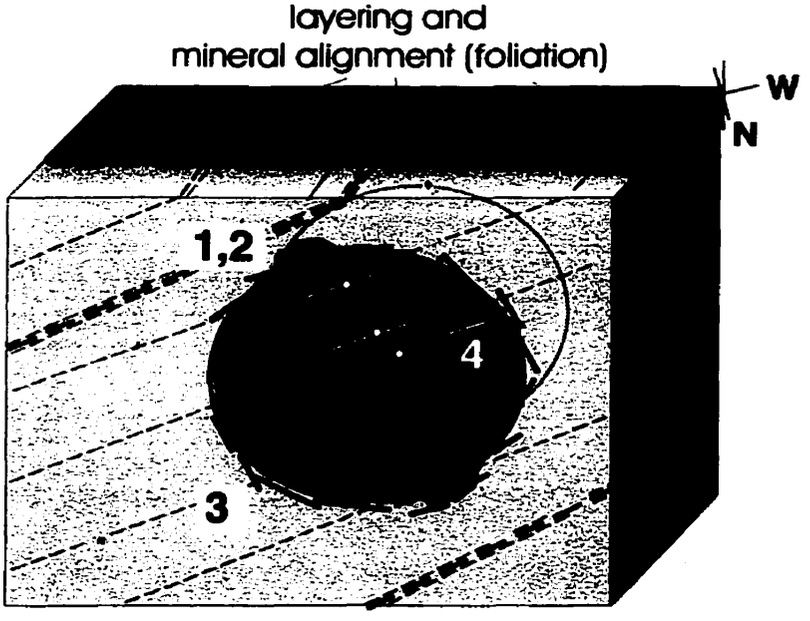



Figure 31. Influence of rock fabric on Excavation Damage - example from the 420 Level.

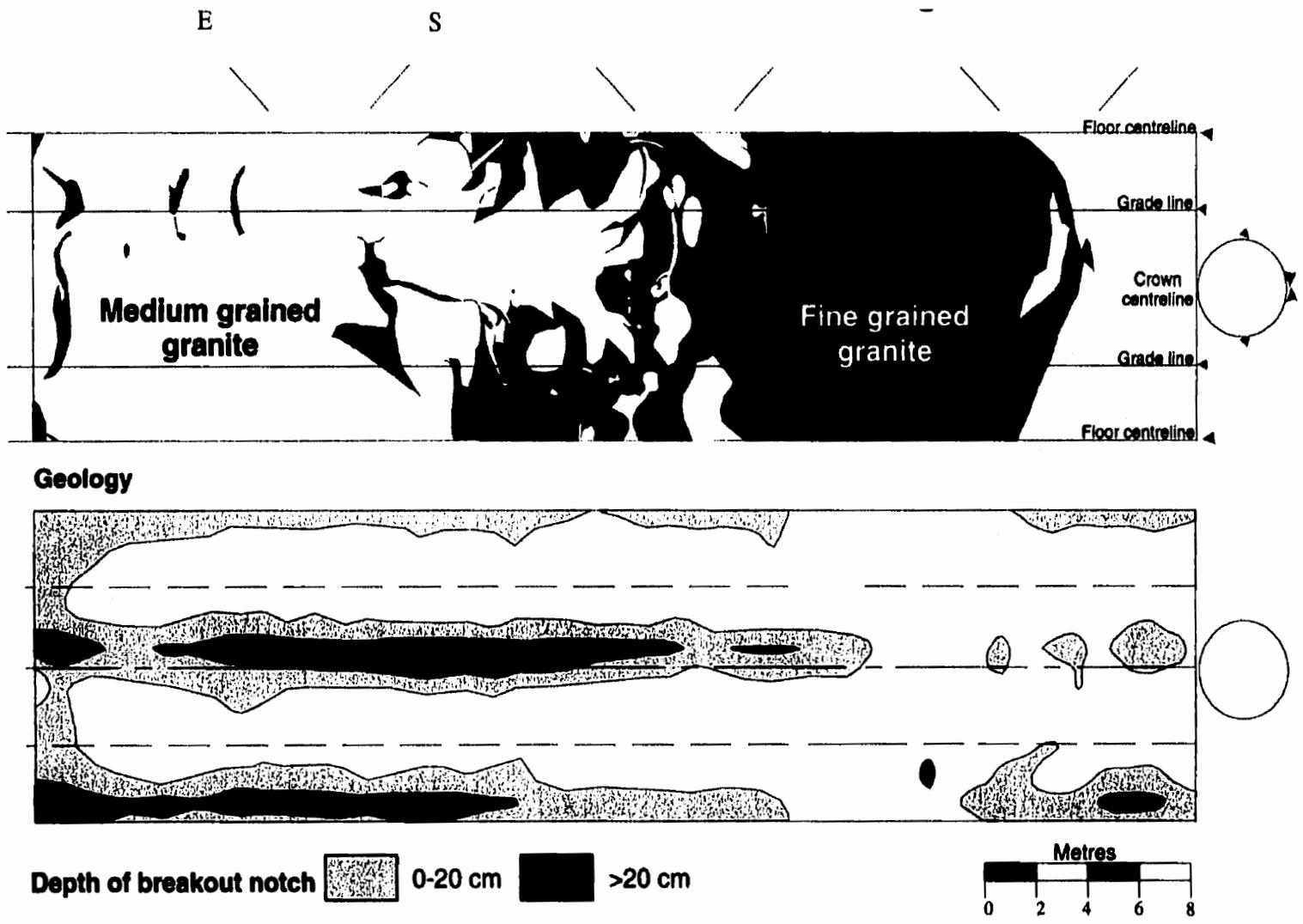



Figure 32. Influence of geology on breakout development within a single line-drilled tunnel (the Mine-by Experiment). Room 415 at the 420 Level of the URL provides an example of the influence of subtle differences in geology on breakout development in granite. The geology, notch depth and construction are shown.

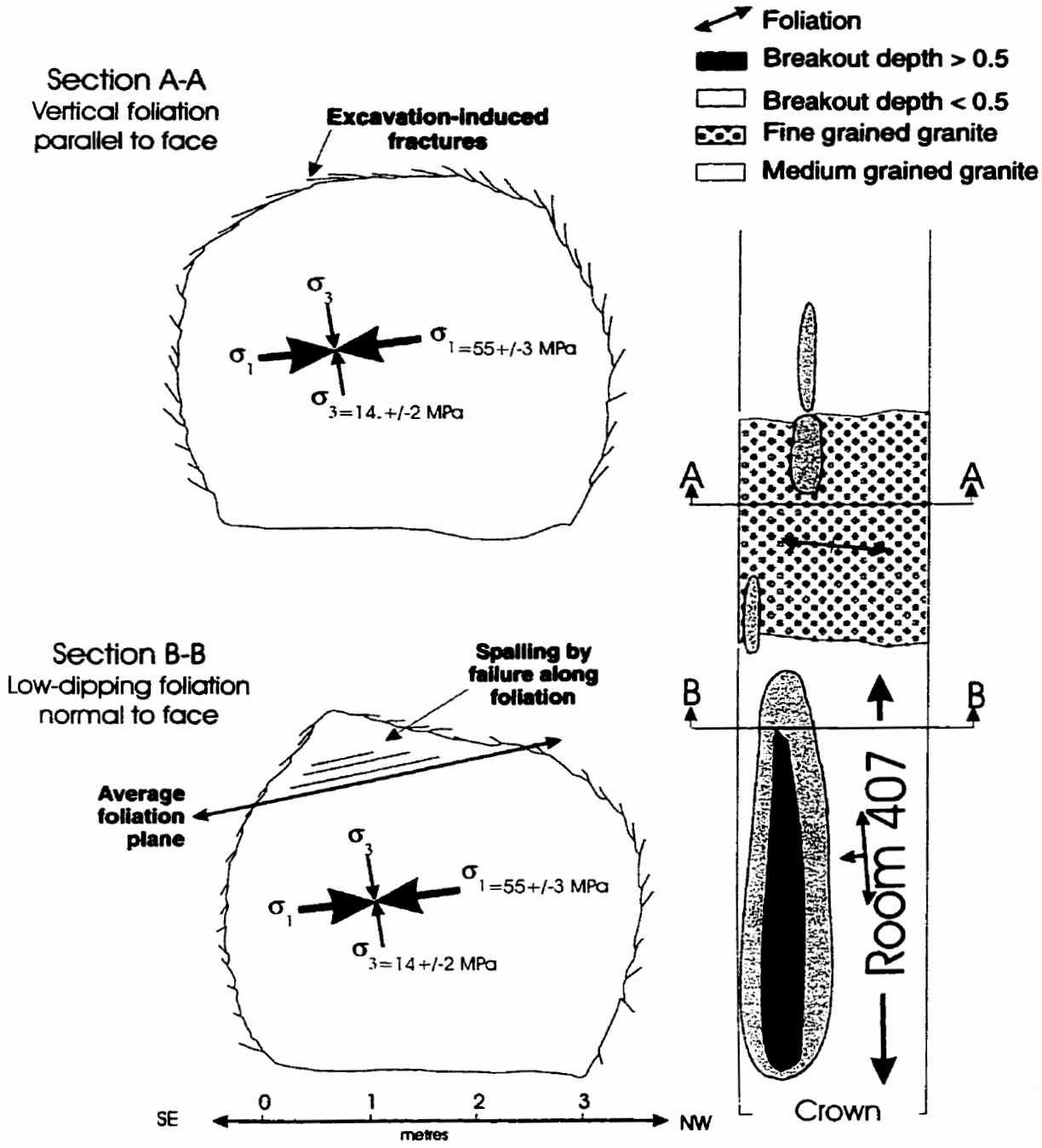



Figure 33. Influence of geology on breakout development within a single drill-and-blast tunnel. Room 407 at the 420 Level of the URL provides an example of the influence of subtle differences in geology on breakout development in granite where conventional excavation is used. The tunnel profiles developed in the fine grained and medium grained granites are shown in A and B, respectively.

Figure 34. Influence of geology on breakout development within a single drill-and-blast tunnel. Room 421 at the 420 Level of the URL provides an example of the influence of subtle differences in geology on breakout development in the fine grained granite where conventional excavation is used. In the previous figure it was noted that a biotite foliation also exists in the fine grained granite, but it is subvertical and normal to the tunnel axis. As such, its presence does not influence breakout development for this excavation orientation and geometry". In Room 421 however, the tunnel axis is parallel to the strike of the foliation, and spalling now occurs along this structure. A similar situation occurred in the shaft at the 300 Level, with spalling developed along this foliation where it paralleled the shaft wall (see Figure 35 and section 5.7 following).

## 5.7. Fabric Control of Excavation - Induced Fractures about the Vertical Shaft at the 300 Level

The 300 Level Shaft Station (Figure 35) is located in unfractured grey granite below Fracture Zone 2. The grey granite here is cut by three generations of auto-intrusive dykes, the most noteworthy of which, a granodiorite swarm, is the predominant lithology at this stop. The host phase of the batholith is represented by angular inclusions of granite and by variously disrupted blocks between the dykes.

Special efforts were made during the excavation of the 300 Level Shaft Station to optimize the amount and quality of exposure available in order to characterize the damage zone around the circular access shaft. Prior to construction of the 300 Level Shaft Station, the shaft was excavated to a depth of about 15 m, or 3 shaft diameters below the floor of the station to accommodate any time-dependent behaviour in the formation and propagation of the excavation damage, and to avoid the effect of stress concentration. The 300 Level Shaft Station access station was then excavated using controlled pilot and slash techniques for the walls, floor, and crown. This was done to duplicate the quality of excavation characteristic of the shaft, and to avoid inadvertent excavation of damaged rock. Two surfaces were then made available for observation of the excavation damage zone about the shaft perimeter. These were: 1) a temporary 90° brow formed by the intersection of the shaft walls, 2) the station crown (once photographed, this was trimmed to a more stable 45° slope), and 3) the station floor. A third area, referred to as the station lip and lip pocket, provided a sample of the damage zone beneath the floor of the station as seen from the shaft walls.

The results of the excavation damage zone mapping are shown in Figure 35. Although it may be possible to distinguish in outcrop certain specific fractures that originated either through excavation-induced stress or through blasting (chiefly on the basis of location and pattern), most fractures are likely to be the result of a combination of processes.

Damage distribution may be correlated with rock fabric and with the principal stress orientations. In Figure 35, note the preferential development of induced fractures in the fine grained granite, on the east side of the shaft. As was the situation in Figure 34, the fine grained granite contains a subvertical foliation which, where parallel to the excavation surface, facilitates spalling. This does not occur in the medium grained granite, where the foliation is subhorizontal and therefore not favourably oriented with respect to the local stresses to induce spalling.

Clearly, while the damage is initiated by stress release about the shaft profile, the geometry and extent of this damage is again, very dependent upon the presence of a suitably oriented fabric such as layering or foliation.

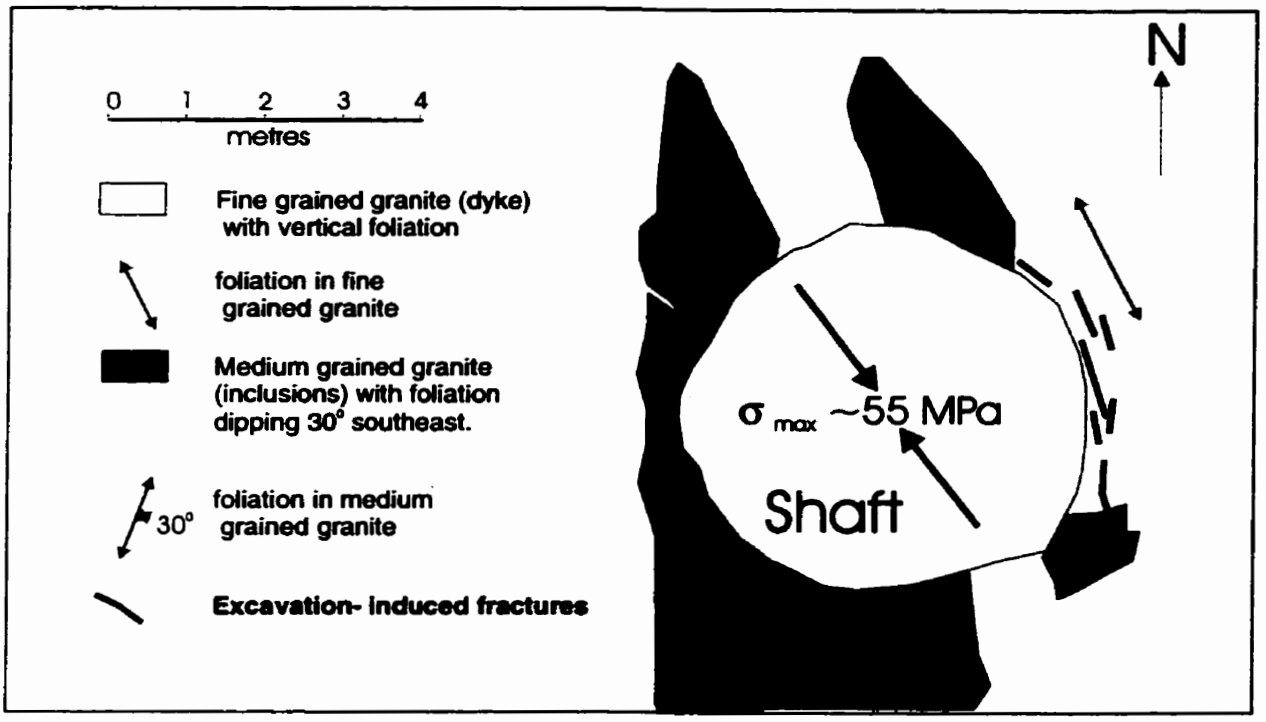



Figure 35. Simplified geological map showing the excavation damage at the 300 Level station.

## 5.8. Conclusion of Field Investigations into Fabric Control of Excavation - Induced Fractures

Within the allegedly homogeneous Lac du Bonnet Batholith, there is ample evidence at all scales, that even subtle variations in rock type, and in the degree and orientation of foliations and other fabrics have clearly influenced subsequent brittle deformation, whether this was natural or induced by excavations. This influence is present in excavations with vertical or horizontal axes, and in tunnels excavated by different methods. At compressive stress concentrations around excavations, breakout is far more extensive where the foliation coincides with the direction of expected slabbing. Rock fabric and its influence on rock strength are therefore significant to the testing and interpretation of rock mass response, and to experiment and excavation design.

# Chapter 6

# SAMPLE COLLECTION FOR ROCK PROPERTIES TESTING

Three locations (sample areas) were chosen to further examine the interaction between rock fabric and excavation damage, through the collection of a suite of samples for rock properties testing. The three locations selected provide a representative sample of the range of the compositional and textural heterogeneity present in the Lac du Bonnet granite, and their affect on rock strength and rock mass response. These are:

1. a fine grained granite dyke with subvertical flow banding,
2. a medium grained and weakly to moderately layered granite forming the main mass of the batholith, and
3. a generally coarse grained and leucocratic unit that occurs as sills and recrystallized zones in the main phase.

For the sake of convenience, these are referred to as the fine, medium and coarse grained granites, respectively.

The three case studies were sited in newly excavated tunnels at the 420 Level (420 m depth of the URL) where the damage due to stress relief can be observed in the process of developing. The sites are in close proximity to each other. They are distant from faults or major fractures, and are in the same in-situ stress-domain. As such, they are subject to the same stress magnitudes and orientations.

The general location of the 420 Level, and the names of the various horizontal tunnels and inclined ramps referred to in this text are shown in Figure 36. The location of the samples, and the local geology at each sample location are shown in Figures 37, 38 and 39.

At each site, oriented core has been obtained from an array of boreholes. The arrays were designed to intersect the most prominent foliation / microcrack set at 90°, and then at 20°-30° increments until the core axis was parallel to the foliation (as noted earlier in Table 5 and 6, p. 24 and p. 27). The core was logged for foliations and microcracks, and then cut to provide oriented samples for uniaxial compression tests. Strength anisotropy and the influence of foliation were then systematically investigated using rock mechanics tests that are sensitive to planar anisotropy (Point Load and Brazilian tests).

## 6.1. Coarse Grained Granite

The geology, tunnel profile, excavation damage, and sampling array at the coarse grained granite study location in Room 418 is shown in Figure 40. This tunnel (Room 418) trends 055°, plunges at 15% (06°), and is parallel to $\sigma_2$. The tunnel section is oval, with

its major axis rotated 11° from horizontal to align with the $\sigma_1$ direction. The rockmass here contains a variety of pervasive and non-pervasive foliations including decimetre to dekametre-scale layering, and a parallel mineral foliation (chiefly biotite). These are deformed or crosscut by boudinage structure and ductile faults, and by fine grained granite dykes, semi-ductile faults, and mineral alignments. Spalling along compositional layering and biotitic schlieren, was well developed in the area of the floor and crown. Compositional heterogeneity also appears to be responsible for a pop-up developed at the floor at the array location. This pop-up (1.95 x 0.3 x 0.03 m) between two drill traces popped up (i.e. buckled) at the array. Unlike similar pop-ups at this level, (e.g. Figure 32, p. 144) this one was oriented such that the long axis of the buckled slab was aligned with the tunnel axis rather than tangential to the tunnel.

## 6.2. Medium Grained Granite

The geology, tunnel profile, excavation damage, and sampling array at the medium grained granite study location in Room 417 is shown in Figure 41. This tunnel trends 235°, plunges 00°, and is parallel to $\sigma_2$. Half barrels, an indication of controlled quality blasting, were evident in the sidewalls following blasting, but were lost by spalling over a 2 to 3 m section of the roof where notch development occurred. Blast-induced fractures radiate away from blastholes on the tunnel face, but those dipping 10°-40° to the southeast are much larger and more frequent than other directions. Tensile fractures oriented radial to the tunnel periphery in the tensile sidewall region are also present. This orientation is coincident with the $\sigma_1$ / $\sigma_2$ plane (or the normal to $\sigma_3$), and as noted in Chapter 4, is coincident with the preferred orientation of natural microcracks.

## 6.3. Fine Grained Granite

The geology, tunnel profile, excavation damage, and sampling array at the fine grained granite study location in Room 421 is shown in Figure 42. Room 421 trends 235°, plunges at (07°), and is parallel to $\sigma_2$. The tunnel section is oval, with its major axis horizontal. At this location, the fine grained granite has a tendency to break in a blocky fashion, with block surfaces being defined by a well-developed subvertical flow banding, and by low-dipping excavation-induced fractures. No natural fractures were encountered. Blast-induced fractures on the tunnel face radiate away from blastholes but those dipping 10°-40° to the southeast are much larger and more frequent than other directions. This orientation is coincident with the $\sigma_1$ - $\sigma_2$ plane (or normal to $\sigma_3$). It is also parallel to the average orientation of the thrust faults above this level, and to one of the microcrack sets developed at this level. Half-barrels were evident in the sidewalls following blasting, but were removed by spalling from a 2.5 to 3 m section of the roof in each round.

This space intentionally left blank for effective pagination.

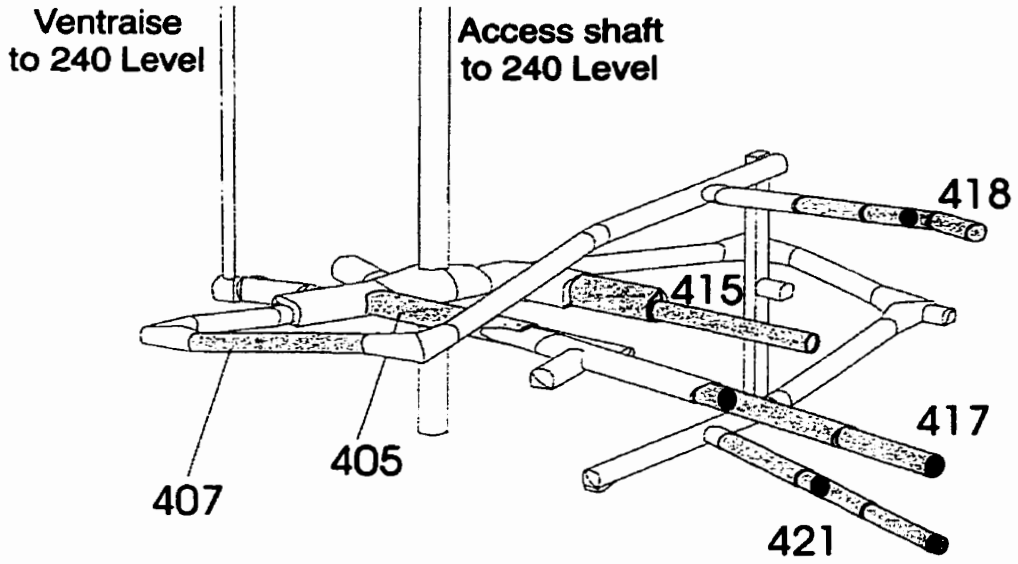



Figure 36. General location of the 420 Level, and the names of rooms referred to in this text. The black ellipses represent the sample locations.

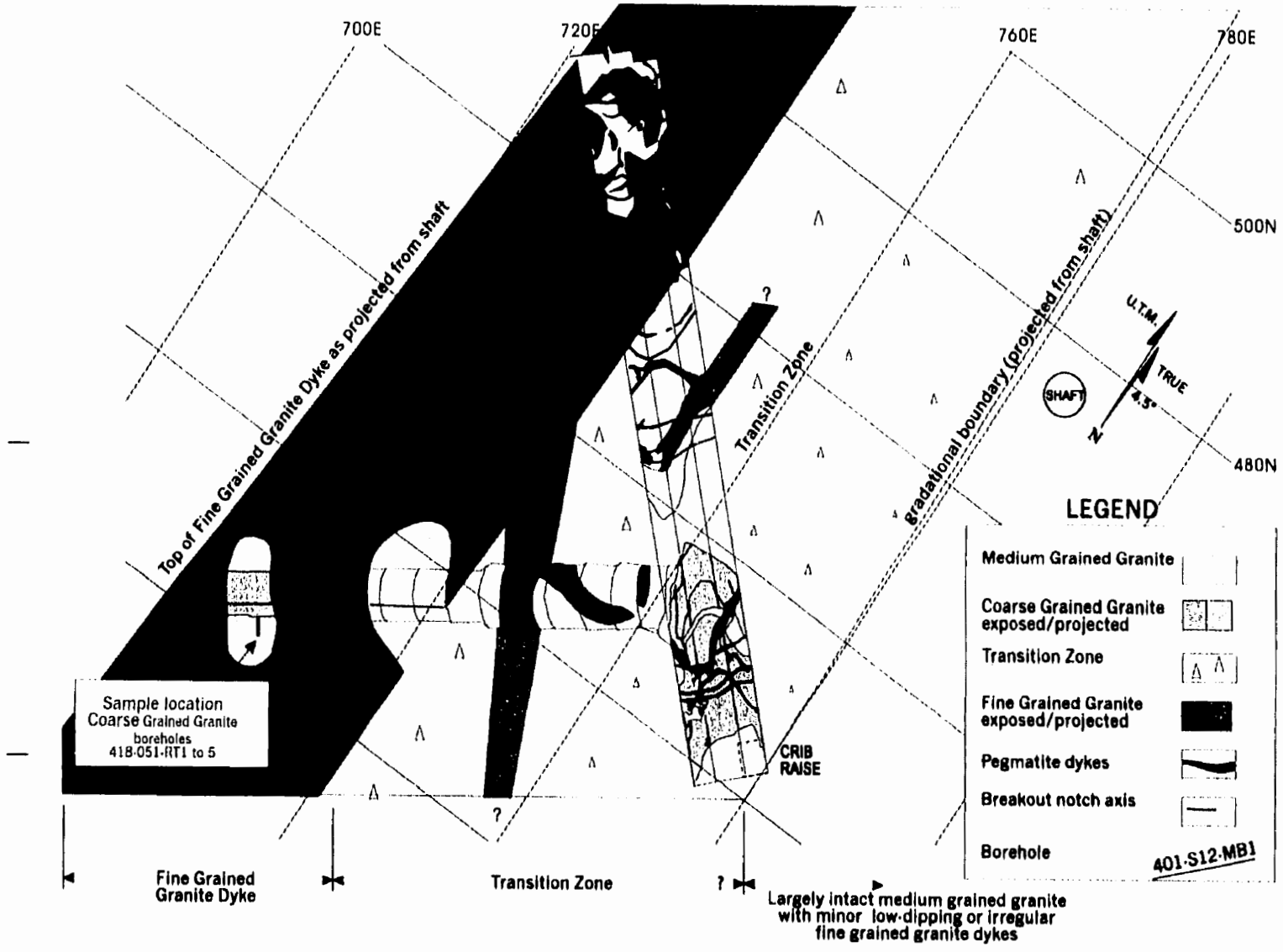



**Figure 37. 420 Level geology, upper ramp level (depth 354 m).**

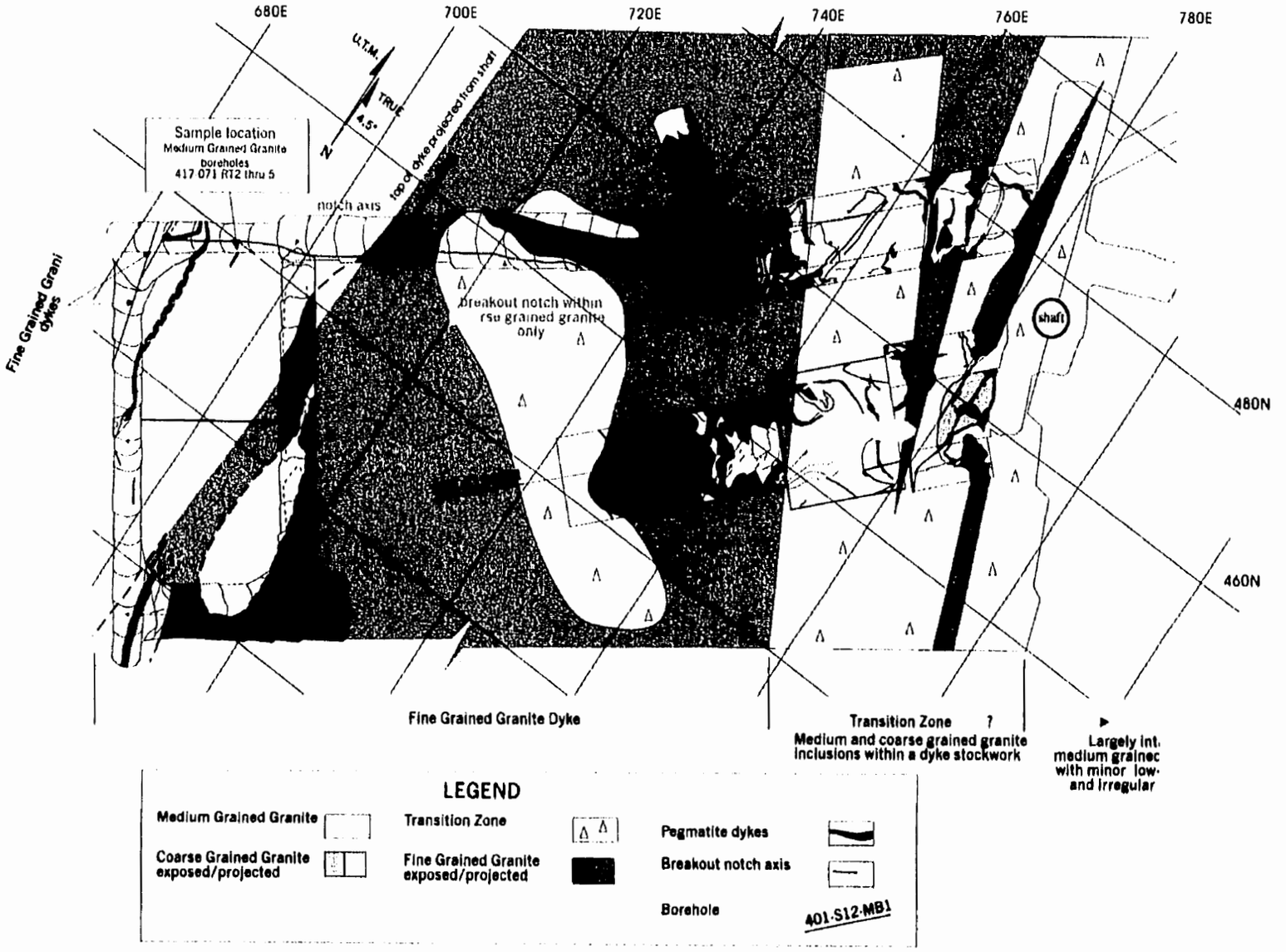



**Figure 38. 420 Level geology, main level (depth 415 m).**

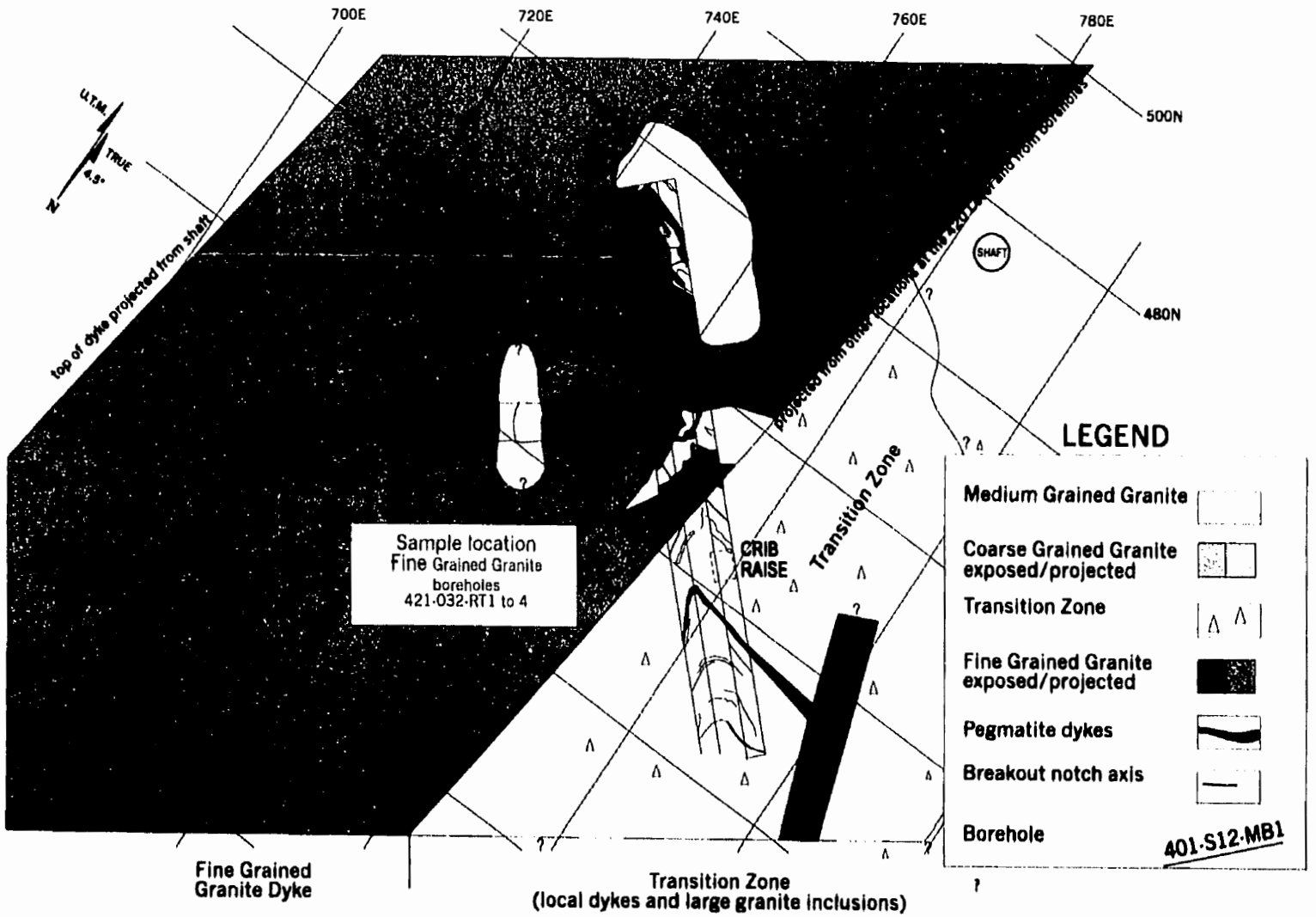



Figure 39. 420 Level geology, lower level, (depth 437 m).

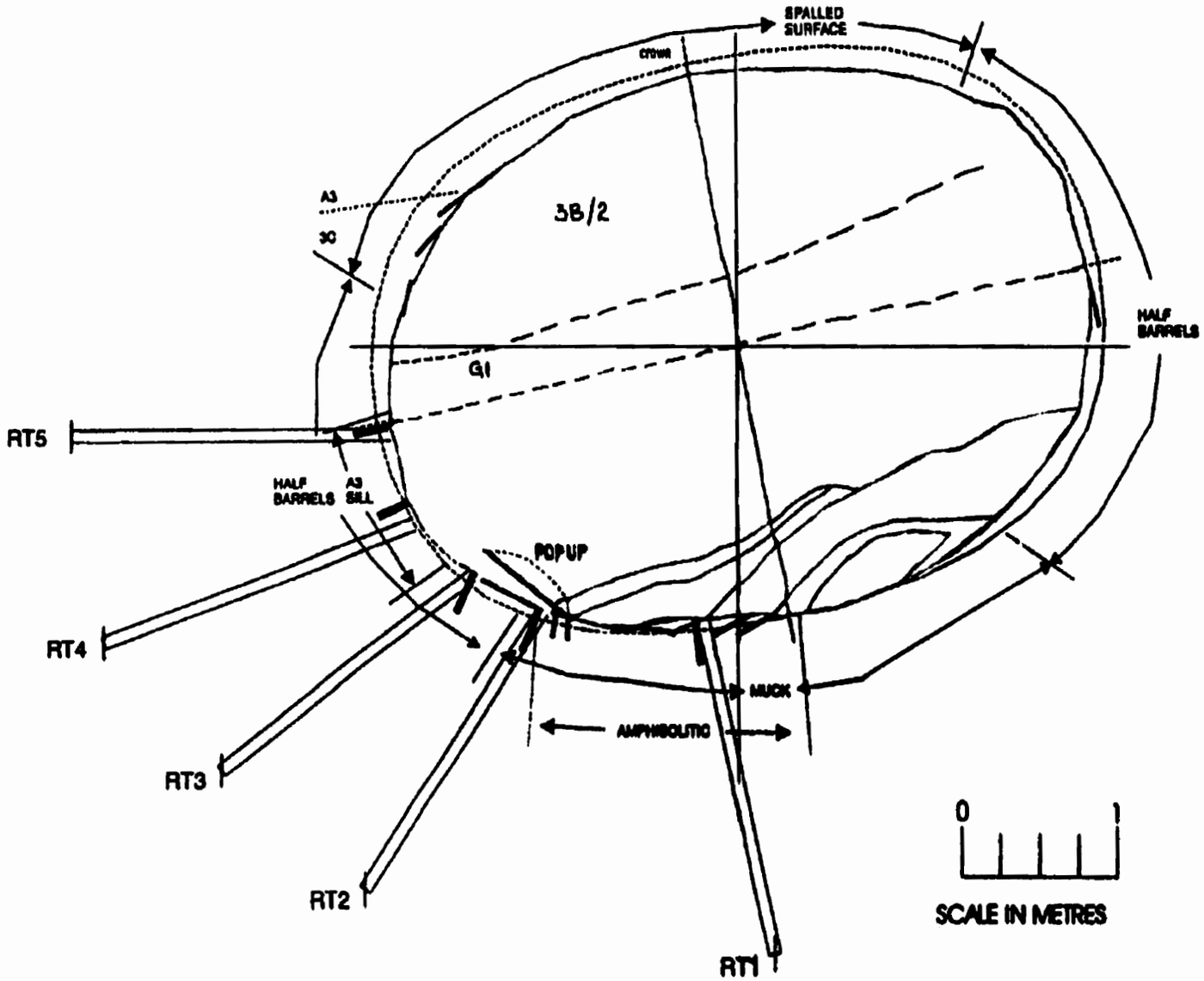



Figure 40. The geology, tunnel profile, excavation damage, and sampling array at the coarse grained granite location in Room 418.

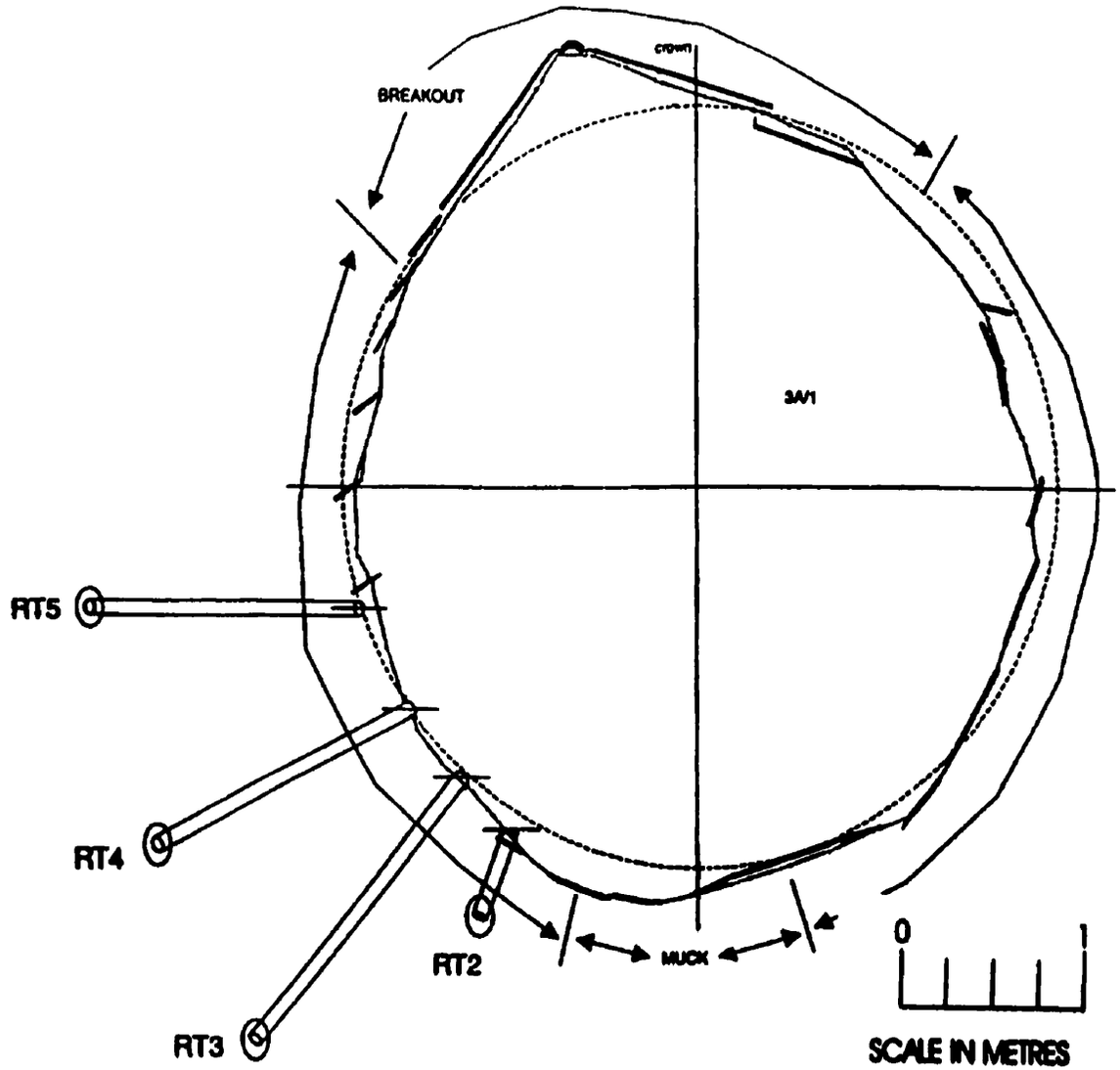



Figure 41. The geology, tunnel profile, excavation damage, and sampling array at the medium grained granite location in Room 417.

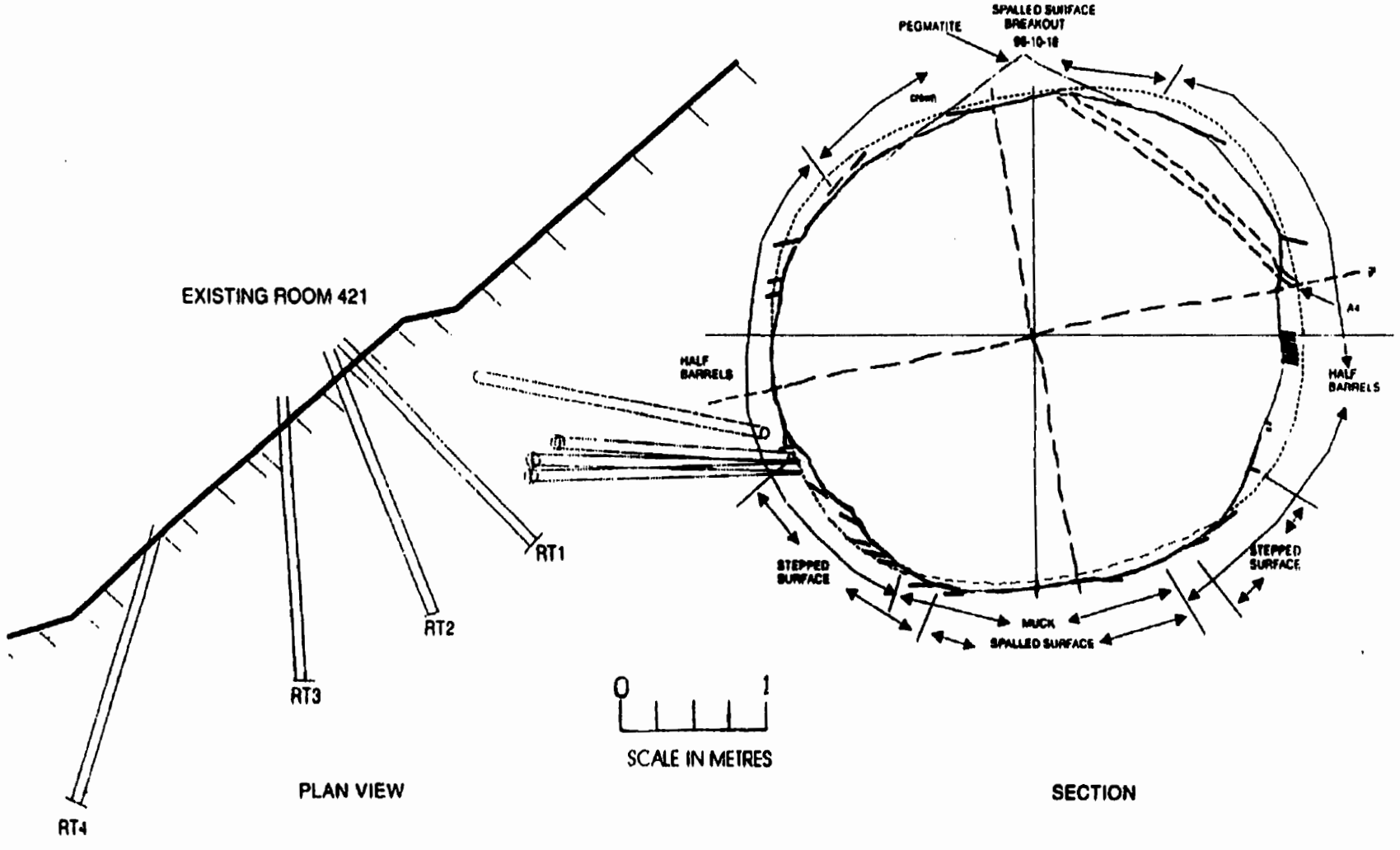



Figure 42. The geology, tunnel profile, excavation damage, and sampling array at the medium grained granite sample site, Room 421.

# Chapter 7

# POINT LOAD TESTING

## 7.1. Introduction

This section summarizes Point Load testing conducted to identify variation in tensile strength with the direction of loading, and its relationship to rock fabric within the granite of the Lac Du Bonnet Batholith. Three texturally distinct end-members of the granite of the Lac Du Bonnet Batholith were tested: 1) fine grained granite dykes with subvertical flow banding, 2) a medium grained and weakly to moderately layered granite forming the main mass of the batholith, and 3) a generally coarse grained and leucocratic unit that occurs as sills and recrystallized zones in the main phase. For the sake of convenience, these are referred to as the fine, medium and coarse grained granites, respectively.

## 7.2. Specimen Preparation

Samples used in these tests were obtained from the boreholes listed in Table 11. All samples are segments of HQ- sized core (46 mm diameter), the choice of specimen diameter being largely dictated by budget constraints. Specimens were prepared in accordance with the procedure described in (Coates and Gyenge, 1966; Gyenge, 1980), and by the International Society for Rock Mechanics (Brown, 1981). The end surfaces of each of the specimens were ground flat to ensure that they were parallel to each other and perpendicular to the axis of the specimen.

A total of 105 oriented specimens were selected as summarized in Table 11. Their size and density statistics are summarized in Table 12. The mean bulk density ranged from 2.54 to 2.72 gm/cm$^3$; these corresponding to the leucocratic and melanocratic layers in the coarse grained granite. Density variations showed no correlation with distance from the excavation.

Each specimen was retained after loading to failure for detailed mapping of their fabric and the test-induced fractures formed. The orientations of foliations, natural microcracks and the test-induced fractures, all relative to the core reference line, were recorded and these were then converted to true orientations. The fabric / fracture maps for each specimen and data tables for each specimen are provided in Appendix A.

The orientation of all fabric elements and test-induced fractures in the samples were required so that the compressive and tensile strength results could be compared to the results from tunnel mapping of rock structure, to the natural microcrack investigations, and to the in-situ stress data and rock structure. Accordingly, each specimen was retained after loading to failure for detailed mapping of their fabric and the test-induced fractures. The orientations of foliations, natural microcracks and the test-induced fractures, relative to the core reference line, were recorded and these were then converted to true orientations. The angles necessary to convert relative orientations to true orientations, given the trend and plunge of the borehole are shown in Figures 43 to 45.

**Table 11.**
**Sample and testing summary**

| Rock Type | Borehole | Orientation | | test type and number | | |
|---|---|---|---|---|---|---|
| | | trend | Plunge | Uniaxial compression | Brazilian | Axial Point Load |
| Coarse grained granite | 418-051-RT1 | 324.90 | 77.10 | | *Discing* | |
| | 418-051-RT2 | 143.61 | 54.76 | 1 | 8 | 15 |
| | 418-051-RT3 | 143.40 | 35.40 | 1 | 8 | 6 |
| | 418-051-RT4 | 143.56 | 19.70 | 1 | 7 | 6 |
| | 418-051-RT5 | 143.38 | 0.40 | 1 | 4 | 12 |
| Medium grained granite | 403-014-MB2 | | | 1 | | |
| | 417-071-RT1 | Not drilled | | | | |
| | 417-071-RT2 | 190.43 | 59.44 | | *Discing* | |
| | 417-071-RT3 | 190.20 | 39.48 | | *Discing* | |
| | 417-071-RT4 | 189.10 | 19.50 | 1 | 6 | 6 |
| | 417-071-RT5 | 189.70 | -0.20 | 1 | 4 | 8 |
| Fine grained granite | 421-032-RT1 | 133.99 | 0.23 | 1 | 4 | 10 |
| | 421-032-RT2 | 157.61 | 0.46 | 1 | 6 | 7 |
| | 421-032-RT3 | 175.81 | 0.80 | 1 | 5 | 10 |
| | 421-034-RT4 | 199.05 | 0.00 | 1 | 6 | 13 |

**Table 12.**
**Sample Properties**

| Granite type [1] | Number of samples | Diameter (mm) | | | Thickness (mm) | | | Average T/D Ratio | Density gm/cm$^3$ | | | Point Load Strength Index (MPa) [2] | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | min. | -mean- | max | min. | -mean- | max | | min. | -mean- | max | min. | -mean- | max |
| Fine | 43 | 44.75 | 44.87 | 44.96 | 18.30 | 20.37 | 23.34 | 0.55 | 2.61 | 2.64 | 2.67 | 16.20 | 21.32 | 27.00 |
| Medium | 16 | 44.68 | 44.73 | 44.80 | 19.40 | 20.73 | 22.45 | 0.46 | 2.60 | 2.62 | 2.63 | 12.90 | 16.32 | 22.58 |
| Coarse | 41 | 44.50 | 44.80 | 45.10 | 18.60 | 20.50 | 22.80 | 0.45 | 2.54 | 2.65 | 2.72 | 13.20 | 15.00 | 18.00 |

[1] Listed by average grain size, these are the abbreviated names. The full rock names are: 1) fine grained and flow banded granodiorite dykes, 2) the homogeneous to weakly banded main phase granite, and 3) the coarse grained xenolithic leucocratic granite.

[2] These are averages based on all boreholes. The Point Load Index (and the tensile strength) is in fact anisotropic, and varies with respect to the orientation of the loading axis relative to one or more fabric elements within each rock type.

## 7.3. Methodology

The most commonly used laboratory methods to study tensile strength, are the point-loading and the line-loading (or Brazilian) methods. In the point-loading method, core discs are loaded axially between a pair of indentors while in line-loading method; the disc is placed on edge and loaded in the direction of its diameter between two flat platens. In the line-load method, the diametral fracture plane is constrained by the position of the loading indentors, while in the point-load test, the diametral fracture plane is free to take any orientation. In practice, failure tends to be controlled by heterogeneities in the rock, such as foliations or microcracks, which are close to or parallel to the plane of failure under this loading geometry. Heterogeneities well outside this plane, such as incipient core discing, do not affect the test.

According to Lajtai (1980) "... The Brazilian test seems to yield a more accurate definition of both the tensile strength and its variation with direction ... (and is the more commonly applied tool for this application in rock mechanics).... The point load test is more suitable for the determination of the minimum value of tensile strength, and ...(in determining if there are preferred directions of fracture)...".

The same approach was applied in this study: with Point Load tests to determine if there are preferred directions of fracture in the massive rocks of the LDBB, and Brazilian tests to determine if there are variations in tensile strength with the direction of loading.

The Point Load test was also used to provide a relative ranking of the tensile strength of the various rock types, although this was a secondary objective of this experiment. This ranking is based on calculation of the "Point Load Strength Index" (Goodman, 1989). The "point load index" is given by the following:

$$\text{Point Load strength index, } I_s = P/D^2 ,$$

*where* $P$ = *the load to break the specimen, and*

$D$ = *the distance between the platen contacts.*

The methodology used in this study followed the recommendations laid out in Brown (1981), Goodman (1989), and Gyenge (1980). The tests were performed in the CANMET Bell's Corner's Laboratory. The testing machine incorporated a loading system, a system for measuring the load required to break the specimen and a system for measuring the distance between the two platen contact points. Two spherically truncated conical platens (an upper and lower), transmitted the load to the specimen. The load measuring system incorporated a maximum-indicating device so that the reading was retained in the case of sudden specimen failure.

There are two variants of the Point Load Test: the diametral test and the axial test. In the diametral test, the specimen is inserted in the test machine and the conical platens advanced to make contact along a core diameter, ensuring that the distance, L, between the contact point and the nearest free end is at least 0.7 D. In the axial test, the specimens

are loaded parallel to the core axis. In both cases, the inclination of layering, foliation or other plane of weakness is recorded with respect to the line of loading.

For this project only axial loading was conducted, as the length of core available was insufficient to allow for multiple samples for both diametral and axial loading, as well as Brazilian and uniaxial compression tests. Sampling was also limited by the fact that there is sufficient fabric heterogeneity at the core-scale to make it difficult to limit samples (for a given test type) to a single texture/ composition / fabric domain.

## 7.4. Determination of Orientation

The orientation of all fabric elements and test-induced fractures in the samples were recorded so that the compressive and tensile strength results could be compared to the results from tunnel mapping of rock structure, to the natural microcrack investigations, and to the in-situ stress data. and rock structure.

The upper figure shows the angles necessary to convert relative orientations to true orientations, given the trend and plunge of the borehole

- This space intentionally left blank for effective pagination -

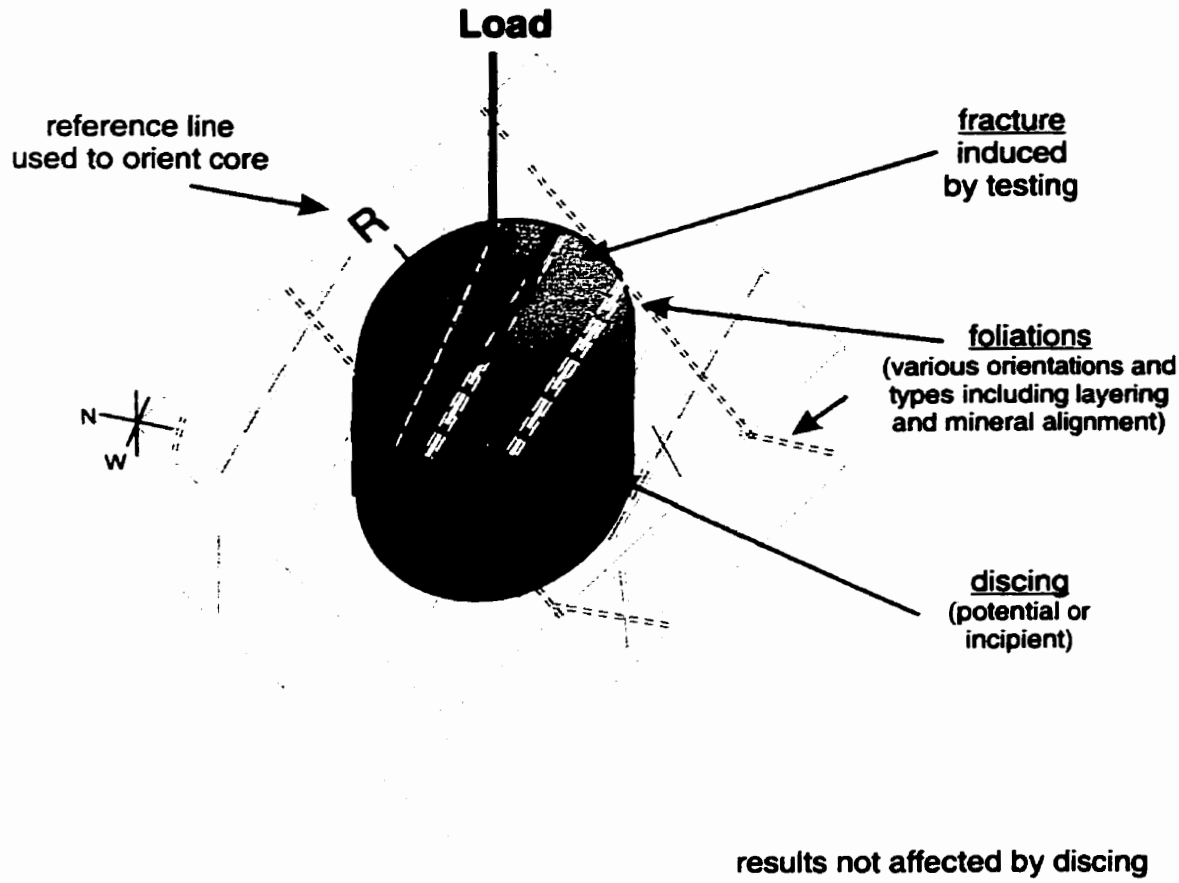



Figure 43. Point Load testing, rotation of loading direction relative to foliation. Rock cylinders were placed vertically in a loading frame and subjected to axial point loading as shown. The orientation of the test-induced fracture, and the fabrics present were recorded, and converted to give the true orientation.

- This space intentionally left blank for effective pagination -

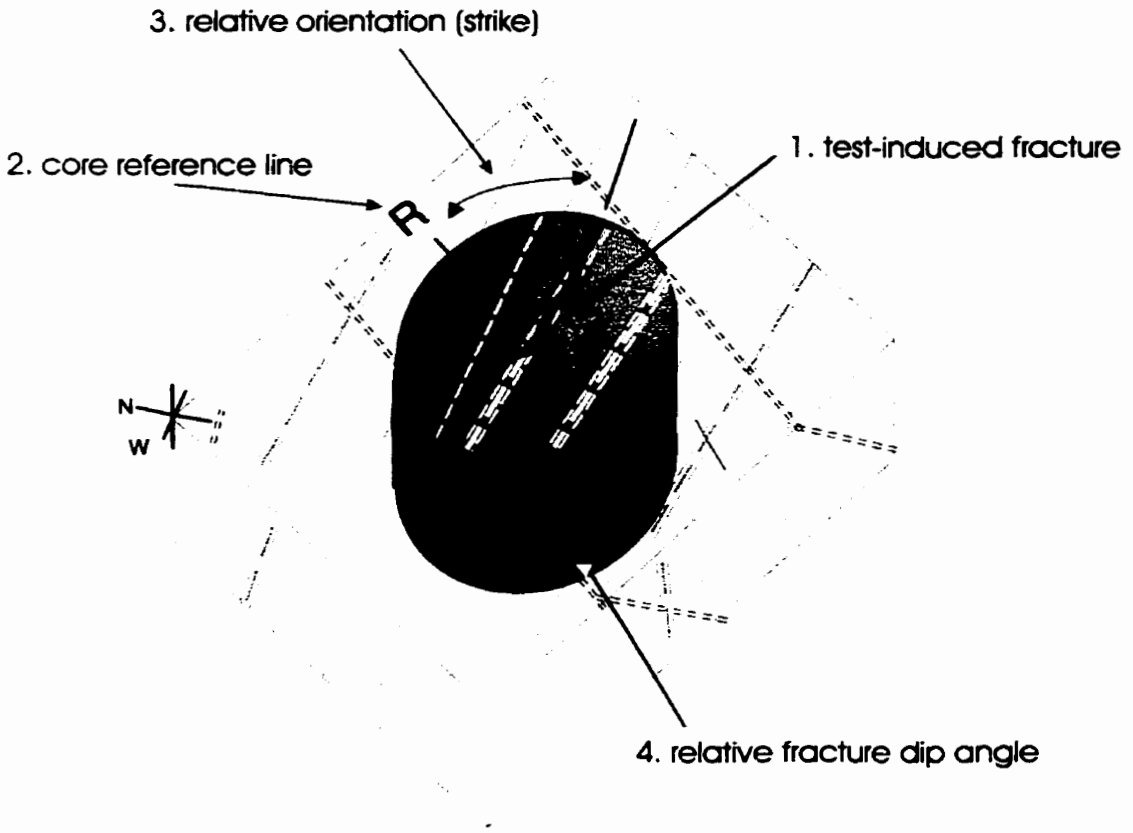



Figure 44. Method for recording orientation of test-induced fracture.

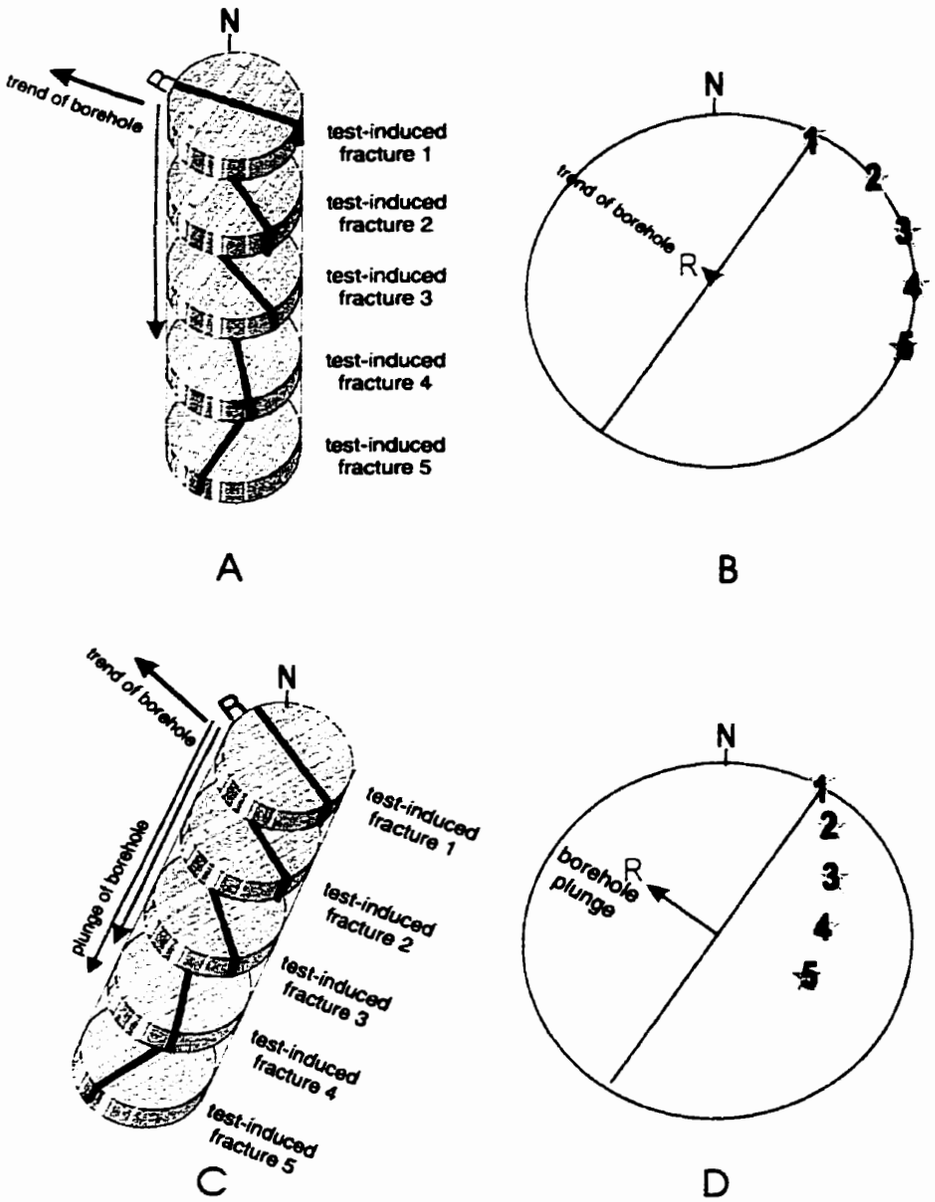



Figure 45. Conversion of relative orientations to true orientations. In each case, the "R" and red arrow represent the reference line painted on the up-dip side of the core, and oriented to show the trend and plunge directions. A: A series of samples with test-induced fractures. In this view, the core is aligned along the trend of the borehole, but "held" in a vertical orientation so that it plunges 90°. The strike of each fracture is recorded relative to the reference line. B: The data from A plotted on a lower hemisphere equal area projection, with the reference line and the poles to the induced fractures indicated. C: As A, but with the core rotated to the correct plunge for this borehole. D: Lower hemisphere equal area projection of the data in C. The reference line "R" rotates from its position in B, to the location shown here. The poles to each fracture rotate in the same direction and orientation.

Table 13.
Average Point Load Strength Index by Rock Type and Borehole.

| | Fine Granite | | | | Medium | | Coarse Granite | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | RT1 | RT2 | RT3 | RT4 | RT4 | RT5 | RT2 | RT3 | RT4 | RT5 |
| Count | 10.0 | 10.0 | 10.0 | 13.0 | 6.0 | 14.0 | 15.0 | 6.0 | 6.0 | 11.0 |
| Mean | 21.8 | 19.9 | 23.3 | 20.5 | 16.4 | 16.2 | 15.0 | 14.4 | 15.3 | 15.1 |
| Median | 21.3 | 20.3 | 23.4 | 20.6 | 15.7 | 16.1 | 15.3 | 14.3 | 15.2 | 14.1 |
| Maximum | 24.8 | 22.2 | 27.0 | 23.2 | 21.0 | 19.7 | 16.8 | 15.9 | 17.5 | 18.0 |
| Minimum | 19.9 | 16.2 | 19.3 | 17.3 | 13.9 | 13.8 | 13.4 | 13.4 | 13.4 | 13.2 |
| Skewness | 0.9 | -0.9 | -0.2 | -0.2 | 1.7 | 0.7 | 0.0 | 1.0 | 0.3 | 0.7 |
| Kurtosis | -0.5 | 0.0 | -1.2 | -0.1 | 3.5 | 0.6 | -1.4 | 1.5 | -1.1 | -1.1 |
| Std. dev. | 1.7 | 1.9 | 2.8 | 1.6 | 2.4 | 1.5 | 1.2 | 0.9 | 1.5 | 1.8 |

Table 14.
Average Point Load Strength Index by Rock Type and Borehole.

| | Fine Granite | Medium | Coarse Granite |
|---|---|---|---|
| Maximum | 27.0 | 21.0 | 18.0 |
| + *1 std.dev* | 23.4 | 18.3 | 16.3 |
| Mean | **21.4** | **16.3** | **15.0** |
| - *1 std.dev* | 19.4 | 14.3 | 13.6 |
| Minimum | 16.2 | 13.8 | 13.2 |
| Skewness | -0.1 | 1.2 | 0.5 |
| Kurtosis | -0.4 | 2.0 | -0.5 |
| Std. dev. | 2.0 | 2.0 | 1.3 |

## 7.5. Summary of Point Load Testing Results

### 7.5.1. Presentation

The data from each rock type are presented in the following format and order:

1. summary tables for each rock type, which lists the boreholes, the samples from each borehole, the sample densities, and the tensile strengths recorded (Table 13 and 14).
2. a summary figure of the average point load indices grouped by grain size to indicate the relative strength of each rock type (Figure 46). The resultant ranking is consistent with that determined from the uniaxial compression test data, but much of the variability shown is in fact direction dependent.
3. For each rock type and borehole, a pair of histograms which show the strike of the test-induced fractures, and the strike of the foliations in the same samples (Figures 47 to 51).

Additional information is provided in Appendix A. Drawings of the foliations and test-induced fractures in each of the samples. The axial point load testing results, including the sample dimensions, load at failure, the index and the index normalized to the average sample size are also included.

This space intentionally left blank for effective pagination.

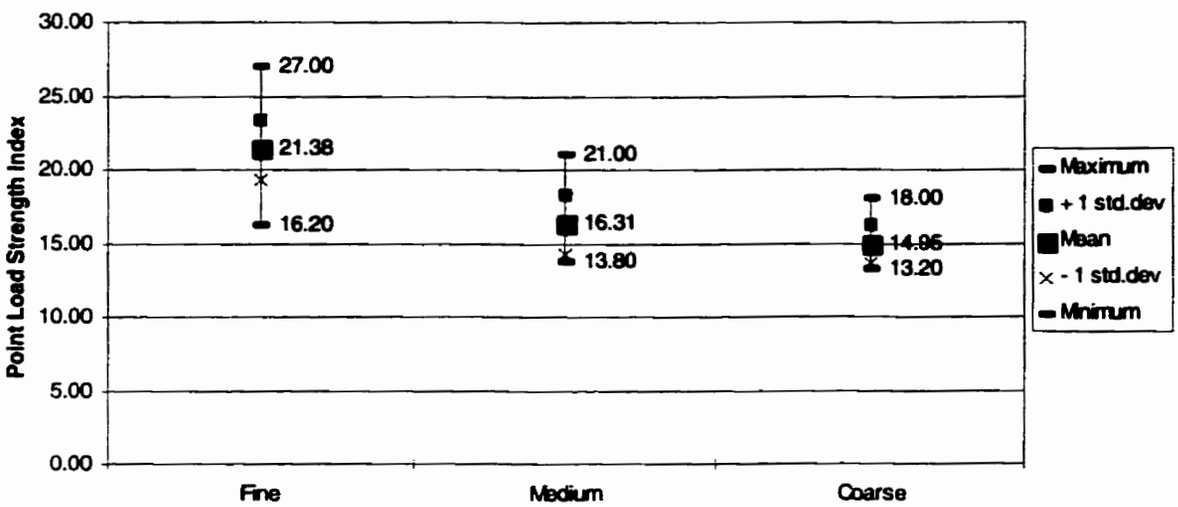



Figure 46. Comparison of Point Load Strength Indices by rock type (relative average grain size).

### 7.5.2. General Observations.

The Point Load testing indicated that failure was controlled by heterogeneities in the rock, such as foliations or the preferred orientation of natural microcracks, which are close to or parallel to the plane of failure under the various loading geometries tested. This affect was present to some degree in all three of the rock types tested, but varies systematically with the rock heterogeneity, and with the relative orientation of the loading direction. The Coarse Grained Granite has the most repeated tendency to split along dominant (low-dipping) fabric elements, with the lowest average tensile strength and Point Load strength index. The Medium Grained granite has a lesser tendency to split along dominant (and other) fabric elements, but with a comparable average tensile strength and Point Load strength index. The Fine Grained Granite has the least repeated tendency to split along any fabric, which in this case alternated between the steeply-dipping flow banding and the plane correlating with the preferred orientation of natural microfractures, and with the highest average tensile strength and Point Load strength index.

This systematic variation reflects the relative strength of the low-dipping fabric, which comprises a mixture of coplanar compositional layering, mineral alignment, and natural microcracks. All of these elements are present and well developed in the coarse granite, the weakest and most mechanically anisotropic of the group. Where this fabric is less well developed or absent, as in the medium and fine grained granites, this is reflected by an increase in rock strength and isotropy.

### 7.5.3. Influence of Grain Size on Tensile Strength

As noted earlier in this presentation, the Point Load test provides a relative index of the tensile strength of the various rock types (the Point Load Index, Goodman, 1989), although this was a secondary objective of this experiment. The result, shown in Figure 46 and in Tables 13 and 14, is consistent with the uniaxial compression testing described in Chapter 9, in that relative strength decreases with increasing grain size.

### 7.5.4. Influence of Fabric on Tensile Strength

The principal objective of applying the Axial Point Load Test in this study was to determine if there were preferred orientations of failure, and if these existed, their correlation with rock fabric. In the axial point configuration, the fracture is free to form in any orientation. If however, the rock is anisotropic, there is a tendency for failure to follow one or more preferred orientations that generally correspond with some fabric element such as natural microfractures, mineral alignment, etc. The results of this testing for the coarse, medium and fine grained granites are presented for each rock type in the following sections.

### 7.5.5. Coarse Grained Granite, Borehole 418-051-RT 2, 3, 4, 5.

The results of point load testing in each of the 4 boreholes from the coarse grained granite are shown in Figure 47, and in Tables 13 and 14. A pair of histograms for each borehole shows the strike of test-induced fractures, and the strike of the foliations in the same samples. This granite is marked by a well developed low dipping foliation referred to for convenience as the large scale layering, but in fact a mixture of metre to millimetre scale compositional layering, mineral alignment, schlieren (in some cases showing evidence of slip) and natural microfractures in quartz.

In this rock type, there is a very clear tendency for the tensile failure to be localized by the foliations present in each sample. Which foliation is responsible is dependent upon the orientation of the borehole (and sample) axis. In shallow dipping boreholes (< 45 °), axial point loading of the core produces fractures following the strongest fabric. In steeply dipping boreholes, the layering is not favourably oriented with respect to axial loading, and the plane of failure then "jumps" to other, steeper dipping foliations (this transition is reflected in the values of the average tensile strength), and is shown schematically in Figure 48.

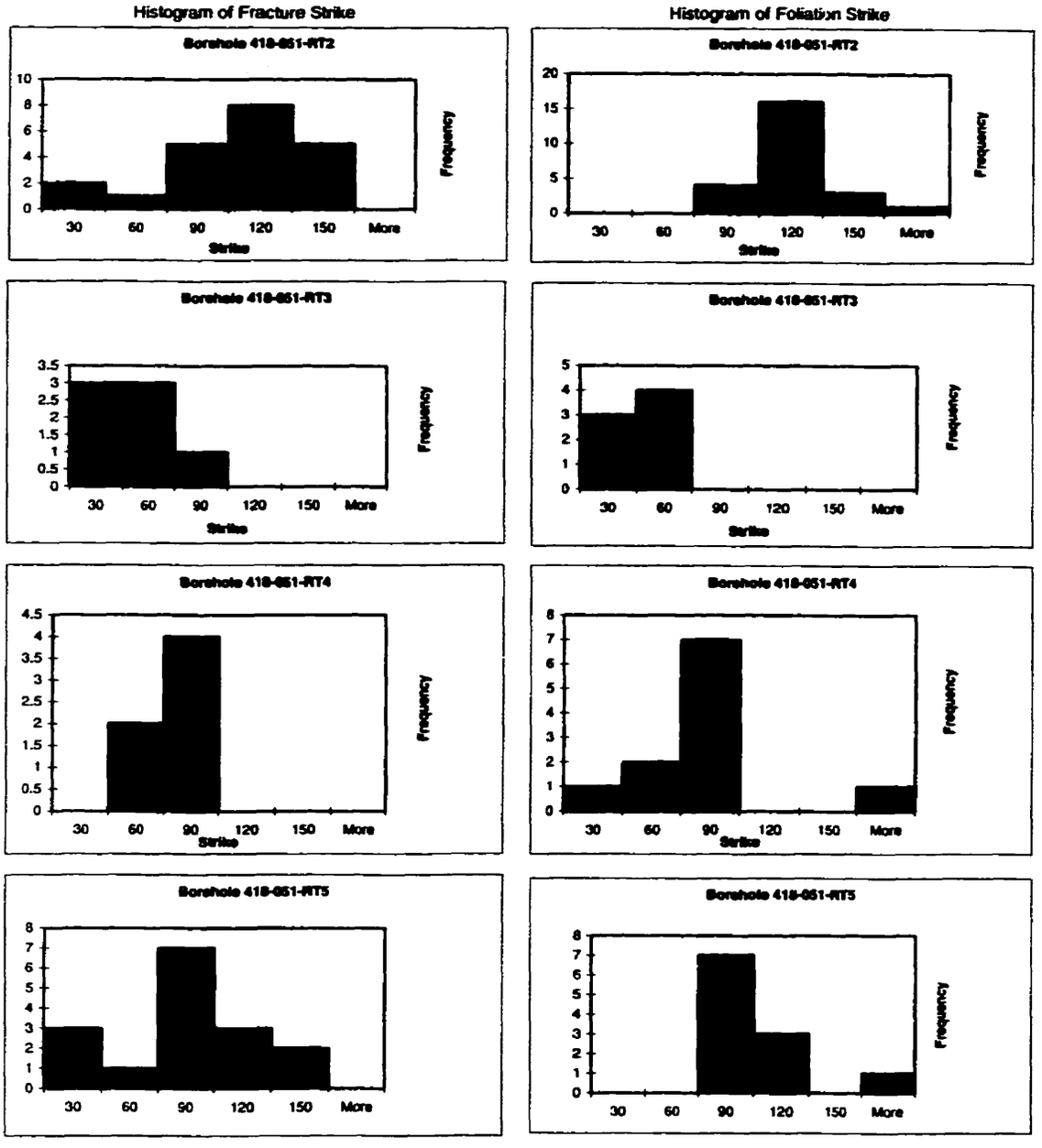



Figure 47. Point Load test results from the coarse grained granite. For each borehole, a pair of histograms shows the strike of test-induced fractures and the strike of the foliations in the same samples.

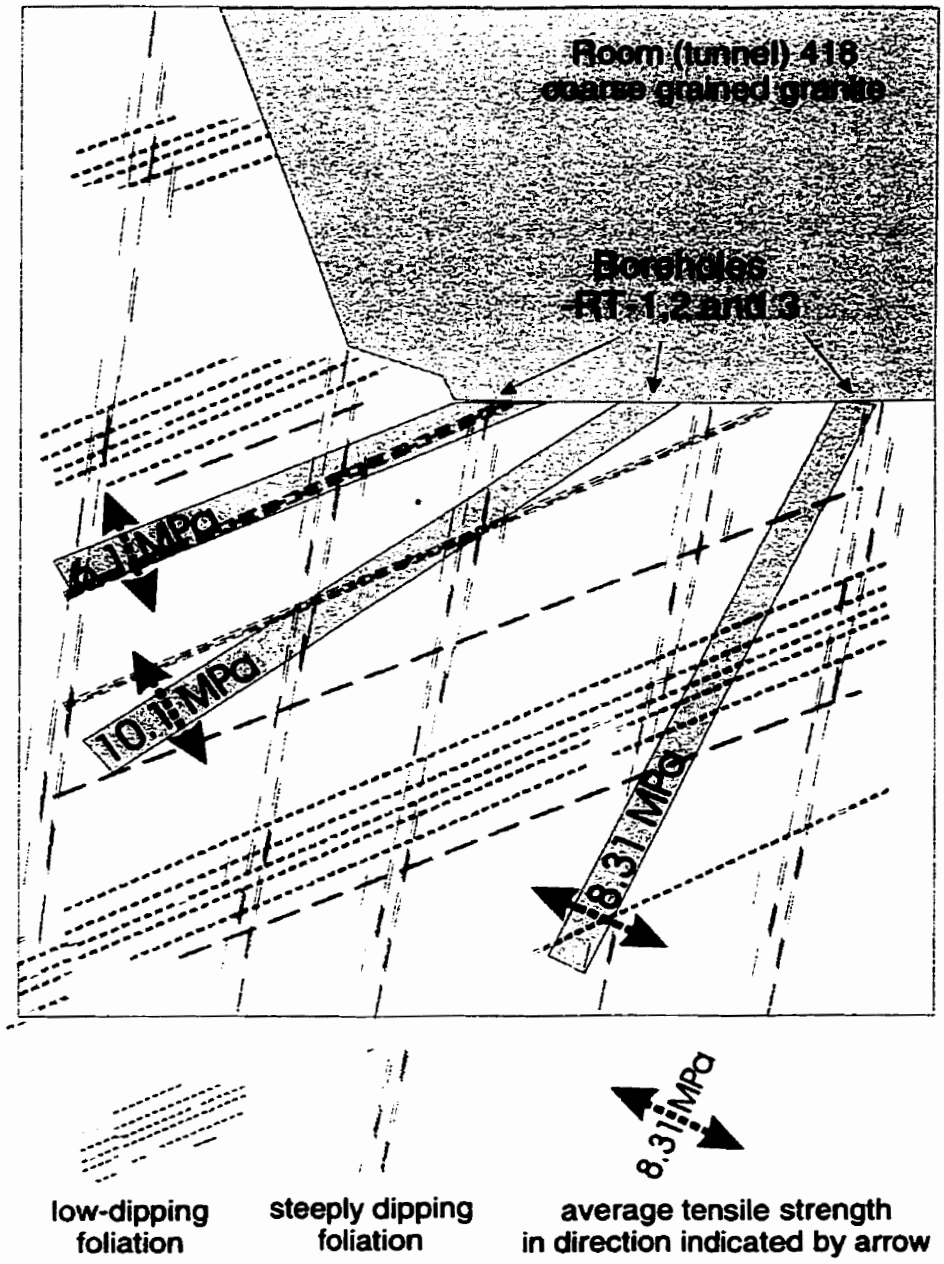



Figure 48. Schematic representation of the variation in tensile strength in the coarse grained granite and its correlation with the fabrics present. In shallow dipping boreholes (< 45 °), axial point loading of the core produces fractures following the strongest fabric (referred to for convenience as the large scale layering, but in fact a mixture of metre to millimetre scale compositional layering, mineral alignment, schlieren (in some cases showing evidence of slip) and natural microfractures in quartz). In steeply dipping boreholes, the layering is not favourably oriented with respect to axial loading, and the plane of failure then "jumps" to other, steeper dipping foliations (this transition is reflected in the values of the average tensile strength). These relationships are shown schematically above.

### 7.5.6. Medium Grained Granite, Borehole 417-071-RT 4, 5.

The results of point load testing in each of the 2 boreholes from the medium grained granite are shown in Figure 49, and in Tables 13 and 14. A pair of histograms for each borehole shows the strike of test-induced fractures, and the strike of the foliations in the same samples. This granite is marked by a weakly developed low dipping foliation referred to for convenience as the large scale layering, but which is in fact a mixture of metre - scale compositional layering, mineral alignment, and natural microfractures in quartz.

In this rock type, there is also a tendency for the tensile failure to follow the direction of the large scale layering, but the correlation is not as strong as in the coarse grained, well-foliated granite, and the load required to break the rock along this plane is, on average, greater.

- This space intentionally left blank for effective pagination -

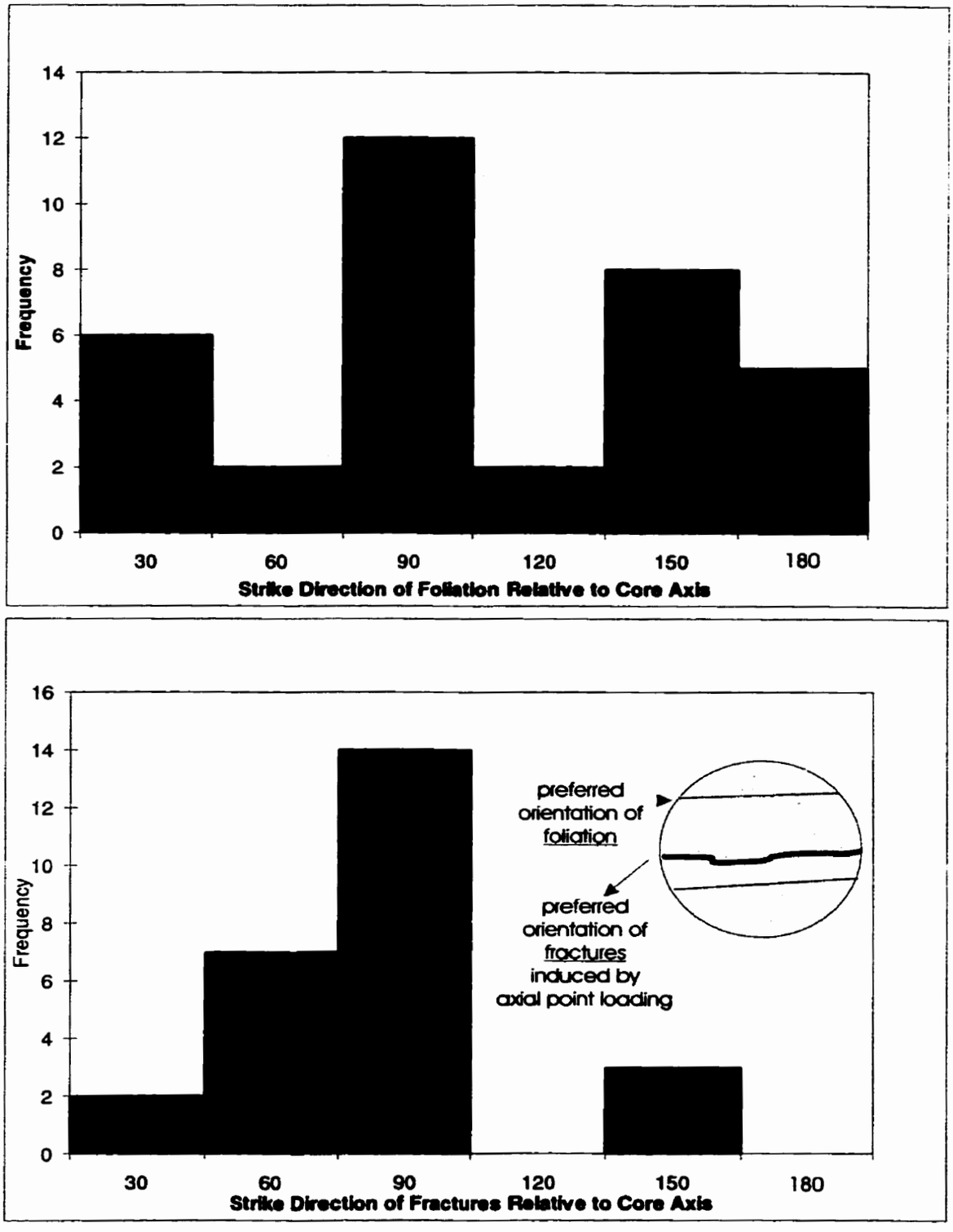



Figure 49. Point Load test results from the medium grained granite. This pair of histograms shows the strike of the foliations (upper) and the strike of the test-induced fractures (lower) from the same samples.

### 7.5.7. Fine Grained Granite, Borehole 421-032-RT1, 2, 3, 4.

The results of point load testing in each of the 4 boreholes from the fine grained granite are shown in Figure 50, and in Tables 13 and 14. A pair of histograms for each borehole shows the strike of test-induced fractures, and the strike of the foliations in the same samples. This granite is marked by a well developed subvertical flow banding, and a low-dipping natural microcrack set.

In this rock, the tendency is for splitting to occur in the plane of the natural microcracks, and to a lesser extent, along the subvertical layering. This is consistent with the Brazilian tests.

The links between rock structure and the preferred orientations of the fractures induced by the point load tests are also shown in Figures 51 and 52. The former is a summary plot of poles to fractures induced by axial point loading in each of the specimens taken from the 4 boreholes, each oriented at different angles to the foliation. The latter is a lower hemisphere equal area plot of poles to natural fractures (fluid inclusion or alteration-filled planes in quartz) obtained by universal stage mapping of thin section from the granodiorite. The correlation between the test-induced fractures (Figure 51) and the natural microcracks (Figure 52) is obvious.

The pattern of point load induced fractures is dominated by the low-dipping to subhorizontal fractures (the vertical sets are quantitatively minor in both outcrop and in specimens). The orientation of the low-dipping fractures is consistent with that of the natural microcracks (Figure 52).

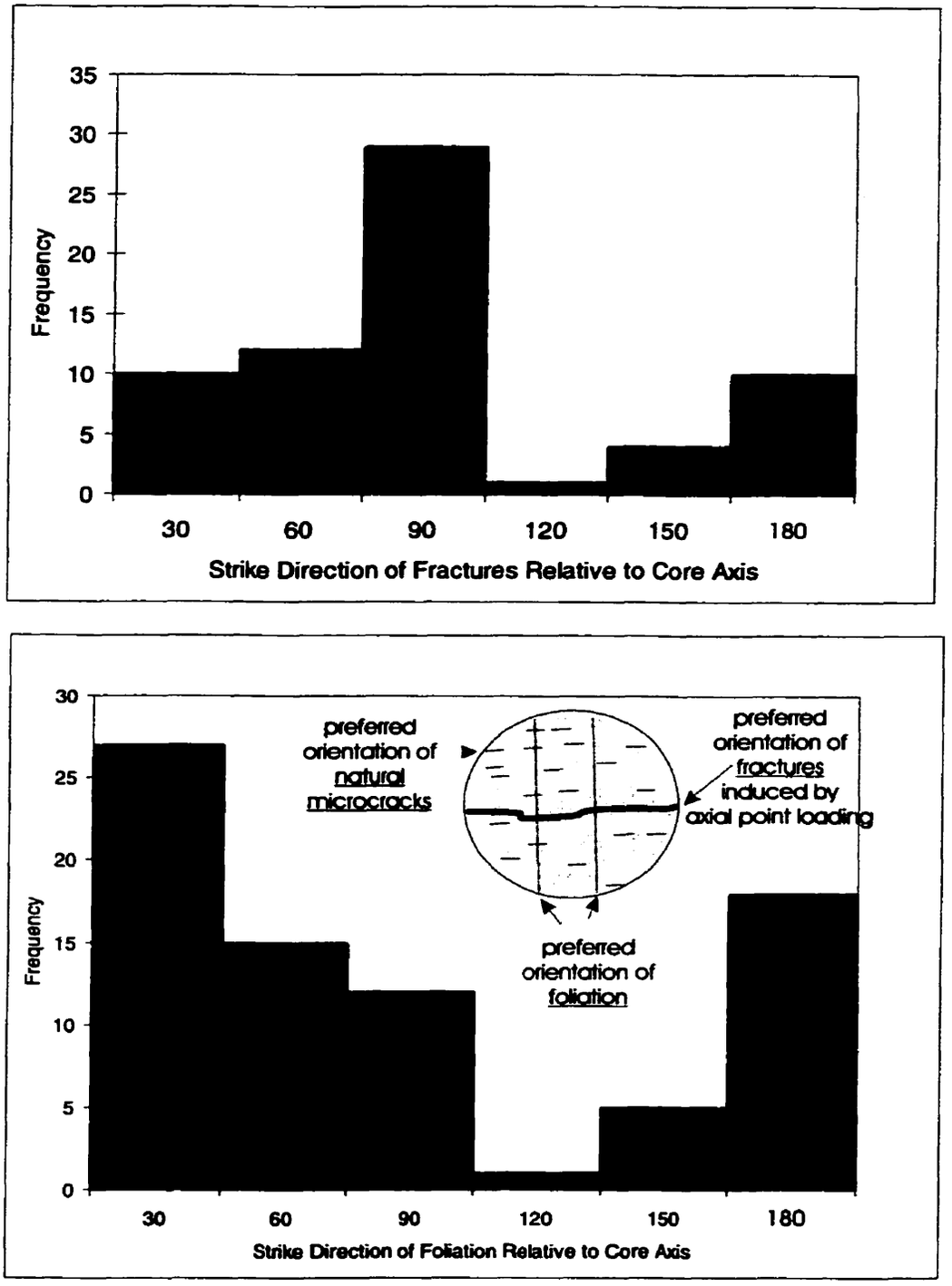



Figure 50. Point Load test results from the fine grained granite. The histograms show the strike of the test-induced fractures (upper), and the strike of the foliations (lower) from the same samples. This granite hosts a well developed subvertical flow banding and a low-dipping natural microcrack set. Fractures induced by axial point loading preferentially follow the (low-dipping) plane of the natural microcracks (which strike ~90° in this 2-D view), the subvertical foliation being ignored. This relationship between the natural microcrack set and the point-load induced fractures is displayed to better effect in Figure 51 following.

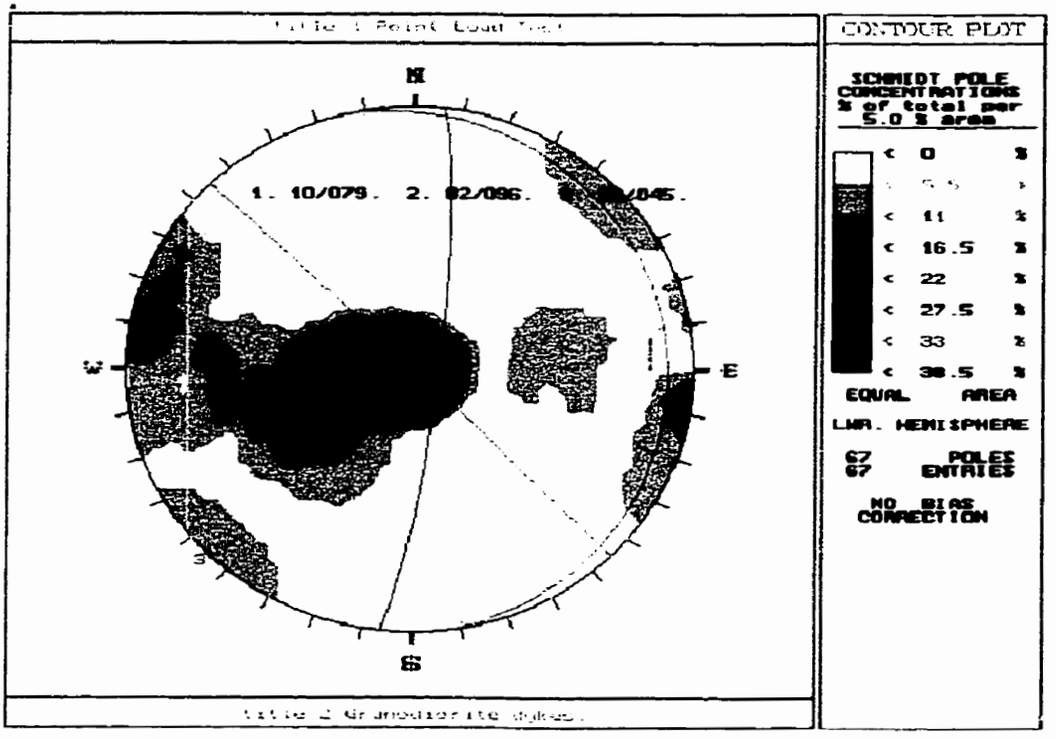



Figure 51. Lower hemisphere equal area plot of poles to fractures induced by axial point loading, in fine grained granite from boreholes 421-032-RT1,2,3 and 4.

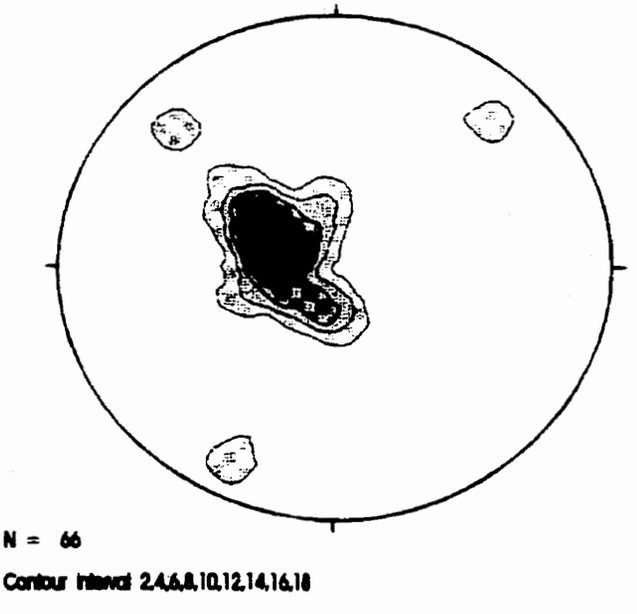



Figure 52. Lower hemisphere equal area plot of poles to natural fractures (fluid inclusion or alteration-filled planes in quartz) obtained by universal stage mapping of thin section from the granodiorite. The induced fractures in Figure 51 share the same general orientation as the natural fractures shown here.

# Chapter 8

# BRAZILIAN TESTS

The Brazilian test method is a convenient but indirect method for estimating the tensile strength (Goodman, 1989). It has been found that a rock core will split along the diameter and parallel to the cylinder axis when line loaded on its side in a compression machine. The reason for this can be demonstrated by examining the stress inside a disk loaded at opposite sides of a diametral plane. In such a configuration the tensile strength is calculated from:

$$Brazilian Tensile\ strength,\ \sigma_{tB} = \frac{2P}{\pi Dt} \qquad \text{Equation 1}$$

where P = load at failure, D = diameter of the specimen, and t = thickness of the specimen

The "Brazilian tensile strength - $\sigma_{tB}$" is estimated from the peak compression load. The actual cause of failure may in fact be more complex, reflecting the interaction of small-scale flaws, compression parallel to the eventual rupture plane, and the horizontal tension induced by this loading.

The Brazilian test is an indirect but popular method, as it is much easier to perform this test than to arrange the precise alignment and end preparation required for a direct tensile test (Goodman, 1989). According to Goodman (1989), the estimates it provides are

generally higher than those produced by direct tension tests, and the difference between estimates increases with increasing grain size and heterogeneity. This is attributed to the effect of grain-scale (and larger) discontinuities, in which "...*larger fissures in the coarse rocks weaken a direct tension-specimen more severely than they weaken a splitting tension specimen* .... , The ratio of Brazilian to direct tensile strength has been found to vary from one to more than ten as the length of preexisting fissures grows larger.

In this study, values for tensile strength were also calculated using the modification suggested by Chen (1993), which includes a correction needed for modulus values owing to the presence of microcracks. The correct tensile strength can be approximated by the following function:

$$\sigma_t = 1.023k^{0.252}\left(\frac{P}{\pi R}\right) \qquad \text{Equation 2}$$

where k = ratio of Young's moduli in tension and compression

(the value used was that of the Westerly granite (0.24) given on p. 79 of Chen, 1993)

A comparison of values calculated using both equations is given in Table 15. The results are consistent with the general observation, with the coarser rocks being most strongly affected.

Table 15.
Comparison of Calculated Brazilian Tensile Strengths

| Average tensile strength | Fine | Medium | Coarse |
|---|---|---|---|
| From equation 1 | 7.0 MPa | 5.0 MPa | 6.0 MPa |
| From equation 2 | 2.4 MPa | 1.25 MPa | 1.5 MPa |
| Ratio (2/1%) | 34% | 25% | 25% |

The testing methodology used in this study followed the recommendations laid out in Brown (1981); Goodman (1989); Gyenge (1980), and in the Pit Slope Manual Supplement 3-5 (Gyenge and Ladanyi, 1977). All specimens were prepared in accordance with the procedure described in the Pit Slope Manual Supplement 3-5 (Gyenge and Ladanyi, 1977) and the method suggested by the International Society for Rock Mechanics (Brown, 1981).

Rock cylinders were placed vertically in a loading frame and subjected to line-loading in compression along a preset diameter at a known angle to some reference line or foliation (Figure 53), and then repeated through a series of angles (Figure 54). These angles were recorded, and converted to give the true orientation of the tensile strength vector so determined (Figure 55). The tests were performed by the author, at the CANMET Laboratories at Bells Corners, Ottawa. The compression machine used included a load measuring system for indicating the failure load, and a maximum-load indicating device, i.e. one which retains the maximum load so that the latter can be recorded after specimen failure loading platens.

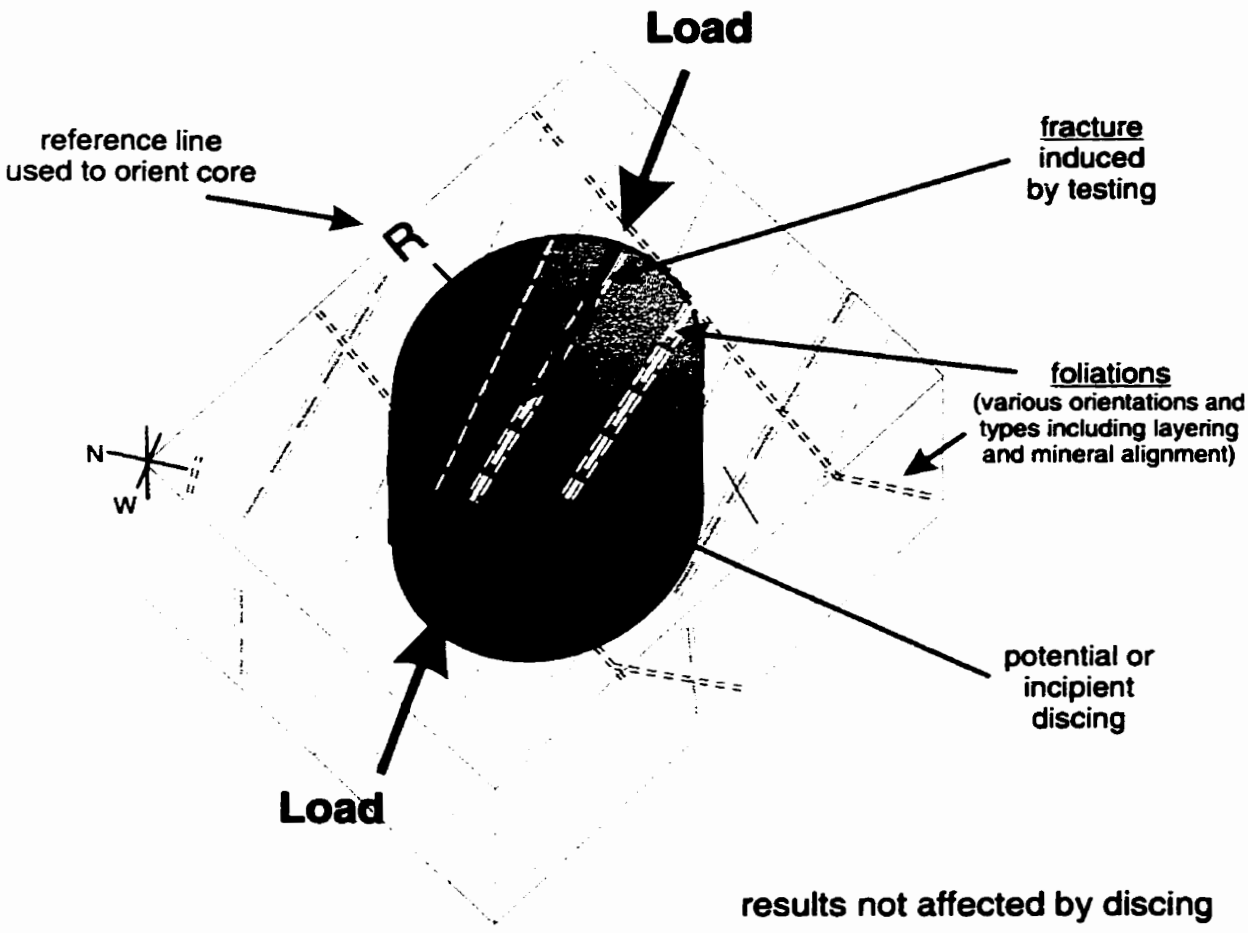



Figure 53. Configuration of loading for the Brazilian test as seen on a single specimen.

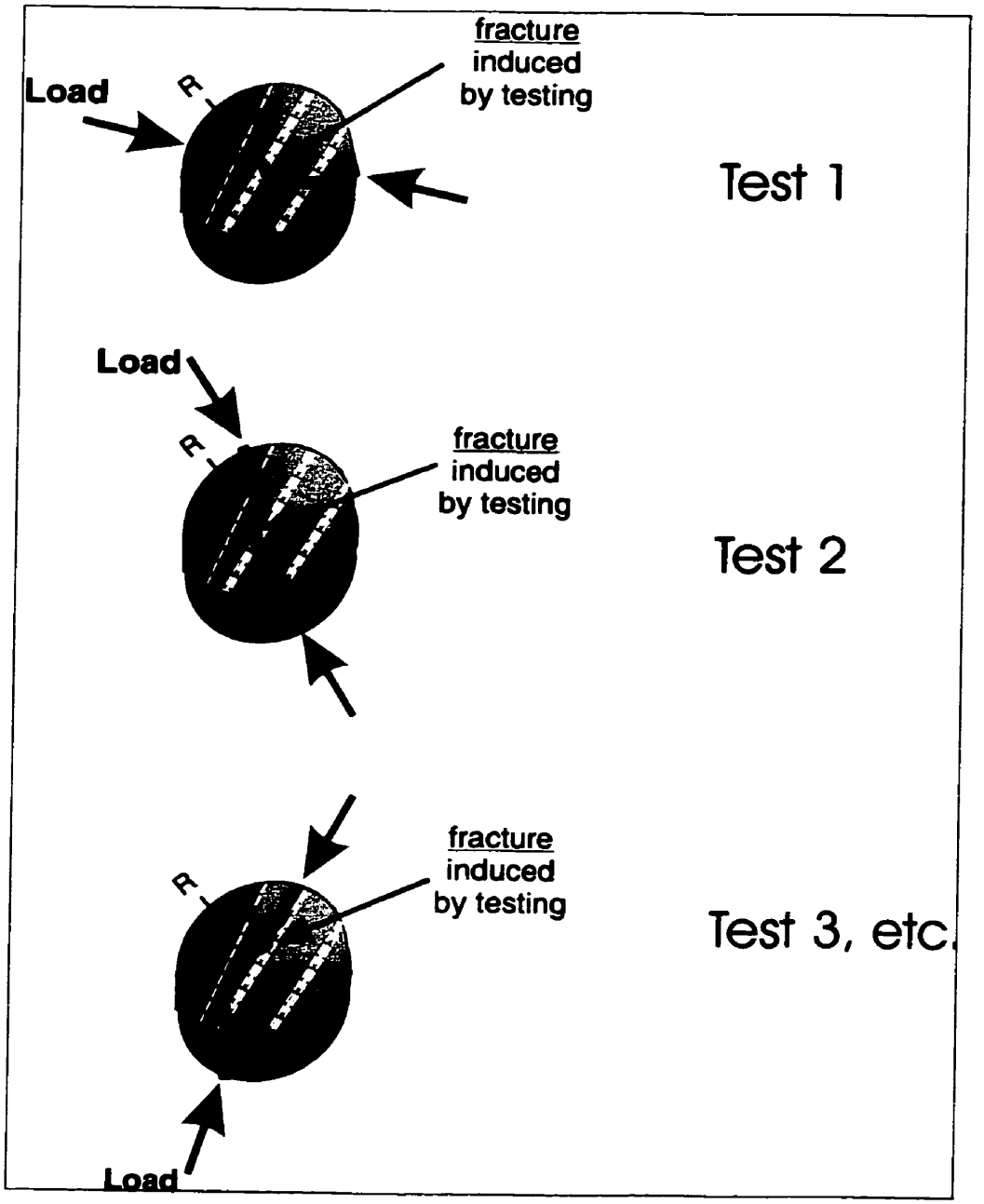



Figure 54. Brazilian tests, rotation of loading direction relative to foliation.

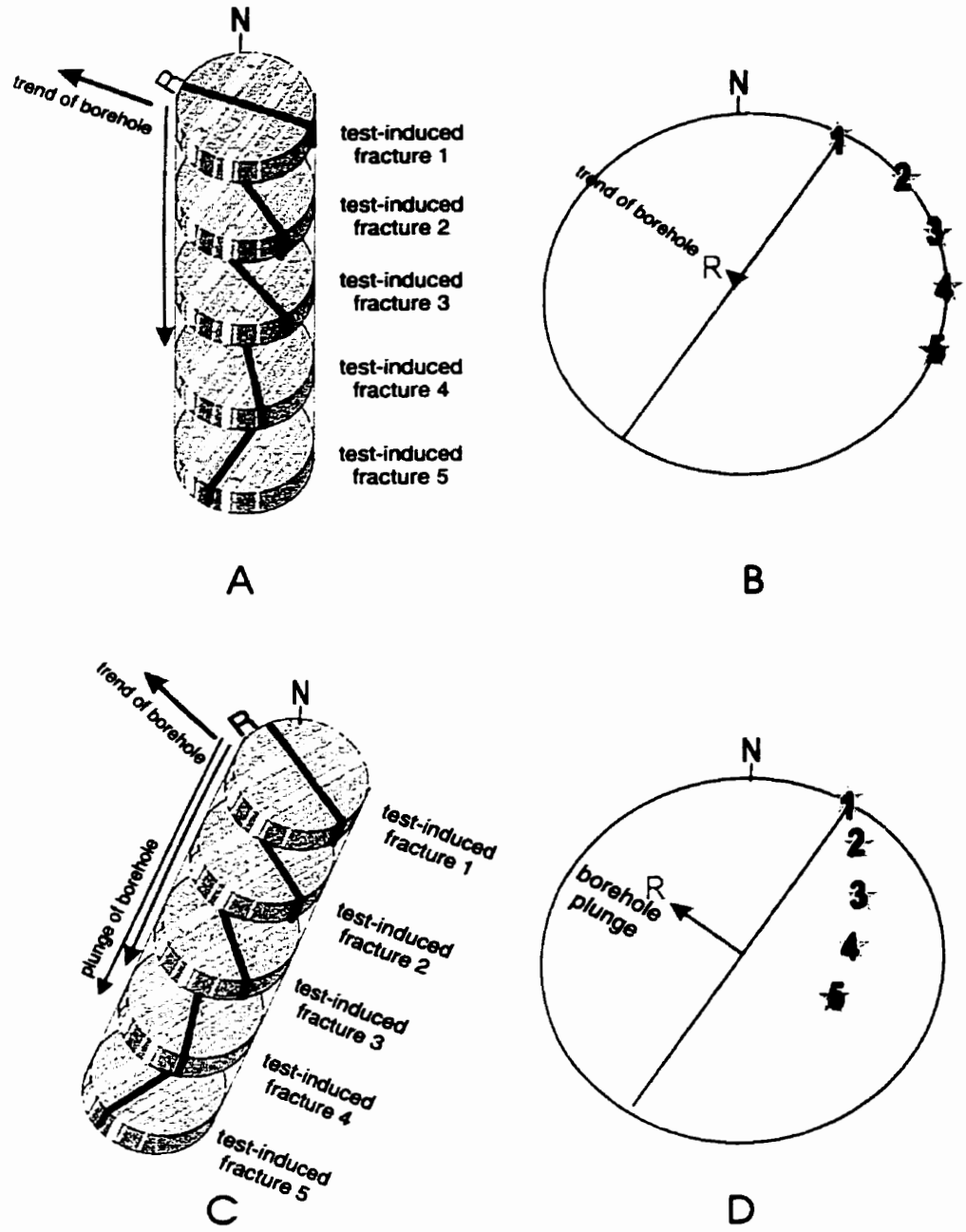



Figure 55. Conversion of relative orientations to true orientations. In each case, the "R" and red arrow represent the reference line painted on the up-dip side of the core, and oriented to show the trend and plunge directions. A: A series of samples with test-induced fractures. In this view, the core is aligned along the trend of the borehole, but "held" in a vertical orientation so that it plunges 90°. The strike of each fracture is recorded relative to the reference line. B: The data from A plotted on a lower hemisphere equal area projection, with the reference line and the poles to the induced fractures indicated. C: As A, but with the core rotated to the correct plunge for this borehole. D: Lower hemisphere equal area projection of the data in C. The reference line "R" rotates from its position in B, to the location shown here. The poles to each fracture rotate in the same direction and orientation.

## 8.1. Rock Type and Specimen Preparation

Samples used in these tests were obtained from the boreholes listed in Table 16. All samples are segments of HQ- sized core (46 mm diameter), the choice of specimen diameter being largely dictated by budget constraints. Specimens were prepared in accordance with the procedure described in the Pit Slope Manual Supplement 3-5 (Gyenge and Ladanyi, 1977) and the method suggested by the International Society for Rock Mechanics (Brown, 1981). The end surfaces of each of the specimens were ground flat to 0.015 mm to ensure that they were parallel to each other and perpendicular to the axis of the specimen.

The average dimensions and bulk densities of samples from each of the three rock types are listed in Table 17. The mean bulk density ranged from 2.54 to 2.72 gm/cm$^3$; these corresponding to the leucocratic and melanocratic layers in the coarse grained granite. Specimen lengths varied from 22.03 to 28.36 mm and diameters varied from 44.43 to 44.91 mm, with length to diameter ratios ranging from 0.56 to 0.58.

- This space intentionally left blank for consecutive pagination -

## Table 16
## Sample and testing summary

| Granite Type | Borehole | Orientation | | test type and number | | |
|---|---|---|---|---|---|---|
| | | trend | Plunge | uniaxial compression | Brazilian | Axial point load |
| Coarse grained | 418-051-RT1 | 324.90° | 77.10° | | *Discing* | |
| | 418-051-RT2 | 143.61° | 54.76° | 1 | **8** | 15 |
| | 418-051-RT3 | 143.40° | 35.40° | 1 | **8** | 6 |
| | 418-051-RT4 | 143.56° | 19.70° | 1 | **7** | 6 |
| | 418-051-RT5 | 143.38° | 0.40° | 1 | **4** | 12 |
| Medium grained | 403-014-MB2 | 225° | 00° | 1 | | |
| | 417-071-RT1 | not drilled | | | | |
| | 417-071-RT2 | 190.43° | 59.44° | | *Discing* | |
| | 417-071-RT3 | 190.20° | 39.48° | | *Discing* | |
| | 417-071-RT4 | 189.10° | 19.50° | 1 | **6** | 6 |
| | 417-071-RT5 | 189.70° | -0.20° | 1 | **4** | 8 |
| Fine grained | 421-032-RT1 | 133.99° | 0.23° | 1 | **4** | 10 |
| | 421-032-RT2 | 157.61° | 0.46° | 1 | **6** | 7 |
| | 421-032-RT3 | 175.81° | 0.80° | 1 | **5** | 10 |
| | 421-034-RT4 | 199.05° | 0.00° | 1 | **6** | 13 |

## Table 17.
## Sample Properties

| Granite type | Number of samples | Diameter mm min. –mean– max | Thickness Mm min. –mean– max | Average T/D ratio | Density gm/cm$^3$ min. –mean– max | Tensile Strength (MPa) [1] min. –mean– max |
|---|---|---|---|---|---|---|
| Fine grained | 21 | 44.82 **44.86** 44.91 | 22.24 **24.93** 28.36 | 0.56 | 2.61 **2.64** 2.67 | 4.62 **2.4** 9.14 |
| Medium grained | 10 | 44.73 **44.77** 44.80 | 22.56 **25.86** 28.06 | 0.58 | 2.60 **2.62** 2.63 | 3.47 **1.25** 7.15 |
| Coarse grained | 27 | 44.43 **44.60** 44.81 | 22.03 **25.05** 27.55 | 0.56 | 2.54 **2.65** 2.72 | 3.71 **1.5** 10.12 |

[1] These are averages based on all boreholes. Tensile strength is in fact anisotropic, and varies with respect to the orientation of the loading axis relative to one or more fabric elements within each rock type.

## 8.2. Brazilian Test Results

The data for each rock type are presented in the following format and order:

- a summary table that lists the boreholes, the samples from each borehole, the sample densities, and the tensile strengths recorded (Tables 18 to 20).
- a summary figure that depicts the variation in average tensile strength between boreholes (Figure 56, 60, 63). This was done to look for directional anisotropy that might not be obvious at the hand-specimen scale of testing.
- an x-y plot of sample density versus their respective tensile strengths (Figures 57, 61, and 64). This was done to see if any variations in tensile strength could be due to sample damage. If the damage zone around the excavation affected the samples, it would be expected that such damage would lower both the rock density and the tensile strength, and that the specimens affected would be closest to the excavation (or to the known stress concentrations about the excavation).
- an x-y plot of the Brazilian tensile strengths of each sample against the inclination of the loading axis relative to the preferred orientation of the dominant fabric element(s) (Figures 58,62,65). Insets on these plots graphically illustrate the orientation of the loading axis (arrows) relative to the flow banding as seen in a cross section of the core. The best-fit linear regression curve, the equation for the curve, and the $R^2$ correlation coefficient are included.

## 8.3. General Observations.

The following observations, though qualitative, illustrate the differences in response to loading which correlate with rock fabric.

1. 23% of the coarse grained granite samples displayed audible cracking well before failure. When loaded parallel to the foliation / mineral banding, failure was often marked by a simple weak "pop", with the sample splitting into two pieces along the foliation plane. In contrast, failure in the medium and fine grained granites was sudden and sharp. These observations are consistent with those from tunnel mapping, and highlight the importance of the foliations as weakness planes.
2. The load rate to failure increased much more slowly in the coarse (and heterogeneous) granite than in the medium and fine grained granites. It is suspected that this behaviour is attributable to a greater degree of crack closure under loading in the coarse grained granite, than occurred in the other finer grained rock types.

## 8.4. Fine Grained Granite, Borehole 421-032-RT1, 2, 3, 4.

Brazilian test results for the Fine Grained Granite from boreholes 421-032-RT1, 2, 3, 4 are summarized in Table 18. This summary includes, for each sample, the loading axis angles relative to structure, and two calculated tensile strengths for each sample. The first strength (max) represents the value at failure. The second strength (min.) includes values in which audible cracking (fracture initiation) preceded failure. As this occurred in only one sample for this rock type, this data set is nearly identical to the first, and is not further discussed.

Summary statistics for the Brazilian tests from the entire borehole array within the Fine Grained Granite are shown in Figure 56. The array of boreholes was arranged in a horizontal plane, and oriented to intersect the subvertical flow banding at angles of approximately 60°, 45°, 10°, and 00°, respectively. The data from each borehole includes loading orientations ranging from 00° to 90° to the foliation, but have been averaged together to look for directional anisotropy at the metre-scale. This was done to test for systematic variation within the horizontal plane, or put in another way - to examine the affect of borehole orientation when sampling for rock properties testing.

The results indicate that there is significant direction -dependent anisotropy.
To determine if this trend is a function of varying degrees of sample damage, the tensile strength for each sample was plotted against its density (Figure 57). The tensile strength of the sample suite decreases with increasing density, which is the opposite of what would be expected if the sample was damaged by discing and enhanced microcracking. This result indicates that the direction anisotropy seen is not due to damage.

The influence of the subvertical flow banding on tensile strength is examined in Figure 58. The tensile strength is somewhat lower when the loading axis is parallel to the flow banding, but the tendency for this is weak, and the data curve could be subdivided into linear segments. This is believed to reflect the influence of additional fabric elements in the rock, most notably the natural microcracks. At this location in the URL, these microcracks are subhorizontal or have a shallow (<30°) southeast dip. Variation in the tensile strength relative to the preferred orientation of the microcracks is shown in

Figure 59. **In this case, the data spread is much less, and the anisotropy is much greater - i.e. the tensile strength parallel to the natural microcrack set is ~ 65% of that measured perpendicular to the microcracks.**

- This space intentionally left blank for effective pagination -

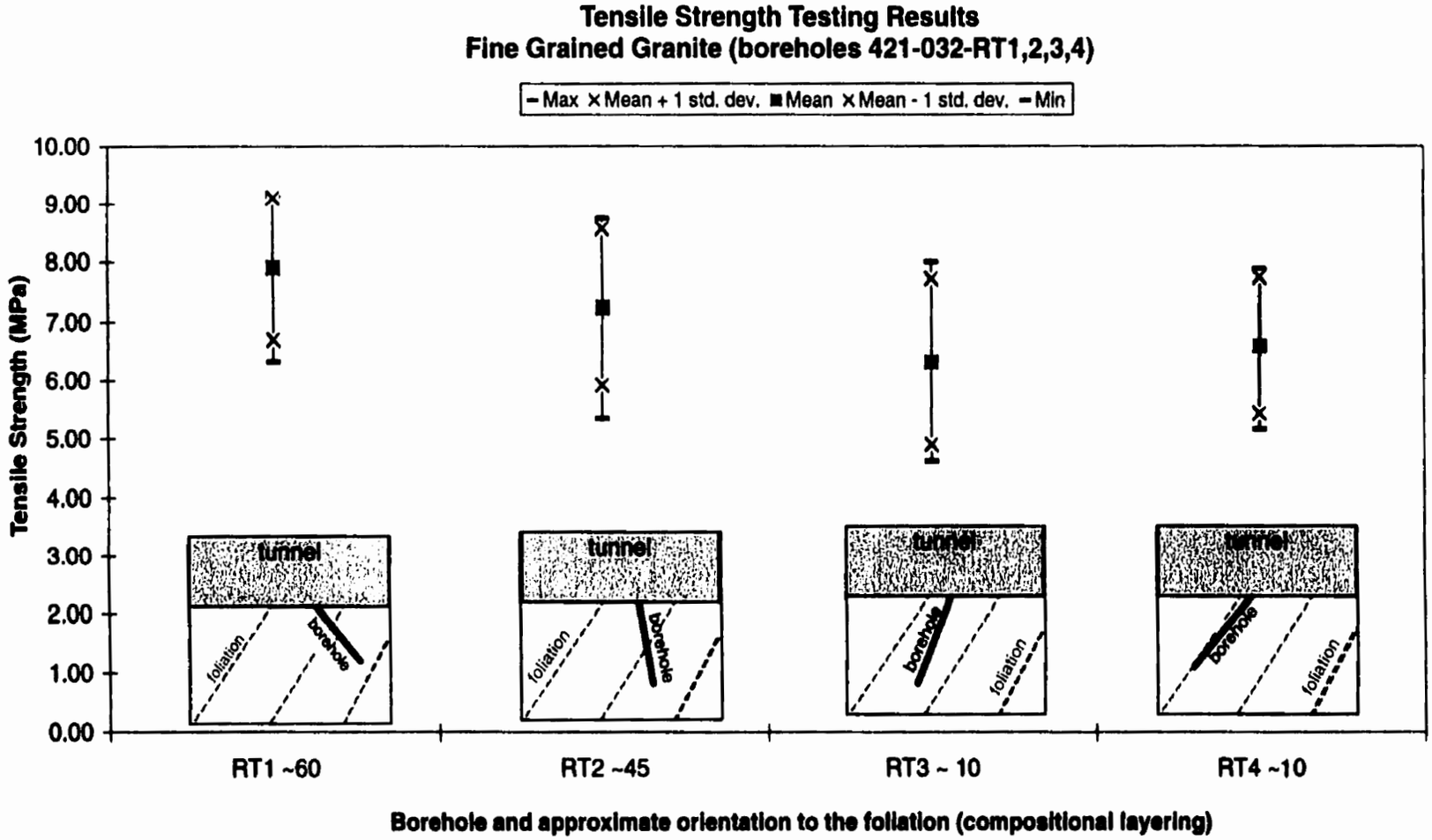



Figure 56. Summary statistics for the Brazilian tests from the entire borehole array within the Fine Grained Granite. The array includes boreholes 421-032-RT1, 2, 3, 4. These were arranged in a horizontal plane, and oriented to intersect the subvertical flow banding at angles of ~60°, ~45°, ~ 10°, and ~00°, respectively. The data from each borehole includes loading orientations ranging from 0° to 90° to the foliation in each borehole. This was done to test for systematic variation within the horizontal plane, or put in another way - to examine the affect of borehole orientation when sampling for rock properties testing. The results indicate that there is significant direction -dependent anisotropy.

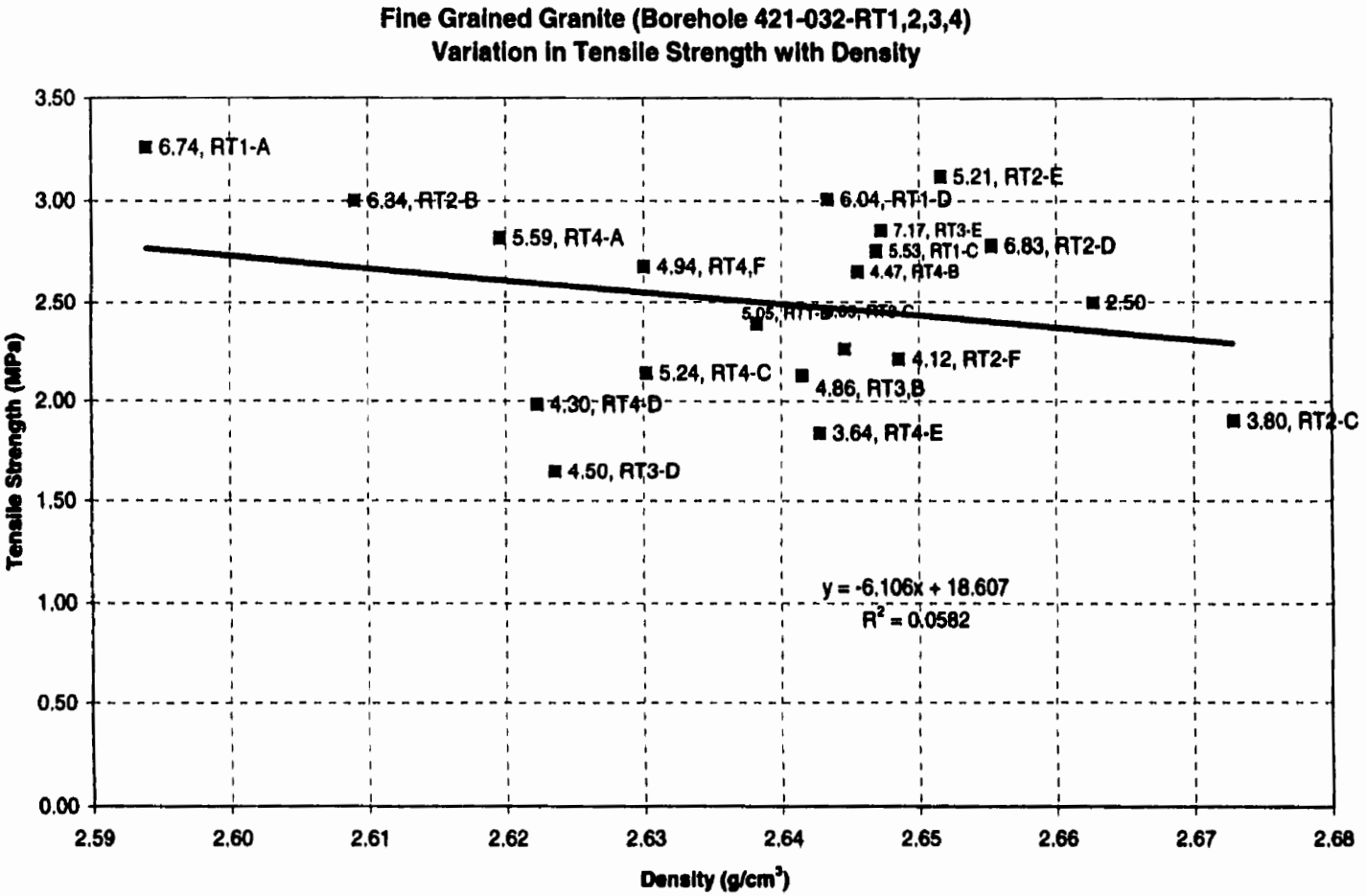



Figure 57. Brazilian test results for the Fine Grained Granite from boreholes 421-032-RT1, 2, 3, and 4, showing the variation in the estimated tensile strengths of each sample plotted against the corresponding sample density. If the damage zone around the excavation affected the samples, it would be expected that such damage would lower both the rock density and the tensile strength, and that the specimens affected would be closest to the excavation. Instead, there is no correlation between density and location (as given by the sample numbers). Tensile strength actually decreases with increasing sample density, which is the opposite of what would be expected if damage was present.

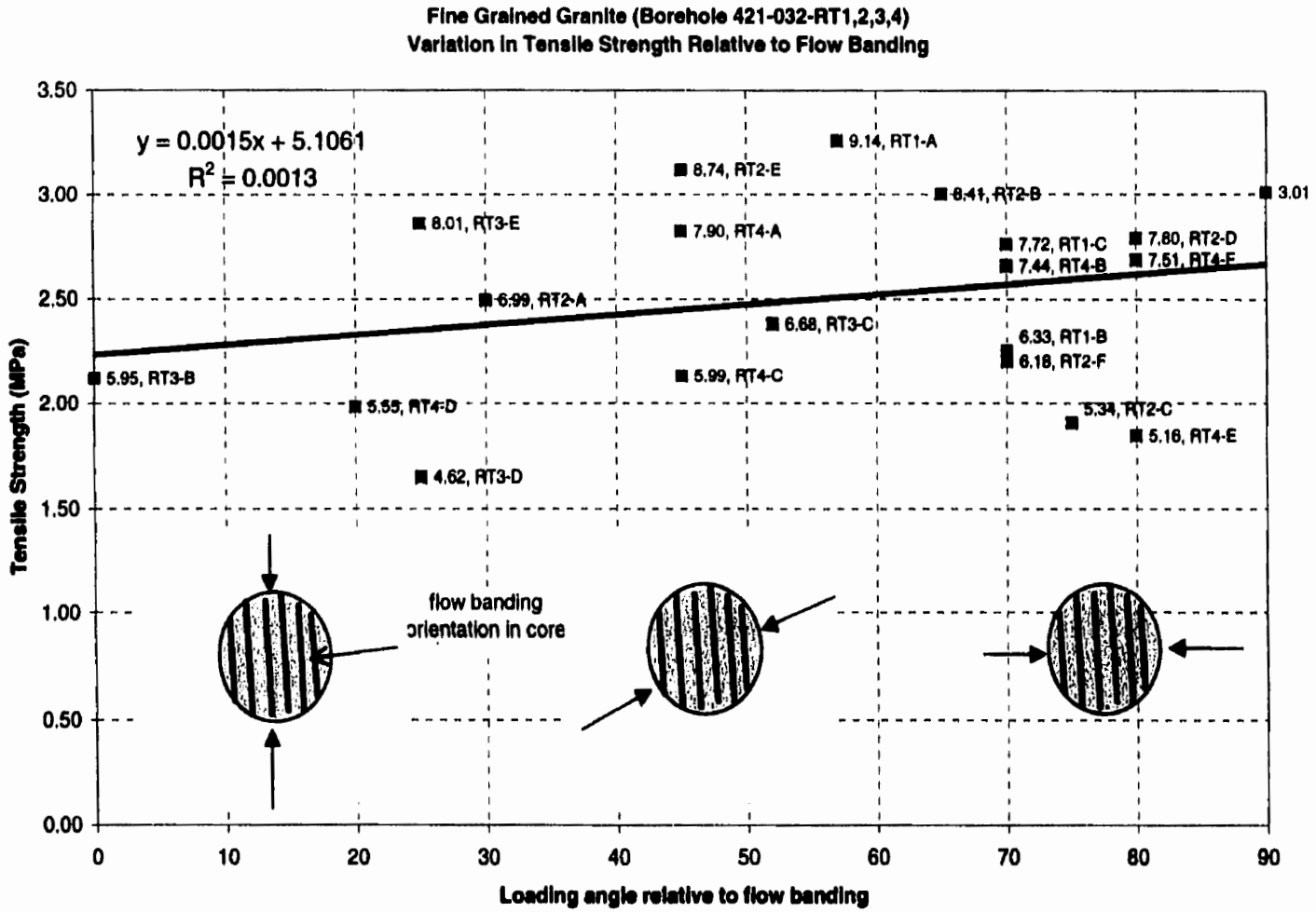



Figure 58. Brazilian test results for the Fine Grained Granite from boreholes 421-032-RT1, 2, 3, and 4. Variation in the estimated tensile strengths of each sample against the inclination of the loading axis relative to the flow banding. The three insets show the orientation of the loading axis (arrows) relative to the flow banding as seen in a cross section of the core. The best-fit linear regression curve, the equation for the curve, and the $R^2$ correlation coefficient are included. In this example, the tensile strength is somewhat lower when loading is done parallel to the flow banding, but the trend is dispersed.

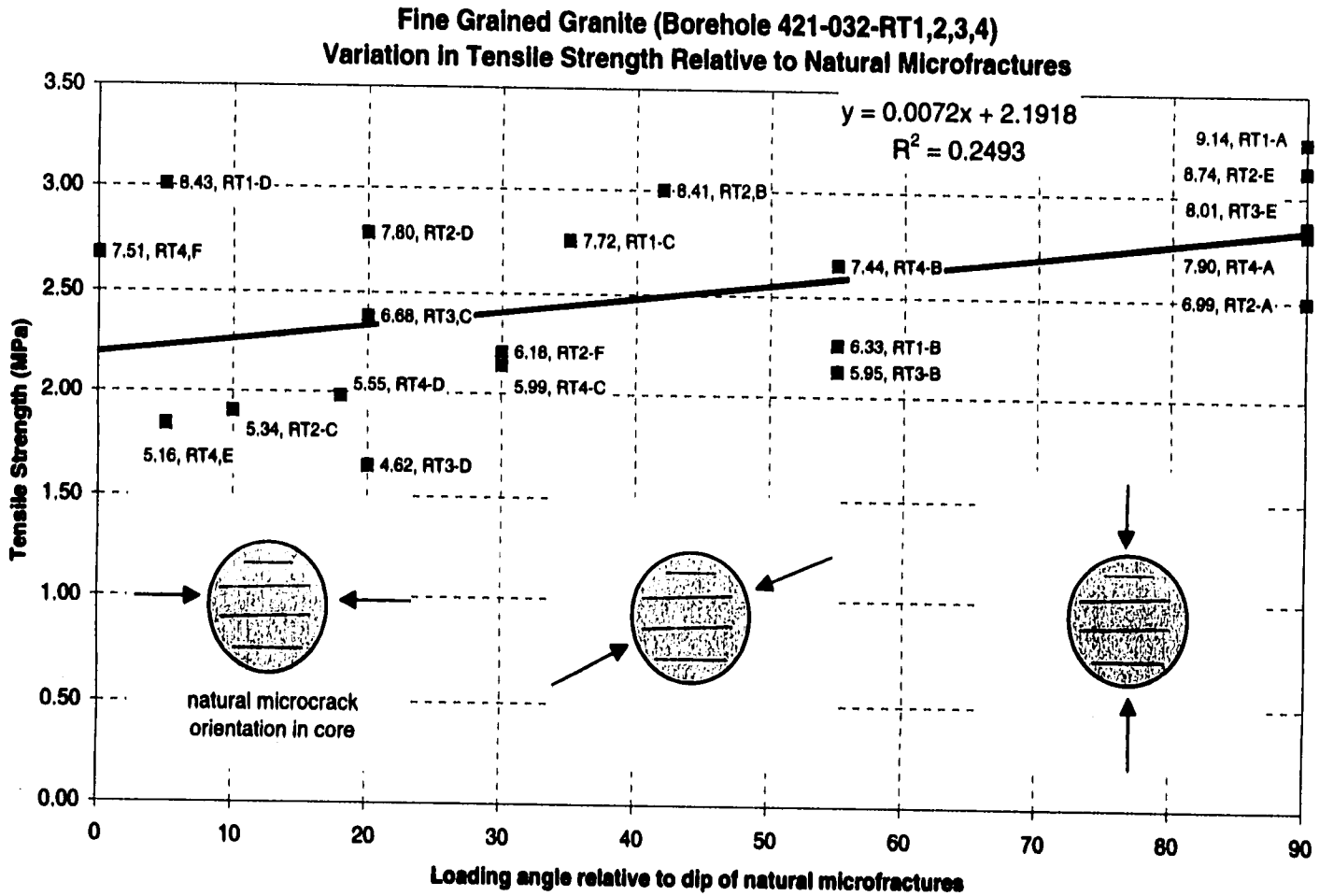



Figure 59. Brazilian test results for the Fine Grained Granite from boreholes 421-032-RT1, 2, 3, and 4. Variation in the estimated tensile strengths of each sample against the inclination of the loading axis relative to the preferred orientation of natural microcracks. The three insets show the orientation of the loading axis (arrows) relative to the microcracks as seen in a cross section of the core. The best-fit linear regression curve, the equation for the curve, and the $R^2$ correlation coefficient are included. In this example, the tensile strength parallel to the natural microcracks is ~ 65% of that perpendicular to the microcracks.

## 8.5. Medium Grained Granite, Borehole 417-071-RT 4, 5.

Brazilian test results for the Medium Grained Granite from boreholes 417-071-RT 4, 5 are summarized in Table 19. This summary includes, for each sample, the loading axis angles relative to structure, and two calculated tensile strengths for each sample. The first strength (min.) includes values in which audible cracking (fracture initiation) preceded failure, while the second strength (max) represents the value at failure. Three samples were affected by this phenomenon. These samples do not appear to be anomalous in shape (finishing), texture or in terms of their orientation and location. Summary statistics for the Brazilian tests from the entire borehole array within the Medium Grained Granite are shown in Figure 60. The array of boreholes was arranged in a vertical plane, and oriented so as to intersect the very inconspicuous low-dipping layering at angles of ~45°, and ~00°. The data from each borehole includes loading orientations ranging from 00° to 90° to the foliation in each borehole, and have been averaged together as if the rockmass was massive. This was done to test for systematic variation within the horizontal plane, or put in another way - to examine the effect of borehole orientation when sampling for rock properties testing. Although this array was less extensive than the arrays in the fine grained and coarse grained granites, the results confirm a similarly oriented direction -dependent anisotropy.

To determine if this trend is a function of varying degrees of sample damage, the tensile strength for each sample was plotted against its density (Figure 61). The tensile strength of the sample suite decreases with increasing density, which suggests that the direction anisotropy seen is not due to damage. There is no correlation between density and

location (as given by the sample numbers). Further, tensile strength decreases with increasing sample density, which is the opposite of what would be expected if damage was present. This evidence suggests that the observed variations in tensile strength with orientation are not the result of sample damage.

The influence of the low-dipping fabrics (an inseparable combination of compositional layering, mineral alignment and natural microcracks) on tensile strength is examined in Figure 62. Tensile strength parallel to the preferred orientation of these structures is up to 70% of that measured perpendicular to these structures.

Table 18.
Brazilian test results
Fine Grained Granite from boreholes 421-032-RT1, 2, 3, 4.

| Borehole | Sample | Density | Loading angle relative to low-dipping compositional layering | Minimum Tensile Strength (Crack Initiation) (MPa) | Maximum Tensile Strength (Failure) (MPa) |
|---|---|---|---|---|---|
| RT1 | D | 2.64 | 90 | | 3.01 |
| | A | 2.59 | 57 | | 3.26 |
| | B | 2.64 | 70 | | 2.26 |
| | C | 2.65 | 70 | | 2.76 |
| RT2 | A | 2.66 | 30 | | 2.50 |
| | B | 2.61 | 65 | | 3.00 |
| | C | 2.67 | 75 | | 1.91 |
| | D | 2.66 | 80 | 1.34 | 2.79 |
| | F | 2.65 | 70 | | 2.21 |
| | E | 2.65 | 45 | | 3.12 |
| RT3 | A | 2.64 | 0 | | |
| | B | 2.64 | 0 | | 2.12 |
| | C | 2.64 | 52 | | 2.38 |
| | D | 2.62 | 25 | | 1.65 |
| | E | 2.65 | 25 | | 2.86 |
| RT4 | A | 2.62 | 45 | | 2.82 |
| | B | 2.65 | 70 | | 2.66 |
| | C | 2.63 | 45 | | 2.14 |
| | D | 2.62 | 20 | | 1.98 |
| | E | 2.64 | 80 | | 1.84 |
| | F | 2.63 | 80 | | 2.68 |

Figure 08-08.cdr

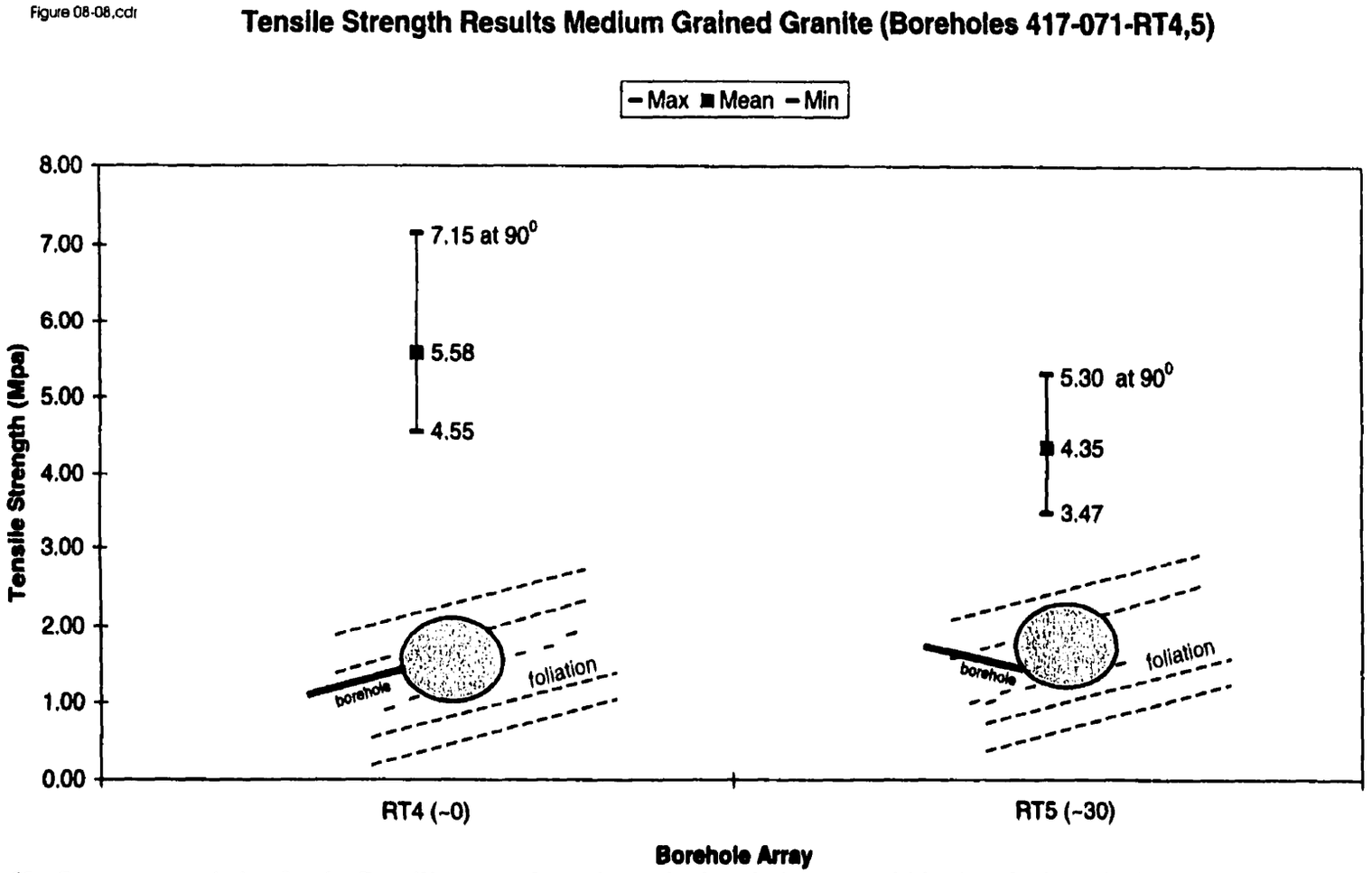



Figure 60. Summary statistics for the Brazilian tests from the entire borehole array within the Medium Grained Granite. The array includes boreholes 417-071-RT4 and 5. These were arranged in a vertical plane, and oriented to intersect the subvertical flow banding at angles of ~45°, and ~00°, respectively. The data from each borehole includes loading orientations ranging from 0° to 90° to the foliation in each borehole. This was done to test for systematic variation within the horizontal plane, or put in another way - to examine the affect of borehole orientation when sampling for rock properties testing. The results indicate that there is significant direction -dependent anisotropy.

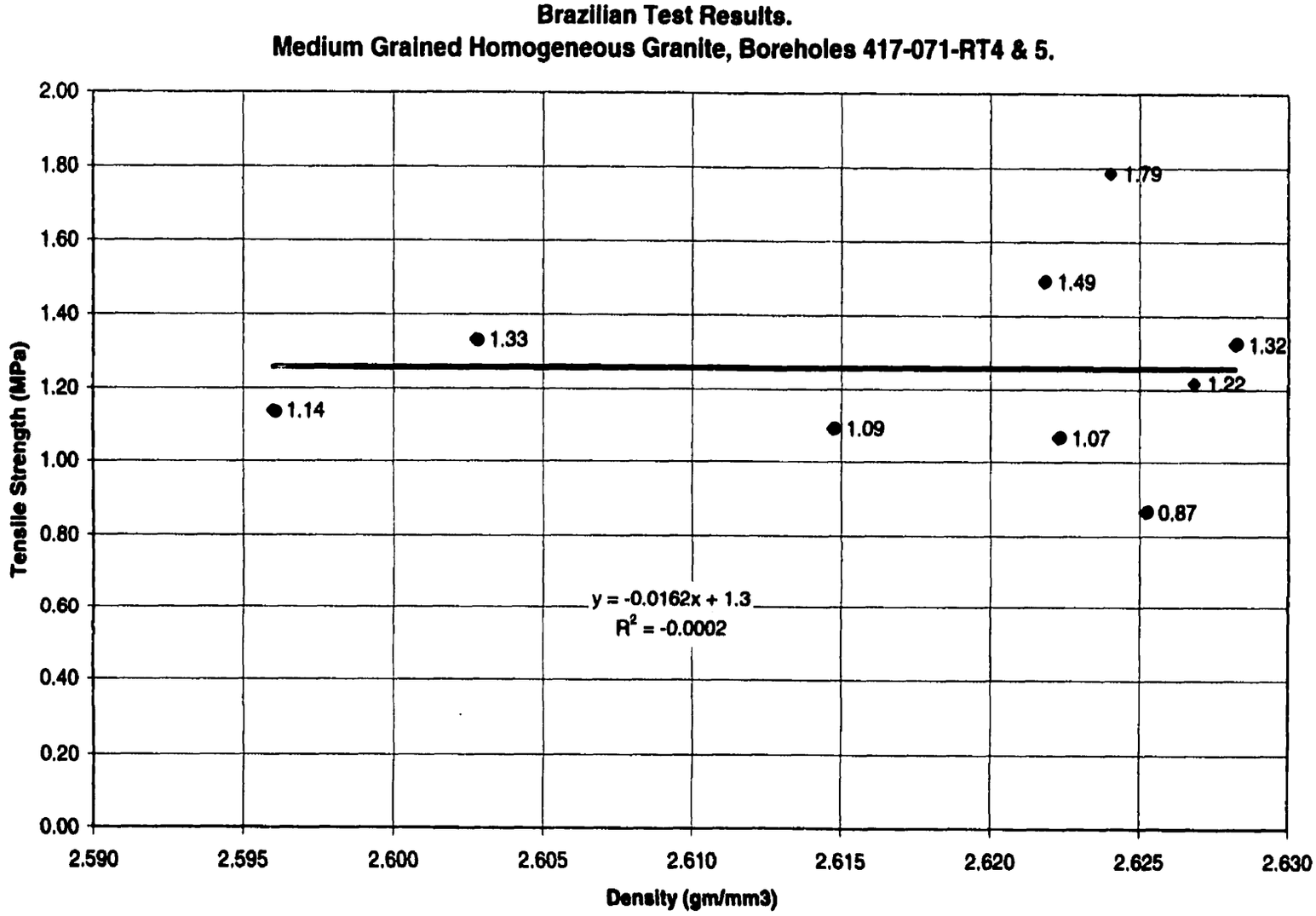



Figure 61. Brazilian test results for the Medium Grained Granite from boreholes 417-071-RT4 and 5. Variation in the estimated tensile strengths of each sample against the corresponding sample density. There is no correlation between density and location (as given by the sample numbers). Tensile strength decreases with increasing sample density, which is the opposite of what would be expected if damage was present.

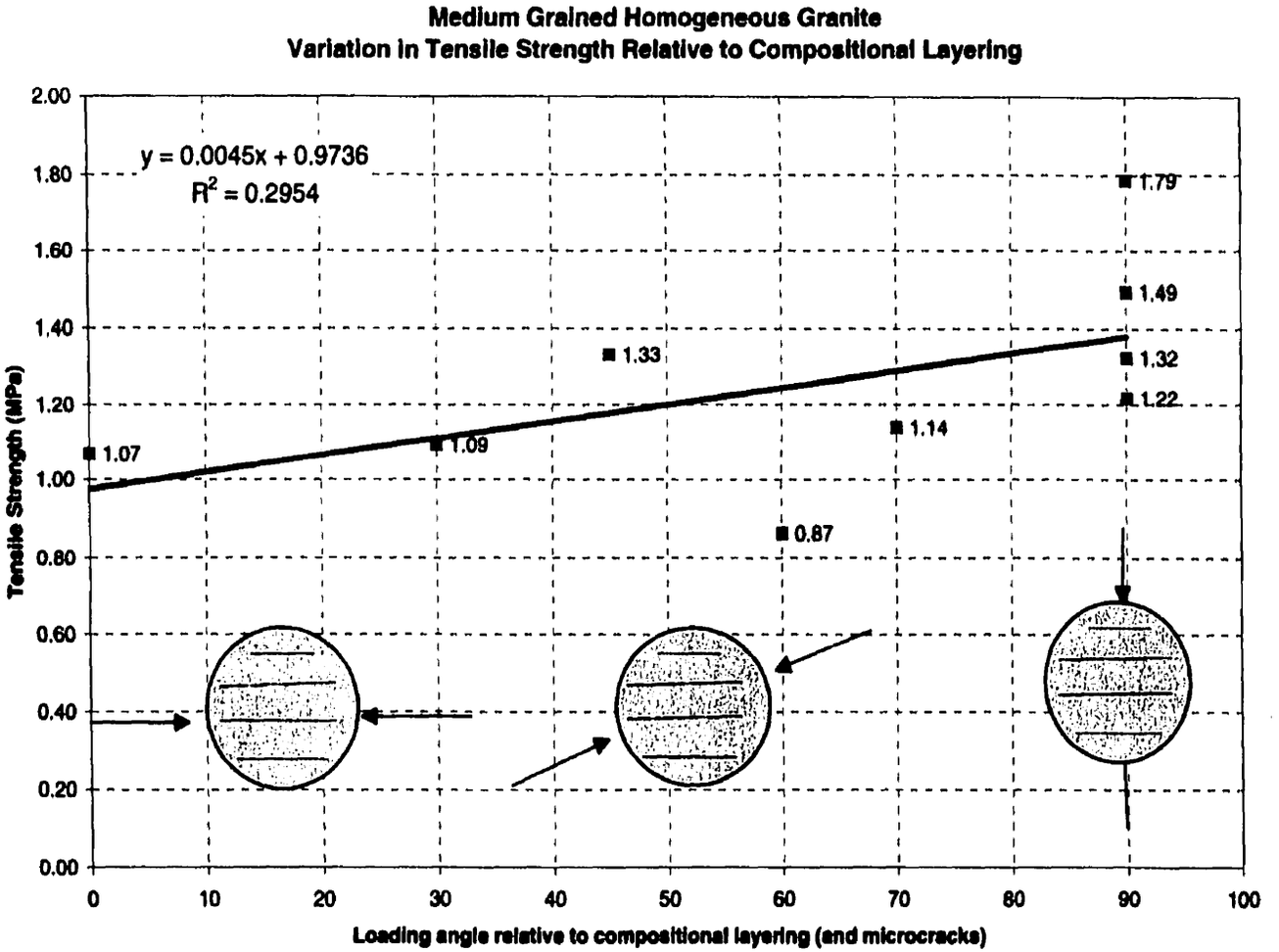



Figure 62. Brazilian test results for the Medium Grained Granite from boreholes 417-071-RT4, and 5. Variation in the estimated tensile strengths of each sample against the inclination of the loading axis relative to the compositional layering. The three insets show the relative orientation of the loading axis (arrows) relative to the flow banding as seen in a cross section of the core. The best-fit linear regression curve, the equation for the curve, and the $R^2$ correlation coefficient are included.

## 8.6. Coarse Grained Granite, Borehole 418-051-RT 2, 3, 4, 5.

Brazilian test results for the Coarse Grained Granite from boreholes 417-071-RT 2, 3, 4, 5 are summarized in Table 20. This summary includes, for each sample, the loading axis angles relative to structure, and two calculated tensile strengths for each sample. The first strength (min.) includes values in which audible cracking (fracture initiation) preceded failure. Six samples were effected. These samples do not appear to be anomalous in shape (finishing), texture or in terms of their orientation and location.

Summary statistics for the Brazilian tests from the entire borehole array within the Medium Grained Granite are shown in Figure 63. The array of boreholes was arranged in a vertical plane, and oriented so as to intersect the low-dipping fabric elements, an inseparable mixture of compositional layering, mineral foliation, and natural microcrack, at angles of ~90°, ~30°, ~ 15°, and ~00°, respectively. The data from each borehole includes loading orientations ranging from 0° to 90° to the foliation in each borehole, but have been averaged together here as if the rockmass was massive. This was done to test for systematic variation at the metre-scale, or put in another way - to examine the affect of borehole orientation when sampling for rock properties testing. The results indicate that there is significant direction -dependent anisotropy.

To determine if this trend is a function of varying degrees of sample damage, the tensile strength for each sample was plotted against its density (Figure 64). The tensile strength of the sample suite increases with increasing density, which is the opposite of what is seen in the other rock types. However, this does not appear to be attributable to damage,

but rather to the marked differences in composition and degree of banding/ foliation that occur at the array scale. This conclusion is supported by the fact that there is no decline in sample density and strength with proximity to either the borehole collar, or to areas of known stress concentration.

The influence of the low-dipping compositional layering on tensile strength is presented in Figure 65. The tensile strength is lower when the loading axis is parallel to the low-dipping fabric. The scatter in the data reflects the compositional heterogeneity noted earlier, and is indicated in this figure by the symbols used for each data point. Strength anisotropy in this rock type is marked. The tensile strength varies from a low of approximately 1 MPa in the plane of the layering, to approximately 2 MPa normal to this plane. (4 and 10 MPa, respectively using Equation 1 from p. 180).

Audible crack initiation was heard more frequently in this rock type than in the others types combined. The samples affected do not have any obvious characteristic in common. The phenomena occurred over the whole range of borehole orientations, loading angles. There does not appear to be any correlation with distance from the excavation, and any stress release phenomena that would be associated with proximity. Incipient discing is not a factor as any discing fractures would be unfavorably oriented for propagation under the diametric line loading of Brazilian tests.

Table 19.
Brazilian Test Data
Medium Grained Granite (Borehole 417-071-RT4, 5)

| Borehole | Sample | Density | Loading angle relative to low-dipping compositional layering | Minimum Tensile Strength (Crack Initiation) (MPa) | Maximum Tensile Strength (Failure) (MPa) |
|---|---|---|---|---|---|
| RT1, 2,3 | - not suitable due to core discing - | | | | |
| RT4 | A | 2.62 | 90° | | 1.49 |
| RT4 | B | 2.60 | 90° | | 1.79 |
| RT4 | C | 2.60 | 70° | 0.91 | 1.14 |
| RT4 | D | 2.63 | 45° | | 1.33 |
| RT4 | A' | 2.62 | 90° | 1.03 | 1.22 |
| RT5 | A | 2.63 | 90° | | 1.32 |
| RT5 | B | 2.61 | 30° | | 1.09 |
| RT5 | C | 2.62 | 0° | 0.87 | 1.07 |
| RT5 | D | 2.63 | 60° | | 0.87 |

Table 20.
Brazilian Test Data.
Coarse Grained Granite (Borehole 418-051-RT2,3,4,5)

| Borehole | Sample | Density | Loading angle relative to low-dipping compositional layering | Minimum Tensile Strength (Crack Initiation) (MPa) | Maximum Tensile Strength (Failure) (MPa) |
|---|---|---|---|---|---|
| RT1 | - not suitable due to core discing - | | | | |
| RT2 | A | 2.67 | 90° | | 1.70 |
| | B | 2.70 | 75° | | 1.75 |
| | C | 2.68 | 90° | | 2.07 |
| | D | 2.67 | 0° | | 1.49 |
| | E | 2.66 | 45° | | 1.69 |
| | F | 2.68 | 0° | | 1.11 |
| | G | 2.63 | 10° | | 1.17 |
| | H | 2.67 | 0° | 0.60 | 1.39 |
| RT3 | B | 2.70 | 45° | | 1.45 |
| | C | 2.69 | 75° | 0.67 | 2.01 |
| | D | 2.67 | 90° | | 2.53 |
| | E | 2.70 | 85° | | 2.49 |
| | F | 2.71 | 0° | | 1.88 |
| | A' | 2.68 | 55° | | 1.56 |
| | G | 2.65 | 60° | 0.60 | 1.65 |
| RT4 | A | 2.61 | 50° | | 0.96 |
| | B | 2.59 | 45° | | 1.21 |
| | C | 2.60 | 45° | | 1.04 |
| | D | 2.56 | 15° | | 1.33 |
| | E | 2.60 | 20° | | 1.53 |
| | F | 2.60 | 38° | 0.62 | 0.88 |
| | G | 2.61 | 52° | 0.54 | 1.42 |
| RT5 | A | 2.62 | 0° | | 0.93 |
| | B | 2.54 | 25° | 0.52 | 1.18 |
| | C | 2.65 | 45° | | 1.52 |
| | D | 2.64 | 85° | | 1.65 |

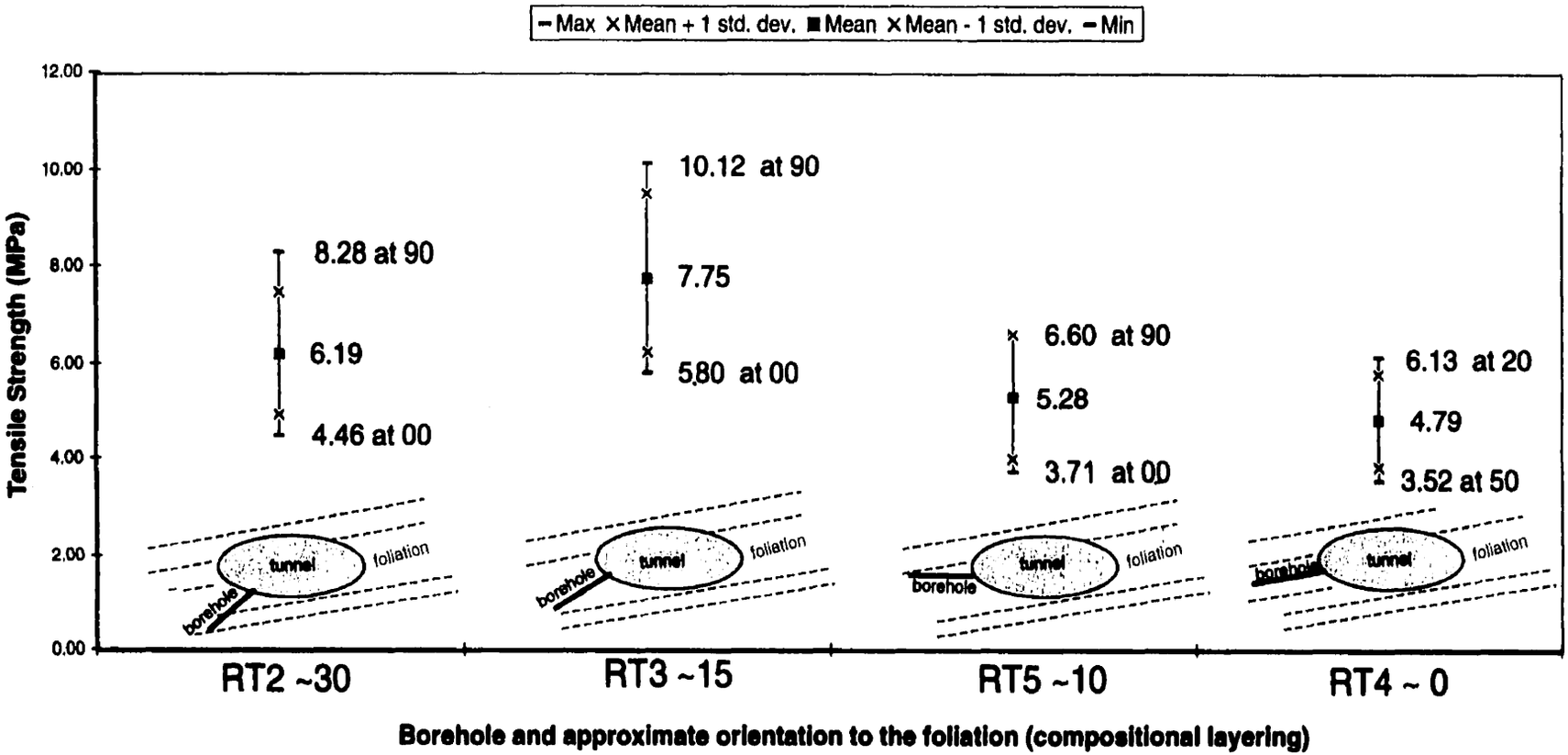



Figure 63. Summary statistics for the Brazilian tests from the entire borehole array within the Coarse Grained Granite. The array includes boreholes 418-051-RT1, 2, 3, 4. These were arranged in a vertical plane, and oriented so as to intersect the compositional layering at angles of ~90°, ~30°, ~15°, ~ 0°, and ~10°, respectively as shown by the schematic illustrations - (the tunnel profile is represented by the ellipse, a xenolithic zone by the grey shading, the layering by the light grey dashed lines, and the boreholes by the solid lines Note: the near-vertical borehole RT1 was not useable due to discing). The data from each borehole includes loading orientations ranging from 0° to 90° to the foliation in each borehole. This was done to test for systematic variation within the vertical plane, or put in another way - to examine the affect of borehole orientation when sampling for rock properties testing. The results indicate that there is significant direction -dependent anisotropy in tensile strength.

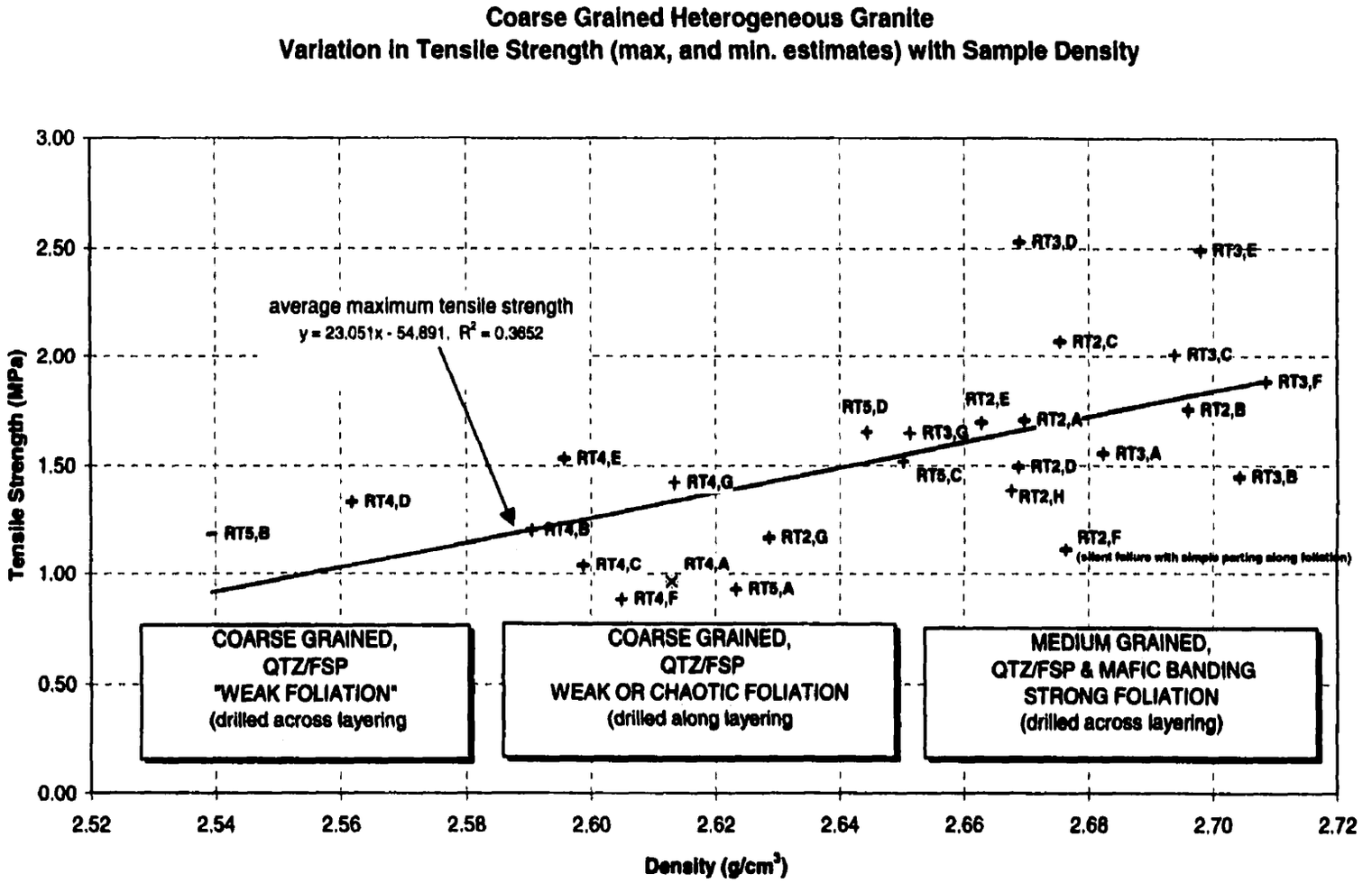



Figure 64. Brazilian test results for the Coarse Grained Granite from boreholes 418-051-RT2, 3, 4, and 5. Variation in the estimated tensile strengths of each sample against the corresponding sample density. This suite differs from the fine and medium grained granites in that tensile strength increases with increasing density. This does not however, suggest sample damage, as there is no correlation between strength and the depth of the sample from the borehole collars. Instead, this relationship appears to be due to a combination of factors, including the orientation of the borehole relative to the layering, and to differences in rock composition and degree of fabric development as indicated in the boxes above.

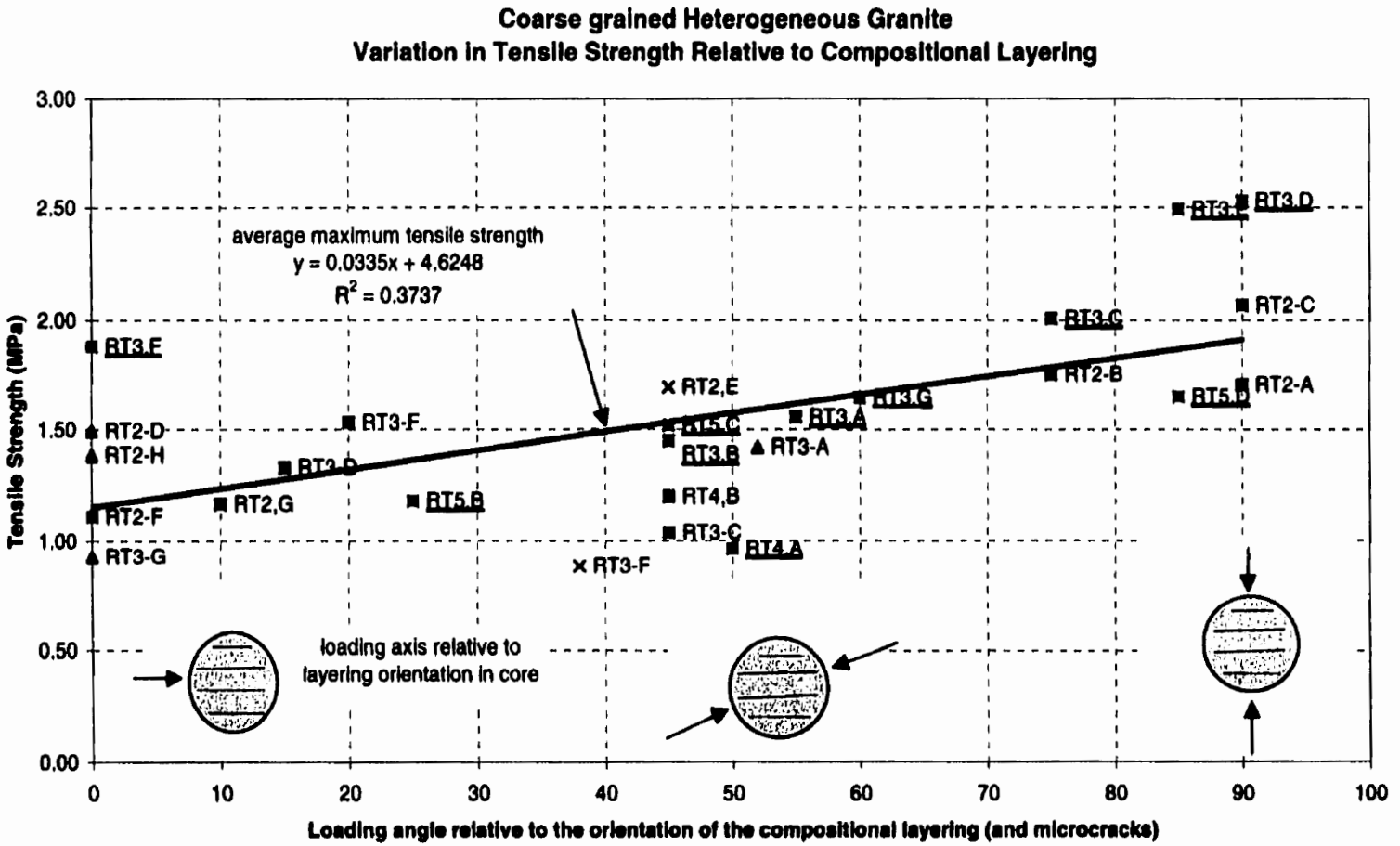



Figure 65. Brazilian test results for the Coarse Grained Granite from boreholes 417-071-RT2, 3, 4, and 5. Variation in the estimated tensile strengths of each sample against the inclination of the loading axis relative to the low-dipping compositional layering. The three insets show the relative orientation of the loading axis (arrows) relative to the flow banding as seen in a cross section of the core. The best-fit linear regression curve, the equation for the curve, and the $R^2$ correlation coefficient are included. Tensile strength is lowest when splitting occurs along the plane of the foliation. The scatter in the data reflects the compositional heterogeneity noted earlier, and is indicated in this figure by the following symbols (■- sample has obvious compositional banding and mineral foliation, ▲ - sample has mineral alignment only, X - sample appears nearly massive or has a chaotic or folded foliation).

## 8.7. Summary and Conclusions from the Brazilian Tests

The Brazilian tests reveal that tensile strength varies with average grain size, and is also anisotropic, being dependent upon the orientation of loading relative to the fabrics in each granite.

The tensile strength in the fine grained granite varies with the orientation of the sampling borehole, and on the orientation of loading relative to the fabric in each sample. Directional anisotropy in strength is most noticeable in the vertical plane. In this case, the tensile strength measured in the plane of natural microcracks is approximately 65% of that measured normal to the microcracks. In the horizontal plane, the tensile strength is lowest when loading is parallel to the flow banding, and greatest when normal to this structure.

The medium grained granite shows directional anisotropy in tensile strength, which may be correlated with the rock mass fabric. Tensile strength parallel to the preferred orientation of the low-dipping layering and natural microcracks is ~ 70% of that measured perpendicular to these structures.

The coarse grained granite is the most heterogeneous of the rock types tested. A low-dipping fabric dominates its structure at the tunnel and the hand specimen scales. This fabric includes a largely concordant assemblage of compositional and textural layering, an alignment of biotite and feldspar laths, and a preferred orientation of natural microcracks in quartz. This fabric induces a marked directional anisotropy, with the

Brazilian tensile strength parallel to the layering being ~ 40% of that measured normal to this plane.

The correction for calculation of the Brazilian test recommended by Chen (1993) reduces the calculated strengths by approximately 50%, but does not affect the presence of directional anisotropy in tensile strength, or their correlations with rock fabric.

# Chapter 9

# UNIAXIAL COMPRESSION TESTS

## 9.1. Introduction

This chapter summarizes the uniaxial compression testing conducted to define strength anisotropy related to the textural and structural variations in the Lac Du Bonnet Batholith. The purpose of this test is to measure the ultimate compressive strength of a cylindrical rock sample loaded axially. The axial deformation of the specimen is also measured to obtain information on the deformation and failure characteristics of the rock substance to its ultimate strength.

## 9.2. Rock Type and Specimen Preparation

As noted earlier, three distinct textural varieties of the Lac Du Bonnet granite, differing in type and orientation of foliation were identified and selected for testing. Eleven samples were obtained from the boreholes listed in Chapter 1. All samples are segments of HQ-sized core (46 mm diameter), the choice of specimen diameter being largely dictated by budget constraints. Specimens were prepared in accordance with the procedure described in the Pit Slope Manual Supplement 3-5 (Gyenge and Ladanyi, 1977; and Gyenge, 1980) and the method suggested by the International Society for Rock Mechanics (Brown, 1981). The end surfaces of each of the specimens were ground flat to 0.015 mm to ensure that they were parallel to each other and perpendicular to the axis of the specimen.

## 9.3. Test and Apparatus And Procedure

For each sample, axial and lateral strains induced during compression were provided by the attached strain gauges. All samples were loaded, parallel to their long axis, in the testing apparatus at the Canadian Mine Technology Laboratory, Mining Research Laboratories, CANMET, Energy, Mines and Resources Canada in Ottawa. For a detailed description of the testing apparatus and its operation, the reader is referred to Gorski (1987). The data records for each specimen were saved as ASCII text files, and included the time (sec), load (MPa), and corresponding axial and lateral strains. Loading was continuous rather than cycled. As the peak strength of the specimen was approached, loading was closely monitored to prevent rapid failure.

## 9.4. Presentation Of Results

The data obtained in this experiment were compiled as follows:

1. Stress-strain data tables for each sample, which list the axial, lateral and volumetric strains recorded for each strain gauge by the data acquisition system, under accumulating load. Volumetric strain was computed by using the following equation:

$$\varepsilon_v = \varepsilon_a + 2\varepsilon_c$$

where $\varepsilon_v$ = volumetric strain, $\varepsilon_a$ = axial strain, and $\varepsilon_c$ = circumferential strain.

2. Stress-strain diagrams for each sample, showing the axial, lateral and volumetric strains under accumulating loads.
3. Summary tables listing the tensile strength (from the Brazilian tests), and the values for crack closure, the onset of dilatancy (yield stress), and the peak load failure

obtained by inspection of the stress-strain curves, as described in the following section.

4. Summary charts depicting the tensile strength (from the Brazilian tests), and the values for crack closure, the onset of dilatancy (yield stress), and the peak load failure obtained by inspection of the stress-strain curves, as described in the following section.

The data readings (1) and stress-strain diagrams (2) are provided in Appendix C, while the summary tables (3) and summary charts (4) are presented in this Chapter.

## 9.5. Interpretation of Stress-strain Curves

Figure 66 illustrates the general stress-strain characteristics for the Lac du Bonnet granite under uniaxial compression. When an applied load exceeds the rock strength, bonds break and separate and a fracture (or microcrack) is formed with an attendant release of energy. For brittle rocks, under tensile conditions, the initiation of cracking coincides with failure. However, under compressive loading, crack initiation occurs well before failure. Bieniawski (1967a,b) identified five stages of crack development on brittle rock (also Goodman, 1989).

1. Crack closure: Grain boundaries act as, or may contain, microcracks. Microcracks can also exist within grains especially if stress-relief or previous loading has occurred. This first stage of loading involves the closure of microcracks which are not parallel to the applied stress. This increases the modulus of elasticity of the rock and produces a non-linear stress-strain curve. This region is more prominent for tests performed at low confining pressures than those performed at high confining

pressures, typical for rock samples obtained from the 420 Level of the URL, due to the presence of stress-induced microcracks in those samples.

2. Linear elastic deformation: This stage follows crack closure, and is noted by a linear "straight line segment of both the axial and lateral strain curves, the sample approximating linear elastic behavior.

3. Crack initiation: During the active loading of each specimen, "... a change from a deformation that is elastic and therefore recoverable to one that involves permanent deformation occurs at a point where the initially linear stress versus lateral strain, or stress versus volumetric strain curve deflects from its initial, linear trend ..... From this point on, the rate of lateral strain increases above what can be attributed to purely elastic deformation. The point of deflection is the "crack initiation point". More than one crack initiation point can occur in a given specimen, due to the contribution of different mineral phases, or to grain or sub-grain scale heterogeneities..." (Lajtai, 1982).

4. Stable crack propagation: Existing cracks extend and new cracks can nucleate. As the cracks develop parallel to the major principal stress, the strain gauges oriented along the specimen axis (axial) still produce a linear response while those oriented laterally produce a non-linear response. The volumetric strains also become non-linear at this time. The point of crack initiation varies with rock type ad structure, but a range of 40-50% of the peak strength is common for brittle rock.

5. Unstable crack growth: Also referred to as the onset of dilatancy, this point in the loading curve, occurs at approximately 80% of the peak strength. It is marked by a relatively sudden change in the pattern of propagating cracks, by reversal of the sign

of volumetric strain, with a rapid increase in volume. There is permanent lateral (circumferential) and volumetric strain on loading.

6. Failure: In this region, crack formation, crack growth and sliding on existing crack interfaces results in a decrease of the slope of each of the strain curves. Failure of the specimen is the peak stress achieved which is easily noted by a decrease in axial strain on the stress-strain curve. The final post-peak stage is denoted by decreasing stress (i.e. not load-bearing).

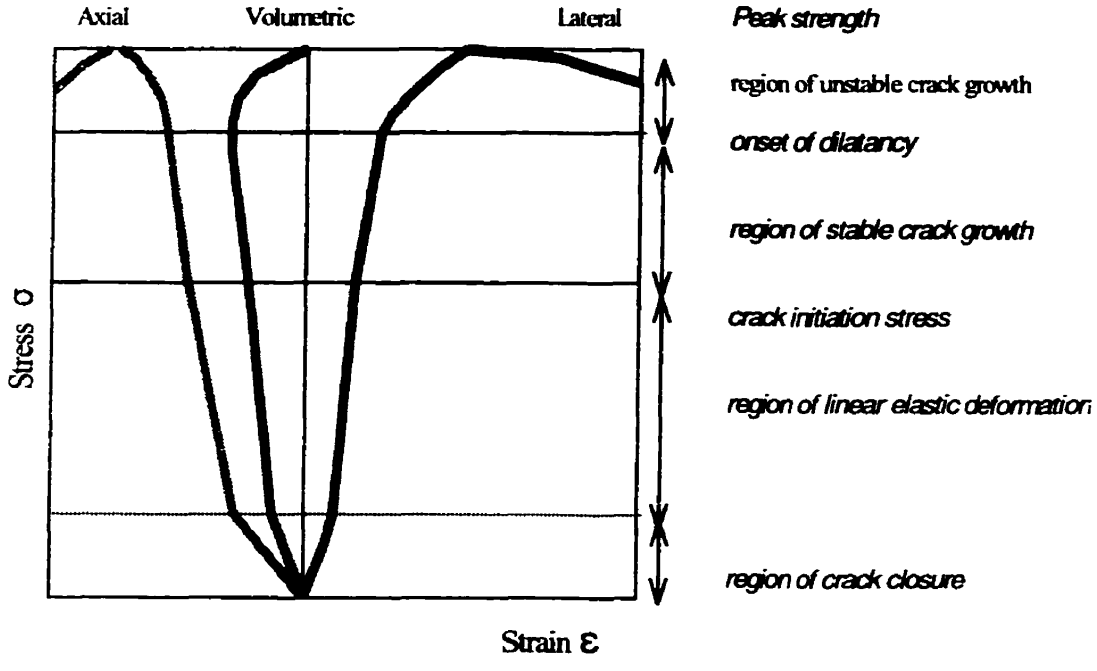



Figure 66. Idealized stress-strain diagram for a brittle rock with the stages of brittle deformation as represented by the axial, volumetric and lateral stress-strain curves.

## 9.6. Summary of Observations

The results of the uniaxial compression tests are presented in the summary charts which show: 1) the influence of grain size, and 2) the influence of loading angle relative to the dominant fabric element. Crack closure, the onset of stable crack growth (also referred to as the onset of dilatancy), and the peak load at failure all vary in a systematic manner with the average grain size of the rock, and with rock mass fabric.

### 9.6.1. Influence of Grain Size

The relative minimum rock strength data are listed in Table 21, and are shown in Figures 67, 68 and 69 for each of the fine, medium and coarse grained granites. The stresses corresponding with crack closure, the onset of dilatancy and the peak load failure were determined from the stress / strain curves in Appendix C. The data for the very coarse granite is from Everitt (1995).

Strength decreases as the grain size and heterogeneity of the granite increases (Figure 67). This trend is shown best by the minimum estimates for each of the various measures of strength. This trend is obscured by the specimen-scale heterogeneity in the coarser granites, where the mean and maximum strengths are more variable (Figures 68 and 69). In other words, depending on the location and orientation of the borehole, the coarsest granite can be either weaker than, or stronger than, the medium grained granite.

### 9.6.2. Influence of Rock Fabric

Anisotropy of the various measures of compressive strength as a function of rock fabric is presented in Table 22 and in Figures 70, 71 and 72. The "measures" referred to here include: the peak load at failure, the onset of dilatancy, the crack initiation point, the

crack closure stress, and the average tensile strength. The average tensile strength was obtained from the Brazilian tests. The values for crack closure, crack initiation, the onset of dilatancy and the peak load failure were determined from inspection of the individual stress / strain curves in Appendix C. There is insufficient data to review the influence of fabric in the medium grained granite on compressive strength.

Anisotropy of compressive strength related to rock fabric, is most strikingly displayed by the fine grained granite (Figure 70). In this rock, each of the various measures of rock strength decrease together in a systematic manner as the orientation of the sampling borehole, (and therefore the axis of the test specimen) rotates towards parallelism with the strike of the rock unit contact. The weakest direction correlates with the strike of the dykes and of rare subvertical fractures found at this level.

Directional anisotropy is also evident in the coarse grained granite, although its relationship to the rock fabric requires more explanation. The variation, with orientation, of the compressive (and tensile) strengths for the coarse grained granite is shown in Figure 71. Inspection of this figure indicates that samples are weakest in compression when the centimetre- to decimetre –scale layering intersects the sample axis at an angle between 20° and 40°. The rock is stronger (at the specimen scale) when the sample is within the plane of layering, or sharply inclined to it. This interpretation is consistent with field observations and with the Point Load and Brazilian tests, in that the contacts between layers provide more clearly defined and continuous weakness planes, than those structures that would promote axial splitting of the specimens (i.e. grain alignment or

natural microcracks) at the test-specimen scale. Similar results were obtained from the tests for the medium-grained granite (Figure 72).

For this coarsely layered medium grained granite, the various measures of strength were lowest when the layering, and the subparallel natural microcracks, were inclined 30 degrees to the loading axis.

### 9.6.3. Summary and Conclusions from Uniaxial Compression Tests

The uniaxial compression tests conducted in this study reveal that the compressive strength varies with average grain size, and with the direction of loading relative to the dominant fabric present in each sample. Directional anisotropy, related to different types of fabric elements, exists to a significant degree in each of the three rock types. The minimum estimates for peak load strength for the fine, medium, coarse, and very coarse grained granites are 168, 155, 125, and 120 MPa, respectively.

Table 21
Summary tables of peak load failure and other strength measures (in MPa)
as determined by manual inspection of the stress-strain curves.
(tensile strength data from Brazilian tests).

Minimums

| | Very coarse granite | Coarse granite | Medium granite | Fine granite |
|---|---|---|---|---|
| Peak load failure | 120 | 125 | 155 | 168 |
| Onset of dilatancy | 80 | 92 | 96 | 100 |
| Microcrack closure | 20 | 16 | 22 | 40 |
| Tensile strength | 3.5 | 3.7 | 3.5 | 4.6 |

Means

| | Very coarse granite | Coarse granite | Medium granite | Fine granite |
|---|---|---|---|---|
| Peak load failure | | 158 | 161 | 193 |
| Onset of dilatancy | | 110 | 100 | 128 |
| Microcrack closure | | 18 | 23 | 40 |
| Tensile strength | | 6.8 | 5 | 6.3 |

Maximums

| | Very coarse granite | Coarse granite | Medium granite | Fine granite |
|---|---|---|---|---|
| Peak load failure | | 183 | 166 | 210 |
| Onset of dilatancy | | 128 | 104 | 150 |
| Microcrack closure | | 20 | 24 | 40 |
| Tensile strength | | 10.1 | 7.1 | 9.1 |

Table 22
Uniaxial Compression Testing Summary by Rock Type and Borehole (in MPa)
(tensile strength data from Brazilian tests)

**Fine grained granite**

| Minimums | 421-032-RT1 | 421-032-RT2 | 421-032-RT3 | 421-032-RT4 |
|---|---|---|---|---|
| | *75** | *50** | *35** | *10** |
| Peak load failure | 210 | 196 | 168 | 120 |
| Onset of dilatancy | 150 | 136 | 108 | 88 |
| Crack initiation | 130 | 120 | 82 | 60 |
| Microcrack closure | 40 | 40 | 40 | 28 |
| Average tensile strength | 9.1 | 8.7 | 8 | 7.9 |

**Medium grained granite**

| Minimums | 403-014-MB2 | 417-071-RT4 | 417-071-RT5 |
|---|---|---|---|
| | *53** | *0** | *30** |
| Peak load failure | 166 | 163 | 155 |
| Onset of Dilatancy | 92 | 104 | 95 |
| Crack initiation | 75 | 90 | 60 |
| Microcrack closure | 23 | 32 | 24 |
| Average Tensile strength | 5 | 5 | 5 |

**Coarse grained granite**

| Minimums | 418-051-RT2 | 418-051-RT3 | 418-051-RT4 | 418-051-RT5 |
|---|---|---|---|---|
| | *40** | *20** | *0** | *15** |
| Peak load failure | 140 | 125 | 183 | 183 |
| Onset of Dilatancy | 92 | 96 | 128 | 126 |
| Crack initiation | 72 | 84 | 96 | 100 |
| Microcrack closure | 16 | 16 | 20 | 20 |
| Tensile strength | 4 | 4 | 4 | 4 |

Note: The angle between the borehole and the dominant rock structure is indicated in italics below each borehole name.

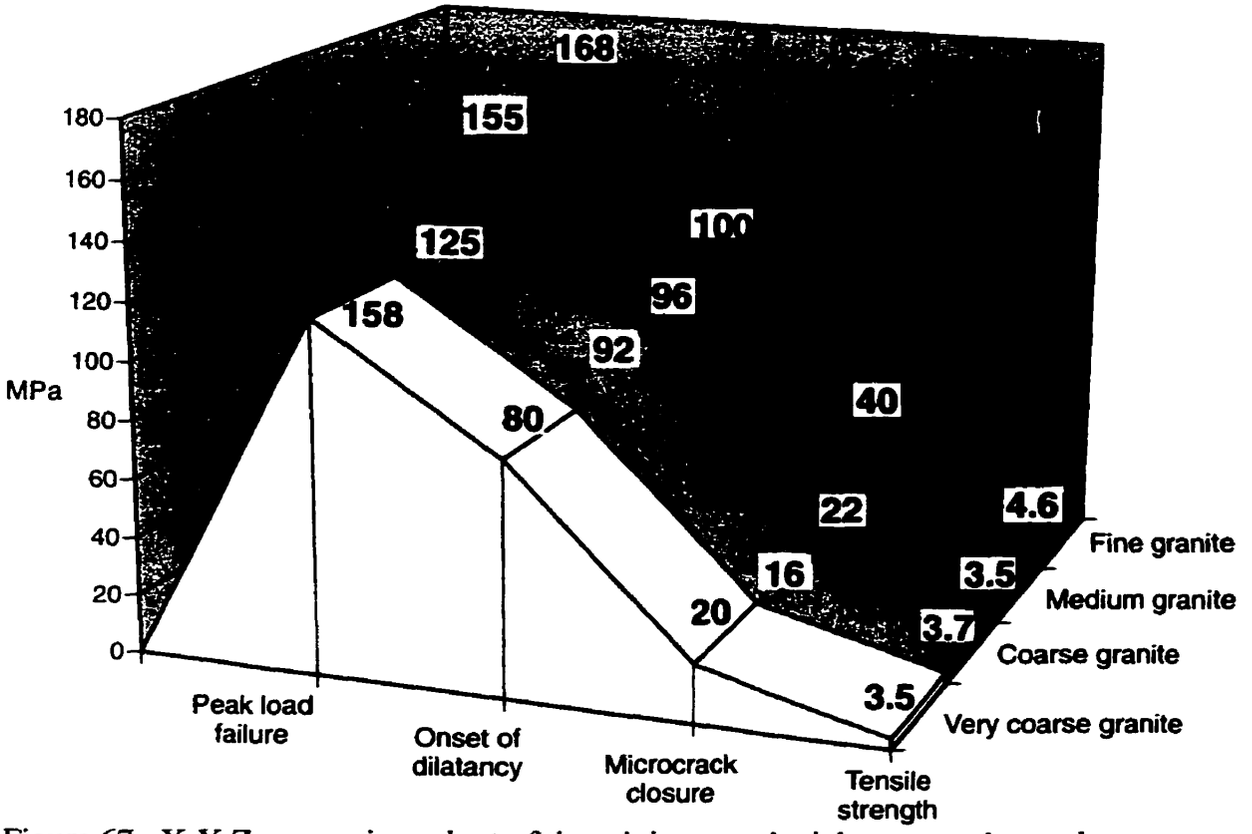



Figure 67. X-Y-Z comparison chart of the minimum uniaxial compression and (Brazilian) tensile strengths of the 3 rock types. The stresses corresponding with crack closure, the onset of dilatancy and the peak load failure were determined from the stress / strain curves in Appendix C. The data for the very coarse granite is from Everitt (1995). The figure clearly indicates that the various measures of strength decrease as the grain size and heterogeneity of the granite increase. Compare with Figures 68 and 69 showing the mean and maximum Uniaxial Compression and (Brazilian) Tensile Strengths, respectively.

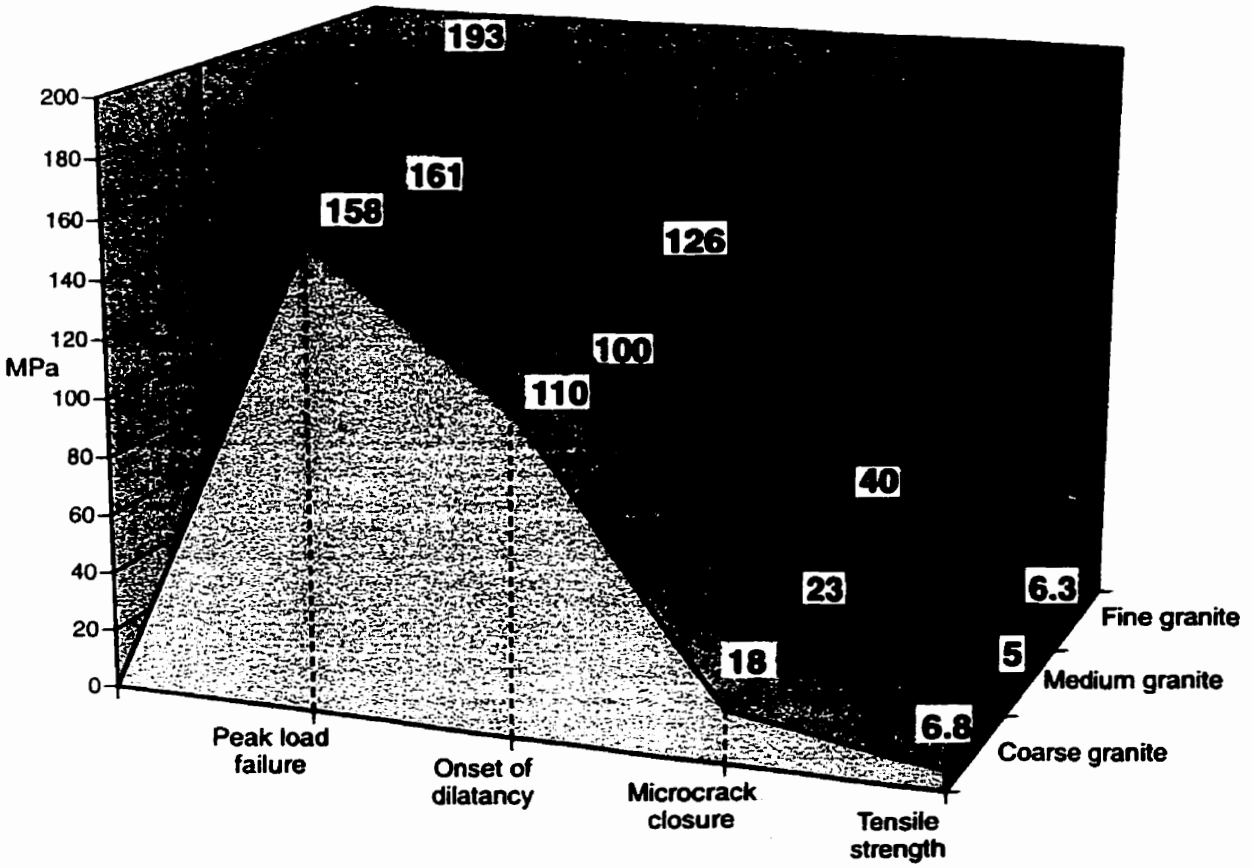



Figure 68. X-Y-Z comparison chart of the mean uniaxial compression and (Brazilian) tensile strengths of the 3 rock types. The stresses corresponding with crack closure, the onset of dilatancy and the peak load failure were determined from the stress / strain curves in Appendix C. The data for the very coarse granite is from Everitt (1995). The figure clearly indicates that the various measures of strength decrease as the grain size and heterogeneity of the granite increase. Compare with Figures 67 and 69 showing the minimum and maximum Uniaxial Compression and (Brazilian) Tensile Strengths, respectively.

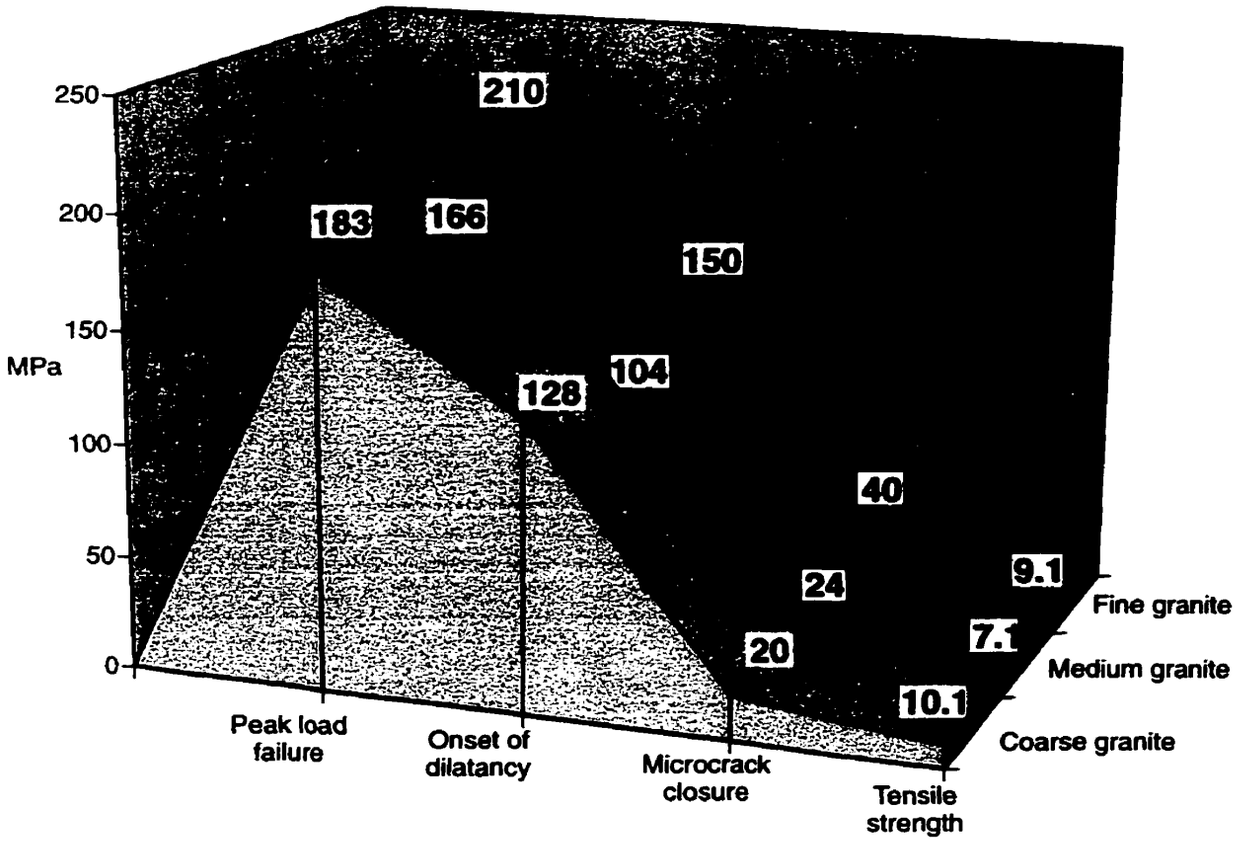



Figure 69. X-Y-Z comparison chart of the maximum uniaxial compression and (Brazilian) tensile strengths of the 3 rock types. The stresses corresponding with crack closure, the onset of dilatancy and the peak load failure were determined from the stress / strain curves in Appendix C. The data for the very coarse granite is from Everitt (1995). As was shown in Figure 67 (minimum values), the various measures of strength decrease as the grain size and heterogeneity of the granite increase. This trend is obscured by the specimen-scale heterogeneity in the coarser granites, where the mean and maximum strengths are more variable (Figure 68 and 69). In other words, depending on the location and orientation of the borehole, the coarsest granite can be either weaker than, or stronger than, the medium grained granite.

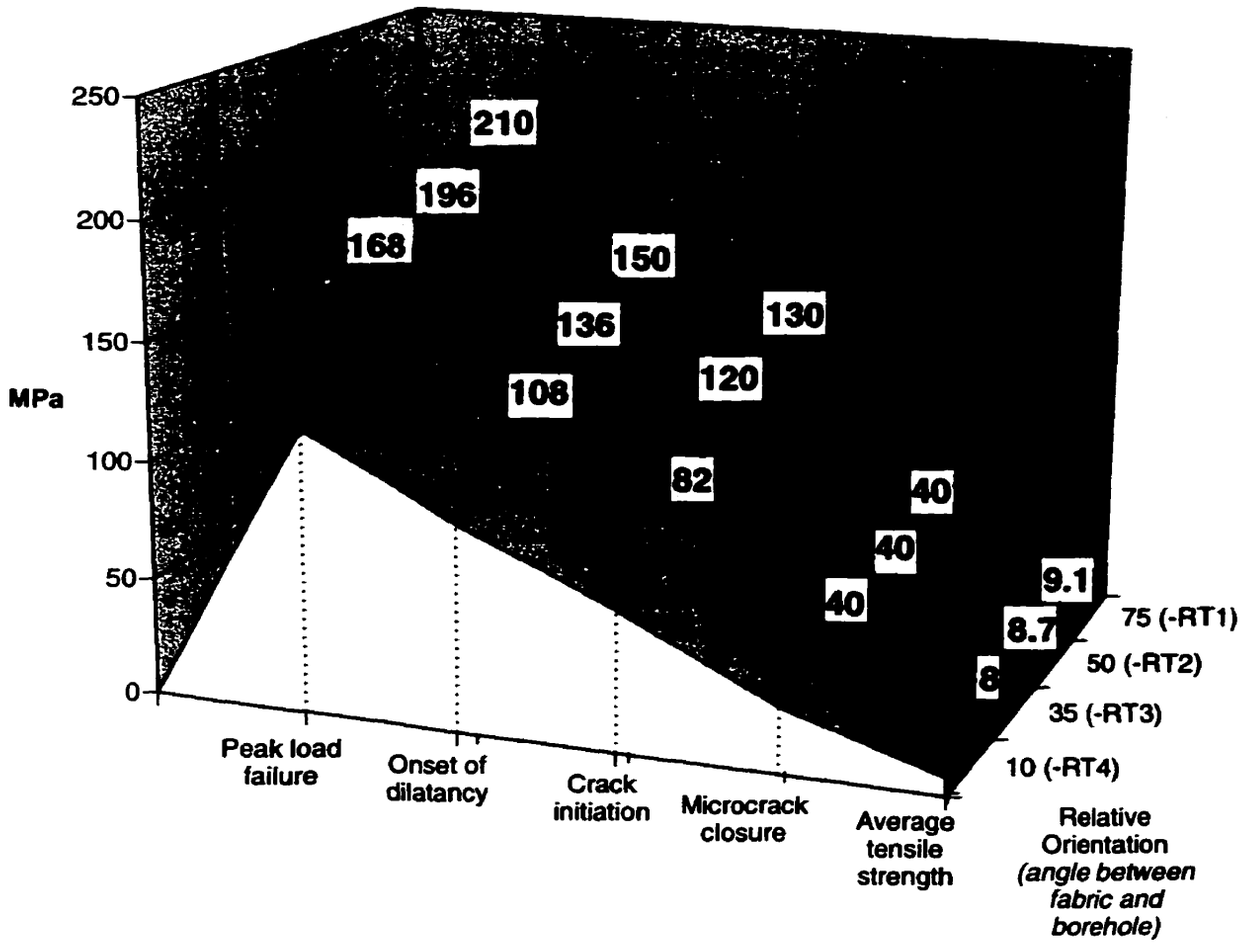



Figure 70. Anisotropy of compressive strength related to rock fabric, is most strikingly displayed by the fine grained granite. This figure compares the variation in the various strength estimates with orientation. Orientation refers to the angle between the strike of the dyke and the sampling boreholes, which ranges from 10°, 35°, 50°, and 75° in boreholes –RT4, - RT3, - RT2 and - RT1, respectively. The values for the peak load failure, onset of dilatancy, crack initiation, and microcrack closure are derived from the uniaxial compression data (stress /strain curves) in Appendix C. The tensile strength is from the Brazilian tests. In this rock, each of the various measures of rock strength decrease together in a systematic manner as the orientation of the sampling borehole rotates towards parallelism with the strike of the rock unit contact. The weakest direction correlates with the strike of the dykes and of rare subvertical fractures found at this level.

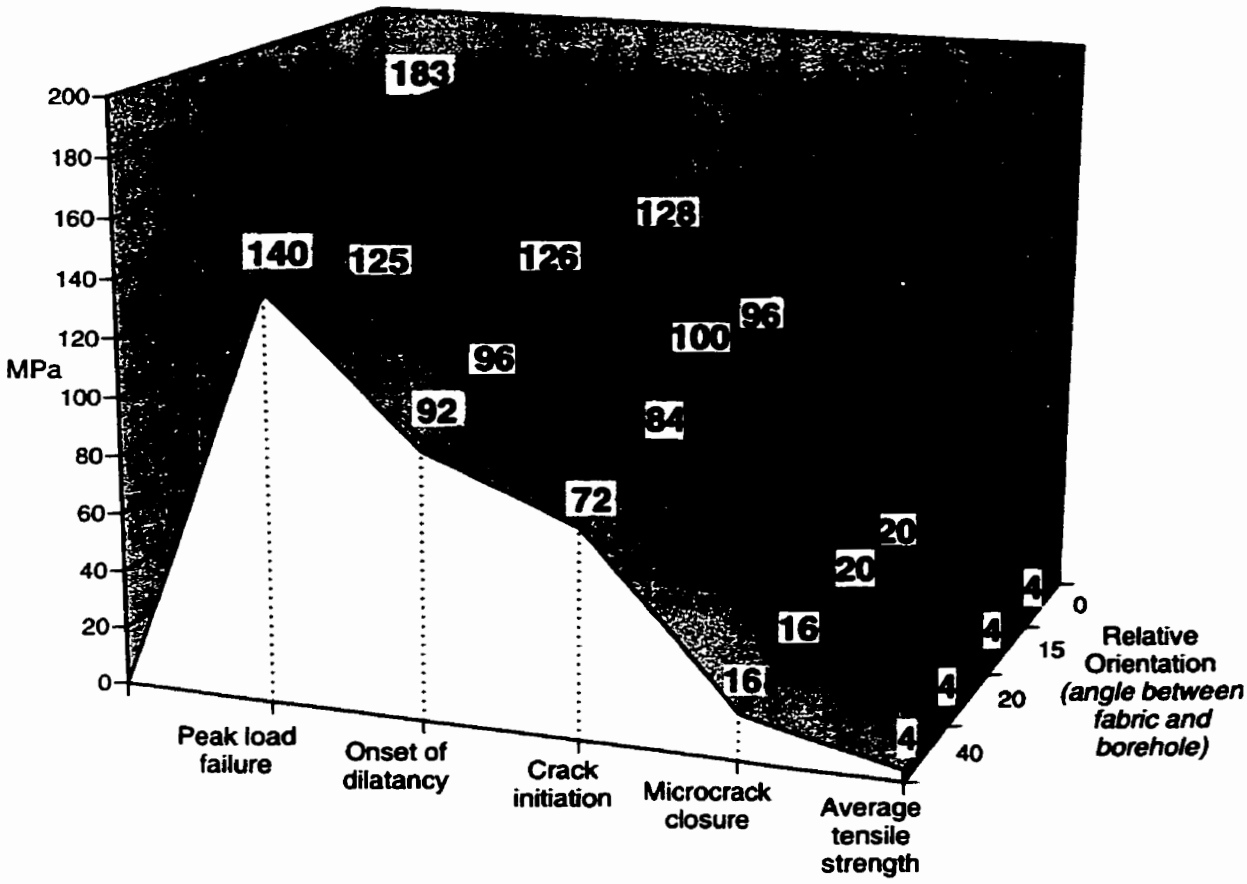



Figure 71. Influence of low-dipping fabric on rock strength for the coarse grained granite. This figure compares the variation in the various strength estimates with orientation. The values for the peak load failure, onset of dilatancy, crack initiation, and microcrack closure are derived from the uniaxial compression data (stress /strain curves) in Appendix C. The tensile strength is from the Brazilian tests.

The data presented in this figure clearly indicate that directional anisotropy is also evident in the coarse grained granite. In the coarse grained granite, samples are weakest in compression when the centimetre- to decimetre –scale layering intersects the sample axis at an angle between 20° and 40°. The rock is stronger (at the specimen scale) when the sample is within the plane of layering, or sharply inclined to it.

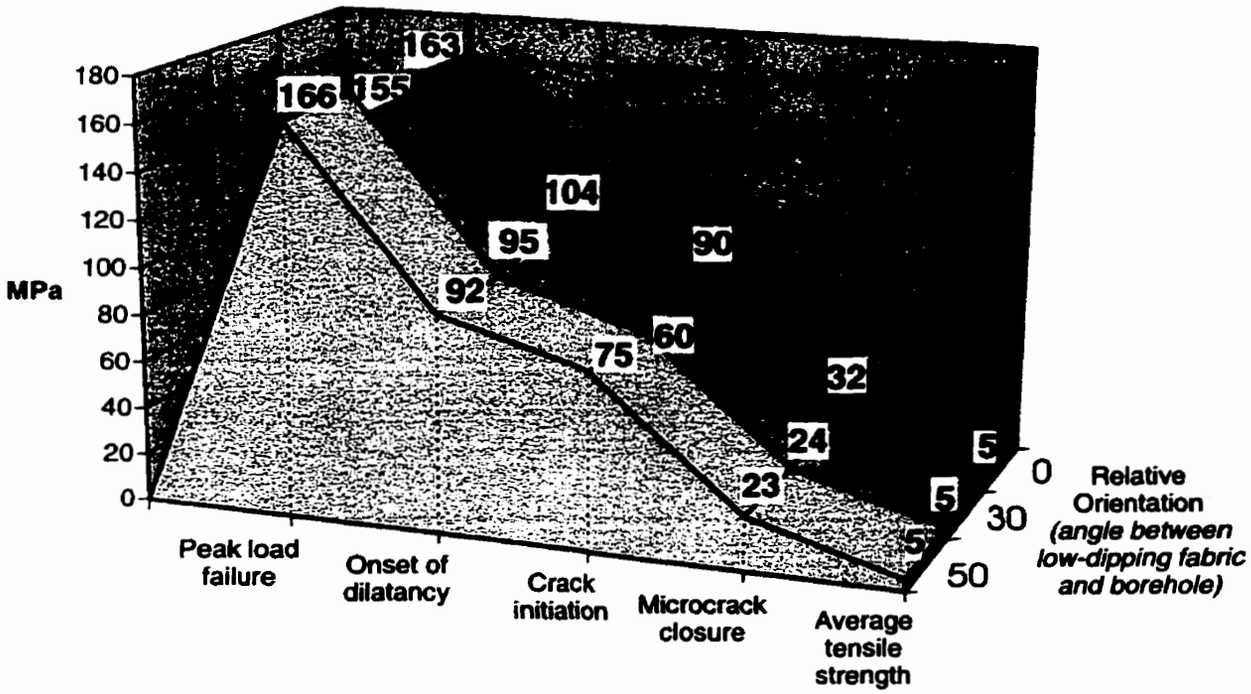



Figure 72. Influence of low-dipping fabric on rock strength in the medium grained granite. X-Y-Z comparison chart of the variation, with orientation, of the estimated average strengths for the medium grained granite. The values for the peak load failure, onset of dilatancy, crack initiation, and microcrack closure are derived from the uniaxial compression data (stress /strain curves) in Appendix C. The tensile strength is from the Brazilian tests. Samples are weakest in compression when the centimetre- to decimetre – scale layering intersects the sample axis at 30°. The rock is stronger (at the specimen scale) when the sample is within the plane of layering, or sharply inclined to it.

# Chapter 10

# SUMMARY AND CONCLUSIONS

The purpose of this thesis was to investigate and document the contribution of rock fabric on excavation damage development. The hypothesis, based on observation of failure in a variety of rocks at the URL, was that failure occurred at strengths lower than expected, because of fabric-induced strength anisotropy. The hypothesis was tested by undertaking a combined laboratory and field investigation. Although the different excavation response of texturally distinct granite varieties at the URL has been noted, previous investigations assumed that granite (at the URL site and elsewhere) could be treated as a continuous intact (unfractured) mass that is homogenous and isotropic in strength.

The conclusion from this investigation is that the Lac du Bonnet Batholith is not, in fact homogeneous, but contains structural elements and distinct textural varieties which differ significantly in their response to rock properties testing and to excavation.

Data collected in this study demonstrates unequivocally, and for the first time, the highly heterogeneous internal structure characteristic of allegedly homogenous batholiths, and the influence of this heterogeneity on rock mass response. Evidence for anisotropic behaviour of the rockmass is demonstrated to:

1. have occurred on the geological time-scale, and its influence is still present in recent excavations, and
2. at a wide range of scales, from the localization of fault zones along low-dipping layering at the kilometre-scale, to the localization of failure around tunnels by separation along this and other weakness directions.

## 10.1. Outline of Investigation

The hypothesis was tested through a detailed and combined program of mapping and rock properties testing. The geological setting and the general geology of the study site were described, with emphasis on the factors affecting the development of the ancient (natural) fractures. This description includes a synopsis of the site-scale patterns of natural microfractures, an extensive effort conducted as part of this investigation. To investigate the role of rock fabric on rock strength and damage development, three case studies were selected in the intact rock at the 420 Level of the Underground Research Laboratory (URL). The case studies included fine, medium and coarse grained varieties of the LDBB granite. Foliations of various types and orientations are present at each sample site, and borehole arrays were designed at each of the three locations to provide a wide range of intersection angles for rock properties testing.

The rock properties testing conducted was intended to define the influence of foliation and preferred microcrack orientations in creating strength anisotropies, which in turn influence excavation damage development. Techniques sensitive to planar anisotropy in rock were identified from the literature. It is recognized that fabric is just one of the factors contributing to lowering strength from what is seen under lab conditions. This report summarizes:

1) what was learned from the URL that is useful in future site characterization,
2) how geological features in granite influence excavation damage (link between geology and engineering), and
3) the role of rock fabric on rock mass response (including rock properties testing and in-situ stress measurement).

A relative ranking of rock strength correlating with average grain size is presented. The results are compared with field observations of excavation damage development. Directional anisotropy in rock strength is compared to the orientation of known fabric elements such as natural microcracks directions and some foliations. The results are consistent with tunnel mapping observations regarding preferential development of excavation damage in some rock types and in certain directions.

### 10.2. Identification of Anisotropy in the "Homogeneous" Lac du Bonnet Batholith

This investigation has revealed two basic errors in the literature regarding the structure of the LDBB and its influence on excavation response:

1. The Lac du Bonnet Batholith is a type example of a homogeneous granite pluton in the Canadian Shield (e.g. Beakhouse, 1991; Card and Poulsen, 1998).
2. With regard to excavation damage development, the LDBB has been used as textbook example of a homogeneous and isotropic granite (e.g. Hoek E., Kaiser P.K. and Bawden W.F., 1995. Support of Underground Excavations in Hard Rock).

Neither of these assumption is correct. At the URL, textural and compositional heterogeneities control: 1) the location of large-scale fault zones, 2) the extent, distribution and infilling characteristics of lesser-scale fractures between the fault zones, and 3) the extent and density of excavation-induced fractures. It has also been demonstrated by the testing program conducted here, that crack initiation and propagation are influenced by rock fabric and grain structure. Finally, breakout development and other types of excavation damage for example vary in response to textural and structural variations in the granite present at the tunnel- and site-scales.

### 10.3. Natural Microcracks

The site characterization included a compilation and synopsis of the studies of natural (i.e. ancient) microfracturing in quartz at the URL. This study was undertaken in three stages: 1) an initial reconnaissance survey, 2) a more detailed reconnaissance survey, and 3) a study of samples from the 420 Level only to correlate with the samples taken for geomechanical testing.

The results of this study indicate that the orientations and domains of ancient (natural) microcracking are consistent with the patterns and domains defined for the macro-scale fractures at this site. Microfractures are more abundant, more strongly oriented with more sets above the basal thrust fault (Fracture Zone 2), than in the sparsely fractured rock below the fault. Natural microfractures in the sparsely fractured rock below FZ2 are low dipping. In the moderately fractured rock above the fault, the microfracture pattern comprises a low dipping set and two subvertical sets. The low-dipping set in the sample from the 240 Level parallels the average orientation for Fracture Zone 2, as well as the compositional layering, which this fault follows. The low dipping set in the sample from the 130 Level parallels the orientation for Fracture Zone 3 (based on an average taken over a 500 $m^2$ area), but not the local compositional layering, which deviates here from the regional norm. This is consistent with other URL and surface work, which showed that the low-dipping fracture zones are fabric-controlled where the layering is near the theoretical failure plane, but not where the layering varies substantially in orientation.

Average microfracture lengths are dependent on depth from surface and grain size. The mean crack length decreases with depth from surface. In the moderately fractured rock above Fracture Zone 2, microfractures equal or exceed the grain size, while in the sparsely fractured rock at and below the 240 Level, the microfractures are usually less than the grain size.

The rock properties testing results indicate that the various foliations, compositional layering and natural microfractures influence rock properties, in particular inducing anisotropy to tensile and compressive strength.

### 10.4. Influence of Grain Size on Rock Strength

Strength, (whether expressed as tensile strength, crack closure strength, crack initiation stress etc.) decreases as the average grain size and heterogeneity of the granite increase. This trend is shown best by the minimum estimates for each type of strength. The mean and maximum strengths are much more variable, being affected by the relative size and orientation of specimen-scale heterogeneities. These influences are most pronounced in the coarser granites.

### 10.5. Influence of Rock Fabric on Compressive Strength

The testing program has also demonstrated that there is a directional anisotropy to the compressive and tensile strengths in each of the three phases of the LDBB granite, and that this anisotropy is controlled by the different fabrics present in each case. Directional anisotropy of strength and the influence of subvertical fabric on the compressive strength was noted in the fine grained granite. The compressive strength of the fine grained granite is dependent upon the orientation of the sampling borehole relative to the strike of the dyke. Its strength parallel to the dyke contact is 40% less than that measured (near-) normal to this direction.

Directional anisotropy is also evident in the medium and coarse grained granites, although its relationship to the rock fabric is more complex. Samples are weakest in compression when the centimetre- to decimetre– scale layering intersects the sample axis at an angle between 20° and 40°. The rock is stronger (at the specimen scale) when the sample is within the plane of layering, or sharply inclined to it. This interpretation is consistent with field observations and with the Point Load and Brazilian tests, in that the

contacts between layers provide more clearly defined and continuous weakness planes, than those structures that would promote axial splitting of the specimens (i.e. grain alignment or natural microcracks).

### 10.6. Influence of Grain Size on Tensile Strength

The Point Load test provides a relative index of the strength of the various rock types (the Point Load Index, Goodman, 1989), although this was a secondary objective of this experiment. The "point load strength index" is defined as follows:

$$\textit{Point Load strength index, } I_s = P/D^2$$

$$\textit{where } P = \textit{the load to break the specimen,}$$

$$\textit{and } D = \textit{the distance between the platen contacts.}$$

The test results are consistent with the uniaxial compression testing, in that relative strength decreases with increasing grain size. However, the principal objective of the Point Load Test was to identify weakness planes and their relative strengths. The result of this application is summarized in the following section.

### 10.7. Fabric-induced Anisotropy of Tensile Strength

In the Point Load test configuration, the fracture is free to form in any orientation parallel to the load. If however, the rock is anisotropic, there is a tendency for failure to follow one or more preferred orientations which generally correspond with some fabric element such as natural microfractures, mineral alignment, etc. In the case of the URL granite, there is a very clear tendency for the tensile failure to be localized by the foliations present in each sample. Which foliation is responsible varies with the sample, being dependent upon the orientation of the sampling borehole relative to the local fabric or fabrics.

In shallow dipping boreholes (< 45°), axial point loading of the core produces fractures following the strongest fabric (referred to for convenience as the large scale layering, but which is in fact a mixture of metre- to millimetre-scale compositional layering, mineral alignment, schlieren (in some cases showing evidence of slip) and natural microfractures in quartz. In steeply dipping boreholes, the layering is no longer favourably oriented with respect to axial loading, and the plane of failure then "jumps" to other, steeper dipping foliations. The presence of "preferred failure planes" is marked by significant directional variations in the average tensile strength (as defined by the Brazilian tests). In the coarse grained granite for example, the tensile strength can be 60% and 80% of the strength obtained when the sample is forced to break oblique to the primary and secondary foliations, respectively.

### 10.8. Influence of Natural Microcracks on Tensile Strength

In the case of the fine grained granite there are two distinct fabric elements to consider - a subvertical flow banding, and a low-dipping microcrack set. While the microcrack set is also present in the medium and coarse granite, it and its effects are easily masked by the subparallel foliation. In the fine grained granite, the two structures are nearly perpendicular to each other, and their affect on rock strength can be examined separately. The Brazilian tests indicate the tensile strength within the horizontal plane is somewhat lower when the loading axis is parallel to the flow banding, but the tendency for this is weak, and the scatter in the data is considerable. For testing done for the vertical plane, the data spread is much less, and the anisotropy is greater - with the tensile strength parallel to the natural microcrack set being approximately 65% of that normal to this

direction. In this rock unit, it appears that the natural microcracks in quartz have a greater influence on the tensile strength than the centimetre-scale compositional layering.

### 10.9. Summary of Variations by Rock Type

The coarse grained granite is the most heterogeneous of the types tested. A low-dipping fabric dominates its structure at the tunnel and the hand specimen scales. This fabric includes a largely concordant assemblage of compositional and textural layering, an alignment of biotite and feldspar laths, and a preferred orientation of natural microcracks in quartz. Point Load and Brazilian tests indicate that this fabric imparts a mechanical anisotropy, with the tensile strength in the plane of the fabric being about 40% of that determined for the normal to this plane. This anisotropy is obvious underground, where it facilitates scaling, and the development of excavation-induced fractures where these are developed in an orientation comparable to the predominant foliation. There is also a very clear tendency for the tensile failure to be localized by other foliations. Which foliation is responsible is dependent upon the orientation of the sampling borehole.

In shallow dipping boreholes (< 45°), axial point loading of the core produces fractures following the strongest fabric (the large scale layering as defined above). In steeply dipping boreholes, the layering is not favourably oriented with respect to axial loading, and the plane of failure then "jumps" to other, more steeply dipping foliations. The transition from one foliation plane to another is reflected by fluctuations in the tensile strength as noted in the preceding section.

The properties of the fine grained granite are influenced by two distinct fabric elements - a subvertical flow banding, and a low-dipping microcrack set. While the microcrack set is also present in the medium and coarse granite, it and its effects are easily masked by the subparallel foliation. In the fine grained granite, the two structures are nearly perpendicular to each other, and their influence on rock mass response can be examined separately. Brazilian tests indicate that the tensile strength is somewhat lower when the loading axis is parallel to the subvertical flow banding, but the tendency for this is weak, and the scatter in the data is considerable. A stronger correlation is shown by the influence of the low-dipping natural microcrack set, with the tensile strength parallel to the natural microcrack set being approximately 65% of that normal to this direction. In this rock type, it appears that natural microcracks in quartz have a greater influence on the tensile strength than centimetre -scale compositional layering.

The Point Load testing indicated that failure was controlled by heterogeneities in the rock, such as foliations or the preferred orientation of natural microcracks, which are close to or parallel to the plane of failure under this loading geometry. This tendency was present to some degree in all three of the rock types tested, but varies systematically with the degree of heterogeneity. The coarse grained granite has the most repeated tendency to split along dominant (low-dipping) fabric elements, with the lowest average tensile strength and Point Load strength index. The medium grained granite has a lesser tendency to split along dominant (and other) fabric elements, but with a comparable average tensile strength and Point Load strength index. The fine grained granite has the least repeated tendency to split along any fabric, which in this case alternated between the

steeply-dipping flow banding and the plane correlating with the preferred orientation of natural microfractures, and with the highest average tensile strength and Point Load strength index.

This systematic variation reflects the relative strength of the low-dipping fabric, which comprises a mixture of co-planar compositional layering, mineral alignment, and natural microcracks. All of these elements are present and well developed in the coarse granite, the weakest and most mechanically anisotropic of the group. Where this fabric is less well developed or absent, as in the medium and fine grained granites, this is reflected by an increase in rock strength and isotropy at the specimen –scale.

At the tunnel-scale, these differences in mechanical properties result in a differences in excavation damage including: whether breakouts develop or not, and the rate at which they develop, varying from hours in the case of the coarse grained granite, to days or weeks in fine and more homogeneous granites.

### 10.10. Application of Results

The geological investigations described in this thesis have the following significance with respect to the selection of a site for a disposal facility.

1. It must be recognized that even "simple" batholiths such as the Lac du Bonnet Batholith have complex internal structure. This structure is a product of the mode of intrusion, cooling history, and later deformation.
2. The rock fabric produced by these events exerts a significant control on the location, extent and orientation of natural (i.e. ancient) fractures. This association therefore

provides a guide to predicting fracture occurrence in the rock mass. It also provides an indication of the potential rock mass response to excavation.

3. Subtle variations in granite fabric can lead to significant differences in the rock mass response to excavation. Foliations and preferred orientations of natural microcracks may reduce strength, and act as weakness planes, in one or more directions. This anisotropy will influence the development and propagation of excavation induced fractures and breakouts.
4. The results presented in this thesis highlight the risk of arbitrarily correlating excavation stability with lithology, or in assuming that there is a typical behaviour from any given rock type. The Lac du Bonnet Batholith granite is not isotropic, and excavation response was shown to vary with the orientation and shape of the excavation relative to the rock structure. Granite types that resisted damage development, including breakout, in the vertical shaft, were more susceptible to damage in the horizontal tunnels. It is to be expected that similar results will occur in other granites.

Based on these findings, the following recommendations are made with regard to the site characterization program required to design, construct and monitor a disposal facility:

1) The geology program for site characterization in granites must include a single program of mapping, core logging and the compilation of the results through the construction of geological models.
2) This effort must be done in a comprehensive and consistent manner rather than as a series of separate follow-on investigations funded by individual experiments. The

latter may produce spurious results or overlook significant geological influences, and in the process, be more expensive over the duration of the facility.

3) The geology program should define the lithology and fabric of the rockmass as well as the natural fractures and types of excavation damage.

4) The effort on lithology should focus on identifying the litho-structural domains that correlate with the intrusion history of the pluton, and which influence the ancient and recent rock mass response (i.e. the natural and induced fractures, respectively). These litho-structural domains vary in texture and composition (as in a dyke which changes from fine to coarse grained as its dimensions increase). Characteristic rock units should then be identified, and subjected to standard petrography, as well as the analysis of fabric and natural microcracks.

5) This effort should only then be followed by the selection of a suite of representative rock types for rock properties testing. As this study has shown, colour is not an appropriate criterion on which to base a program of geotechnical investigation and facility design.

6) There should be a proper core logging program, not just of the current borehole, but of the entire array of boreholes in view of the decametre-scale layering and directional sampling biases identified in this study. Each borehole should be logged according to a standard procedure, which should include a complete analysis of the rock structure.

7) An alternative approach to defining the in-situ properties of highly stressed rock at depth, would be to identify the granite varieties present at the intended test location and depth, and then to obtain samples of these same varieties where they are exposed

at or near the surface. In adopting this approach, it would be necessary to determine if the granites at surface and at depth differ substantially in the characteristics of their natural microcracks.

8) Efforts to determine the in-situ stresses by overcoring or hydrofracing, and to interpret the results of ground penetrating seismic and radar surveys must incorporate the structural and mechanical anisotropy of the granites.

a) Overcore results are dependent upon orientation of a borehole relative to foliations and natural microcrack sets. In the sample itself, strain cells can be adversely affected by their positioning with respect to textural heterogeneities such as coarsely crystalline feldspars. As such, instrumentation placement should be based on a thorough logging, not just of a single borehole, but of the entire array of boreholes which have been designed to provide a range of samples with which to test for directional anisotropy.

b) Determination of in-situ stress by the hydrofracturing method assumes that the rockmass is isotropic, that the tensile fracture created by the test is normal to the least principal stress direction, and that this fracture is not deflected by fabric anisotropy. However, careful excavation at the URL through fractures formed by hydrofracture tests revealed that hydrofracture orientations were nearly always following either the low-dipping compositional layering or the steeply dipping pegmatite dykes (an incipient fracture direction in much of the URL). As such the results from these and similar hydrofracture tests at this site are suspect. Local-scale structures must be considered in the siting of hydrofrac tests and in their interpretation.

9) Ground penetrating radar and portable seismic surveys have been used in attempts to quantify excavation damage about a tunnel. Hayles et al. (1991) found that the pattern of anomalies could not be interpreted without some knowledge of the variations in rock composition and rock fabric. It is recommended that applications of ground penetrating radar and portable seismic be done in an area of a tunnel selected for optimum geology, and that the structure and microstructure be understood prior to running an experiment. Attempts to fill in the geology later to address difficulties in interpretation may not be successful.

These efforts will provide:

- more realistic models of batholith shape, internal structure, and the distribution of both natural and induced fractures,
- better (i.e. more representative) sampling for rock properties testing.
- better understanding of in-situ rock strength, and
- a better understanding of the rock mass response to different excavation geometries and methods.

# REFERENCES

ASTM. 1974. Annual Book of ASTM Standards "Standard method of test for direct tensile strength of rock core specimens"; Designation: D2936-71; American Society for Testing and Materials.

ASTM. 1989. Annual book of ASTM Standards, Section 4: Construction, Volume 04.08: Soil and Rock; Building Stones; Geotextiles. American Society for Testing and Materials, 1916 Race Street, Philadelphia, PA 19103-1187, USA.

Baek, H. 1997. Effect of the rock fabrics on the physical properties of a granite. Geological Society of America (GSA), Abstracts with Programs - Geological Society of America vol. 29 no. 6; p. 447.

Balk, R. 1937. Structural behaviour of igneous rocks: Geological Survey of America, Memoir 5, 177 pages.

Bawden, W.F. 1982. In situ stress determination by hydraulic fracturing in well bore URL-6. A report prepared by Geotechnical Resources Ltd. for Atomic Energy of Canada Limited.

Beakhouse, G.P. 1977. A subdivision of the western English River Subprovince. Canadian Journal of Earth Sciences 14: 1481-1489.

Beakhouse, G.P. 1991. Winnipeg River Subprovince. In: Geology of Ontario, (eds. P.C. Thurston, H.R. Williams, R.H. Sutcliffe and G.M. Stott). Ontario Geological Survey, Special Volume 4, Part 1, 279-302.

Behrestaghi, M.H.N., Rao, K.S., and Ramamurthy, T. 1996. Engineering geological and geotechnical responses of schistose rocks from dam project areas in India, Engineering Geology 44(1-4) 183-201.

Benn, K., Ham, N.H., Pignotta, G.F., and Bleeker, W. 1998. Emplacement and deformation of granites during transpression: magnetic fabrics of the Archean Sparrow pluton, Slave Province, Canada. In: Special Issue: Extraction, Transport and Emplacement of Granitic Magmas, edited by K. Benn, A.R. Cruden and E.W. Sawyer, Journal Of Structural Geology Vol. 20 (9-10), 1247-1259.

Bergeron, M. 1991a. Granite Industry in Canada, 1990 Summary Industrial Minerals Division, Mineral Policy Sector of Energy Mines and Resources Canada.

Bergeron, M. 1991b. Granite Industry in Canada, Directory. Mineral Policy Sector, Industrial Minerals Division, MP-4825-2.

Bergeron, M. 1991c. Granite industry in Canada. 1991 Directory. Energy Mines and Resources, Mineral Policy Sector, Industrial Minerals Division, MP4825-2.

Bieniawski, Z.T. 1967a. Stability concept of brittle fracture propagation in rock, Eng. Geol. Vol. 2, p. 149-162.

Bieniawski, Z.T. 1967b. Mechanism of brittle fracture of rock, Int. J. Rock Mech. Min. Sci. Vol. 4. p. 395-430.

Bieniawski, Z.T. 1972. Propagation of brittle fracture in rock. Proceedings, 10th Symposium Rock Mechanics, AIME, New York, 409-427.

Bieniawski, Z.T. 1975. The Point Load test in geotechnical practice. Engineering Geology 9, 1-11.

Bouchez, J.L. 1997. Granite is never Isotropic: An introduction to AMS studies of granitic rocks. In: Granite: From segregation of melt to emplacement fabrics, eds. J.L Bouchez, D.H.W. Hutton, and W. E. Stephens, p. 95-112. Kluwer Academic, Dordrecht.

Breaks, F.W. 1991. English River subprovince. In Geology of Ontario, Thurston, P.C., Williams, H.R., Sutcliffe, R.H., and Stott, G.M. (eds.). Ontario Geological Survey Special Volume 4, Part 1, 239-278.

Broch, E. 1974. The influence of water on some rock properties. In: Advances in Rock mechanics, Proc. 3rd Congress, International Society of Rock Mechanics. Denver, 1974, Part A, pages 33-38.

Broch, E. 1983. Estimation of Strength Anisotropy Using the Point-Load Test. International Journal of Rock Mechanics, Min. Sci. & Geomech. Abstr. Vol. 20, No. 4, 181-187.

Broch, E. and Franklin, J.F. 1972. The Point Load Test. International Journal of Rock Mechanics, Min. Sci. 9, 669-697.

Brown, A. and Everitt, R.A. 1984. Hydrofracing experiment in shaft. 2 p. plus 1 map. Memo #URL-GEN-M007 (URL-GEN-M00 Series).

Brown, A., Soonawala, N.M., Everitt, R.A., and Kamineni, D.C. 1989a. Geology and geophysics of Underground Research Laboratory site, Lac du Bonnet batholith, Manitoba. Canadian Journal of Earth Sciences, 26, 404-425.

Brown, A., Kamineni, D. C., and Everitt, R.A. 1989b. Deformation in the Lac du Bonnet Batholith, Man., Canada. In: Geological Association of Canada/Mineralogical Association of Canada Annual Meeting, Montreal, PQ, Canada, May 15-17, 1989, Program with Abstracts, vol. 14, p. 54.

Brown, E.T. 1981. Rock Characterization testing and monitoring. ISRM Suggested methods for determining the uniaxial compressive strength of rock materials and the point load strength index"; International Society for Rock Mechanics; Commission on Standardization of Laboratory and Field Tests; Committee on Laboratory Tests; Document no. 1; October 1972.

Card, K.D. and Poulsen, K.H. 1998. Geology and Mineral Deposits of the Superior province of the Canadian Shield. Chapter 2 In: Geology of the Superior province of the Canadian Superior and grenville Provinces and Precambrian Fossils in north America, (co-ord) S. Lucas; Geological Survey of Canada, no. 7, p.13-194. (also Geological Survey of America, The Geology of North America, v. C-1).

Carter, B.J. 1992. Size and Stress gradient effects on fracture around cavities. Rock Mechanics and Rock Engineering, 25, (3), 167-186.

Cerny, P., Fryer, B.J., Longstaffe, F.J., and Tammemagi, H.Y. 1987. The Archean Lac du Bonnet Batholith, Manitoba: Igneous history, metamorphic effects, and fluid overprinting. Geochimica et Cosmochimica Acta, 51: 421-438.

Chandler, N.A., Read, R.S., and Martin C.D. 1996. In situ stress measurement for nuclear fuel waste repository design. In: proceedings of the 2nd North American Rock mechanics Symposium: NARMS'96. A regional conference of ISRM/ Montreal/ Quebec/ Canada. 19-21 June 1996. 929-936.

Chen, R. 1993. In situ and laboratory study of potash deformation, with reference to Saskatchewan potash. University of Manitoba, Winnipeg, MB, Canada. Doctoral Thesis. 280p.

Chen, Y., Nishiyama, T., Kita, H., and Sato, T. 1997. Correlation between microfracture type and splitting planes of Inada granite and Kurihashi granodiorite. Journal of the Japan Society of Engineering Geology. v. 38:4. p. 196-204.

Chernis, P.J. 1981. Scanning Electron Microscope Study of the Microcrack Structure of a Granite Sample from Pinawa, Manitoba. Atomic Energy of Canada Limited Technical Record TR-173.

Chernis, P.J. 1984. Comparison of the Pore-Microcrack Structure of Shallow and Deep Samples of the Lac Du Bonnet Granite. Atomic Energy of Canada Limited Technical Record TR-223.

Christiansson, R. and Hamberger, U. 1991. Blasting Damage Investigations in Access Ramp Section 0/256 – 0/565 M. No. 1. Tunnel Excavation and Geological Documentation. Äspö Hard Rock Laboratory Progress Report 25-91-12.

Clarke, D.B., Henry, A.S., and White, M.A. 1998. Exploding xenoliths and the absence of `elephants' graveyards' in granite batholiths. In: Special Issue: Extraction,

Transport and Emplacement of Granitic Magmas, edited by K. Benn, A.R. Cruden and E.W. Sawyer, Journal Of Structural Geology Vol. 20 (9-10), 1325-1343.

Cloos, E. 1946 Lineation: Geological Survey of America, Memoir 18, 122p.

Coates, D.F. and Gyenge, M. 1966. "Suggested test specifications on plate load testing and conventional uniaxial compression testing techniques for rock mechanics"; American Society for Testing and Materials; Special technical publication no. 402; p. 34-34.

Colback, P.S.B. and Wild, B.L. 1965. The influence of moisture content on the compressive strength of rock, Proc. 3rd Canadian Rock Mech. Symp., Toronto, 66-83.

Corthesy, R., Gill, D.E., and Leite, M.H. 1993. Doorstoppers Stress Measurements at the 420 Level of the URL. C.D.T. Project P1596, L'Ecole Polytechnique de Montreal.

Cruden, A.R. 1990. Flow and Fabric Development during the diapiric rise of magma. Journal of Geology, 98, 681-698.

Cruden, A.R. 1998. On the emplacement of tabular granites. Journal of the Geological Society of London, 155, 852-862.

Cruden, A.R. and Lanueau, P. 1994. Structure, magnetic fabric and emplacement of the Archean Lebel Stock, SW Abitibi Greenstone Belt, Journal of Structural Geology, 16, No. 5, 667-691.

Cruden, A.R., Davis D., Menard T., and Robin, P.-Y.F. 1997. Structural and Geochronological relationships between the Winnipeg River and the Wabigoon Subprovinces: Implications for the Terrane Accretion Model. In: Western Superior transect, Lithoprobe Report 63, University of British Columbia.

Cruden, A.R., Koyi H., and Schmeling H. 1995. Diapiric basal entrainment of mafic into felsic magma. Earth and Planetary Science Letters, 131, 321-340.

Cruden, D.M. 1983. Long-term behaviour in compression of Lac Du Bonnet granite. Atomic Energy of Canada Limited Technical Record TR-211.

Dale, T.N. 1923. The commercial granites of new England. I, U.S. Geol. Sur. Bull., 738, 1-97.

Davis, D.W., Corfu, F., and Krogh, T.E. 1986. High precision U-Pb geochronology and implications for the tectonic evolution of the Superior Province. In Workshop on tectonic evolution of greenstone belts. Edited by M.J. deWit and L.D. Ashwal. Lunar and Planetary Institute, Houston, TX, Technical Report 86-10, p. 77-79.

Davis, G.H. 1984. Structural geology of rocks and regions. John Wiley & Sons.

Djouadi, T., Gleizes, G., Ferré, E., Bouchez, J.L., Caby, R., and Lesquer, A. 1997. Oblique magmatic structures of two epizonal granite plutons, Hoggar, Algeria: late-orogenic emplacement in a transcurrent orogen. In: Special Issue: The Adolphe Nicolas Volume , edited by D. Mainprice, F. Boudier and J.-L. Bouchez, [Editorial] , Tectonophysics, Volume 279, Issue 1-4, 15-October-1997 , Tectonophysics Vol. 279 (1-4), 351-374.

Doe,T.W. 1987. Hydraulic fracturing stress measurements in the shaft probe hole at the Underground Research Laboratory, Manitoba . Golder Associates Inc. report to Battelle Memorial Institute, Office of Waste Technology Development.

Doe,T.W. 1989. Hydraulic fracturing stress measurements in horizontal holes at the 240-and 420-m Levels of the Underground Research Laboratory. Golder Associates Inc. report to Atomic Energy of Canada Limited.

Donath, F.A. 1964 Strength variation and deformational behaviour in anisotropic rocks. In: W. Judd (Ed.), State of Stress in the Earth's Crust, Elsevier, New York, p. 281-300.

Duevel, B. and Haimson, B. 1997. Mechanical Characterization Of Pink Lac Du Bonnet Granite: Evidence Of Nonlinearity And Anisotropy. International Journal of Rock Mechanics and Mining Sciences 1997. Vol. 34, No. 3-4, 117.

Duncan, S.E.J. 1987. Mesoscopic scale fractures, microscopic fractures and tensile strength characteristics of the Shield Granite: Southeastern Manitoba. Unpublished M.Sc. thesis, University of Manitoba, Department of Geological Sciences, Winnipeg, Manitoba. 136p.

Engelder, T. 1993. Stress Regimes in the Lithosphere. Princeton University Press. 457p.

Everitt, R.A. 1993a. Orientation and Distribution of Inclusion-Bearing Microcracks in core from the 420m Level of the Underground Research Laboratory, Pinawa, Manitoba. Prepared by Don H. Rousell, Department of Geology, Laurentian University, Sudbury, Ontario, P3E 2C6. URL Library Identifier 01.73.03.

Everitt, R.A. 1993b. Petrography of Granitic Rock Samples from the 420m Level of the Underground Research Laboratory, Pinawa, Manitoba. Prepared by Dana Kelly, David S. Peck and Richard S. James, Department of Geology, Laurentian University, Sudbury, Ontario, P3E 2C6. URL Library Identifier 01.73.03.

Everitt, R.A. 1995. Status of Pegmatite Testing Study. Fracture Initiation and Propagation by Compressive Loading and Unloading in Component Mineral Phases of the Lac du Bonnet batholith. Unpublished internal memorandum. URL-01.84.41, R-3002-07-00.

Everitt, R.A. 1998. The Influence of Rock Fabric on Excavation Damage in the Lac Du Bonnet granite. Report for DS-3, Rock Engineering Studies, URL-G-4200, Memo AGB-98-008.

Everitt, R.A. 2000. Synopsis of natural microcrack work at the Underground Research Laboratory. Unpublished report.

Everitt, R.A. and Brown A. 1986. Subsurface Geology of the Underground Research Laboratory. An Overview of Recent Developments. 25 p. 14 Tables and Figures In: Proceedings of the 20th Information Meeting of the CNFWMP held in Winnipeg, Canada, October, 1985. AECL Technical Record TR-375. p. 146-181.

Everitt, R.A. and Brown A. 1996. Geological Mapping of AECL Research's Underground Research Laboratory. A Cross Section of Thrust faults and Associated Fractures in the Roof Zone of an Archean batholith. Proceedings of Fractured and Jointed rock masses, a regional conference of the International

Society for Rock Mechanics. June 1-6th 1992. Granlibakken, California. Volume 1 , p. 1-11.

Everitt, R.A. and McArthur, J.R. 1994. Petrofabric analysis of natural microfractures in quartz of the Lac du Bonnet Batholith. Unpublished AECL Internal memorandum.

Everitt, R.A. and McGregor, C. 1990. Pilot Petrofabric Study of Natural Microfractures in core from AECL Research's Underground Research Laboratory. Unpublished.

Everitt, R.A. and Passmore, W. 1993. Shaft Extension Data Report - Wall Topography and Geology map. Unpublished AECL internal memorandum.

Everitt, R.A. and Read, R.S. 1989. Geology of the 240 Level of the Underground Research Laboratory - Volume I - General Geology. AECL Technical Record TR-491-1, URL-EXP-006-R001.

Everitt, R.A. and Rousell, D.H. 1994. Orientation and Distribution of Inclusion-Bearing Microfractures in core from the 420m Level of the Underground Research Laboratory, Pinawa, Manitoba. Prepared by Don H. Rousell, Department of Geology, Laurentian University, Sudbury, Ontario. URL Library Identifier 01.73.03.

Everitt, R.A. and Woodcock, D.R. 1994. Contribution to the HFT Characterization Summary Report. Mapping Results for the 600 mm Heated Failure Test Borehole. Unpublished AECL Internal Memorandum. URL-01,84.41, AGB-94-190.

Everitt, R.A. and Woodcock, D.R. 1999 Geology of the Deep Stress Measurement Borehole. OPG Nuclear Waste Management Report 06819-REP-01200-10012-R00. August 1999. 146p.

Everitt, R.A., Brown, A., Davison, C.C., Gascoyne, M., and Martin, C.D. 1990. Regional and local setting of the Underground Research Laboratory. In Proceedings of the Symposium on Unique Underground Structures, Denver, Colorado, p. 64-1 to 64-23.

Everitt, R.A., Brown, D. A., Boychuk, D.M., and Gann, P. 1994. Excavation damage and organic growth in a 1.2m diameter borehole. Poster presentation at the 5th Annual International High-Level Radioactive Waste Management Conference & Exposition. May 22-26th 1994.

Everitt, R.A., Chapman, A.E., Northrop, P., and Brown, A. 1987. Geology of the Underground Research Laboratory Ventilation Raise. Atomic Energy of Canada Limited, Technical Record TR-404.

Everitt, R.A., Chernis, P.J., and Good, D.H. 1989b. Geology of the Room 209 Floor. Contributions to the Excavation Damage Assessment Experiment. 5 p., 4 figures, 4 tables and 2 maps.

Everitt, R.A., Chernis, P.J., Good, D.H., and Grogan, A. 1989. Mapping of the Excavation Damage Zone around the Circular Access Shaft at Atomic Energy of Canada Limited's Underground Research Laboratory. In: Excavation Response in Geological Repositories for Radioactive Waste. Proceedings of an NEA Workshop, April 1988, Winnipeg, Canada, p. 271-282.

Everitt, R.A., Gann, P., and Boychuk, D.M. 1993. Mine-by Experiment : Part 7 - Geological Setting and General Geology. AECL Technical Record RC-1080.1. This report and the following supplements complete the geological characterization of the 420 Level for the Mine-by Experiment and the Connected permeability Experiment.

Everitt, R.A., Gann, P., Brown, A., Woodcock, D.R., Sikorsky, R.I., and Ejeckam, R. B. 1998. Subhorizontal Layering in the "Homogeneous" Lac du Bonnet Batholith at the Underground Research Laboratory, South Eastern Manitoba. In: Special Symposium on Extraction, Transport and Emplacement of Granitic Magmas, Ottawa '97 GAC-MAC Annual Meeting. Journal of Structural Geology. Volume 20, No. 9/10, p. 1291-1304.

Everitt, R.A., Martin, C.D., Brown , A., and Gann, P. 1994. Pop-Ups And Fractures In Granite At The Canital Quarry, Manitoba. Presented at the Geological Association of Canada - Mineralogical Association of Canada Joint Annual Meeting, May 16-18th, 1994.

Everitt, R.A., McMurry, J., Brown, A., and Davison, C. 1996. Geology of the Lac du Bonnet Batholith, Inside and Out: AECL's Underground Research Laboratory, Southeastern Manitoba. Field Excursion B-5: Guidebook,. Geological Association of Canada - Mineralogical Association of Canada Joint Annual Meeting 1996, Winnipeg, Manitoba, Canada. 1996 May 30. Copyright Geological Association of Canada, Winnipeg '96 Committee.

Feves, M. and Simmons G. 1976. Effects of Stress on Cracks on Westerly Granite. Bulletin of the Seismological Society of America. Vol. 66, (5), p. 1755-1765.

Friedman, M. and Bur, T.R. 1974. Investigations of the relations among residual strain, fabric, fracture and ultrasonic attenuation and velocity in rocks. International Journal of Rock Mechanics and Mining Science and Geomechancial Abstracts, Vol. 11, 221-234.

Gerow, M.C. and Bellinger, J.A. 1990. Northwestern region industrial minerals program-1989; in Report of Activities 1989, Resident Geologists, (eds. K.G. Fenwick, P.E. Giblin and A.E. Pitts). Ontario Geological Survey, Miscellaneous Paper 147, 161-179 p.

Gleizes, G., Leblanc, D., Santana, V., Olivier, P., and Bouchez, J.L. 1998. Sigmoidal structures featuring dextral shear during emplacement of the Hercynian granite complex of Cauterets--Panticosa (Pyrenees). In: Special Issue: Extraction, Transport and Emplacement of Granitic Magmas , edited by K. Benn, A.R. Cruden and E.W. Sawyer, Journal of Structural Geology, Volume 20, Issue 9-10, 1229-1245.

Goodman, R.E. 1989. Introduction to Rock Mechanics. 2nd edition. John Wiley and Sons.

Gorski, B. 1987. Post-failure uniaxial strength determinations using a servo-hydraulic test system. Division Report MRL 87-33 (INT), CANMET, Energy, Mines and Resources, Canada, 77p.

Guermani, A. and Pennacchioni, G. 1998. Brittle precursors of plastic deformation in a granite: an example from the Mont Blanc massif (Helvetic, western Alps). In: Special Issue: Structures and Properties of High Strain Zones in Rocks , edited by E.H. Rutter, A. Boriani, H. Brodie and L. Burlini, Journal of Structural Geology, Volume 20, Issue 2-3, -February-1998 , Journal of Structural Geology Vol. 20 (2-3) p. 135-148.

Gyenge, M. 1980. Pit Slope Manual Supplement 3-1- Laboratory Classification Tests; CANMET (Canada Centre for Mineral and Energy Technology, formerly Mines Branch, Energy, Mines and Resources Canada), CANMET Report 77-25; 31p; May 1977. Reprinted May 1980.

Gyenge, M. and Herget G. 1977. Laboratory tests for design parameters. Mining Research Laboratories Pit Slope Manual Canada Centre for Mineral and Energy Technology, Ottawa. Supplement 3-2, p. 1-74. (1977).

Gyenge, M. and Ladanyi B. 1977. Pit Slope Manual Supplement 3-5 – sampling and specimen preparation. CANMET, Report 77-29, 30p.

Haimson, B.C. 1982. Hydrofracturing in situ stress measurements in the Lac du Bonnet batholith drillholes URL-1 and WN-4. A report prepared for Atomic Energy of Canada Limited.

Haimson, B.C. 1990. Stress measurements in the Sioux Falls Quartzite and the state of stress in the midcontinent. In: Proceedings of the 31st U.S. Symposium on Rock Mechanics, Golden, Colorado, June 18th-20th 1990. 31, 397-404.

Haimson, B.C. 1991. Hydraulic fracture stress measurements at the 420 Level, Underground Research Laboratory, AECL. A report to Atomic Energy of Canada Limited.

Haimson, B.C. 1992. Hydraulic fracturing stress measurements in three holes , Room 415, URL. A report to Atomic Energy of Canada Limited.

Hayles, J. G., Stevens, K. M., Serzu, M. H., and Lodha, G. S. 1991. Developments in the Mini-charts system and field trial results in measuring excavation damage in granite. In: Fourth International Symposium on Borehole Geophysics for Minerals, Geotechnical and Groundwater Applications, Toronto, ON, Canada, August 18-22, 1991, International Symposium on Borehole Geophysics for Minerals, Geotechnical, and Groundwater Applications, vol. 4, p. 267-306.

Hayles, J.G., Serzu, M.H., Tomsons D.K., Everitt. R.A., and Lodha G.S. 1994. Cross-hole seismic hole geophysical surveys for characterizing an area of moderately fractured rock. Presented at the CSEG Convention in Calgary Alberta May 10-12th, 1994.

Heuze, F.E. 1980. Scale effects in the determination of rock mass strength and deformability. Rock Mechanics, 12, (3)

Hill, J.R. 1989a. Dimension stone inventory of Islington I.R. No.29; Independent Consultant Report for Rat Portage Indian Band, Kenora, Ontario, January, 1989, 23 p.

Hill, J.R. 1989b. Dimension stone inventory of Swan Lake I.R. No.29; Independent Consultant Report for Rat Portage Indian Band, Kenora, Ontario, January,1989, 22 p.

Hill, J.R. 1990a. Phase II Follow-up Dimension Stone Evaluation of English River I.R. 21 prepared by Northwood Geoscience for The Assin Partnership, Kenora, Ontario, March, 1990, 17 p.

Hill, J.R. 1990b. Phase II Follow-up Dimension Stone Evaluation of Swan Lake I.R. 29 and Islington I.R. 29 prepared by Northwood Geoscience for The Assin Partnership, Kenora, Ontario, March, 1990, 24 p.

Hinz, P., Landry, R.M., and Gerow, M.C. 1994. Dimension stone occurrences and deposits in Northwestern Ontario. Ontario Geological Survey, Open File Report 5890.

Hodgson, K. and Cook, N.G.W. 1970. The effect of size and stress gradient on the strength of rock. Proc. 2nd Congress International Society of Rock Mechanics, Belgrade, Vol. 2, 1970. Paper 3-5.

Hoek, E. and Brown, E.T. 1980. Underground Excavations in Rock. Institute of Mining and Metallurgy, London. 5. Support of Underground Excavations in Hard Rock. A.A. Balkema.

Hoek, E., Kaiser, P.K., and Bawden, W.F. 1995. Support of Underground Excavations in Hard Rock, A.A. Balkema.

Homand-Etienne, F. and Sebaibi, A. 1996. Study of microcracking of the Lac du Bonnet granite. Eurock'96. Barla (ed.), 1996 Balkema, Rotterdam. P. 1353-1360.

Hooker, V.E. and Bickel, D.L. 1974. Overcoring equipment and techniques used in rock stress determination: Information Circular of the U.S. Bureau of Mines, No. 8618.

Hoskins, J.R. and Horino, F.G. 1969. The influence of spherical head size and specimen diameter on the uniaxial compressive strength of rocks. United States Bureau of Mines Report of Investigations, 7234, 16p.

IUGS (International Union of Geological Sciences) Subcommission on the Systematics of Igneous Rocks. 1973. Plutonic rocks, classification, and nomenclature. Geotimes, 18(10): 26-30.

Jaeger, J.C. 1960. Shear Failure of Anisotropic Rocks. Geol. Magazine, 97, 65-72.

Jaeger, J.C. and Cook, N.G.W. 1979. Fundamental of rock mechanics, 3rd edition, Chapman and Hall, London, 593p.

Jang, B., Wang, H.F., Ren X., and Kowallis B.J. 1989. Precambrian paleostress from microcracks and fluid inclusions in the Wolf River batholith of central Wisconsin. Geological Society of America Bulletin, 101: 1457-1464.

Kennedy, M.C. and Shelock, E.J. 1988. Northwestern Region industrial minerals program-1987; in Report of Activities, 1987, Resident Geologists, (ed. C.R. Kustra). Ontario Geological Survey, Miscellaneous Paper 138, 161-174.

Kornfalt, K., Wikman, H., Nordlund, E., and Chunlin, L. 1991. Blasting Damage Investigations in Access Ramp Section 0/256 – 0/565 M. No. 4. Optical

examination of microcracks in thin sections of core samples and acoustic emission of core samples. Äspö Hard Rock Laboratory Progress Report 25-91-15.

Kronenberg, A.K., Russell, J.E., and Carter, N.L. 1990. Anisotropic yielding of rocks at high temperatures and pressures. Final report (Progress report) Texas A and M Univ., College Station. Center for Tectonophysics. Corp. Source Codes: 004736025; 9504864 Sponsor: Department of Energy, Washington, DC. Report No.: DOE/ER/13711-T1 14 Oct 90 15p. Sponsored by Department of Energy, Washington, D.C. Contract No.: FG05-87ER13711.

Kudo, Y., Hashimoto, K, Sano, O., and Nakagawa, K. 1987. Relation between physical anisotropy and microstructures of granitic rock in Japan. In: Proceedings of the Sixth Congress of the International Society for Rock Mechanics, edited by G. Herget and S. Vongpaisal, p. 429-432, A.A. Balkema, Rotterdam, The Netherlands.

Kulander, S.L.D. and Ward B.J. Jr. 1990. Fractured Core Analysis: Interpretation, logging and use of natural and induced fractures in core by. in: Methods in Exploration Series, n.8, American Association of Petroleum Geologists. 1990.

Kusumi, H.; Mine, Y.; and Nishida, K. 1997. Effect of structural anisotropy on deformation properties of granite under cyclic loading. Journal of the Society of Materials Science, Japan v. 46 no. 9 Sep 1997. p 1029-1034.

Lacey, J.K. 1989. Building stone inventory of the Sudbury Resident Geologist's Area. Ontario Geological Survey, Open File Report 5721.

Lajtai, E.Z. 1980. Tensile Strength and its Anisotropy measured by Point – and Line-Loading of Sandstone. Engineering geology, 15, 163-171.

Lajtai, E.Z. 1982. The fracture of Lac du Bonnet Granite. Unpublished report to Atomic Energy of Canada Limited, 126 p.

Lajtai, E.Z. and Dzik, E. 1996. Searching for the Damage Threshold in Intact Rock. In: Proceedings, North American Rock mechanics Symposium.

Lajtai, E.Z., Carter, B.J., and Scott Duncan, E.J. 1991. Mapping the state of fracture around cavities. Eng. Geol., 31, p. 277-289.

Lajtai, E.Z., Schmidtke, R.H., and Bielus, L.P. 1987. The effect of water on the time-dependent deformation and fracture of a granite. International Journal of Rock Mechanics and Mining Sciences & Geomechanics Abstracts, vol. 24, no. 4, p. 247-255.

Lang, P.A., Babulic, P.J., Bilinsky D.M., Everitt, R.A., Spinney, M.H., Kozak, E.T., and Davison, C.C. 1990. Underground Research laboratory Room 209 Instrument Array: Measured Response to Excavation, Volumes 1 and 2, Atomic Energy of Canada Limited Report AECL 9566-3.

Lang, P.A., Everitt, R.A., Kozak, E.T., and Davison, C.C. 1988. Underground Research Laboratory Room 209 Instrument Array: Pre Excavation Information for Modellers. 1988 December, AECL 9566-1.

Lau, J.S.O. and Gorski, B. 1991a. The post- failure behavior of the Lac du Bonnet Grey Granite . Divisional Report MRL 91-079(TR). Canada Mine Technology

Laboratory, Mining Research Laboratories, CANMET, Energy, Mines and Resources Canada, Ottawa, Ontario.

Lau, J.S.O. and Gorski, B. 1991b. The post- failure behavior of the Lac du Bonnet Pink Granite . Divisional Report MRL 91-103(TR). Canada Mine Technology Laboratory, Mining Research Laboratories, CANMET, Energy, Mines and Resources Canada, Ottawa, Ontario.

Lau, J.S.O. and Gorski, B. 1991c. The post- failure behavior of the Lac du Bonnet Granodiorite . Divisional Report MRL 91-098(TR). Canada Mine Technology Laboratory, Mining Research Laboratories, CANMET, Energy, Mines and Resources Canada, Ottawa, Ontario.

Leeman, E.R. 1967. The borehole deformation type of rock stress measuring instrument. The International Journal of Rock Mechanics and Mining Science. Vol. 4, (1), p. 23-34.

Leeman, E.R. 1971. The CSIR "Doorstopper" and triaxial rock measuring instruments: Rock Mechanics, Vol. 3, p. 25-50.

MacLeod, I.N. 1980. A gravity survey of Lac de Bonnet Batholith research area RA-3, Lac du Bonnet, Manitoba. Applied Geoscience Branch, Whiteshell Laboratories, Pinawa, Manitoba, unpublished report.

Marmont, C. 1993. Exploration guidelines and opportunities for dimensional stone in central Ontario. Ontario Geological Survey, Open File Report 5853, 83 p.

Martin C.D. and Kozak, E.T. 1992. Flow measurements in the excavation-disturbed zone of Room 209,. Eurock'92. Thomas Telford. London. Paper 69. p. 402-407.

Martin, C.D. 1989a. Characterizing in-situ stress domains at the Underground Research Laboratory. In Proceedings of the 42nd Canadian Geotechnical Conference, Winnipeg, MB, 1989, 61-74.

Martin, C.D. 1989b. Failure observations and in situ stress domains at the Underground, Research Laboratory. In: International Symposium of Rock Mechanics, Pau, France, Aug. 28-31, 1989, A. A. Balkema, Rotterdam, Netherlands, p. 719-726.

Martin, C.D. 1990. Characterizing in-situ stress domains at the AECL Underground Research Laboratory. Canadian Geotechnical Journal 27: p. 631-646.

Martin, C.D. 1993. The strength of massive Lac du Bonnet Granite around underground openings. University of Manitoba, Winnipeg, MB, Canada. Doctoral Thesis. 304 p.

Martin, C.D., Chandler N.A., and Read, R.S. 1996. The role of convergence measurements in characterizing a rock mass. Canadian Geotechnical Journal, vol. 33, 363-370.

Martin, C.D., Read, R.S., and Lang, P.A. 1990. Seven years of in situ stress measurements at the URL; an overview. In: Proceedings of 31st U.S. Symposium on Rock Mechanics; Contributions and Challenges, Golden, CO, United States, June 18-20, 1990, vol. 31, p. 15-26.

Martino, J.B. 1989. Distometer Convergence Readings in the Shaft Extension of Atomic Energy of Canada Limited's Underground Research Laboratory. Atomic Energy of Canada Limited Technical Report TR-482.

Martino, J.B. 1994. A comparison of fracture growth in intact and Micro-fractured Lac du Bonnet granite. A Masters Thesis submitted to the Department of Civil and Geological Engineering, University of Manitoba.

Martinsen, M. and Wingen, P.T. 1991. Granite dimension stone quarry pre-feasibility study. Prepared for the Sudbury East Municipal Association (SEMA).

Mayo, E.B. 1941. Deformation in the interval Mt. Lyell – Mt. Whitney, California: Geological Survey of America Bulletin, Volume 52, p. 1001-1084.

McCrank, G.F.D. 1985. A Geological Survey of the Lac du Bonnet Batholith, Manitoba. Atomic Energy of Canada Limited Technical Report AECL-7816.

McCrank, G.F.D., Misiura, J.D., and Brown, P.A. 1981. Plutonic Rocks in Ontario. Atomic Energy of Canada Limited, TR-114. Also In: Geological Survey of Canada, G.S.C. Paper 80-23.

McLamore, R.T. 1966. Strength-deformation characteristics of anisotropic sedimentary rocks, Ph.D. thesis, University of Texas, Austin.

McRitchie, W.D. 1971. The petrology and environment of the acidic plutonic rocks of the Wanipigow - Winnipeg Rivers region, southeastern Manitoba. In Geology and geophysics of the Rice Lake region, southeastern Manitoba (Project Pioneer). Edited by W. D. McRitchie and W. Weber. Manitoba Mines Branch, Publication 71-1, p. 7-61.

Mesri, G. and Gibala, R. 1972. Engineering properties of a Pennsylvanian shale, Proceedings, 13th Symposium on Rock Mechanics (ASCE), p. 57-75.

Mogi, K. 1962. The influence of the dimensions of specimens on the fracture strength of rocks. Bulletin Earthquake Research Institute, Tokyo University, Vol. 40, 175-185.

Nur, A. and Simmons, G. 1970. The origin of small cracks in igneous rocks, Int. J. Rock Mech. Min, Sci., vol. 7, p. 307-314.

Obert, L., Windes, S.L., and Duvall, W.I. 1946. Standardized tests for determining the physical properties of mine rock. U.S. Bureau of Mines Report of Investigations, 3891.

Olsson, O. and Conterra AB 1991. Blasting Damage Investigations in Access Ramp Section 0/256 – 0/565 M. No. 2. Geophysical Investigations in Boreholes. Äspö Hard Rock Laboratory Progress Report 25-91-13.

Ontario Ministry of Northern Development and Mines. 1989. Building stone inventory of the Sudbury resident geologist's area. Ontario Geological Survey, Open File Report 5721.

Ontario Ministry of Northern Development and Mines. 1990. Ontario catalogue, dimensional stone.

Ontario Ministry of Northern Development and Mines. 1992. Ontario dimension stone producers directory.

Osmani, I.A. 1991. Proterozoic mafic dike swarms in the Superior Province of Ontario. In: Geology of Ontario, (eds. P.C. Thurston, H.R. Williams, R.H. Sutcliffe and G.M. Stott). Ontario Geological Survey, Special Volume 4, Part 1, 661-682.

Ouchterlony, F., Matsuki, K., Takahashi, H., and Hashida, T. 1991. Experiences from Fracture Toughness Testing of Rock According to the ISRM Suggested Methods. Sponsor: Tohoku Univ., Sendai (Japan). Report No.: DS-1991: 2 Prepared in cooperation with Tohoku Univ., Sendai (Japan).

Papertzian, C. and Farrow, D. 1995. Dimension stone: A guide to prospecting and developing, Ontario Geological Survey, Open File Report 5920.

Park, H.D. 1996. The anisotropy of the granite revealed by tensile tests. 30th international geological congress, Beijing, China, Aug. 4-14, 1996. International Geological Congress, Abstracts vol. 30, Vol. 3 p. 398.

Peng, S. and Johnson, A.M. 1972. Crack growth and faulting in cylindrical specimens of Chelmsford granite. International Journal of Rock Mechanics, vol. 9, p. 37-86.

Plumb, R., Engelder, T., and Yale, D. 1984. Near-surface in situ stress, 3, Correlation with microcrack fabric within the New Hampshire granites, Journal of Geophysical Research. Vol. 89, 9350-9364.

Pratt, H.R., Black, A.D. Brown, W.S., and Brace, W.R. 1972. The effect of specimen size ion the mechanical properties of unjointed diorite. Intnl. J. Rock Mech. Min. Sci., Vol. 9, 513-529.

Price, N.J. and Cosgrove, J.W. 1994. Analysis of geological structures. Cambridge University Press. 502 p.

Protodiakonov, M.M. and Koifman, M.I. 1963. The scale effect in investigations of rock and coal. Proc. 5th Congress International Bureau Rock Mechanics, Leipzig, 1963.

Rawlence, L. 1990. Borehole breakouts: a review of the theory and a discussion of their investigation using seismic methods, #RP005AECL.

Read, R.S. 1994. Interpreting excavation-induced displacements around a tunnel in highly stressed granite. Ph.D. thesis, Department of Civil and Geological Engineering, University of Manitoba, Winnipeg. Manitoba.

Read, R.S. and Martin, C.D. 1990. The Underground Research Laboratory mine-By Experiment - a research perspective on tunnel design. 8th Canadian Tunneling Conference, Oct. 31 - Nov. 02, 1990. Vancouver, B.C.

Read, R.S. and Martin, C.D. 1991. Mine-by Experiment final design report. Atomic Energy of Canada Limited Research Report, AECL-10430.

Read, R.S. and Martin, C.D. 1992. Monitoring the excavation-induced response of granite. In Proceedings of the 33rd U.S. Symposium on Rock Mechanics, Santa Fe, NM. Edited by J.R. Tillerson and W.R. Wawersik, p. 201-210.

Read, R.S. and Martin, C.D. 1996. Technical summary of AECL's Mine-by Experiment, Phase 1: Excavation response. Atomic Energy of Canada Limited Report, AECL-11311.

Read, R.S. and Martino, J.B. 1996. In situ thermal testing at AECL's Underground Research Laboratory. 1487-1494. In: Proceedings of the 2nd North American Rock Mechanics Symposium: Narms' 96, Montreal, Quebec, Canada, June 19-21, (M. Aubertin, F. Hassani, and H. Mitri, Eds.). A. A. Balkema: Rotterdam.

Read, R.S., Martin, C.D., and Dzik, E.J. 1995. Asymmetric borehole breakouts at the URL. In Proceedings of the 35th U.S. Symposium on Rock Mechanics, Lake Tahoe, NV. Edited by J.J.K. Daemen and R.A. Schultz, p. 879-884.

Rehrig, W.A. and Heidrick, T.R. 1972. Regional fracturing in Laramide Stocks of Arizona, Economic Geology. v. 67, #2, 198-213.

Reik, G. A. and Currie, J.B. 1974. A study of relations between rock fabric and joints in sandstone. Canadian Journal of Earth Sciences, Vol. 11, 1253-1268.

Reitan, P.H. and Murphy, J.W. 1983. Influence of metamorphic fabric and residual stresses on joint orientations in basement rock, Adirondack Mountains, New York. In: International Basement tectonics Association Publication No. 4, p.21-34. Oslo, Norway, August 10-14th (1981).

Ren, X., Kowallis, B.J., and Best, M.G. 1989. Paleostress history of the basin and range province in western Utah and eastern Nevada from healed microfracture orientations in granites. Geology, v. 17, p. 487-490.

Richter, D. and Simmons, G. 1977. Microcracks in crustal igneous rocks: Microscopy, In: Heacock, J.G., ed., The Earth's crust: American Geophysical Union monograph, 20: 149-179.

Riller, U., Cruden A.R., and Schwerdtner, W.M. 1996. Magnetic Fabric And Microstructural Evidence For A Tectonothermal Overprint Of The Early Proterozoic Murray Pluton, Central Ontario, Journal Of Structural Geology, V18, N8 (AUG), P1005-1016.

Roman-Berdiel, T., Pueyo-Morer, E-L, and Casas-Sainz, A.M. 1995. Granite emplacement during contemporary shortening and normal faulting: structural and magnetic study of the Veiga Massif (NW Spain), Journal of Structural Geology Vol. 17 (12), 1689-1706.

Rousell, D.H. 1981. A correction factor for contouring planar microfabric data. Tectonophysics, 77, T19-T21.

Rousell, D.H. and Everitt, R.A. 1981. Jointing in the Sudbury Basin, Ontario. In: Proceedings of the Third International Conference on Basement Tectonics, D.W. O'Leary and W.L. Earle, eds. Basement Tectonics Committee, Inc., Denver, Colorado. 381-391.

Rousell, D.H. and Everitt, R.A., 1977. Fracture Analysis of the Sudbury Basin. Abstr. of Paper, Geol. Assoc. Can-Min. Assoc. Can. Annual Meeting, Vancouver, B.C. p. 18.

Sano, O., Kudo, Y., and Mizuta, Y. 1992. Experimental determination of elastic constants of Oshima Granite, Barre Granite, and Chelmsford Granite. Journal of Geophysical Research, B, Solid Earth and Planets vol. 97 no. 3; p. 3367-3379.

Schmidtke, B.E. 1986. Industrial minerals of Southeast Manitoba. Manitoba Dep. Energy and Mines, Winnipeg, MB, Canada, Report of field activities, 1986, (Eds.) Bannatyne, B. B. (editor), Manitoba. Dep. Energy and Mines, Winnipeg, MB, Canada, Report of Field Activities - Mineral Resources Division (Winnipeg), vol. 1986, p. 134-138.

Scholz, C.H. and Koczynski, T.A. 1979. Dilatancy anisotropy and the response of rock to large cyclic loads. Journal of Geophysical Research. Vol. 84, 5525-5534.

Schulmann, K. and Mlcoch, B. 1996. High-temperature microstructures and rheology of deformed granite, Erzgebirge, Bohemian Massif. Journal of Structural Geology, Volume 18, Issue 6, -June-1996, p. 719-733.

Simmons, G., Todd, T., and Baldridge, W.S. 1974. Toward a quantitative relationship between elastic properties and cracks in low porosity rocks, Am. J. Sci., vol. 275, p.318-345.

Simmons, G.R. 1988. Geotechnical Research at Atomic Energy of Canada Limited's Underground Research Laboratory, in: Proceedings of the International Symposium on Uranium and Electricity - The Complete Nuclear Fuel Cycle, Saskatoon, Canada, 1988 September 18-21 (1988), p. 4-7 to 4-20.

Simmons, G.R. 1990a. Introduction to the Underground Research Laboratory. In: Proceedings of the International Symposium on Unique Underground Structures, Denver, Colorado, June 12-15, 1990. p. 63-1 to 63-20.

Simmons, G.R. 1990b. Operating Phase Experiments for Atomic Energy of Canada Limited's Underground Research Laboratory. In: Proceedings of the International Symposium on Unique Underground Structures, Denver, Colorado, June 12-15, 1990. p. 67-1 to 67-19.

Simmons, G.R. and Soonawala N.M. (eds.) 1982. Underground Research Laboratory Experimental Program, Atomic Energy of Canada Limited Technical Record, TR-153.

Sprunt, E.S. and Brace, W.F. 1974. Direct observation of microcavities in crystalline rocks, Int. J. Rock, Mech. Min, Sci. Geomech. Abstr., vol. 11, p. 139-150.

Stecich, J.P., Chin I.R., and Heidbrink F.D. 1992. Testing for thin stone veneers on buildings. Dimensional Stone Magazine, November 1992. p. 28-51.

Stone, D., Kamineni, D. C., and Brown, A. 1984. Geology and fracture characteristics of the Underground Research Laboratory lease near Lac du Bonnet, Manitoba. Atomic Energy of Canada Limited Technical Record TR-243, URL-EXP-00001-R01.

Stone, D., Kamineni, D. C., Brown, A. and Everitt, R.A. 1989. A Comparison of Fracture Styles in Two Granite Bodies of the Superior Province. Canadian Journal of Earth Sciences, Volume 26, p. 387-403.

Svab, M. and Lajtai E. Z. 1981. Microstructural control of crack growth in Lac du Bonnet granite. Proceedings of the 5th Canadian Fracture Mechanics Conference, Winnipeg, Canada. p. 219-228.

Talebi, S. and Young, R.P. 1990. Design of a Microseismic System for the URL Mineby Experiment. Provisional Report #RP006AECL, Department of Geological Sciences, Queen's University.

Talebi, S. and Young, R.P. 1992. Microseismic monitoring in highly stressed granite: Relations between shaft-wall cracking and in situ stress. International Journal of Rock Mechanics, Mineral Science & Geomechanics Abstracts 29, 25-34.

Tammemagi, H. Y. 1982. The Lac du Bonnet Batholith (RA 3); geology and rock properties. In: Geotechnical Research; Proceedings of the Seventh Nuclear Fuel Waste Management Information Meeting, Ottawa, ON, Canada, May 5-6, 1980, p. 261. Atomic Energy of Canada Limited Technical Report TR-190.

Tammemagi, H.Y. 1979. Developing the Data for Nuclear Waste Disposal. Geoscience Canada, 1979, vol. 6, no.4., p.205-208.

Tammemagi, H.Y., Kerford, P.S., Requeima, J.C., and Temple, C.A. 1980. A geological reconnaissance study of the Lac du Bonnet Batholith. Atomic Energy of Canada Limited Report AECL-6439.

Taylor, H.A. 1992. Dimension stone; United States Department of the Interior, Bureau of Mines Annual Report, p.1-17.

Thurston, P.C., Williams, H.R., Sutcliffe, R.H., and Stott, G.M., (eds.). 1991. Geology of Ontario. Ontario Geological Survey, Special Volume 4, Part 1.

Thurston, P.C., Williams, H.R., Sutcliffe, R.H., and Stott, G.M., (eds.). 1992a. Geology of Ontario. Ontario Geological Survey, Special Volume 4, Part 2.

Thurston, P.C., Williams, H.R., Sutcliffe, R.H., and Stott, G.M., (eds.). 1992b. Geology of Ontario. Ontario Geological Survey, Special Volume 4, Part 3, Maps and Charts.

Todd, T., Simmons, G., and Baldridge W.S. 1973. Acoustic double refraction in low-porosity rocks. Bull. Seismol. Soc. Am. Vol. 63, p. 20007-2020.

Tomsons, D.K., Lodha, G.S., Street, P.J., and Auger, J.L.F. 1995. A Bouguer gravity anomaly map and geophysical interpretation of the geometry of the Lac du Bonnet Batholith, Whiteshell Research Area, Southeastern Manitoba. Atomic Energy of Canada Limited, Technical Record TR-633, COG-94-096.

United States National Research Council. 1993. Stability, Failure, and Measurements of Boreholes and other Circular Openings. 1993. U.S. National Committee for Rock Mechanics, Geotechnical Board, Commission on Engineering and technical Systems, National Research Council. National Academy Press, Washington, D.C.

Verschuren, C.P., Kingston, P.W., and Calley, W.F. 1989. Criteria for quarry development in southeastern Ontario. Canadian Institute of Mining and Metallurgy Bulletin, February, 55-60 p.

Verschuren, C.P., Papertzian, V.C., Kingston, P.W., and Villard, D.J. 1986. Reconnaissance survey of building stones of eastern and central Ontario. Ontario Geological Survey, Open File Report 5585.

Verschuren, C.P., van Haaften, S., and Kingston, P.W. 1985. Building stones of eastern Ontario. Ontario Geological Survey, Open File Report 5556, 116 p.

Vos, M.A., Smith, B.A., and Stevenato,R.J. 1981. Industrial minerals of the Sudbury area. Ontario Geological Survey, Open File Report 5329.

Wahlgren, C.-H., Cruden, A.R., and Stephens, M.B. 1994. Kineamatics of a major fan-like structure in the eastern part of the Sveconorwegian orogen, Baltic Shield, south-central Sweden. Precambrian Research, 70, p. 67-91.

West, G. 1991. A note on undulatory extinction of quartz in granite. Quarterly Journal of Engineering Geology, 24: p.159-162.

Whitaker, S.H. 1987. Geoscience research for the Canadian Nuclear Fuel Waste Management Program. Radioactive Waste Management and the Nuclear Fuel Cycle, 8: p. 145-196.

Whitaker, S.H., Davison, C.C., Dormuth, K.W., and Scott, J.S. 1986. The Canadian Approach to Site Characterization for a Nuclear Fuel Waste Disposal Vault, Proceedings of A Symposium on Siting, Design and Construction of Underground Repositories for Radioactive Wastes organized by International Atomic Energy Agency, Hannover, 1986 March 3-7th.

Wijesinghe, A.M. 1987. Extended analysis of constant –height hydraulic fractures for the estimation of in-situ crack-opening modulus from bottom hole pressure records, UCID-200995, Lawerence Livermore National Laboratory.

Young, R.P. and Talebi, S. 1989. Manual of source locations of microseismic events induced by shaft extension at the Underground Research Laboratory #RP002AECL Queen's University, Kingston, ON.

Zarifa, I.H. 1998. Correlation of mineralogical and textural characteristics with engineering properties of selected granitic rocks from Turkey. Engineering Geology, Volume 51, Issue 4, 303-317.

# Appendix A

# POINT LOAD TESTS

This appendix presents the samples used for and the data obtained from the Point Load Tests. The tables list the specimen characteristics, and the results of the axial Point Load tests described in this report. The figures summarize the fabric and test-induced fractures observed in each of the samples used in the Point Load testing. The symbols used in the sample maps are explained in Figure A1, while the standard format for the sample maps is shown in Figure A2.

## Table A1

## Axial Point Load Testing Results

## Fine Grained Granite, Borehole 421-031-RT1

| Sample | | | Point Load Index | | Tensile Strength | | Dihedral angle | True orientations | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | | | | | | | | Dip | Dip Direction | |
| fracture / *foliation* | Thickness (mm) | Diameter (mm) | $I_s$ (Mpa) | Normalized $I_s$ (MPa) | (PSI) | (MPa) | loading direction & foliation | | Fracture | Foliation |
| A | 19.92 | 44.88 | 20.18 | 21.28 | 1403.42 | 9.68 | 68 | | | |
| $A_G$ | | | | | | | | 82 | | 200 |
| B | 20.31 | 44.88 | 21.57 | 22.31 | 1559.35 | 10.75 | 60 | | | |
| B | | | | | | | 80 | | | |
| $B_G$ | | | | | | | | 65 | | 195 |
| C | 20.19 | 44.88 | 22.92 | 23.84 | 1637.32 | 11.29 | 59 | | | |
| C | | | | | | | 25 | | | |
| $C_G$ | | | | | | | | 82 | | 200 |
| D | 20.15 | 44.88 | 21.92 | 22.84 | 1559.35 | 10.75 | 85 | | | |
| $D_G$ | | | | | | | | 82 | | 200 |
| E | 20.67 | 44.88 | 20.83 | 21.16 | 1559.35 | 10.75 | 35 | | | |
| E | | | | | | | 85 | | | |
| E | | | | | | | 62 | | | |
| $E_G$ | | | | | | | | 90 | | 200 |
| F | 20.19 | 44.88 | 24.56 | 25.54 | 1754.27 | 12.10 | 35 | | | |
| $F_G$ | | | | | | | | 86 | | 10 |
| G | 20.89 | 44.88 | 19.88 | 19.99 | 1520.37 | 10.49 | 90 | | | |
| G | | | | | | | 62 | | | |
| $G_G$ | | | | | | | | 82 | | 200 |
| H | 20.52 | 44.88 | 20.61 | 21.09 | 1520.37 | 10.49 | 35 | | | |
| $H_G$ | | | | | | | | 90 | | 100 |
| I | 20.62 | 44.88 | 20.93 | 21.31 | 1559.35 | 10.75 | 58 | | | |
| $I_G$ | | | | | | | | 80 | | 188 |
| J | 18.94 | 44.88 | 24.81 | 27.50 | 1559.35 | 10.75 | 72 | | | |
| J | | | | | | | 58 | | | |
| $J_G$ | | | | | | | | 70 | | 171 |
| $J_G$ | | | | | | | | 70 | | 180 |
| Count | 10.00 | 10.00 | 10.00 | 10.00 | 10.00 | 10.00 | 16.00 | 11.00 | | 11.00 |
| Maximum | 20.89 | 44.88 | 24.81 | 27.50 | 1754.27 | 12.10 | 90.00 | 90.00 | | 200.00 |
| | | | | | | | | | | |
| Median | 20.25 | 44.88 | 21.25 | 21.81 | 1559.35 | 10.75 | 61.00 | 82.00 | | 195.00 |
| Minimum | 18.94 | 44.88 | 19.88 | 19.99 | 1403.42 | 9.68 | 25.00 | 65.00 | | 10.00 |
| Std. dev. | 0.54 | 0.00 | 1.74 | 2.33 | 88.99 | 0.61 | 19.78 | 8.23 | | 59.99 |

## Table A2
## Axial Point Load Testing Results
## Fine Grained Granite, Borehole 421-031-RT2

| Sample | | | Point Load Index | | Tensile Strength | | Dihedral angle | True orientations | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | | | | | | | | Dip | Dip Direction | |
| fracture / foliation | Thickness (mm) | Diameter (mm) | $I_s$ (Mpa) | Normalized $I_s$ (MPa) | (PSI) | (MPa) | loading direction & foliation | | Fracture | Foliation |
| A | 21.01 | 44.75 | 20.16 | 20.15 | 1568.42 | 10.82 | | | | |
| $A_{sm}$ | | | | | | | 40 | 86 | | 59 |
| $A_{sm}$ | | | | | | | 40 | 20 | | 332 |
| $A_{sm}$ | | | | | | | 20 | 45 | | 52 |
| B | 20.70 | 44.75 | 20.77 | 21.07 | 1568.42 | 10.82 | | | | |
| $B_{sm}$ | | | | | | | 30 | 80 | | 53 |
| $B_{sm}$ | | | | | | | 64 | 83 | | 209 |
| $B_{sm}$ | | | | | | | 60 | 20 | | 332 |
| $B_{sm}$ | | | | | | | 0 | 50 | | 53 |
| B | | | | | | | | | | |
| $B_{sm}$ | | | | | | | 55 | 80 | | 53 |
| $B_{sm}$ | | | | | | | 62 | 83 | | 209 |
| $B_{sm}$ | | | | | | | 60 | 20 | | 332 |
| $B_{sm}$ | | | | | | | 84 | 50 | | 53 |
| B | | | | | | | | | | |
| $B_{sm}$ | | | | | | | 80 | 80 | | 53 |
| $B_{sm}$ | | | | | | | 80 | 83 | | 209 |
| $B_{sm}$ | | | | | | | 20 | 20 | | 332 |
| $B_{sm}$ | | | | | | | 50 | 50 | | 53 |
| C | 20.46 | 44.75 | 19.66 | 20.18 | 1450.79 | 10.01 | | | | |
| $C_{sm}$ | | | | | | | 75 | 80 | | 80 |
| $C_{sm}$ | | | | | | | 12 | 15 | | 23 |
| $C_{sm}$ | | | | | | | 25 | 33 | | 45 |
| D | 19.87 | 44.75 | 21.41 | 22.63 | 1490.00 | 10.28 | | | | |
| $D_{sm}$ | | | | | | | 15 | 90 | | 59 |
| $D_{sm}$ | | | | | | | 58 | 80 | | 215 |
| $D_{sm}$ | | | | | | | 78 | 20 | | 45 |
| $D_{sm}$ | | | | | | | 45 | 40 | | 45 |
| D | | | | | | | | | | |
| $D_{sm}$ | | | | | | | 50 | 90 | | 59 |
| $D_{sm}$ | | | | | | | 55 | 80 | | 215 |
| $D_{sm}$ | | | | | | | 46 | 20 | | 45 |
| $D_{sm}$ | | | | | | | 75 | 40 | | 45 |
| D | | | | | | | | | | |
| $D_{sm}$ | | | | | | | 55 | 90 | | 59 |
| $D_{sm}$ | | | | | | | 82 | 80 | | 215 |
| $D_{sm}$ | | | | | | | 38 | 20 | | 330 |
| $D_{sm}$ | | | | | | | 10 | 40 | | 45 |
| E | 21.46 | 44.75 | 17.39 | 17.02 | 1411.58 | 9.74 | | | | |
| $E_{sm}$ | | | | | | | 60 | 80 | | 51 |
| $E_{sm}$ | | | | | | | 68 | 82 | | 23 |
| $E_{sm}$ | | | | | | | 5 | 21 | | 51 |
| F | 20.86 | 44.75 | 18.92 | 19.04 | 1450.79 | 10.01 | | | | |
| $F_{sm}$ | | | | | | | 62 | 80 | | 62 |
| $F_{sm}$ | | | | | | | 80 | 82 | | 10 |
| F | 20.86 | 44.75 | 18.92 | 19.04 | 1450.79 | 10.01 | | | | |
| $F_{sm}$ | | | | | | | 62 | 80 | | 62 |
| $F_{sm}$ | | | | | | | 80 | 82 | | 10 |
| $F_{sm}$ | | | | | | | 18 | 32 | | 43 |
| G | 20.01 | 44.75 | 22.22 | 23.32 | 1568.42 | 10.82 | | | | |
| $G_{sm}$ | | | | | | | 80 | 86 | | 57 |
| $G_{sm}$ | | | | | | | 82 | 80 | | 215 |
| $G_{sm}$ | | | | | | | 30 | 30 | | 56 |
| H | 20.07 | 44.75 | 20.44 | 21.38 | 1450.79 | 10.01 | | | | |
| $H_{sm}$ | | | | | | | 65 | 75 | | 60 |
| $H_{sm}$ | | | | | | | 53 | 61 | | 20 |
| $H_{sm}$ | | | | | | | 27 | 38 | | 50 |
| I | 20.96 | 44.75 | 16.20 | 16.24 | 1254.74 | 8.65 | | | | |
| $I_{sm}$ | | | | | | | 76 | 78 | | 57 |
| $I_{sm}$ | | | | | | | 82 | 80 | | 27 |
| $I_{sm}$ | | | | | | | 20 | 60 | | 190 |
| $I_{sm}$ | | | | | | | 45 | 44 | | 60 |
| J | 20.25 | 44.75 | 21.70 | 22.50 | 1568.42 | 10.82 | | | | |
| $J_{sm}$ | | | | | | | 88 | 90 | | 46 |
| $J_{sm}$ | | | | | | | 78 | 74 | | 25 |
| $J_{sm}$ | | | | | | | 35 | 40 | | 163 |
| $J_{sm}$ | | | | | | | 50 | 50 | | 54 |
| Count | 10.00 | 10.00 | 10.00 | 10.00 | 10.00 | 10.00 | 50 | 50.00 | | 50 |
| Maximum | 21.46 | 44.75 | 22.22 | 23.32 | 1568.42 | 10.82 | 88 | 90.00 | | 332 |
| [illegible] | | | | | | | | | | |
| Median | 20.58 | 44.75 | 20.30 | 20.63 | 1470.40 | 10.14 | 55 | 67.50 | | 55 |
| Minimum | 19.87 | 44.75 | 16.20 | 16.24 | 1254.74 | 8.65 | 0 | 15.00 | | 10 |
| Std. dev. | 0.52 | 0.00 | 1.92 | 2.36 | 99.63 | 0.69 | 24 | 25.83 | | 98 |

## Table A3

## Axial Point Load Testing Results

## Fine Grained Granite, Borehole 421-031-RT3

| Sample | | | Point Load Index | | Tensile Strength | | Dihedral angle | True orientations | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | | | | | | | | Dip | Dip Direction | |
| fracture / *foliation* | Thickness (mm) | Diameter (mm) | $I_s$ (Mpa) | Normalized $I_s$ (MPa) | (PSI) | (MPa) | loading direction & foliation | | Fracture | Foliation |
| A | 22.03 | 44.87 | 20.63 | 19.66 | 1755.05 | 12.10 | 0 | | | |
| A | | | | | | | 0 | | | |
| A | | | | | | | 0 | | | |
| B | 18.38 | 44.87 | 27.00 | 30.85 | 1599.05 | 11.03 | 0 | | | |
| B | | | | | | | 0 | | | |
| B | | | | | | | 0 | | | |
| C | 19.67 | 44.87 | 24.15 | 25.78 | 1638.05 | 11.30 | 0 | | | |
| D | 19.87 | 44.87 | 23.10 | 24.42 | 1599.05 | 11.03 | 0 | | | |
| E | 18.30 | 44.87 | 26.57 | 30.49 | 1560.05 | 10.76 | 0 | | | |
| F | 19.91 | 44.87 | 23.57 | 24.86 | 1638.05 | 11.30 | 0 | | | |
| F | | | | | | | 0 | | | |
| F | | | | | | | 0 | | | |
| G | 23.34 | 44.87 | 19.60 | 17.64 | 1872.05 | 12.91 | 0 | | | |
| G | | | | | | | 0 | | | |
| H | 18.77 | 44.87 | 25.89 | 28.97 | 1599.05 | 11.03 | 0 | | | |
| H | | | | | | | 0 | | | |
| H | | | | | | | 0 | | | |
| I | 22.02 | 44.87 | 19.27 | 18.38 | 1638.05 | 11.30 | 0 | | | |
| J | 19.81 | 44.87 | 23.24 | 24.64 | 1599.05 | 11.03 | 0 | | | |
| J | | | | | | | 0 | | | |
| J | | | | | | | 0 | | | |
| Count | 10.00 | 10.00 | 10.00 | 10.00 | 10.00 | 10.00 | 21.00 | | | |
| Maximum | 23.34 | 44.87 | 27.00 | 30.85 | 1872.05 | 12.91 | 0.00 | | | |
| [illegible] | [illegible] | [illegible] | [illegible] | [illegible] | [illegible] | [illegible] | [illegible] | [illegible] | [illegible] | [illegible] |
| Median | 19.84 | 44.87 | 23.41 | 24.75 | 1618.55 | 11.16 | 0.00 | | | |
| Minimum | 18.30 | 44.87 | 19.27 | 17.64 | 1560.05 | 10.76 | 0.00 | | | |
| Std. dev. | 1.70 | 0.00 | 2.77 | 4.78 | 93.84 | 0.65 | 0.00 | | | |

NOTE: this borehole is parallel to the most prominent foliation, and the dihedral angles between the fractures and the foliation has been approximated in each case to 0°.

## Table A4

## Axial Point Load Testing Results

## Fine Grained Granite, Borehole 421-031-RT4

| Sample | | | Point Load Index | | Tensile Strength | | Dihedral angle | True orientations | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | | | | | | | | Dip | Dip Direction | |
| fracture *foliation* | Thickness (mm) | Diameter (mm) | $I_s$ (Mpa) | Normalized $I_s$ (MPa) | (PSI) | (MPa) | loading direction & foliation | | Fracture | Foliation |
| A | 20.50 | 44.96 | 20.65 | 21.15 | 1514.96 | 10.45 | | | | |
| $A_{G1}$ | | | | | | | 55 | | 55 | 5 |
| B | 21.35 | 44.96 | 20.50 | 20.16 | 1631.50 | 11.25 | | | | |
| $B_{G1}$ | | | | | | | 55 | | 50 | 20 |
| C | 19.89 | 44.96 | 20.81 | 21.97 | 1437.27 | 9.91 | | | | |
| $C_{G1}$ | | | | | | | 84 | | 46 | 27 |
| D | 20.70 | 44.96 | 20.77 | 21.07 | 1553.81 | 10.72 | | | | |
| $D_G$ | | | | | | | 88 | | 46 | 27 |
| $D_G$ | | | | | | | 60 | | 70 | 200 |
| E | 19.00 | 44.96 | 22.19 | 24.52 | 1398.43 | 9.64 | | | | |
| E | | | | | | | | | | |
| $E_{G1}$ | | | | | | | 65 | | 85 | 43 |
| F | 21.29 | 44.96 | 18.65 | 18.40 | 1476.12 | 10.18 | | | | |
| $F_{G1}$ | | | | | | | 50 | | 40 | 33 |
| $F_G$ | | | | | | | 75 | | 90 | 334 |
| G | 21.81 | 44.96 | 17.30 | 16.66 | 1437.27 | 9.91 | | | | |
| $G_G$ | | | | | | | 32 | | 71 | 133 |
| $G_G$ | | | | | | | 65 | | 55 | 244 |
| H | 20.55 | 44.96 | 20.55 | 21.00 | 1514.96 | 10.45 | | | | |
| $H_{G1}$ | | | | | | | 65 | | 70 | 344 |
| $H_{G2}$ | | | | | | | 48 | | 65 | 57 |
| I | 19.09 | 44.96 | 22.59 | 24.85 | 1437.27 | 9.91 | | | | |
| $I_{M1}$ | | | | | | | 38 | | 80 | 74 |
| J | 19.52 | 44.96 | 21.02 | 22.61 | 1398.43 | 9.64 | | | | |
| $J_{G1}$ | | | | | | | 36 | | 80 | 300 |
| K | 18.56 | 44.96 | 23.25 | 26.31 | 1398.43 | 9.64 | | | | |
| $K_{G1}$ | | | | | | | 20 | | 70 | 290 |
| $K_{G2}$ | | | | | | | 10 | | 76 | 143 |
| L | 21.05 | 44.96 | 19.08 | 19.03 | 1476.12 | 10.18 | | | | |
| $L_{G1}$ | | | | | | | 55 | | 56 | 160 |
| M | 22.34 | 44.96 | 19.17 | 18.02 | 1670.34 | 11.52 | 20 | | | |
| M | | | | | | | 10 | | | |
| $M_G$ | | | | | | | | | 90 | 335 |
| Count | 13.00 | 13.00 | 13.00 | 13.00 | 13.00 | 13.00 | 19.00 | | 18.00 | 18.00 |
| Maximum | 22.34 | 44.96 | 23.25 | 26.31 | 1670.34 | 11.52 | 88.00 | | 90.00 | 344.00 |
| Mean | 20.51 | 44.96 | 20.10 | 21.07 | 1494.87 | 10.28 | 53.11 | | 67.00 | 167.00 |
| Median | 20.55 | 44.96 | 20.65 | 21.07 | 1476.12 | 10.18 | 55.00 | | 70.00 | 138.00 |
| Minimum | 18.56 | 44.96 | 17.30 | 16.66 | 1398.43 | 9.64 | 10.00 | | 40.00 | 5.00 |
| Std. dev. | 1.16 | 0.00 | 1.65 | 2.85 | 87.42 | 0.60 | 23.24 | | 15.62 | 125.34 |

## Table A5

## Axial Point Load Testing Results

## Medium Grained Granite, Borehole 417-071-RT4

| Sample | | | Point Load Index | | Tensile Strength | | Dihedral angle | True orientations | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | | | | | | | | Dip | Dip Direction | |
| fracture / foliation | Thickness (mm) | Diameter (mm) | $I_s$ (Mpa) | Normalized $I_s$ (MPa) | (PSI) | (MPa) | loading direction & foliation | | Fracture | Foliation |
| A | 21.49 | 44.80 | 16 | 16 | 1291 | 9 | | 32 | 135 | |
| B | 21.16 | 44.80 | 14 | 14 | 1095 | 8 | | 20 | 177 | |
| C | 19.53 | 44.80 | 21 | 23 | 1408 | 10 | | 50 | 118 | |
| D | 20.29 | 44.80 | 17 | 17 | 1213 | 8 | | 18 | 194 | |
| E | 20.47 | 44.80 | 15 | 16 | 1135 | 8 | 75 | 22 | 219 | |
| E | | | 15 | 16 | 1135 | 8 | 26 | 45 | 259 | |
| E | | | 15 | 16 | 1135 | 8 | 60 | 48 | 116 | |
| $E_{MI}$ | | | 15 | 16 | 1135 | 8 | | 20 | | 250 |
| F | 21.11 | 44.80 | 15 | 15 | 1213 | 8 | 12 | 20 | 177 | |
| $F_{MI}$ | | | 15 | 15 | 1213 | 8 | | 10 | | 195 |
| Count | 6 | 6 | 10 | 10 | 10 | 10 | 4 | 10 | 8 | 2 |
| Maximum | 21.49 | 44.80 | 21 | 23 | 1408 | 10 | 75 | 50 | 259 | 250 |
| Mean | 20.68 | 44.80 | 16 | 16 | 1197 | 8 | 43 | 29 | 174 | 223 |
| Median | 20.79 | 44.80 | 15 | 16 | 1174 | 8 | 43 | 21 | 177 | 223 |
| Minimum | 19.53 | 44.80 | 14 | 14 | 1095 | 8 | 12 | 10 | 116 | 195 |
| Std. dev. | 0.72 | 0.00 | 2 | 2 | 94 | 1 | 29 | 14 | 50 | 39 |

## Table A6
## Axial Point Load Testing Results
## Medium Grained Granite, Borehole 417-071-RT5

| Sample | | | Point Load Index | | Tensile Strength | | Dihedral angle | True orientations | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | | | | | | | | Dip | Dip Direction | |
| fracture *foliation* | Thickness (mm) | Diameter (mm) | $I_s$ (Mpa) | Normalized $I_s$ (MPa) | (PSI) | (MPa) | loading direction & foliation | | Fracture | Foliation |
| A | 22.39 | 44.68 | 13.76 | 12.90 | 1219.34 | 8.41 | | 30 | 100 | |
| $A_{G1}$ | | | | | | | 70 | 25 | | 15 |
| $A_{G2}$ | | | | | | | 86 | 15 | | 21 |
| $A_{G3}$ | | | | | | | 82 | 55 | | 251 |
| B | 21.42 | 44.68 | 15.52 | 15.21 | 1258.67 | 8.68 | | 0 | 100 | |
| $B_{G1}$ | | | | | | | 25 | 20 | | 100 |
| $B_{G2}$ | | | | | | | 70 | 46 | | 293 |
| B | | | | | | | | 48 | 280 | |
| $B_{G1}$ | | | | | | | 60 | 20 | | 100 |
| $B_{G2}$ | | | | | | | 15 | 46 | | 293 |
| C | 19.59 | 44.68 | 17.39 | 18.64 | 1180.01 | 8.14 | | 5 | 100 | |
| $C_G$ | | | | | | | 30 | 60 | | 5 |
| $C_G$ | | | | | | | 86 | 90 | | 330 |
| D | 22.45 | 44.68 | 15.01 | 14.04 | 1337.34 | 9.22 | | 20 | 100 | |
| $D_{G1}$ | | | | | | | 65 | 80 | | 145 |
| $D_{G2}$ | | | | | | | 75 | 75 | | 184 |
| E | 20.89 | 44.68 | 15.29 | 15.37 | 1180.01 | 8.14 | | 30 | 100 | |
| $E_{G1}$ | | | | | | | 75 | 75 | | 184 |
| $E_{G2}$ | | | | | | | 75 | 75 | | 20 |
| $E_{G3}$ | | | | | | | 88 | 75 | | 213 |
| F | 20.51 | 44.68 | 15.87 | 16.24 | 1180.01 | 8.14 | | 10 | 100 | |
| F | | | | | | | | 45 | 100 | |
| F | | | | | | | | 50 | 280 | |
| G | 19.52 | 44.68 | 17.52 | 18.84 | 1180.01 | 8.14 | | 25 | 100 | |
| $G_1$ | | | | | | | 90 | 77 | | 215 |
| H | 19.40 | 44.68 | 16.55 | 17.92 | 1101.34 | 7.60 | | 90 | 100 | |
| $H_G$ | | | | | | | 65 | 25 | | 100 |
| I | 20.62 | 44.68 | 14.65 | 14.92 | 1101.34 | 7.60 | | 20 | 100 | |
| $I_{G1}$ | | | | | | | 57 | 90 | | 127 |
| J | 20.81 | 44.68 | 16.44 | 16.59 | 1258.67 | 8.68 | | 25 | 100 | |
| $J_G$ | | | | | | | 0 | 25 | | 100 |
| $J_G$ | | | | | | | 75 | 80 | | 295 |
| K | 18.66 | 44.68 | 17.89 | 20.13 | 1101.34 | 7.60 | | 25 | 100 | |
| $K_{G2}$ | | | | | | | 4 | 0 | | 0 |
| $K_{G1}$ | | | | | | | 74 | 75 | | 290 |
| K | | | | | | | | 25 | 100 | |
| $K_{G2}$ | | | | | | | 35 | 0 | | 0 |
| $K_{G1}$ | | | | | | | 75 | 75 | | 290 |
| L | 21.71 | 44.68 | 15.10 | 14.61 | 1258.67 | 8.68 | | 10 | 100 | |
| $L_G$ | | | | | | | 67 | 55 | | 285 |
| $L_{G2}$ | | | | | | | 0 | 10 | | 100 |
| $L_{M1}$ | | | | | | | 65 | 75 | | 100 |
| M | 18.73 | 44.68 | 19.66 | 22.04 | 1219.34 | 8.41 | | 10 | 100 | |
| $M_{G1}$ | | | | | | | 60 | 20 | | 100 |
| $M_{G2}$ | | | | | | | 0 | 90 | | 280 |
| $M_G$ | | | | | | | 80 | 60 | | 280 |
| N | 19.87 | 44.68 | 16.34 | 17.27 | 1140.67 | 7.87 | | 10 | 100 | |
| $N_{G1}$ | | | | | | | 70 | 40 | | 280 |
| $N_{G2}$ | | | | | | | 82 | 88 | | 176 |
| $N_{G3}$ | | | | | | | 10 | 20 | | 100 |
| $N_{G1}$ | | | | | | | 60 | 90 | | 100 |
| | | | | | | | | | | |
| **Count** | 14.00 | 14.00 | 14.00 | 14.00 | 14.00 | 14.00 | 33 | 51 | 18 | 33 |
| **Maximum** | 22.45 | 44.68 | 19.66 | 22.04 | 1337.34 | 9.22 | 90 | 90 | 280 | 330 |
| **Mean** | 20.47 | 44.68 | 16.21 | 16.77 | 1194.05 | 8.23 | 57 | 44 | 120 | 163 |
| **Median** | 20.57 | 44.68 | 16.10 | 16.42 | 1180.01 | 8.14 | 67 | 40 | 100 | 145 |
| **Minimum** | 18.66 | 44.68 | 13.76 | 12.90 | 1101.34 | 7.60 | 0 | 0 | 100 | 0 |
| **Std. dev.** | 1.24 | 0.00 | 1.53 | 2.54 | 70.04 | 0.48 | 29 | 30 | 58 | 106 |

## Table A7
## Axial Point Load Testing Results
## Coarse Grained Granite, Borehole 418-051-RT2

| Sample | | | Point Load Index | | Tensile Strength | | Dihedral angle | True orientations | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | | | | | | | | Dip | Dip Direction | |
| fracture *foliation* | Thickness (mm) | Diameter (mm) | $I_s$ (Mpa) | Normalized $I_s$ (MPa) | (PSI) | (MPa) | loading direction & foliation | | Fracture | Foliation |
| A | 19.6 | 45.1 | 16.2 | 17.3 | 1083.3 | 7.5 | | 56 | 147 | |
| $A_{G1}$ | | | | | | | 60 | 0 | | 0 |
| B | 20.6 | 45.1 | 16.7 | 17.0 | 1238.1 | 8.5 | | 57 | 124 | |
| B | | | | | | | | 57 | 159 | |
| $B_{G1}$ | | | | | | | 65 | 5 | | 228 |
| C | 18.6 | 45.1 | 16.0 | 18.1 | 967.3 | 6.7 | | 56 | 145 | |
| $C_{G1}$ | | | | | | | 90 | 0 | | 0 |
| D | 21.7 | 45.1 | 13.8 | 13.3 | 1122.0 | 7.7 | | 62 | 180 | |
| $D_{G1}$ | | | | | | | 48 | 22 | | 215 |
| $D_{G1}$ | | | | | | | 40 | 18 | | 228 |
| E | 22.1 | 45.1 | 13.6 | 12.9 | 1160.7 | 8.0 | | 56 | 140 | |
| $E_{G1}$ | | | | | | | 45 | | | |
| F | 22.8 | 45.1 | 15.4 | 14.2 | 1392.8 | 9.6 | | 68 | 196 | |
| $F_{G1}$ | | | | | | | 52 | 15 | | 208 |
| $F_{G2}$ | | | | | | | 35 | 80 | | 26 |
| G | 20.7 | 45.1 | 13.6 | 13.8 | 1005.9 | 6.9 | | 58 | 165 | |
| $G_{G1}$ | | | | | | | 45 | 78 | | 240 |
| $G_{G2}$ | | | | | | | 45 | 78 | | 150 |
| H | 21.0 | 45.1 | 14.6 | 14.6 | 1122.0 | 7.7 | | 57 | 154 | |
| $H_{A2}$ | | | | | | | 30 | | | |
| $H_{A2}$ | | | | | | | 45 | | | |
| $H_{G1}$ | | | | | | | 45 | | | |
| I | 21.2 | 45.1 | 13.9 | 13.8 | 1083.3 | 7.5 | | 58 | 160 | |
| $I_{G1}$ | | | | | | | 13 | 72 | | 158 |
| J | 20.9 | 45.1 | 15.3 | 15.4 | 1160.7 | 8.0 | | 52 | 133 | |
| $J_{G1}$ | | | | | | | 16 | 56 | | 165 |
| J* | 20.6 | 45.1 | 16.8 | 17.1 | 1238.1 | 8.5 | | 58 | 160 | |
| $J_{G1}$ | | | | | | | 50 | | | |
| K | 21.9 | 45.1 | 14.9 | 14.3 | 1238.1 | 8.5 | | 61 | 180 | |
| $K_{G1}$ | | | | | | | 45 | 20 | | 230 |
| L | 19.6 | 45.1 | 15.6 | 16.7 | 1044.6 | 7.2 | | 54 | 178 | |
| L | | | | | | | 0 | 54 | | 178 |
| L | | | | | | | 46 | 75 | | 342 |
| M | 20.7 | 45.1 | 15.5 | 15.7 | 1160.7 | 8.0 | | 68 | 197 | |
| $M_{G1}$ | | | | | | | 18 | 88 | | 190 |
| N | 21.6 | 45.1 | 13.4 | 13.0 | 1083.3 | 7.5 | | 56 | 153 | |
| N | | | | | | | | 90 | 53 | |
| N | | | | | | | | 68 | 197 | |
| $N_{G1}$ | | | | | | | 10 | 85 | | 188 |
| O | 19.8 | 45.1 | 16.0 | 17.0 | 1083.3 | 7.5 | | 90 | 54 | |
| $O_{G1}$ | | | | | | | 20 | 90 | | 92 |
| $O_{G2}$ | | | | | | | 35 | 90 | | 92 |
| O | 19.8 | 45.1 | 16.0 | 17.0 | 1083.3 | 7.5 | | 90 | 54 | |
| $O_{G1}$ | | | | | | | 72 | 90 | | 92 |
| $O_{G2}$ | | | | | | | 60 | 90 | | 92 |
| O | 19.8 | 45.1 | 16.0 | 17.0 | 1083.3 | 7.5 | | 90 | 54 | |
| $O_{G1}$ | | | | | | | 55 | 90 | | 92 |
| $O_{G2}$ | | | | | | | 75 | 90 | | 92 |
| O | 19.8 | 45.1 | 16.0 | 17.0 | 1083.3 | 7.5 | | 90 | 54 | |
| $O_{G1}$ | | | | | | | 72 | 90 | | 92 |
| $O_{G2}$ | | | | | | | 60 | 90 | | 92 |
| O | 19.8 | 45.1 | 16.0 | 17.0 | 1083.3 | 7.5 | | 90 | 54 | |
| $O_{G1}$ | | | | | | | 55 | 90 | | 92 |
| $O_{G2}$ | | | | | | | 75 | 90 | | 92 |
| Count | 10 | 10 | 10 | 10 | 10 | 10 | 27 | 43 | 21 | 22 |
| Maximum | 22.8 | 45.1 | 16.7 | 18.1 | 1392.8 | 9.6 | 90.00 | 90 | 197 | 342 |
| Median | 20.9 | 45.1 | 15.0 | 14.4 | 1122.0 | 7.7 | 45.00 | 61 | 154 | 162 |
| Minimum | 18.6 | 45.1 | 13.6 | 12.9 | 967.3 | 6.7 | 0.00 | 0 | 53 | 0 |
| Std. dev. | 1.2 | 0.0 | 1.2 | 1.8 | 119.7 | 0.8 | 21.03 | n/a | n/a | n/a |

## Table A8

## Axial Point Load Testing Results

## Coarse Grained Granite, Borehole 418-051-RT3

| Sample | | | Point Load Index | | Tensile Strength | | Dihedral angle | True orientations | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | | | | | | | | Dip | Dip Direction | |
| fracture *foliation* | Thickness (mm) | Diameter (mm) | $I_s$ (Mpa) | Normalized $I_s$ (MPa) | (PSI) | (MPa) | loading direction & foliation | | Fracture | Foliation |
| A | 20.6 | 44.6 | 14.1 | 14.3 | 1065.3 | 7.3 | | 40 | 110 | |
| $A_{G1}$ | | | | | | | 43 | 65 | | 63 |
| B | 20.5 | 44.6 | 15.9 | 16.2 | 1183.7 | 8.2 | | 48 | 93 | |
| $B_{G1}$ | | | | | | | 18 | 54 | | 75 |
| C | 20.5 | 44.6 | 13.8 | 14.1 | 1025.9 | 7.1 | | 70 | 69 | |
| $C_{G1}$ | | | | | | | 12 | 61 | | 75 |
| D | 20.0 | 44.6 | 14.5 | 15.3 | 1025.9 | 7.1 | | 70 | 69 | |
| $D_{G1}$ | | | | | | | 10 | 63 | | 64 |
| $D_{G1}$ | | | | | | | 34 | 43 | | 70 |
| D | 20.0 | 44.6 | 14.5 | 15.3 | 1025.9 | 7.1 | | 48 | 83 | |
| $D_{G1}$ | | | | | | | 32 | 56 | | 47 |
| $D_{G1}$ | | | | | | | 15 | 56 | | 47 |
| E | 20.7 | 44.6 | 14.6 | 14.8 | 1104.8 | 7.6 | | 55 | 84 | |
| $E_{G1}$ | | | | | | | 20 | | | |
| F | 20.0 | 44.6 | 13.4 | 14.0 | 947.0 | 6.5 | | 65 | 72 | |
| $F_{G1}$ | | | | | | | 5 | | | |
| Count | 7 | 7 | 7 | 7 | 7 | 7 | 9 | 9 | 14 | 7 |
| Maximum | 20.7 | 44.6 | 15.9 | 16.2 | 1183.7 | 8.2 | 43.00 | 43 | 70 | 110 |
| [redacted] | | | | | | | | | | |
| Median | 20.5 | 44.6 | 14.5 | 14.8 | 1025.9 | 7.1 | 18.00 | 18 | 56 | 83 |
| Minimum | 20.0 | 44.6 | 13.4 | 14.0 | 947.0 | 6.5 | 5.00 | 5 | 40 | 69 |
| Std. dev. | 0.3 | 0.0 | 0.8 | 0.8 | 74.6 | 0.5 | 12.64 | 13 | 10 | 15 |

## Table A9

## Axial Point Load Testing Results

## Coarse Grained Granite, Borehole 418-051-RT4

| Sample | | | Point Load Index | | Tensile Strength | | Dihedral angle | True orientations | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | | | | | | | | Dip | Dip Direction | |
| fracture *foliation* | Thickness (mm) | Diameter (mm) | $I_s$ (Mpa) | Normalized $I_s$ (MPa) | (PSI) | (MPa) | loading direction & foliation | | Fracture | Foliation |
| A | 22.3 | 44.5 | 13.4 | 12.7 | 1189.6 | 8.2 | | 35 | 85 | |
| $A_{G1}$ | | | | | | | 25 | 20 | | 160 |
| $A_{G}$ | | | | | | | 75 | 88 | | 10 |
| B | 21.0 | 44.5 | 14.2 | 14.2 | 1110.3 | 7.7 | | 20 | 125 | |
| $B_{G1}$ | | | | | | | | 20 | | 136 |
| C | 19.9 | 44.5 | 17.5 | 18.5 | 1229.2 | 8.5 | | 48 | 73 | |
| $C_{G1}$ | | | | | | | 45 | 20 | | 136 |
| $C_{G2}$ | | | | | | | 62 | 25 | | 322 |
| D | 19.8 | 44.5 | 15.9 | 16.9 | 1110.3 | 7.7 | | 32 | 124 | |
| $D_{G1}$ | | | | | | | 45 | | | |
| $D_{G1}$ | | | | | | | 45 | | | |
| E | 20.7 | 44.5 | 14.5 | 14.7 | 1110.3 | 7.7 | | 27 | 95 | |
| $E_{G1}$ | | | | | | | 15 | 22 | | 122 |
| $E_{G1}$ | | | | | | | 45 | 65 | | 142 |
| $E_{G2}$ | | | | | | | 50 | 30 | | 348 |
| F | 21.0 | 44.5 | 16.2 | 16.2 | 1268.9 | 8.8 | | 28 | 85 | |
| $F_{G1}$ | | | | | | | 0 | 71 | | 79 |
| $F_{G2}$ | | | | | | | 45 | | | |
| Count | 5 | 5 | 5 | 5 | 5 | 5 | 10 | 13 | 5 | 8 |
| Maximum | 21.0 | 44.5 | 17.5 | 18.5 | 1268.9 | 8.8 | 75.00 | 88 | 125 | 348 |
| Mean | | | | | | | | | | |
| Median | 20.7 | 44.5 | 15.9 | 16.2 | 1110.3 | 7.7 | 45.00 | 28 | 95 | 136 |
| Minimum | 19.8 | 44.5 | 14.2 | 14.2 | 1110.3 | 7.7 | 0.00 | 20 | 73 | 10 |
| Std. dev. | 0.6 | 0.0 | 1.3 | 1.7 | 77.3 | 0.5 | 21.31 | 23 | 23 | 116 |

## Table A10

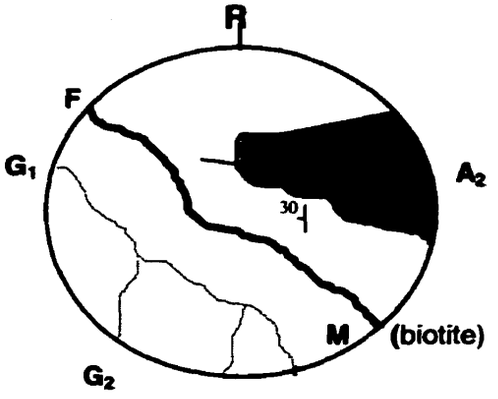
## Axial Point Load Testing Results

## Coarse Grained Granite, Borehole 418-051-RT5

| Sample | | | Point Load Index | | Tensile Strength | | Dihedral angle | True orientations | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | | | | | | | | Dip | Dip Direction | |
| fracture *foliation* | Thickness (mm) | Diameter (mm) | $I_s$ (Mpa) | Normalized $I_s$ (MPa) | (PSI) | (MPa) | loading direction & foliation | | Fracture | Foliation |
| A | 19.1 | 44.6 | 13.8 | 15.2 | 892.1 | 6.2 | | 0 | 0 | |
| $A_{G1}$ | | | | | | | 10 | | | |
| B | 19.7 | 44.6 | 14.1 | 15.1 | 971.1 | 6.7 | | 20 | 87 | |
| B | | | | | | | 15 | 25 | | 90 |
| C | 20.1 | 44.6 | 13.2 | 13.9 | 943.4 | 6.5 | | 25 | 53 | |
| $C_{G1}$ | | | | | | | 8 | 65 | | 90 |
| D | 19.9 | 44.6 | 13.4 | 14.2 | 943.4 | 6.5 | | 15 | 53 | |
| $D_{G1}$ | | | | | | | 72 | 18 | | 54 |
| E | 20.0 | 44.6 | 17.8 | 18.7 | 1263.2 | 8.7 | | 71 | 53 | |
| $E_{G1}$ | | | | | | | 5 | 20 | | 160 |
| F | 20.9 | 44.6 | 14.0 | 14.1 | 1089.5 | 7.5 | | 0 | 0 | |
| $F_{G1}$ | | | | | | | 10 | | | |
| G | | | | | | | 10 | 15 | 233 | |
| G | | | | | | | 35 | 12 | 53 | |
| G | | | | | | | 83 | 10 | 233 | |
| $G_{G1}$ | | | | | | | | 13 | | 230 |
| H | 21.8 | 44.6 | 14.0 | 13.5 | 1184.2 | 8.2 | | 50 | 233 | |
| $H_{G1}$ | | | | | | | 45 | | | |
| I | 20.5 | 44.6 | 18.0 | 18.5 | 1342.1 | 9.3 | | 30 | 53 | |
| $I_{G1}$ | | | | | | | 10 | 28 | | 75 |
| J | 20.4 | 44.6 | 16.1 | 16.6 | 1184.2 | 8.2 | | 5 | 233 | |
| $J_{G1}$ | | | | | | | 5 | 8 | | 187 |
| K | 20.2 | 44.6 | 15.3 | 15.9 | 1105.3 | 7.6 | | 15 | 233 | |
| L | 20.6 | 44.6 | 16.8 | 17.1 | 1263.2 | 8.7 | 30 | 77 | 233 | |
| L | 20.6 | 44.6 | 16.8 | 17.1 | 1263.2 | 8.7 | 35 | 5 | 233 | |
| L | 20.6 | 44.6 | 16.8 | 17.1 | 1263.2 | 8.7 | 82 | 45 | 53 | |
| $L_{G1}$ | | | | | | | | 75 | | 228 |
| Count | 7 | 7 | 7 | 7 | 7 | 7 | 15 | 24 | 16 | 8 |
| Maximum | 21.8 | 44.6 | 18.0 | 18.5 | 1342.1 | 9.3 | 83.00 | 77 | 233 | 230 |
| Mean | 20.7 | 44.6 | 15.4 | 15.9 | 1155.1 | 7.9 | 36.07 | 30 | 133 | 150 |
| Median | 20.6 | 44.6 | 16.8 | 17.1 | 1263.2 | 8.7 | 15.00 | 19 | 70 | 125 |
| Minimum | 20.2 | 44.6 | 14.0 | 13.5 | 1105.3 | 7.6 | 5.00 | 0 | 0 | 54 |
| Std. dev. | 0.5 | 0.0 | 1.3 | 1.5 | 77.0 | 0.5 | 28.17 | 24 | 98 | 71 |

**Sample A**

| Structure | Relative Orientation *(strike / dip & dip direction)* |
|---|---|
| F | 100°/90° NE |
| $G_1$ | 100°/20° NE |
| $G_2$ | 190°/60° NW |
| $M_1$ | 190°/60° NW |
| A | 100°/30° E |
| $^{30}$⊣ | strike & dip with dip direction |




Downhole View
*Dwg. No.*

where:

| | |
|---|---|
| R | core reference line |
| F | fracture induced by loading |
| $G_1$ | gneissic foliation type 1 |
| $G_2$ | gneissic foliation type 2 |
| $M_1$ | mineral alignment, type indicated on figure |
| $A_1$ | axial plane of fold |
| $A_2$ | axis of fold |
| | biotite > 10% or some compositional heterogeneity as indicated by text. |

Note: All orientations shown in the drawings are relative to the section shown, with the core reference line representing an interim North. These have been converted to true orientations in the report.

Figure A1. Key to symbols used on the sample maps.

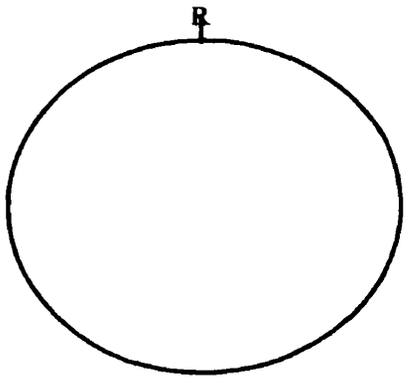
**Rock name** **Test type (Brazilian or Point load)**

**Borehole name (e.g. 401-123-rt2)** **Core size (alphanumeric and in mm).**

Sample a

| Structure | Relative Orientation |
|---|---|
| F | strike (0-360) / dip (0-90) |
| $G_1$ | strike (0-360) / dip (0-90) |
| $G_2$ | strike (0-360) / dip (0-90) |

Comments
Brief description of rock including anything noteworthy about the foliations present and the affect (if any) this had on testing.

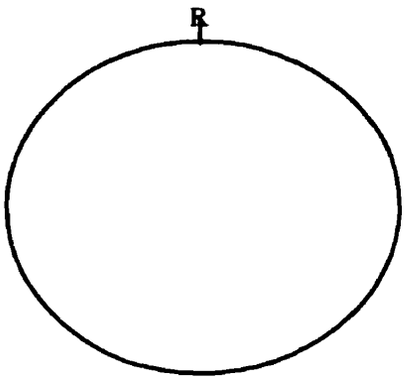



Downhole View
*Dwg. No.*

Sample b

| Structure | Relative Orientation |
|---|---|
| F | strike (0-360) / dip (0-90) |
| $G_1$ | strike (0-360) / dip (0-90) |
| $G_2$ | strike (0-360) / dip (0-90) |

Comments
Brief description of rock including anything noteworthy about the foliations present and the affect (if any) this had on testing.

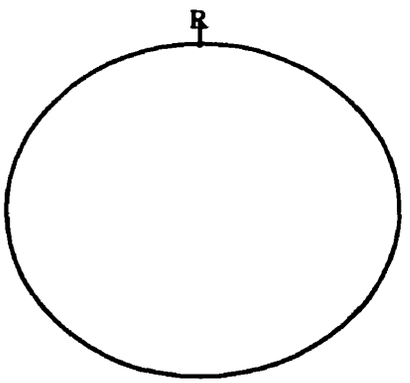



Downhole View
*Dwg. No.*

Sample c

| Structure | Relative Orientation |
|---|---|
| F | strike (0-360) / dip (0-90) |
| $G_1$ | strike (0-360) / dip (0-90) |
| $G_2$ | strike (0-360) / dip (0-90) |

Comments
Brief description of rock including anything noteworthy about the foliations present and the affect (if any) this had on testing.




Downhole View
*Dwg. No.*

Figure A2: Standard format for the fabric maps constructed for each of the samples selected for testing. The rock name, the borehole from which the sample came, the size of the borehole, and the test type are indicated at the top of each sheet. Up to three sample maps are shown per page, at their actual size. The reference line marked on the top of the core is shown, and all views are downhole unless otherwise indicated.

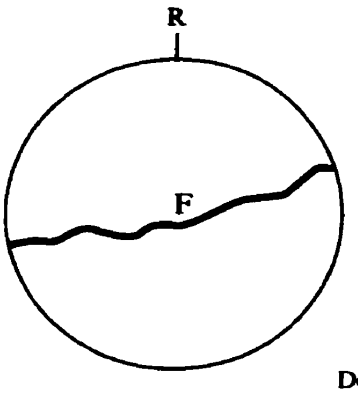
## Medium Grained Granite
Borehole 417-071-RT4

## Test Type (Point Load)
Core Size: NQ-3, 45 mm

### Sample A

| Structure | Relative Orientation |
|---|---|
| F | 065°/90° |

Comments

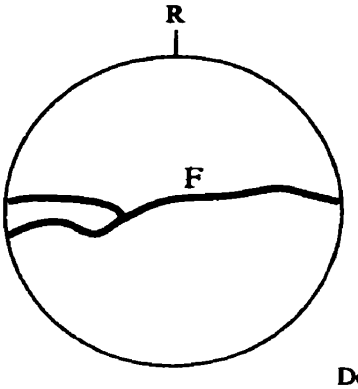



Downhole View
Fig. 2_071

### Sample B

| Structure | Relative Orientation |
|---|---|
| F | 085°/90° |

Comments

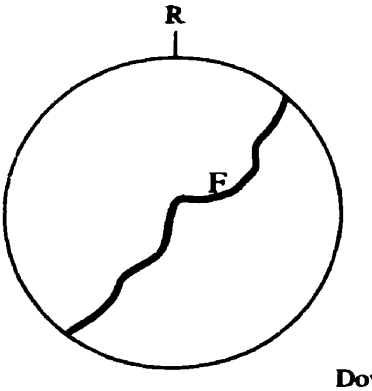



Downhole View
Fig. 2_072

### Sample C

| Structure | Relative Orientation |
|---|---|
| F | 045°/90° |

Comments




Downhole View
Fig. 2_073

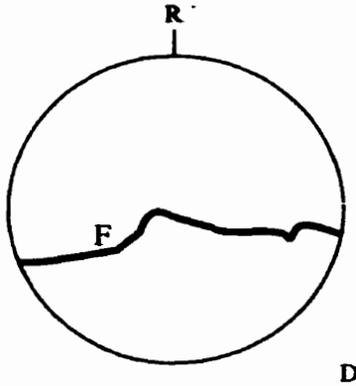
## Medium Grained Granite
Borehole 417-071-RT4

## Test Type (Point Load)
Core Size: NQ-3, 45 mm

### Sample D

| Structure | Relative Orientation |
|---|---|
| F | 090°/90° |

Comments

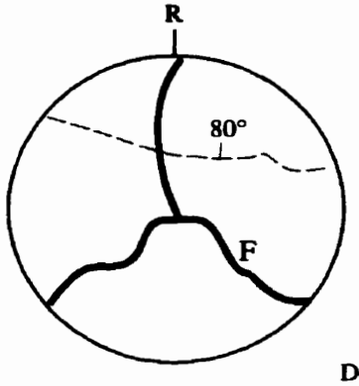



Downhole View
Fig. 2_074

### Sample E

| Structure | Relative Orientation |
|---|---|
| F | 000°/90° |
| | 130°/90° |
| | 225°/90° |
| $M_t$ | 105°/80°N |

Comments
Triple point fracture

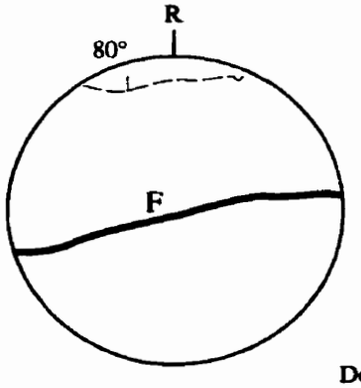



Downhole View
Fig. 2_075

### Sample F

| Structure | Relative Orientation |
|---|---|
| F | 085°/90° |
| $M_t$ | 088°/80°N |

Comments




Downhole View
Fig. 2_076

## Medium Grained Granite
Borehole 417-072-RT5

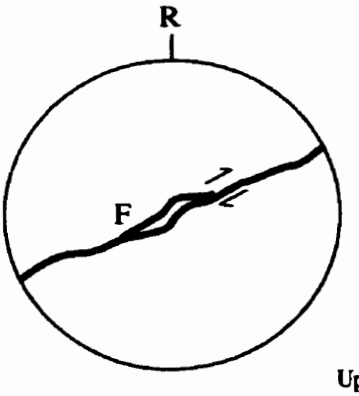
## Test Type (Point Load)
Core Size: NQ-3, 45 mm

Fig. 2_077

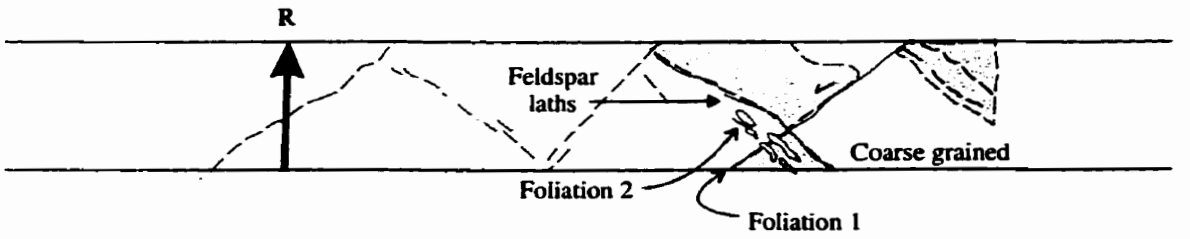
**Sample A**

| Structure | Relative Orientation |
|---|---|
| F | 067°/90° |
| $G_1$ | 080°/25°NW |
| $G_2$ | 025°/35°NW |

Comments



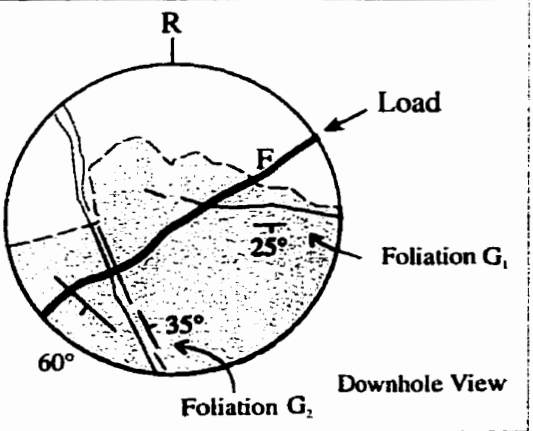

Uphole View




Perimeter View

| Structure | Relative Orientation |
|---|---|

Comments




Downhole View

## Medium Grained Granite
Borehole 417-072-RT5

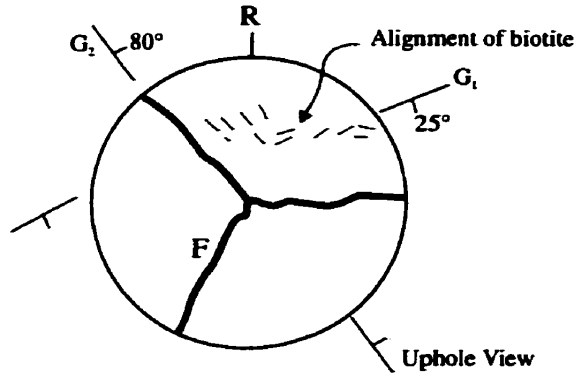
## Test Type (Point Load)
Core Size: NQ-3, 45 mm

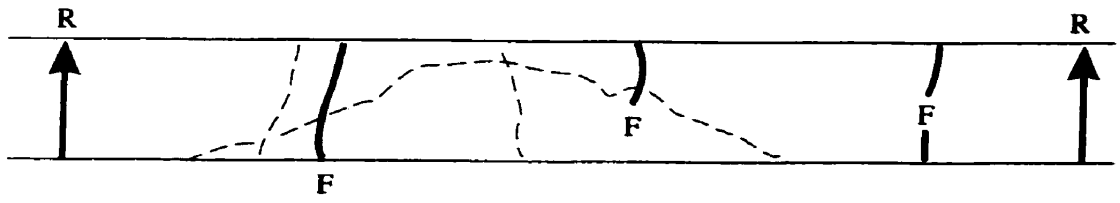
Fig. 2_079

**Sample B'**

| Structure | Relative Orientation |
|---|---|
| F | 090°/90° |
| F | 110°/90° |
| F | 320°/80°NE |
| $G_1$ | 060°/25°SE |
| $G_2$ | 140°/80°NE |

Comments




Perimeter View

## Medium Grained Granite
Borehole 417-072-RT5

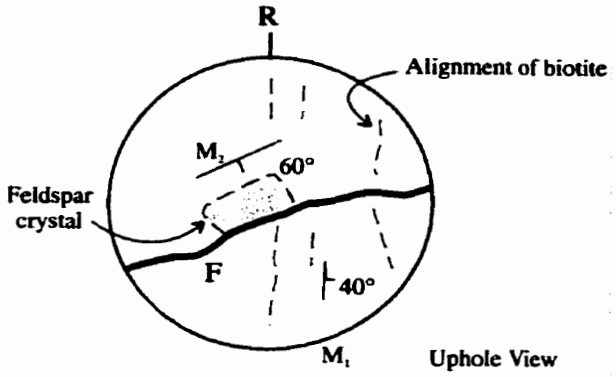
## Test Type (Point Load)
Core Size: NQ-3, 45 mm

Fig. 2_078

Sample C

| Structure | Relative Orientation |
|---|---|
| F | 070°/90° |
| $M_1$ | 180°/40°E |
| $M_2$ | 060°/60°SE |

Comments



Perimeter View

## Medium Grained Granite
Borehole 417-072-RT5

## Test Type (Point Load)
Core Size: NQ-3, 45 mm

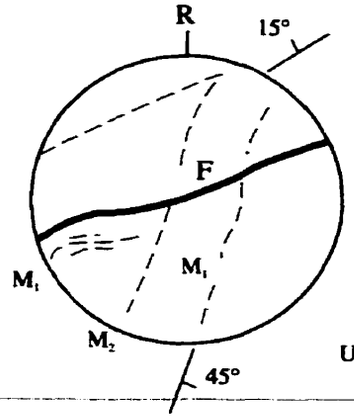
Fig. 2_081

Sample D

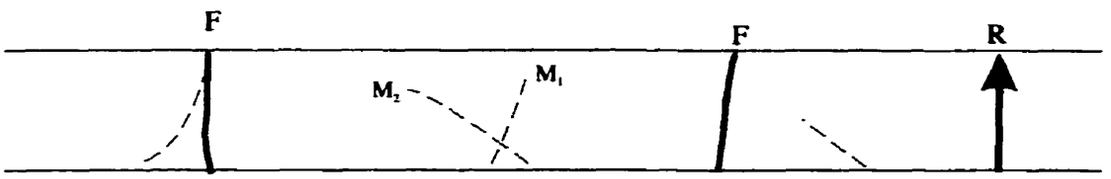
| Structure | Relative Orientation |
|---|---|
| F | 067°/90° |
| $M_1$ | 025°/45°SE |
| $M_2$ | 067°/15°SE |

Comments

Two Foliations;
Relative Orientations as shown




Uphole View




Perimeter View

## Medium Grained Granite
Borehole 417-072-RT5

## Test Type (Point Load)
Core Size: NQ-3, 45 mm

Fig. 2_082

Sample E

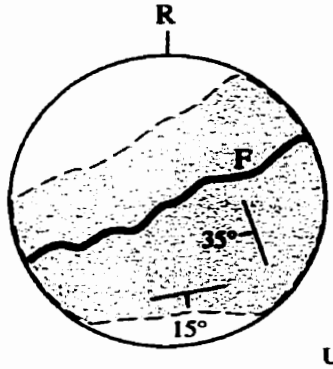
| Structure | Relative Orientation |
|---|---|
| F | 065°/90° |
| $M_1$ | 160°/35°SW |
| $M_1$ | 070°/15°SW |

Comments

Two foliations; orientations as shown

Sample appears to resemble boudin structure with white areas (feldspar) separated by greyish quartzose areas

Steeper foliation is marked by weak banding and offsets of boudins. Fracture as shown.




Uphole View

Perimeter View

## Medium Grained Granite
Borehole 417-072-RT5

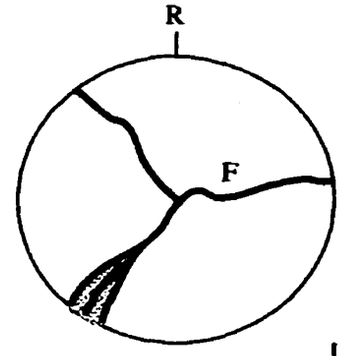
## Test Type (Point Load)
Core Size: NQ-3, 45 mm

Fig. 2_083

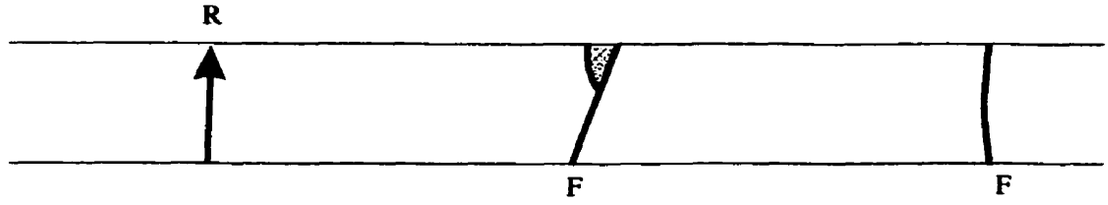
**Sample F**

| Structure | Relative Orientation |
|---|---|
| F | 085°/90° |
| F | 140°/90° |
| F | 040°/90° |

Comments
This sample is essentially massive




Uphole View




Perimeter View

## Medium Grained Granite
**Borehole 417-072-RT5**

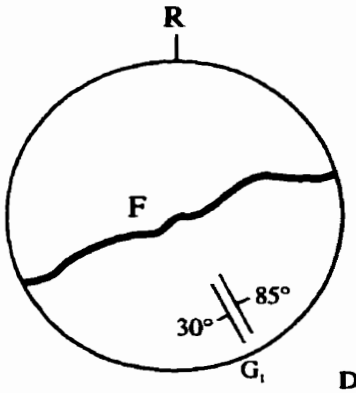
## Test Type (Point Load)
**Core Size: NQ-3, 45 mm**

### Sample G

| Structure | Relative Orientation |
|---|---|
| F | 065°/90° |
| $G_i$ | 145°/85°NE |
| $G_i$ | 145°/30°SW |

Comments

Weak mineral alignment
The shallow dipping foliation is more coarsely defined, with irregular bands which are several grains in width

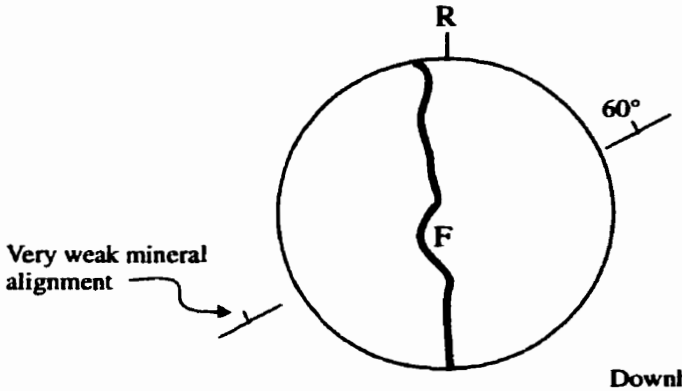



Downhole View
Fig. 2_084

### Sample H

| Structure | Relative Orientation |
|---|---|
| F | 175°/90° |
| $G_i$ | 065°/60°NW |

Comments

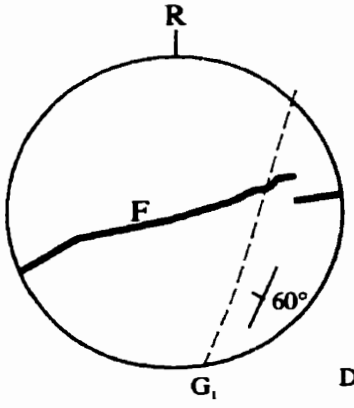



Downhole View
Fig. 2_085

### Sample I

| Structure | Relative Orientation |
|---|---|
| F | 080°/90° |
| $G_i$ | 015°/60°NW |

Comments




Downhole View
Fig. 2_086

## Medium Grained Granite
Borehole 417-072-RT5

## Test Type (Point Load)
Core Size: NQ-3, 45 mm

### Sample J

| Structure | Relative Orientation |
|---|---|
| F | 065°/90° |
| $G_1$ | 170°/75°E |
| $G_2$ | 065°/90°SE |

Comments

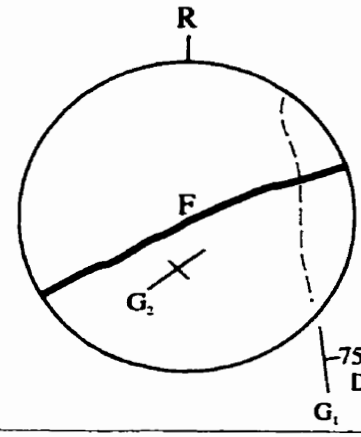


Fig. 2_087

### Sample K

| Structure | Relative Orientation |
|---|---|
| F | 060°/90° |
| $G_1$ | 150°/80°NE |
| $G_2$ | 090°/90° |

Comments

Very weak non-penetrative mineral foliation

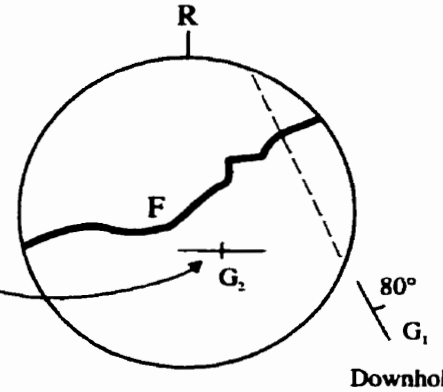


Fig. 2_088

### Sample L

| Structure | Relative Orientation |
|---|---|
| F | 085°/90° |
| $G_1$ | 145°/85°NE |
| $G_2$ | 060°/80°NW |

Comments

Non-penetrative lineament in matrix between feldspar phenocrysts

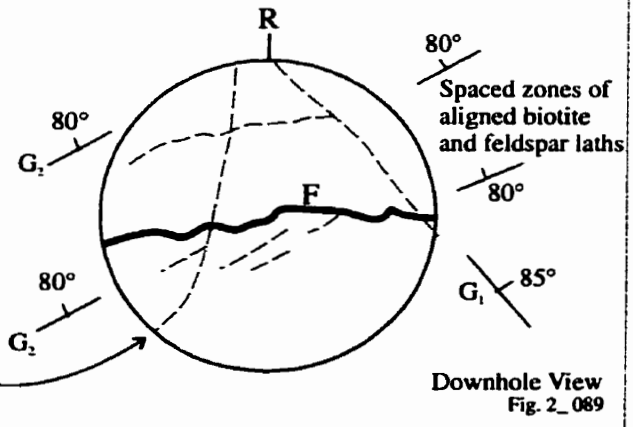


Fig. 2_089

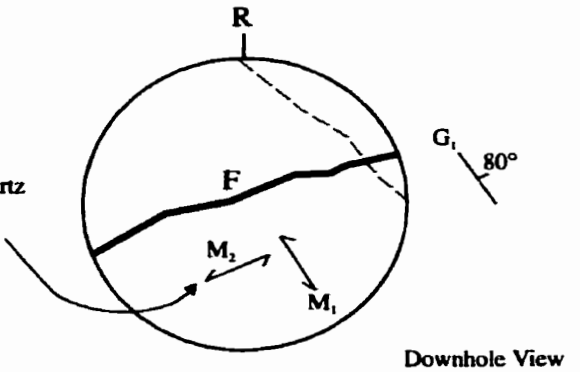
## Medium Grained Granite
Borehole 417-072-RT5

## Test Type (Point Load)
Core Size: NQ-3, 45 mm

### Sample M

| Structure | Relative Orientation |
|---|---|
| F | 070°/90° |
| $G_1$ | 130°/80°NE |
| $M_1$ | 130°/90° |
| $M_2$ | 145°/90° |

Comments: Broken feldspar laths with quartz filling, also alteration veins in feldspar and biotite

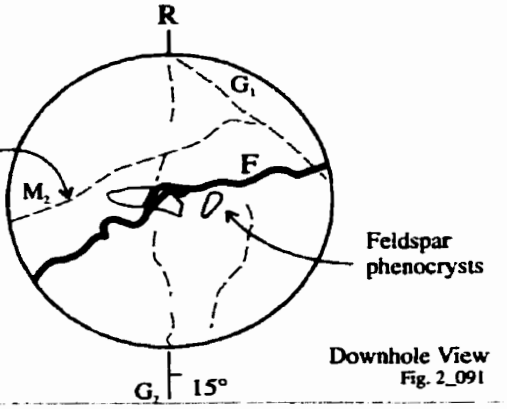



Downhole View
Fig. 2_090

### Sample N

| Structure | Relative Orientation |
|---|---|
| F | 065°/90° |
| $G_1$ | 130°/80°NE |
| $G_2$ | 090°/15° |
| $M_2$ | 065°/90° |

Comments: Weak mineral alignments and bleached alteration along microcracks




Downhole View
Fig. 2_091

## Coarse Grained Granite
Borehole 418-051-RT2

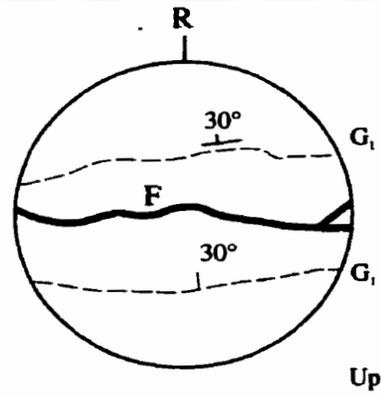
## Test Type (Point Load)
Core Size: NQ-3, 45 mm

Fig. 2_092

**Sample A**

| Structure | Relative Orientation |
|---|---|
| F | 090°/90° |
| $G_t$ | 090°/30°N |

Comments

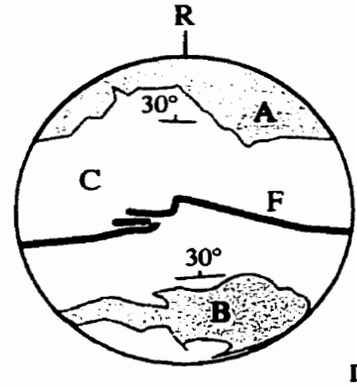



Uphole View

Perimeter View

Structure

Comments




Downhole View

## Coarse Grained Granite
Borehole 418-051-RT2

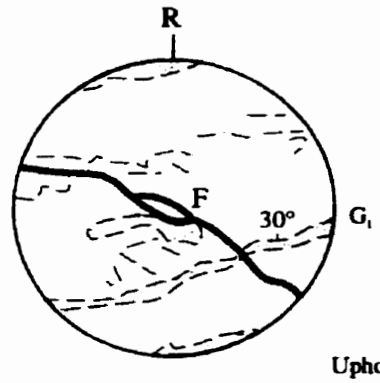
## Test Type (Point Load)
Core Size: NQ-3, 45 mm

Fig. 2_093

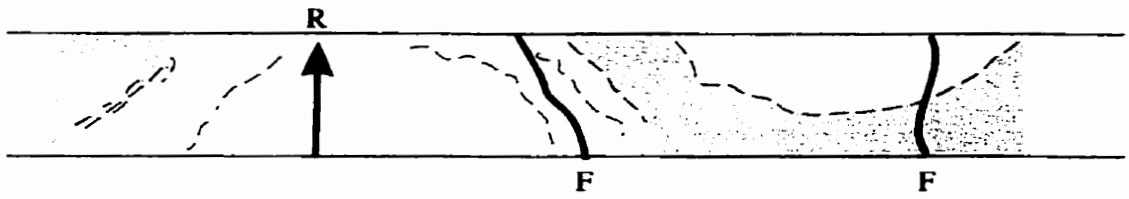
**Sample B**

| Structure | Relative Orientation |
|---|---|
| F | 070°/90° |
| F | 100°/90° |
| $G_1$ | 100°/30°N |

Comments



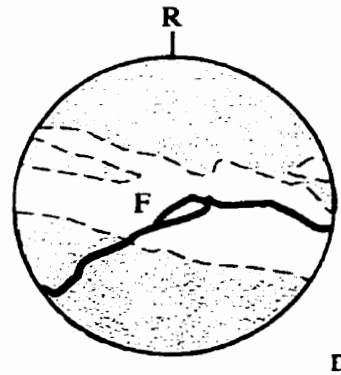

Uphole View




Perimeter View

Structure

Comments




Downhole View

## Coarse Grained Granite
Borehole 418-051-RT2

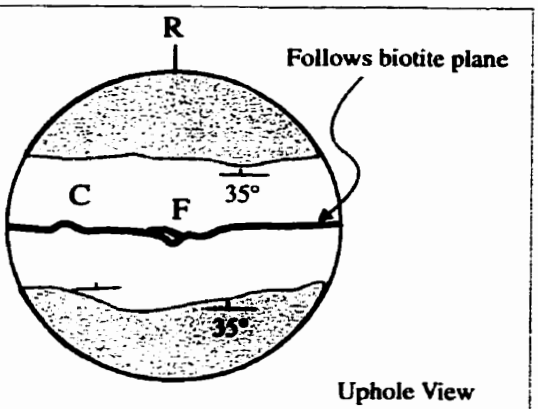
## Test Type (Point Load)
Core Size: NQ-3, 45 mm

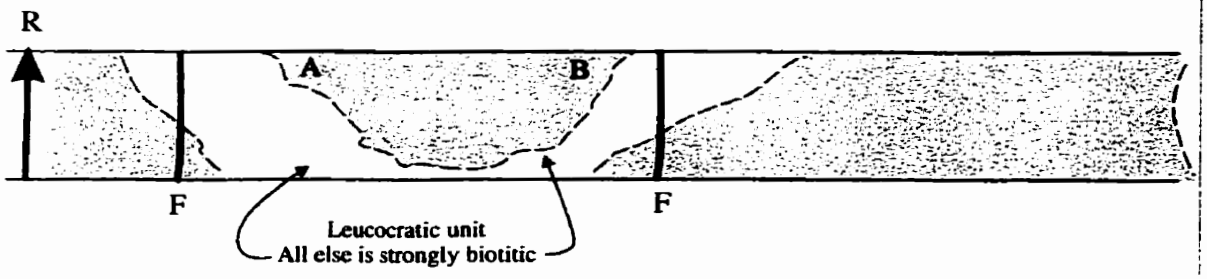
Fig. 2_094

**Sample C**

| Structure | Relative Orientation |
|---|---|
| F | 090°/90° |
| $G_t$ | 090°/35°N |

Comments

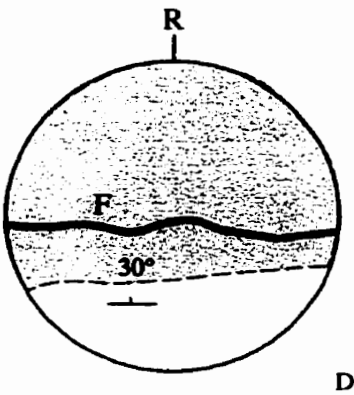



Perimeter View

Structure

Comments



## Coarse Grained Granite
Borehole 418-051-RT2

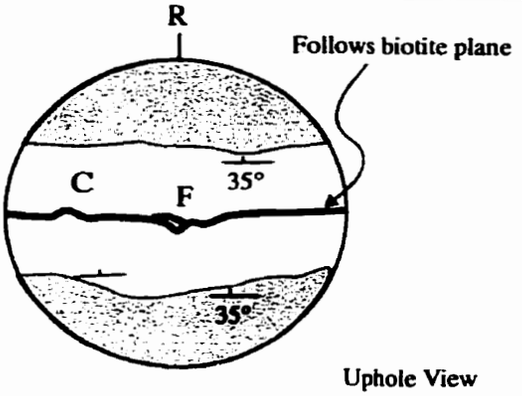
## Test Type (Point Load)
Core Size: NQ-3, 45 mm

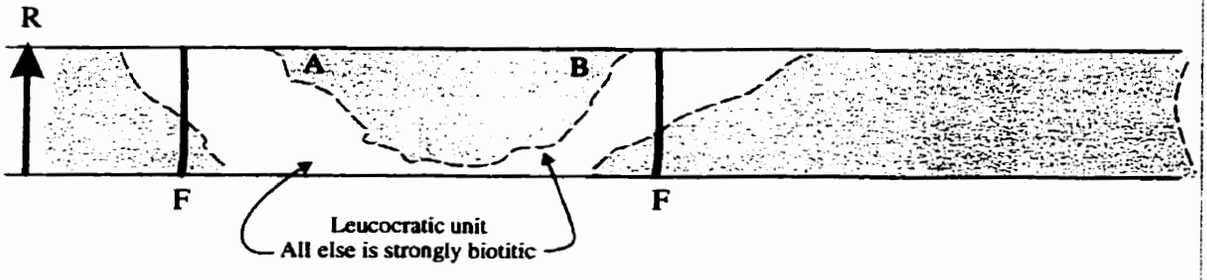
Fig. 2_094

**Sample C**

| Structure | Relative Orientation |
|---|---|
| F | 090°/90° |
| $G_1$ | 090°/35°N |

Comments

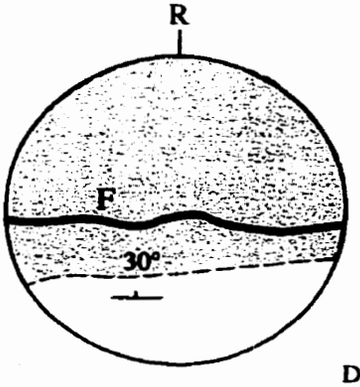



Perimeter View

Structure

Comments



## Coarse Grained Granite
Borehole 418-051-RT2

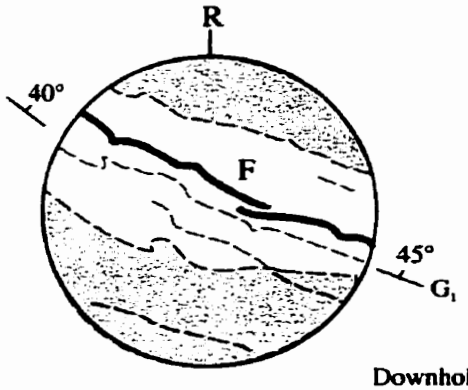
## Test Type (Point Load)
Core Size: NQ-3, 45 mm

### Sample D

| Structure | Relative Orientation | |
|---|---|---|
| F | 120°/90° | Fracture |
| $G_1$ | 120°/45°NE | Gneissosity |
| $G_2$ | 120°/40°NE | Thrusts |

Comments

Sample is a classic fabric type 4 with well demarcated layers of alternating biotite and leucocratic granite. Some of the biotitic lenses are slip surfaces (like miniature thrust faults)



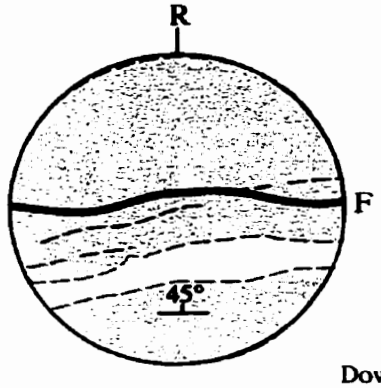

Downhole View
Fig. 2_095

### Sample E

| Structure | Relative Orientation |
|---|---|
| F | 085°/90° |
| $G_1$ | 080°/ 45°N |

Comments

Sample is strongly layered, biotitic schlieren alternating with leucocratic and intermediate layers.

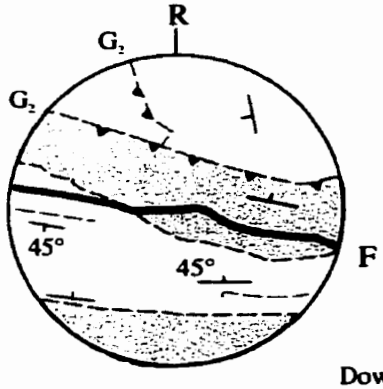



Downhole View
Fig. 2_096

### Sample G

| Structure | Relative Orientation | |
|---|---|---|
| F | 105°/90° | |
| $G_1$ | 105°/45°N | Gneissic Banding |
| $G_2$ | 105°/45°S | Thrusts |

Comments




Downhole View
Fig. 2_098

## Coarse Grained Granite
**Borehole 418-051-RT2**

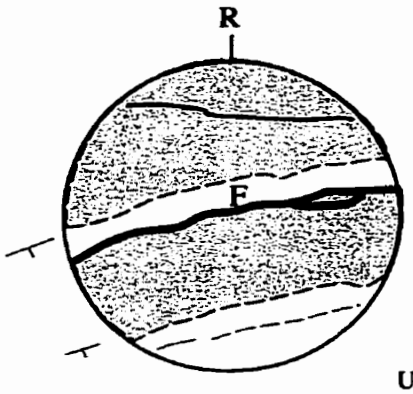
## Test Type (Point Load)
**Core Size: NQ-3, 45 mm**

Fig. 2_099

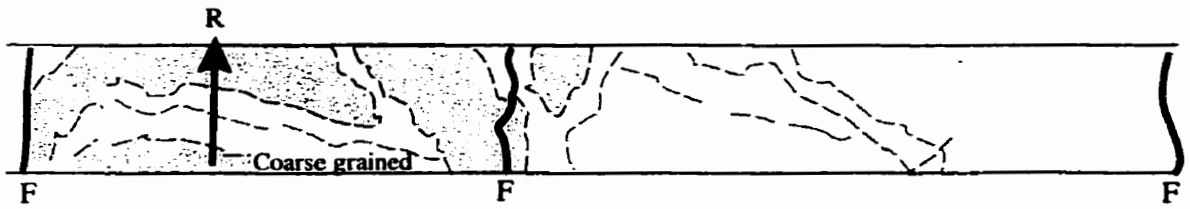
**Sample H**

Structure | Relative Orientation

Comments



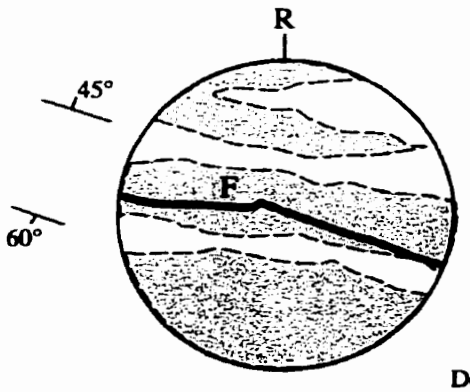

Uphole View




Perimeter View

| Structure | Relative Orientation | |
|---|---|---|
| F | 100°/90° | Fracture |
| $G_1$ | 100°/45°NE | |
| $A_2$ | 100°/60°SW | Fold axial plane |
| $A_2$ | 100°/45°NE | Fold axial plane |

Comments




Downhole View

## Coarse Grained Granite
Borehole 418-051-RT2

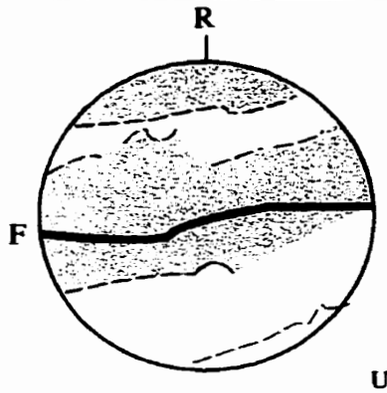
## Test Type (Point Load)
Core Size: NQ-3, 45 mm

Fig. 2_100

Sample I

Structure Relative Orientation

Comments

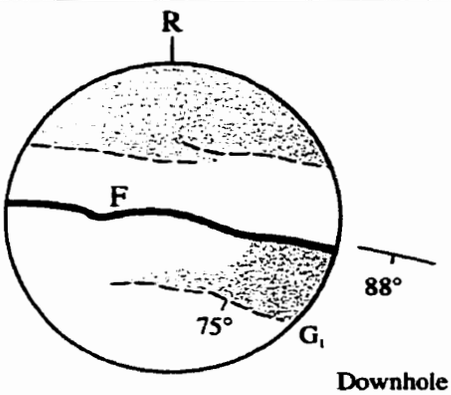



Uphole View

Perimeter View

| Structure | Relative Orientation |
|---|---|
| F | 100°/88°S |
| $G_t$ | 100°/75°S |

Comments




Downhole View

## Coarse Grained Granite
Borehole 418-051-RT2

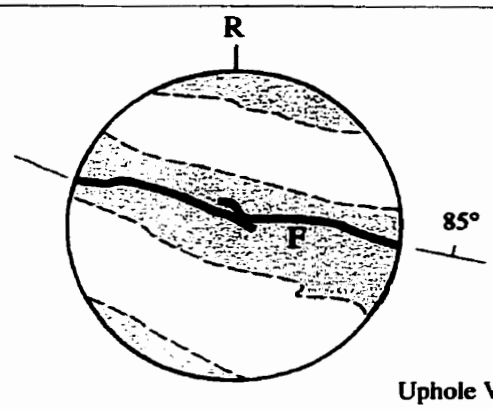
## Test Type (Point Load)
Core Size: NQ-3, 45 mm

Fig. 2_101

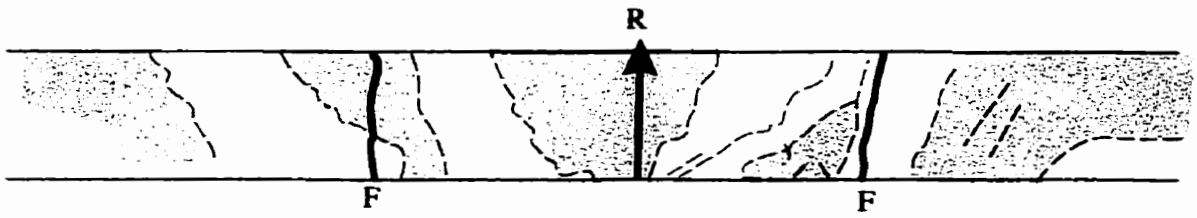
**Sample J**

Structure | Relative Orientation

Comments



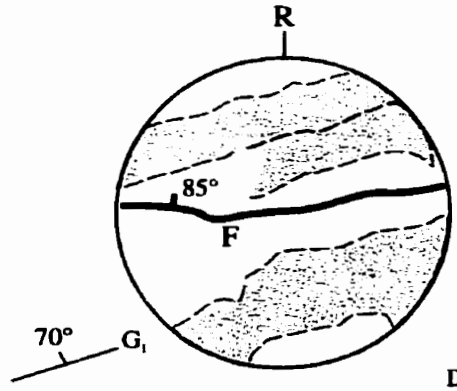

Uphole View




Perimeter View

| Structure | Relative Orientation |
|---|---|
| F | 080°/85°N |
| $G_1$ | 080°/70°N |

Comments




Downhole View

## Coarse Grained Granite
Borehole 418-051-RT2

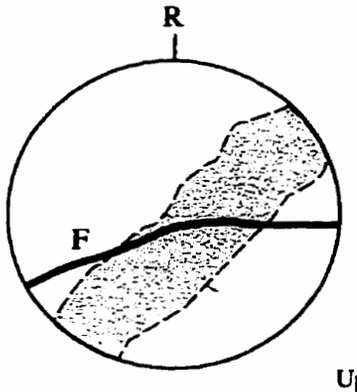
## Test Type (Point Load)
Core Size: NQ-3, 45 mm

Fig. 2_102

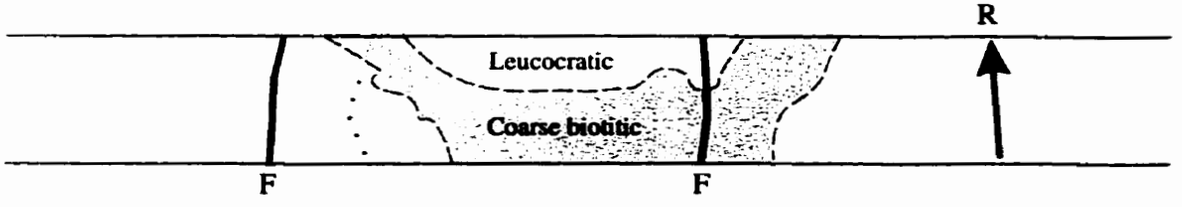
**Sample J***

Structure | Relative Orientation

Comments



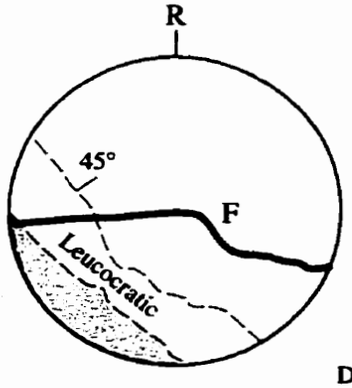

Uphole View




Perimeter View

| Structure | Relative Orientation | |
|---|---|---|
| F | 100°/90° | Fracture |
| $G_1$ | 120°/45°NE | |

Comments




Downhole View

## Coarse Grained Granite
Borehole 418-051-RT2

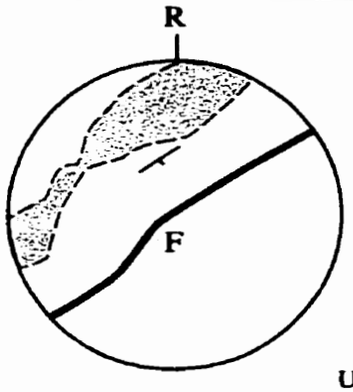
## Test Type (Point Load)
Core Size: NQ-3, 45 mm

Fig. 2_103

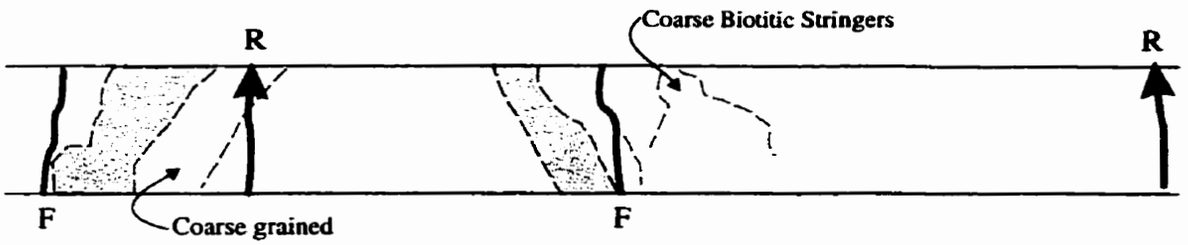
**Sample K**

Structure | Relative Orientation

Comments



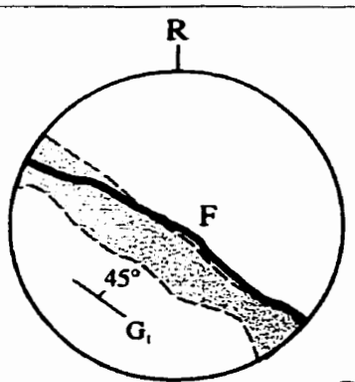

Uphole View




Perimeter View

| Structure | Relative Orientation |
|---|---|
| F | 120°/ 90° |
| $G_t$ | 120°/45°NE |

Comments




Downhole View

## Coarse Grained Granite
Borehole 418-051-RT2

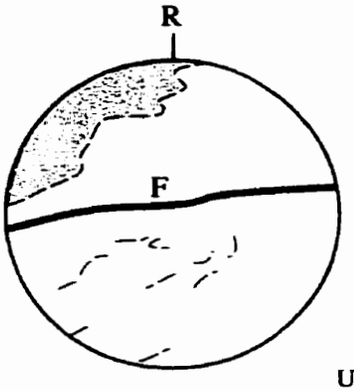
## Test Type (Point Load)
Core Size: NQ-3, 45 mm

Fig. 2_104

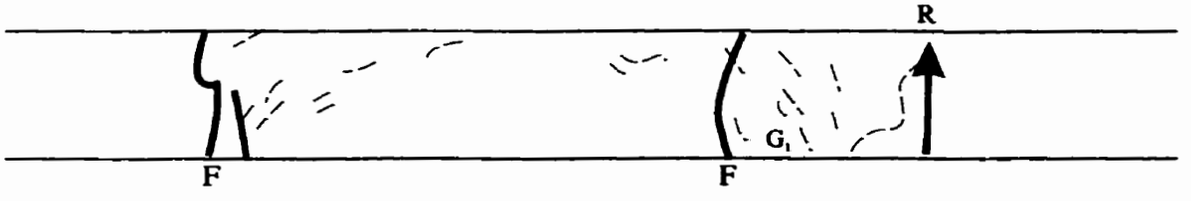
Sample L

Structure Relative Orientation

Comments




Uphole View

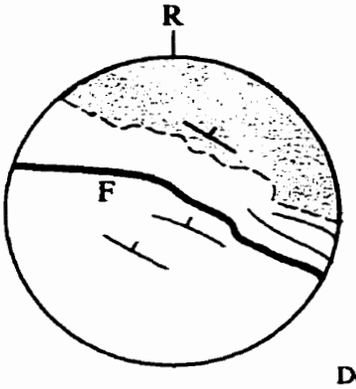



Perimeter View

| Structure | Relative Orientation |
|---|---|
| F | 115°/85°NE |
| $G_1$ | 115°/85°NE |
| $G_2$ | 115°/48°SW |

Comments

Sample contains two distinct gneissic layers, one biotitic, the other coarse to medium grained leucocratic granite. Biotite grains within the latter have a preferred orientation trending same but with opposite dip direction.




Downhole View

## Coarse Grained Granite
**Borehole 418-051-RT2**

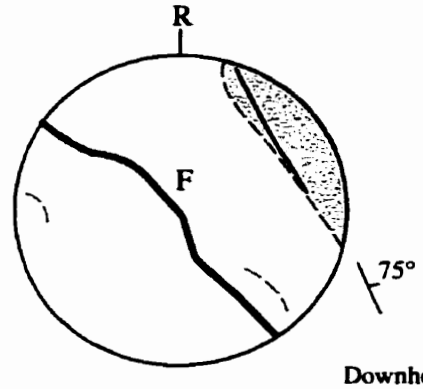
## Test Type (Point Load)
**Core Size: NQ-3, 45 mm**

### Sample M

| Structure | Relative Orientation | |
|---|---|---|
| F | 135°/90° | Fracture |
| $G_1$ | 140°/75°NE | Gneissic foliation |

Comments

Fracture developed in coarse granite leucocratic segment of sample but nearly along foliation. Crack appears to be along grain boundaries.



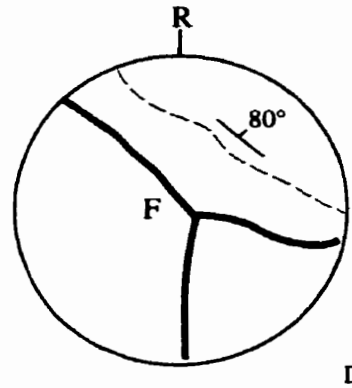

Downhole View
Fig. 2_105

### Sample N

| Structure | Relative Orientation |
|---|---|
| F | 100°/90° |
| F | 180°/90° |
| F | 135°/90° |
| $G_1$ | 135°/80°NE |

Comments

Sample is entirely leucocratic granite with ill-defined banding distinguished by variations in relative abundance of quartz. Coarse grained.



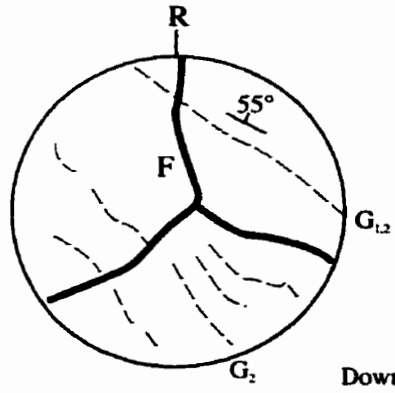

Downhole View
Fig. 2_ 106

### Sample O

| Structure | Relative Orientation |
|---|---|
| F | 105° /90° |
| F | 235°/90° |
| F | 000°/90° |
| $G_1$ | 125°/90° |
| $G_2$ | 105°/55°NE |

Comments

Sample is entirely coarse grained leucocratic granite with no biotitic lenses to impart layering. Instead, layering is indicated by variations in quartz content. Sample may show incipient discing.




Downhole View
Fig. 2_ 107

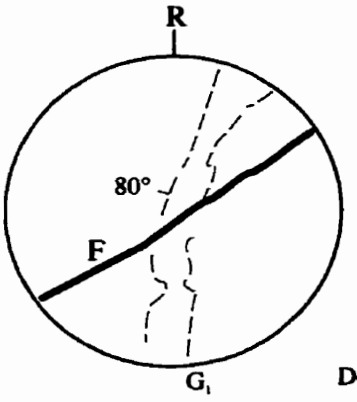
## Coarse Grained Granite
**Borehole 418-051-RT3**

## Test Type (Point Load)
**Core Size: NQ-3, 45 mm**

### Sample A

| Structure | Relative Orientation |
|---|---|
| F | 068°/90° |
| $G_t$ | 025°/80°NW |

Comments

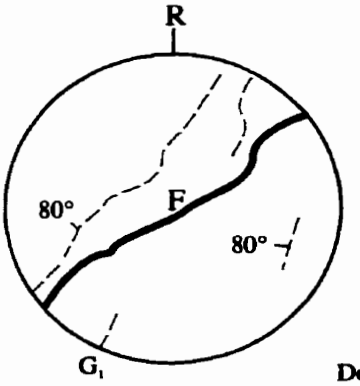



Downhole View
Fig. 2_121

### Sample B

| Structure | Relative Orientation |
|---|---|
| F | 055°/90° |
| $G_t$ | 040°/80°NW |

Comments

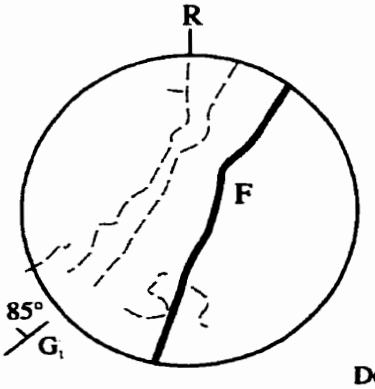



Downhole View
Fig. 2_122

### Sample C

| Structure | Relative Orientation |
|---|---|
| F | 025°/90° |
| $G_t$ | 035°/85°NW |

Comments




Downhole View
Fig. 2_123

## Coarse Grained Granite
**Borehole 418-051-RT3**

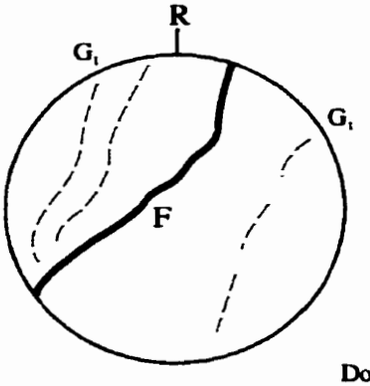
## Test Type (Point Load)
**Core Size: NQ-3, 45 mm**

### Sample D

| Structure | Relative Orientation |
|---|---|
| F | 025°/90° |
| | 055°/90° |
| $G_t$ | 025°/ 81°NW |
| | 055°/75°NW |

Comments

Fracture follows change in foliation



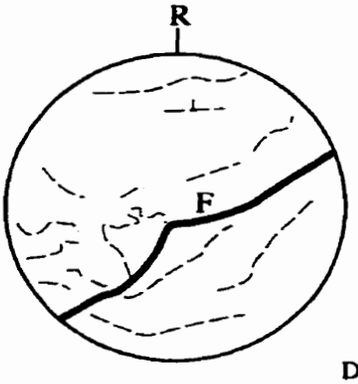

Downhole View
Fig. 2_124

### Sample E

| Structure | Relative Orientation |
|---|---|
| F | 055°/90° |
| F | 082°/90° |
| F | 025°/90° |
| F | 055°/90° |
| $G_t$ | 055°/70°NW |
| $G_t$ | 082°/70°NW |
| $G_t$ | 025°/65°NW |
| $G_t$ | 115°/60°NW |

Comments

Layering in sample folded in a complex manner.

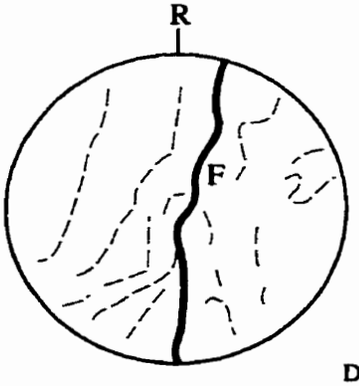



Downhole View
Fig. 2_125

### Sample F

| Structure | Relative Orientation |
|---|---|
| F | 010°/90° to 025°/90° |
| $G_t$ | 010°/85°NW |
| $G_t$ | 030°/85°NW |

Comments




Downhole View
Fig. 2_126

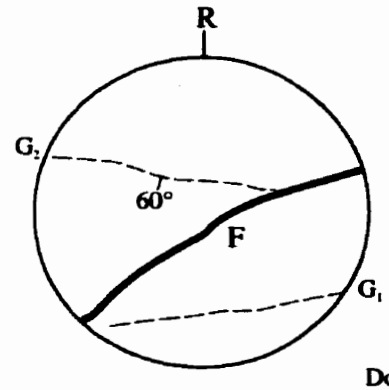
## Coarse Grained Granite
**Borehole 418-051-RT4**

## Test Type (Point Load)
**Core Size: NQ-3, 45 mm**

### Sample A

| Structure | Relative Orientation | |
|---|---|---|
| F | 060°/90° | |
| $G_1$ | 085°/90° | Coarse layering |
| $G_2$ | 165°/60°SW | Boudin plane |

Comments

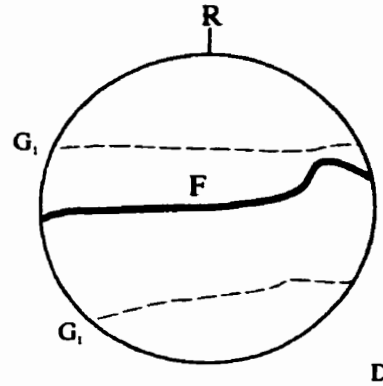



Downhole View
Fig. 2_ 127

### Sample B

| Structure | Relative Orientation |
|---|---|
| F | 085°/90° |
| $G_1$ | 085°/90° |

Comments

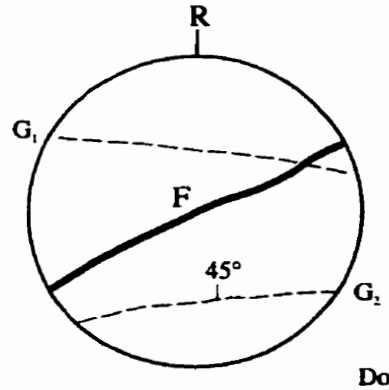



Downhole View
Fig. 2_ 128

### Sample C

| Structure | Relative Orientation |
|---|---|
| F | 045°/90° |
| $G_1$ | 090°/90° |
| $G_2$ | 090°/45°N |

Comments




Downhole View

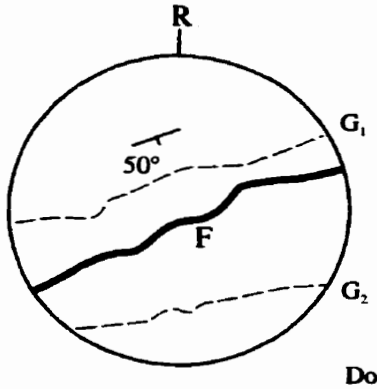
## Coarse Grained Granite
Borehole 418-051-RT4

## Test Type (Point Load)
Core Size: NQ-3, 45 mm

### Sample D

| Structure | Relative Orientation |
|---|---|
| F | 080°/90° |
| $G_1$ | 060°/50°S |
| $G_2$ | 060°/90° |

Comments

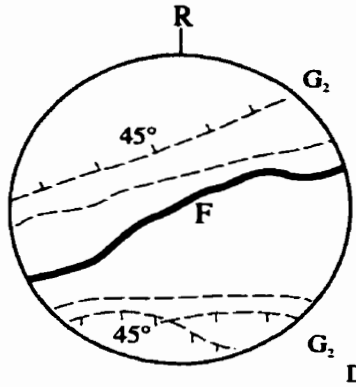



Downhole View
Fig. 2_129

### Sample E

| Structure | Relative Orientation |
|---|---|
| F | 070°/90° |
| $G_1$ | 085°/90° |
| $G_2$ | 070°/45°N |
| $G_1$ | 090°/45°S |

Comments

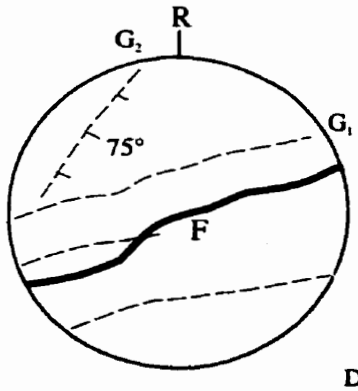



Downhole View
Fig. 2_ 130

### Sample F

| Structure | Relative Orientation |
|---|---|
| F | 070°/90° |
| $G_1$ | 070°/90° |
| $G_2$ | 030°/75°S |

Comments




Downhole View
Fig. 2_ 131

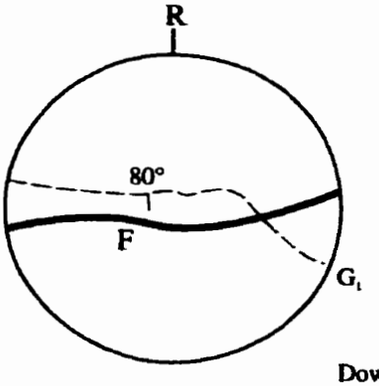
## Coarse Grained Granite
Borehole 418-051-RT5

## Test Type (Point Load)
Core Size: NQ-3, 45 mm

### Sample A

| Structure | Relative Orientation | |
|---|---|---|
| F | 090°/90° | |
| $G_t$ | 090°/80°S | Very irregular at this grain size |

Comments

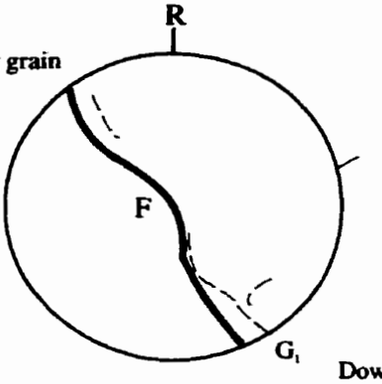



Downhole View
Fig. 2_108

### Sample B

| Structure | Relative Orientation | |
|---|---|---|
| F | 070°/~70°NE | Fracture has broken along grain |
| $G_t$ | 070°/~75°NE | boundaries at the contact |

Comments

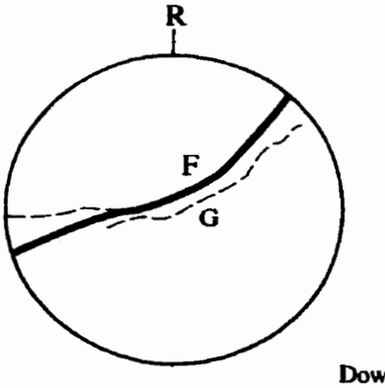



Downhole View
Fig. 2_109

### Sample C

| Structure | Relative Orientation | |
|---|---|---|
| F | 065°/90° | Broke along contact |
| G | 065°/90° | Grain boundary irregularity due to coarse grain size. |

Comments




Downhole View
Fig. 2_110

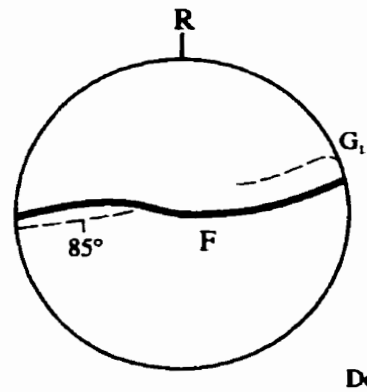
## Coarse Grained Granite
**Borehole 418-051-RT5**

## Test Type (Point Load)
**Core Size: NQ-3, 45 mm**

### Sample D

| Structure | Relative Orientation | |
|---|---|---|
| F | 085°/90° | Follows contact |
| $G_1$ | 080°/85°S | |

Comments

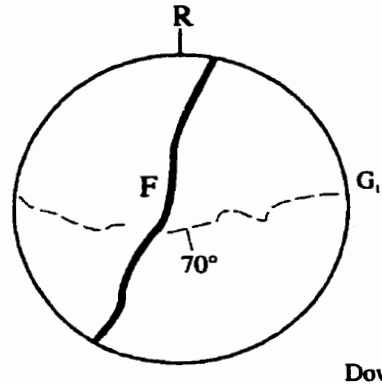



Downhole View
Fig. 2_111

### Sample E

| Structure | Relative Orientation |
|---|---|
| F | 020°/90° |
| $G_1$ | 090°/70°S |

Comments

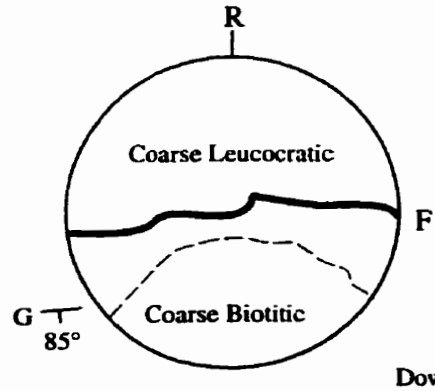



Downhole View
Fig. 2_112

### Sample F

| Structure | Relative Orientation |
|---|---|
| F | 090°/90° |
| G | 090°/85°S |

Comments




Downhole View
Fig. 2_113

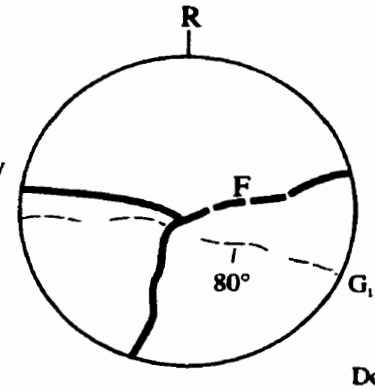
## Coarse Grained Granite
**Borehole 418-051-RT5**

## Test Type (Point Load)
**Core Size: NQ-3, 45 mm**

### Sample G

| Structure | Relative Orientation | |
|---|---|---|
| F | 105°/90° | Runs along contact |
| F | 075°/90° | |
| F | 190°/90° | |
| $G_i$ | 105°/80°S | Very coarsely crystalline, slight variations in mineralogy define layering |

Comments

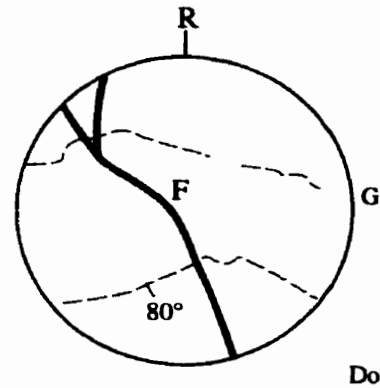



Downhole View
Fig. 2_114

### Sample H

| Structure | Relative Orientation |
|---|---|
| F | 140°/90° |
| $G_i$ | 095°/80°S |

Comments

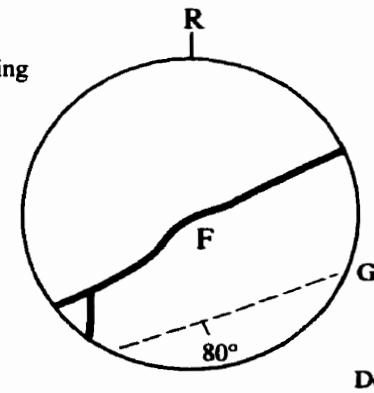



Downhole View
Fig. 2_115

### Sample I

| Structure | Relative Orientation | |
|---|---|---|
| F | 060°/90° | Nearly concordant with layering |
| G | 065°/80°S | |

Comments




Downhole View
Fig. 2_116

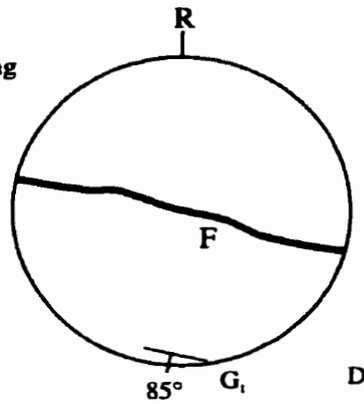
## Coarse Grained Granite
**Borehole 418-051-RT5**

## Test Type (Point Load)
**Core Size: NQ-3, 45 mm**

### Sample J

| Structure | Relative Orientation | |
|---|---|---|
| F | 095°/90° | Concordant with local layering |
| $G_1$ | 095°/85°S | |

Comments



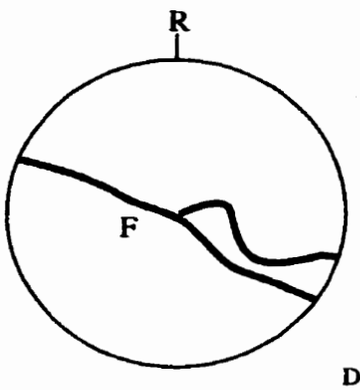

Downhole View
Fig. 2_117

### Sample K

| Structure | Relative Orientation |
|---|---|
| $F_1$ | 115°/90° |

Comments

Sample is massive

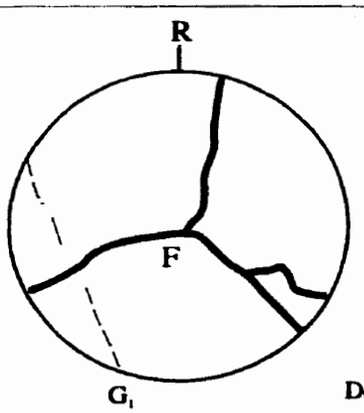



Downhole View
Fig. 2_ 118

### Sample L

| Structure | Relative Orientation |
|---|---|
| F | 010°/90° |
| F | 125°/90° |
| F | 260°/90° |
| G | 160°/90° |

Comments




Downhole View
Fig. 2_ 119

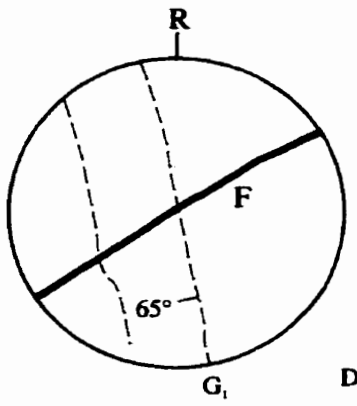
## Fine Grained Granite
Borehole 421-032-RT1

## Test Type (Point Load)
Core Size: NQ-3, 45 mm

### Sample A

| Structure | Relative Orientation |
|---|---|
| F | 055°/90° |
| $G_t$ | 170°/65°SW |

Comments

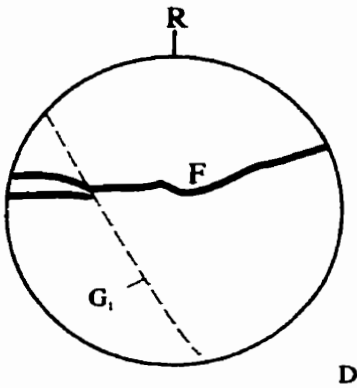



Downhole View
Fig. 2_134

### Sample B

| Structure | Relative Orientation |
|---|---|
| F | 090°/90° |
| F | 070°/90° |
| $G_t$ | 150 °/65°SW |

Comments

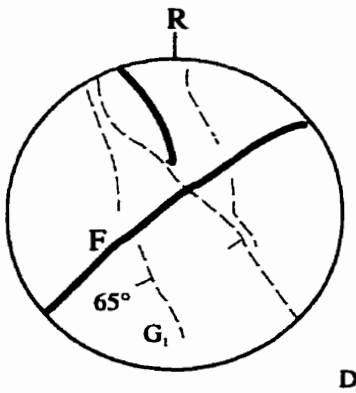



Downhole View
Fig. 2_133

### Sample C

| Structure | Relative Orientation |
|---|---|
| F | 045°/90° |
| F | 175°/90° |
| $G_t$ | 170°/ 65°W |

Comments




Downhole View
Fig. 2_135

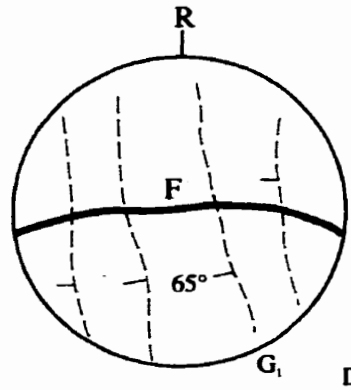
## Fine Grained Granite
**Borehole 421-032-RT1**

## Test Type (Point Load)
**Core Size: NQ-3, 45 mm**

### Sample D

| Structure | Relative Orientation |
|---|---|
| F | 085°/90° |
| $G_t$ | 170°/65°W |

Comments

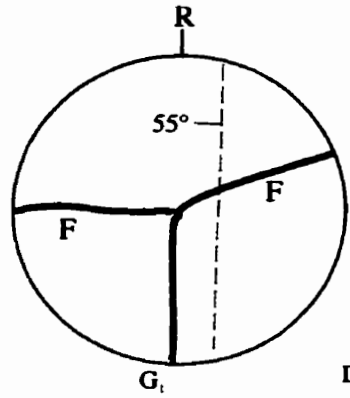



Downhole View
Fig. 2_136

### Sample E

| Structure | Relative Orientation |
|---|---|
| F | 000°/90° |
| F | 090°/90° |
| F | 060°/90° |
| $G_t$ | 005°/55°W |

Comments

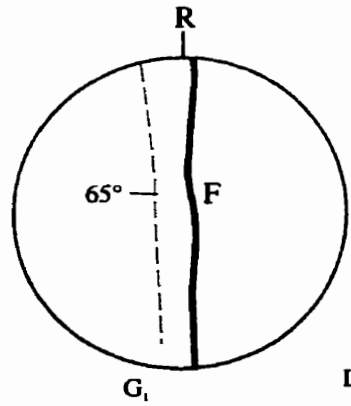



Downhole View

### Sample F

| Structure | Relative Orientation |
|---|---|
| F | 000°/90° |
| $G_t$ | 000°/65°W |

Comments




Downhole View
Fig. 2_137

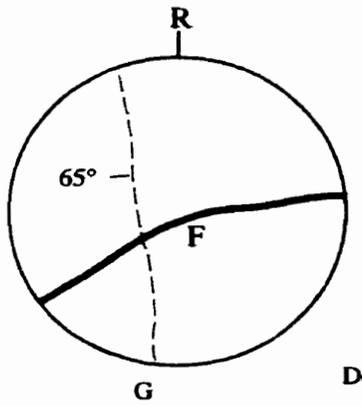
## Fine Grained Granite
Borehole 421-032-RT1

## Test Type (Point Load)
Core Size: NQ-3, 45 mm

### Sample G

| Structure | Relative Orientation |
|---|---|
| F | 085°/90° |
| F | 050°/90° |
| G | 170°/65°W |

Comments

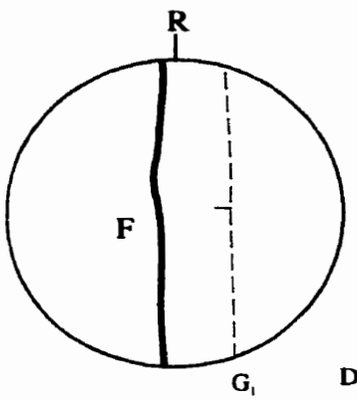



Downhole View
Fig. 2_ 138

### Sample H

| Structure | Relative Orientation |
|---|---|
| F | 120°/90° |
| G | 140°/65°W |

Comments

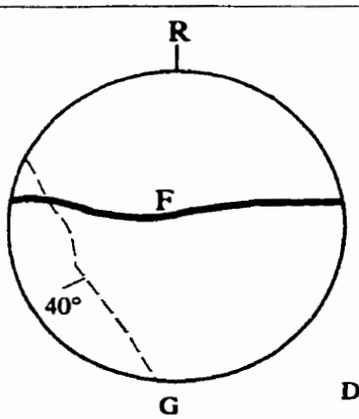



Downhole View

### Sample I

| Structure | Relative Orientation |
|---|---|
| F | 090°/90° |
| G | 145°/40°SW |

Comments




Downhole View
Fig. 2_ 132

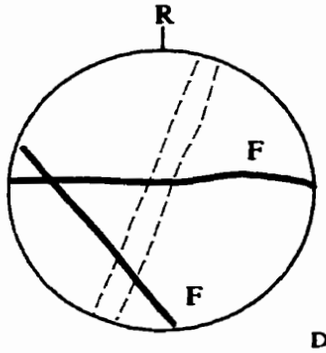
## Fine Grained Granite
Borehole 421-032-RT1

## Test Type (Point Load)
Core Size: NQ-3, 45 mm

### Sample J

| Structure | Relative Orientation |
|---|---|
| F | 085°/90° |
| $G_t$ | 150°/50°W |
| G | 030°/70°W |

Comments

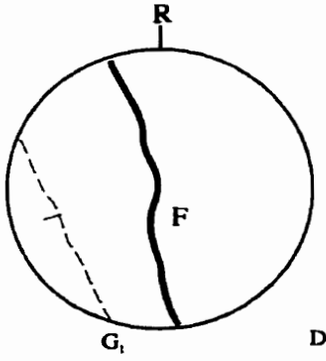


Downhole View

### Sample I-2

| Structure | Relative Orientation |
|---|---|
| F | 170°/90° |
| $G_t$ | 170°/55° |

Comments



Downhole View

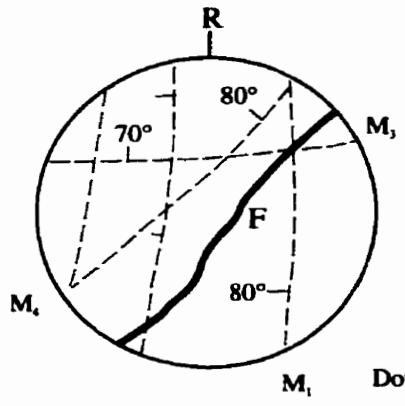
## Fine Grained Granite
**Borehole 421-032-RT2**

## Test Type (Point Load)
**Core Size: NQ-3, 45 mm**

### Sample A

| Structure | Relative Orientation |
|---|---|
| F | 045°/90° |
| $M_1$ | 005°/80°W |
| $M_2$ | |
| $M_3$ | 088°/70°N |
| $M_4$ | 045°/80°NW |

Comments

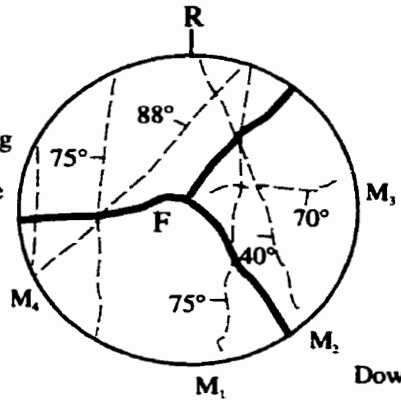



Downhole View
Fig. 2_139

### Sample B

| Structure | Relative Orientation | |
|---|---|---|
| F | 040°/80°NW | |
| F | 140°/80°NE | |
| F | 270°/90° | |
| $M_1$ | 010°/75°W | most prominent banding |
| $M_2$ | 165°/40°W | mineral alignment |
| $M_3$ | 088°/70°S | distinct non-penetrative |
| $M_4$ | 040°/88°W | banding |

Comments

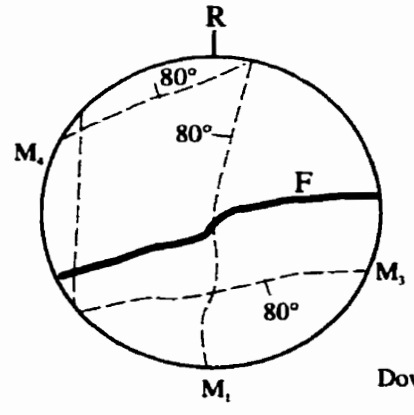



Downhole View
Fig. 2_140

### Sample C

| Structure | Relative Orientation |
|---|---|
| F | 085°/90° |
| $M_1$ | 010°/80°W |
| $M_2$ | |
| $M_3$ | 080°/80°S |
| $M_4$ | 060°/80°NW |

Comments




Downhole View
Fig. 2_141

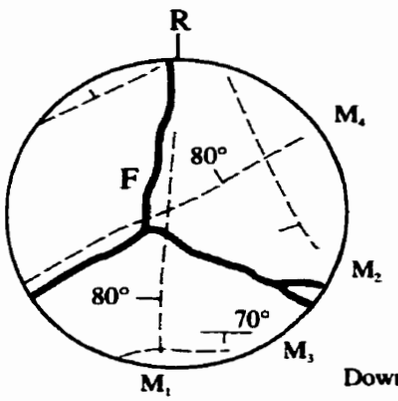
## Fine Grained Granite
**Borehole 421-032-RT2**

## Test Type (Point Load)
**Core Size: NQ-3, 45 mm**

### Sample D

| Structure | Relative Orientation |
|---|---|
| F | 010°/90° |
| F | 120°/90° |
| F | 055°/90° |
| $M_1$ | 000°/80°W |
| $M_2$ | 155°/40°W |
| $M_3$ | 088°/70°S |
| $M_4$ | 055°/80°W |

Comments

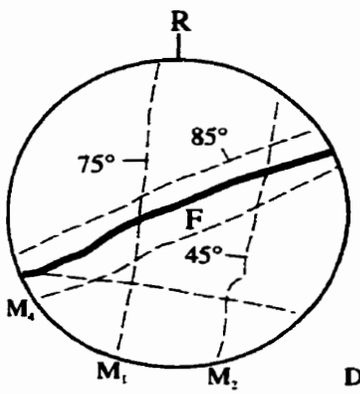



Downhole View
Fig. 2_ 142

### Sample E

| Structure | Relative Orientation |
|---|---|
| F | 070°/90° |
| $M_1$ | 010°/75°W |
| $M_2$ | 010°/45°W |
| $M_3$ | |
| $M_4$ | 070°/85°NW |

Comments

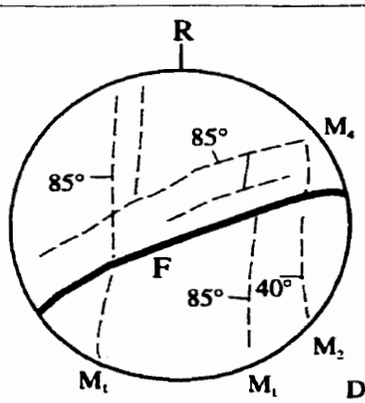



Downhole View
Fig. 2_ 143

### Sample F

| Structure | Relative Orientation |
|---|---|
| F | 070°/85°NW |
| $M_1$ | 010°/85°W |
| $M_2$ | 010°/40°W |
| $M_3$ | |
| $M_4$ | 055°/85°NW |

Comments




Downhole View
Fig. 2_ 144

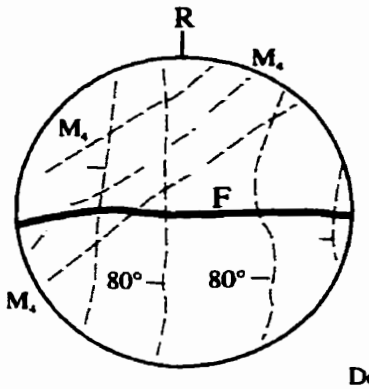
## Fine Grained Granite
Borehole 421-032-RT2

## Test Type (Point Load)
Core Size: NQ-3, 45 mm

### Sample G

| Structure | Relative Orientation |
|---|---|
| F | 090°/90° |
| $M_1$ | |
| $M_1$ | 010°/80°W |
| $M_2$ | 010°/45°W |
| $M_4$ | 060°/85°W |

Comments

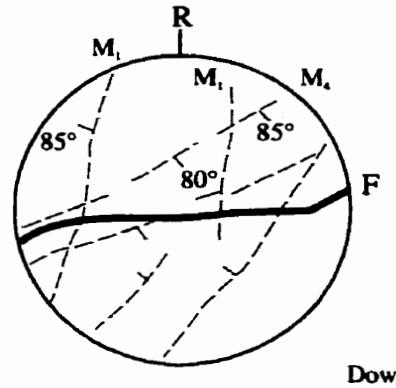



Downhole View
Fig. 2_145

### Sample H

| Structure | Relative Orientation |
|---|---|
| F | 080°/90° |
| $M_1$ | 015°/85°W |
| $M_2$ | 040°/50°NW |
| $M_3$ | |
| $M_4$ | 055°/80°SE |

Comments

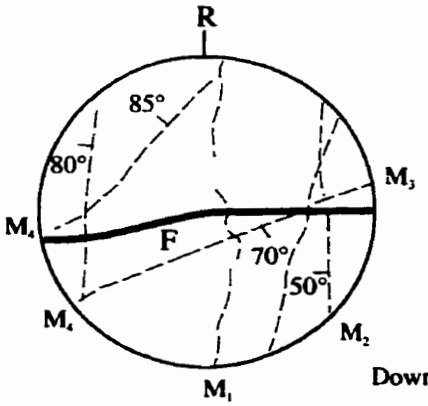



Downhole View
Fig. 2_146

### Sample I

| Structure | Relative Orientation |
|---|---|
| F | 090°/90° |
| $M_1$ | 012°/80°W |
| $M_2$ | 010°/50°W |
| $M_3$ | 085°/70°S |
| $M_4$ | 045°/85°N |

Comments




Downhole View
Fig. 2_147

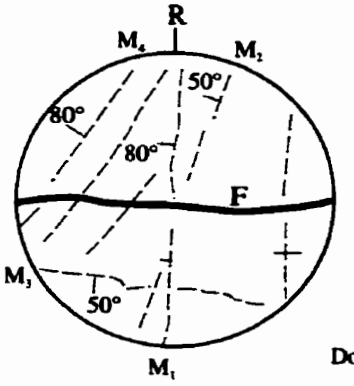
## Fine Grained Granite
Borehole 421-032-RT2

## Test Type (Point Load)
Core Size: NQ-3, 45 mm

**Sample J**

| Structure | Relative Orientation |
|---|---|
| F | 090°/85°S |
| $M_1$ | 000°/80°W |
| $M_2$ | 020°/50°NW |
| $M_3$ | 095°/50°S |
| $M_4$ | 040°/80°NW |

Comments




Downhole View
Fig. 2_148

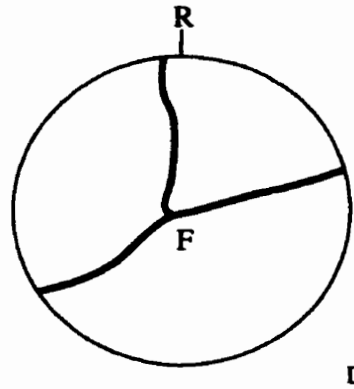
## Fine Grained Granite
Borehole 421-032-RT3

## Test Type (Point Load)
Core Size: NQ-3, 45 mm

### Sample A

| Structure | Relative Orientation |
|---|---|
| F | 000°/90° |
| F | 075°/90° |
| F | 052°/90° |

Comments

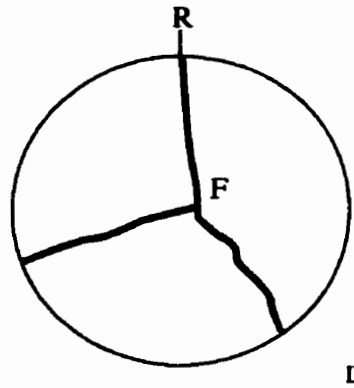



Downhole View
Fig. 2_149

### Sample B

| Structure | Relative Orientation |
|---|---|
| F | 000°/90° |
| F | 145°/90° |
| F | 070°/90° |

Comments

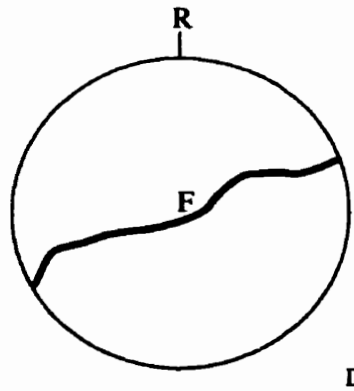



Downhole View
Fig. 2_150

### Sample C

| Structure | Relative Orientation |
|---|---|
| F | 070°/90° |

Comments




Downhole View
Fig. 2_151

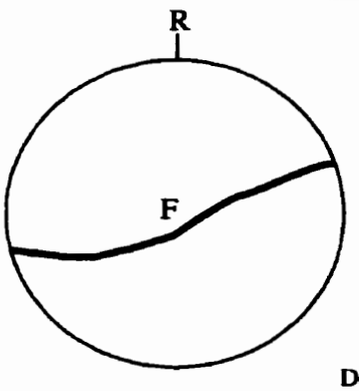
## Fine Grained Granite
**Borehole 421-032-RT3**

## Test Type (Point Load)
**Core Size: NQ-3, 45 mm**

### Sample D

| Structure | Relative Orientation |
|---|---|
| F | 070°/90° |

Comments

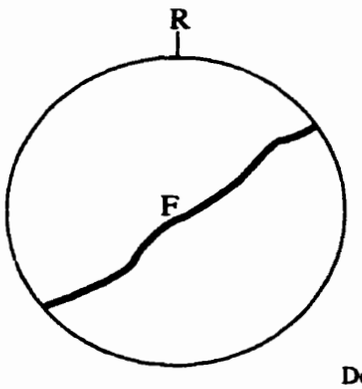



Downhole View
Fig. 2_152

### Sample E

| Structure | Relative Orientation |
|---|---|
| F | 050°/90° |

Comments

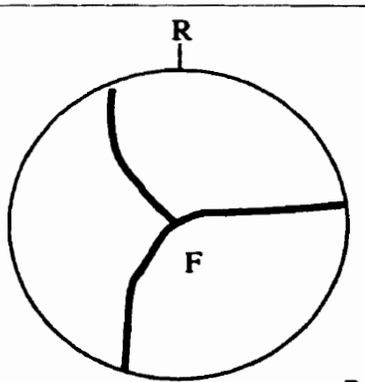



Downhole View
Fig. 2_153

### Sample F

| Structure | Relative Orientation |
|---|---|
| F | 080°/90° |
| F | 020°/90° |
| F | 160°/90° |

Comments




Downhole View

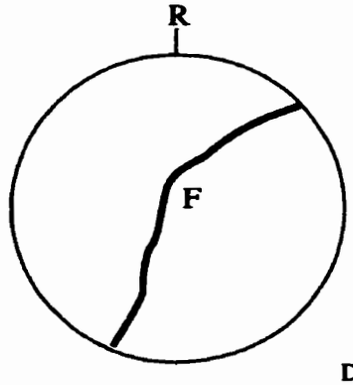
## Fine Grained Granite
**Borehole 421-032-RT3**

## Test Type (Point Load)
**Core Size: NQ-3, 45 mm**

### Sample G

| Structure | Relative Orientation |
|---|---|
| F | 045°/90° |
| F | 020°/90° |

Comments

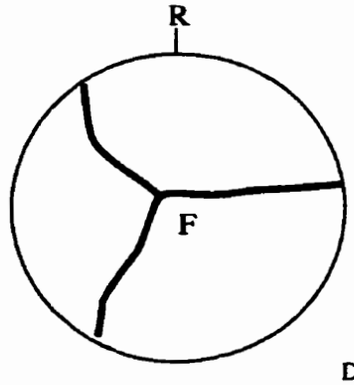



Downhole View

### Sample H

| Structure | Relative Orientation |
|---|---|
| F | 080°/90° |
| F | 045°/90° |
| F | 135°/90° |

Comments

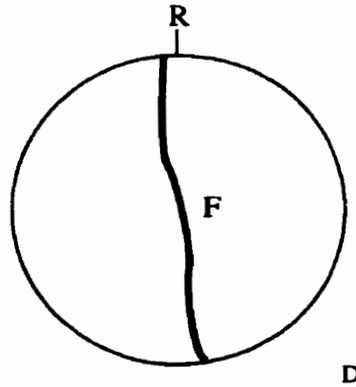



Downhole View

### Sample I

| Structure | Relative Orientation |
|---|---|
| F | 175°/90° |

Comments




Downhole View
Fig. 2_154

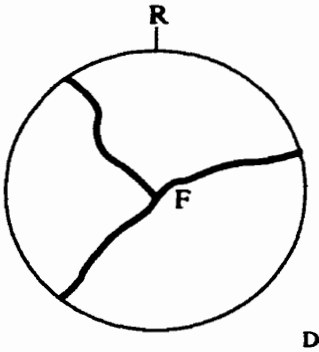
## Fine Grained Granite
**Borehole 421-032-RT3**

## Test Type (Point Load)
**Core Size: NQ-3, 45 mm**

**Sample J**

| Structure | Relative Orientation |
|---|---|
| F | 045°/90° |
| F | 070°/90° |
| F | 145°/90° |

**Comments**




Downhole View
Fig. 2_155

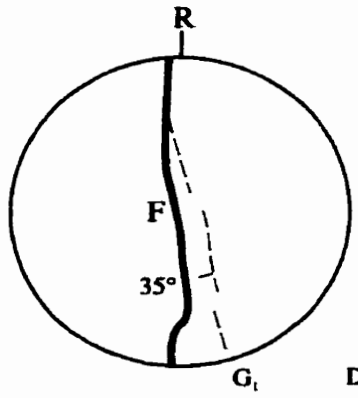
## Fine Grained Granite
Borehole 421-034-RT4

## Test Type (Point Load)
Core Size: NQ-3, 45 mm

### Sample A

| Structure | Relative Orientation |
|---|---|
| F | 180°/90° |
| $G_t$ | 170°/35°W |

Comments

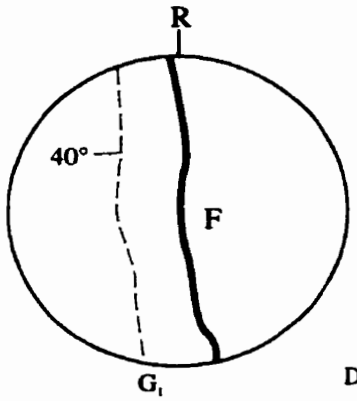



Downhole View
Fig. 2_156

### Sample B

| Structure | Relative Orientation |
|---|---|
| F | 180°/90° |
| $G_t$ | 180°/40°W |

Comments

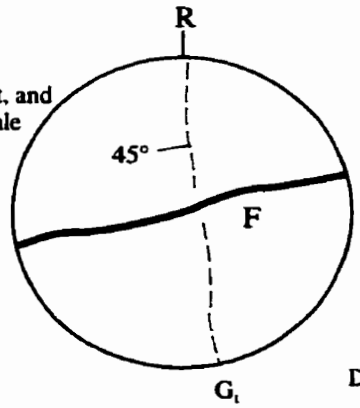



Downhole View
Fig. 2_157

### Sample C

| Structure | Relative Orientation | |
|---|---|---|
| F | 075°/90° | |
| $G_t$ | 175°/45°W | Gneissosity much less distinct, and heterogeneous even at this scale |

Comments




Downhole View
Fig. 2_158

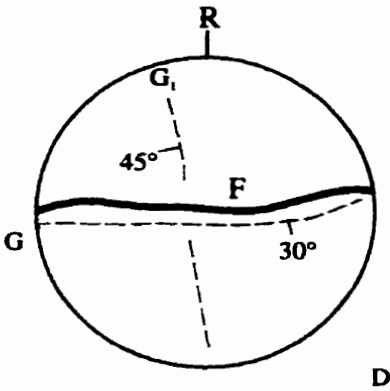
## Fine Grained Granite
**Borehole 421-034-RT4**

## Test Type (Point Load)
**Core Size: NQ-3, 45 mm**

### Sample D

| Structure | Relative Orientation | |
|---|---|---|
| F | 085°/ 90° | |
| $G_t$ | ~170°/45°W | Very weak & ill defined, |
| G | ~090°/30°S | not penetrative |

Comments

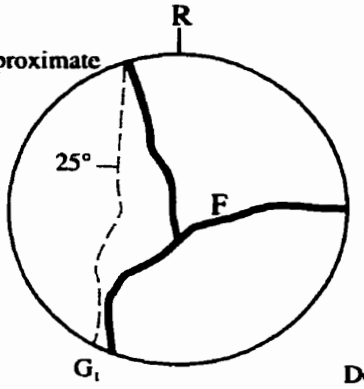



Downhole View
Fig. 2_159

### Sample E

| Structure | Relative Orientation | |
|---|---|---|
| F | ~010°/ 90° | Triple fracture, orientation approximate |
| F | ~090°/ 90° | |
| $G_t$ | ~010°/25°W | Very weak & only partially penetrative |

Comments

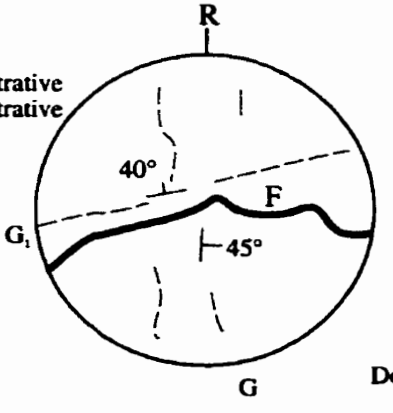



Downhole View
Fig. 2_160

### Sample F

| Structure | Relative Orientation | |
|---|---|---|
| F | 070°-090°/90° | |
| $G_t$ | 075°/40°N | Very ill defined & non-penetrative |
| G | 000°/45°E | Very ill defined & non-penetrative |

Comments




Downhole View
Fig. 2_161

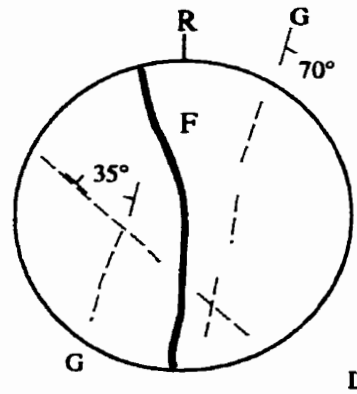
## Fine Grained Granite
**Borehole 421-034-RT4**

## Test Type (Point Load)
**Core Size: NQ-3, 45 mm**

### Sample G

| Structure | Relative Orientation |
|---|---|
| F | 175°/90° |
| G | 020°/70°SE |
| G | 135°/35°NW |

Comments

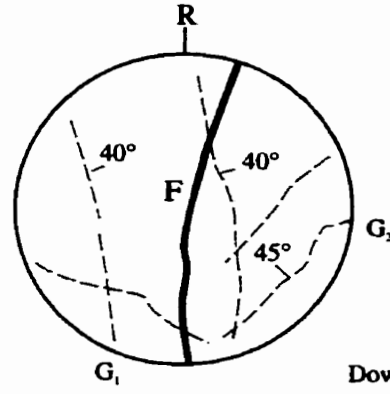



Downhole View
Fig. 2_ 162

### Sample H

| Structure | Relative Orientation |
|---|---|
| F | 015°/90° |
| $G_1$ | 145°/40°NE |
| $G_2$ | 035°/45°NW |

Comments



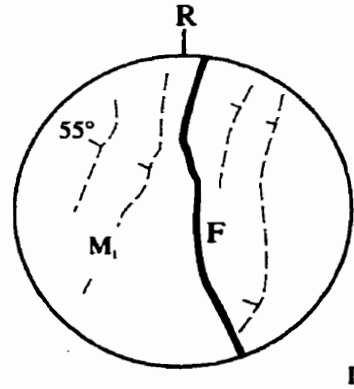

Downhole View
Fig. 2_ 163

### Sample I

| Structure | Relative Orientation |
|---|---|
| F | 175°/90° |
| $M_1$ | 010°/55°W |

Comments

Fracture partially follows foliation



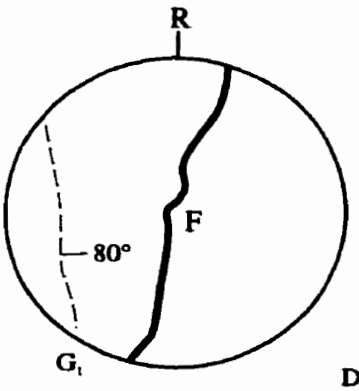

Downhole View
Fig. 2_ 163

## Fine Grained Granite
Borehole 421-034-RT4

## Test Type (Point Load)
Core Size: NQ-3, 45 mm

### Sample J

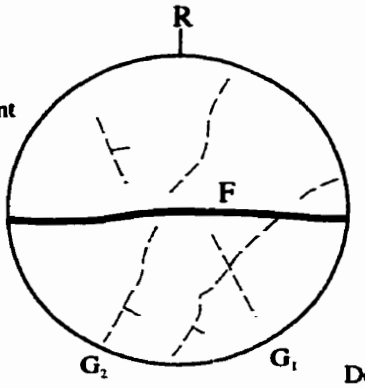
| Structure | Relative Orientation | |
|---|---|---|
| F | 025°/90° | Numerous en echelon steps parallel to G |
| $G_1$ | 170°/80°NE | Sample appears otherwise homogeneous, possible highly irregular foliation approximately normal to core axis |
| Comments | | |




Downhole View
Fig. 2_165

### Sample K

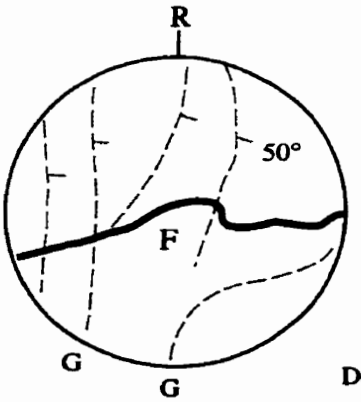
| Structure | Relative Orientation | |
|---|---|---|
| F | 090°/ 90° | |
| $G_1$ | 160°/90°? | Well defined mineral alignment (aggregate alignment in cross section but ill defined around perimeter |
| $G_2$ | 190°/70°SE | Non-penetrative foliation, no offset but appears to cross-cut |
| Comments | | |




Downhole View
Fig. 2_ 166

### Sample L

| Structure | Relative Orientation | |
|---|---|---|
| F | 075°/ 90° | Very irregular, only partially |
| G | 00°-90°SE | Numerous sub-domains marked by grain alignments |




Downhole View

## Fine Grained Granite
Borehole 421-034-RT4

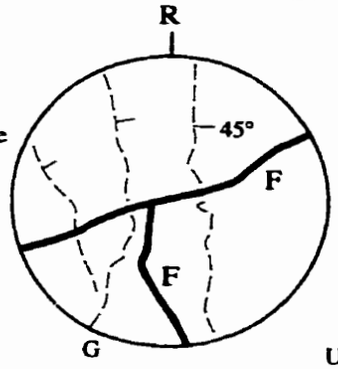
## Test Type (Point Load)
Core Size: NQ-3, 45 mm

Fig. 2_167

Sample M

| Structure | Relative Orientation | |
|---|---|---|
| F | 070°/90° | |
| F | 175°/90° | |
| G | 180°/45°E | Non-penetrative domains with varying grain & grain aggregate orientations. |

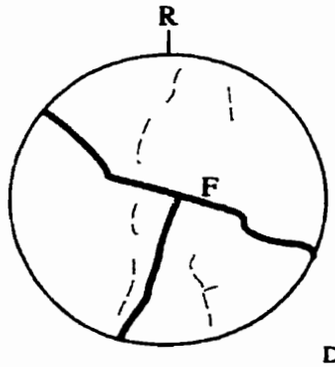
Comments




Uphole View

| Structure | Relative Orientation |
|---|---|
| F | 110°/90° |
| F | 200°/90° |
| F | 130°/90° |
| G | 005°/45° |

Comments




Downhole View

# Appendix B

# BRAZILIAN TESTING

This appendix presents the samples used for and the data obtained from the Brazilian testing. The tables list the specimen characteristics, and the results of the Brazilian tests described in this report. The figures summarize the fabric and test-induced fractures observed in each of the samples used in the Brazilian tests. The symbols used in the sample maps are explained in Figure B1, while the standard format for the sample maps is shown in Figure B2.

**Table B1**
**Brazilian test results**
**Fine Grained Granite from boreholes 421-032-RT1, 2, 3, 4.**

| Borehole | Sample | Density | Loading angle relative to low-dipping compositional layering | Minimum Tensile Strength (Crack Initiation) (MPa) | Maximum Tensile Strength (Failure) (MPa) |
|---|---|---|---|---|---|
| RT1 | D | 2.64 | 90 | | 3.01 |
| | A | 2.59 | 57 | | 3.26 |
| | B | 2.64 | 70 | | 2.26 |
| | C | 2.65 | 70 | | 2.76 |
| RT2 | A | 2.66 | 30 | | 2.50 |
| | B | 2.61 | 65 | | 3.00 |
| | C | 2.67 | 75 | | 1.91 |
| | D | 2.66 | 80 | 1.34 | 2.79 |
| | F | 2.65 | 70 | | 2.21 |
| | E | 2.65 | 45 | | 3.12 |
| RT3 | A | 2.64 | 0 | | |
| | B | 2.64 | 0 | | 2.12 |
| | C | 2.64 | 52 | | 2.38 |
| | D | 2.62 | 25 | | 1.65 |
| | E | 2.65 | 25 | | 2.86 |
| RT4 | A | 2.62 | 45 | | 2.82 |
| | B | 2.65 | 70 | | 2.66 |
| | C | 2.63 | 45 | | 2.14 |
| | D | 2.62 | 20 | | 1.98 |
| | E | 2.64 | 80 | | 1.84 |
| | F | 2.63 | 80 | | 2.68 |

Table B2
Brazilian Test Data
Medium Grained Granite (Borehole 417-071-RT4,5)

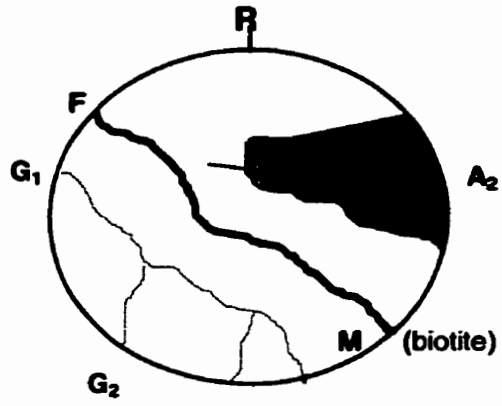
| Borehole | Sample | Density | Loading angle relative to low-dipping compositional layering | Minimum Tensile Strength (Crack Initiation) (MPa) | Maximum Tensile Strength (Failure) (MPa) |
|---|---|---|---|---|---|
| RT1, 2,3 | - not suitable due to core discing - | | | | |
| RT4 | A | 2.62 | 90 | | 1.49 |
| RT4 | B | 2.60 | 90 | | 1.79 |
| RT4 | C | 2.60 | 70 | 0.91 | 1.14 |
| RT4 | D | 2.63 | 45 | | 1.33 |
| RT4 | A' | 2.62 | 90 | 1.03 | 1.22 |
| RT5 | A | 2.63 | 90 | | 1.32 |
| RT5 | B | 2.61 | 30 | | 1.09 |
| RT5 | C | 2.62 | 0 | 0.87 | 1.07 |
| RT5 | D | 2.63 | 60 | | 0.87 |

Table B3
Brazilian Test Data.
Coarse Grained Granite (Borehole 418-051-RT2,3,4,5)

| Borehole | Sample | Density | Loading angle relative to low-dipping compositional layering | Minimum Tensile Strength (Crack Initiation) (MPa) | Maximum Tensile Strength (Failure) (MPa) |
|---|---|---|---|---|---|
| RT1 | - not suitable due to core discing - | | | | |
| RT2 | A | 2.67 | 90 | | 1.70 |
| | B | 2.70 | 75 | | 1.75 |
| | C | 2.68 | 90 | | 2.07 |
| | D | 2.67 | 0 | | 1.49 |
| | E | 2.66 | 45 | | 1.69 |
| | F | 2.68 | 0 | | 1.11 |
| | G | 2.63 | 10 | | 1.17 |
| | H | 2.67 | 0 | 0.60 | 1.39 |
| RT3 | B | 2.70 | 45 | | 1.45 |
| | C | 2.69 | 75 | 0.67 | 2.01 |
| | D | 2.67 | 90 | | 2.53 |
| | E | 2.70 | 85 | | 2.49 |
| | F | 2.71 | 0 | | 1.88 |
| | A' | 2.68 | 55 | | 1.56 |
| | G | 2.65 | 60 | 0.60 | 1.65 |
| RT4 | A | 2.61 | 50 | | 0.96 |
| | B | 2.59 | 45 | | 1.21 |
| | C | 2.60 | 45 | | 1.04 |
| | D | 2.56 | 15 | | 1.33 |
| | E | 2.60 | 20 | | 1.53 |
| | F | 2.60 | 38 | 0.62 | 0.88 |
| | G | 2.61 | 52 | 0.54 | 1.42 |
| RT5 | A | 2.62 | 0 | | 0.93 |
| | B | 2.54 | 25 | 0.52 | 1.18 |
| | C | 2.65 | 45 | | 1.52 |
| | D | 2.64 | 85 | | 1.65 |

**Sample A**

| Structure | Relative Orientation *(strike / dip & dip direction)* |
|---|---|
| F | 100°/90° NE |
| $G_1$ | 100°/20° NE |
| $G_2$ | 190°/60° NW |
| $M_1$ | 190°/60° NW |
| A | 100°/30° E |
| $^{30}$\| | strike & dip with dip direction |




**where:**

| | |
|---|---|
| R | core reference line |
| F | fracture induced by loading |
| $G_1$ | gneissic foliation type 1 |
| $G_2$ | gneissic foliation type 2 |
| $M_1$ | mineral alignment, type indicated on figure |
| $A_1$ | axial plane of fold |
| $A_2$ | axis of fold |
| | biotite > 10% or some compositional heterogeneity as indicated by text. |

**Note:** All orientations shown in the drawings are relative to the section shown, with the core reference line representing an interim North. These have been converted to true orientations in the report.

$^{30}$\|

Downhole View
*Dwg. No*

Figure B1. Key to symbols used on the sample maps.

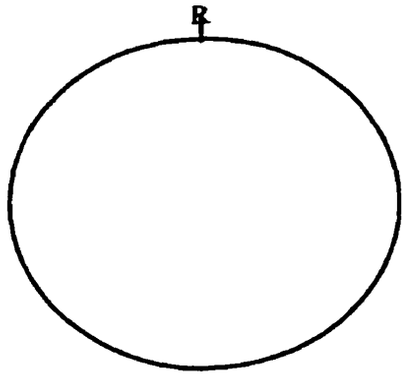
**Rock name** **Test type (Brazilian or Point load)**

**Borehole name (e.g. 401-123-rt2)** **Core size (alphanumeric and in mm).**

Sample a

| Structure | Relative Orientation |
|---|---|
| F | strike (0-360) / dip (0-90) |
| $G_1$ | strike (0-360) / dip (0-90) |
| $G_2$ | strike (0-360) / dip (0-90) |

Comments
Brief description of rock including anything noteworthy about the foliations present and the affect (if any) this had on testing.

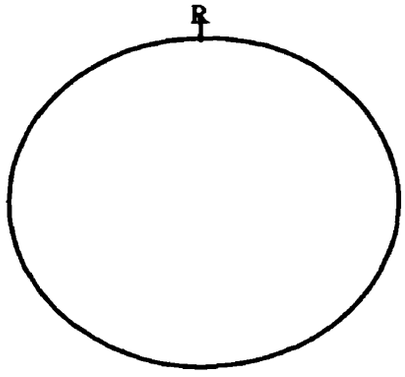



Downhole View
*Dwg. No.*

Sample b

| Structure | Relative Orientation |
|---|---|
| F | strike (0-360) / dip (0-90) |
| $G_1$ | strike (0-360) / dip (0-90) |
| $G_2$ | strike (0-360) / dip (0-90) |

Comments
Brief description of rock including anything noteworthy about the foliations present and the affect (if any) this had on testing.

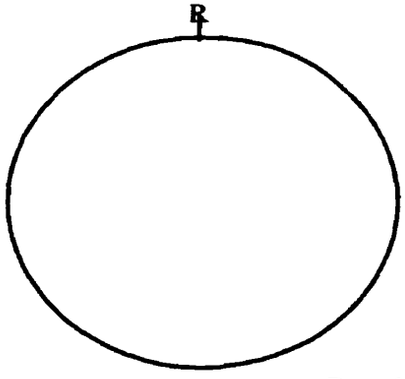



Downhole View
*Dwg. No.*

Sample c

| Structure | Relative Orientation |
|---|---|
| F | strike (0-360) / dip (0-90) |
| $G_1$ | strike (0-360) / dip (0-90) |
| $G_2$ | strike (0-360) / dip (0-90) |

Comments
Brief description of rock including anything noteworthy about the foliations present and the affect (if any) this had on testing.




Downhole View
*Dwg. No.*

Figure B2: Standard format for the fabric maps constructed for each of the samples selected for testing. The rock name, the borehole from which the sample came, the size of the borehole, and the test type are indicated at the top of each sheet. Up to three sample maps are shown per page, at their actual size. The reference line marked on the top of the core is shown, and all views are downhole unless otherwise indicated.

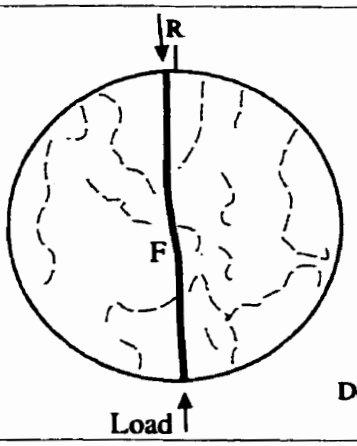
## Medium Grained Granite
**Borehole 417-071-RT4**

## Test Type (Brazilian)
**Core Size: NQ-3, 45 mm**

### Sample A'

| Structure | Relative Orientation | |
|---|---|---|
| L | 000°/26° | Aligned mineral aggregates. Probably in the other samples as well. |
| F | 180°/90° | |

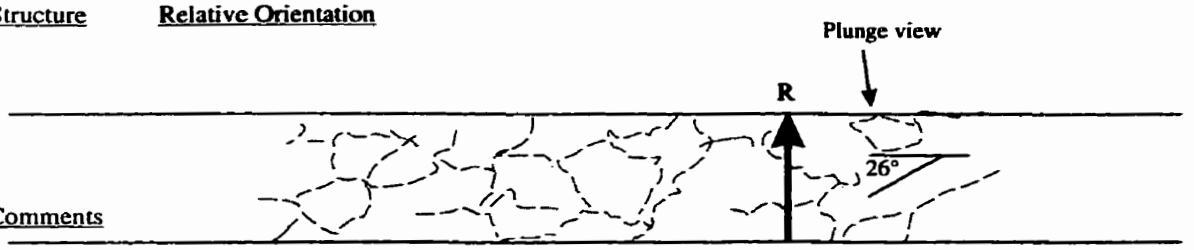
Comments




Downhole View
Fig. 2_066

| Structure | Relative Orientation |
|---|---|

Comments

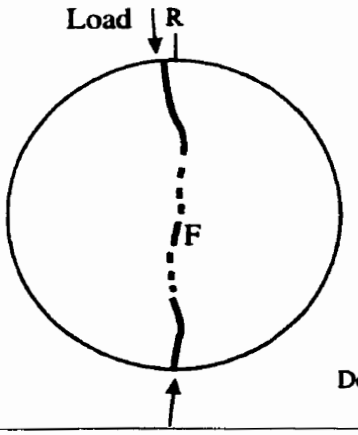



Downhole View

### Sample B

| Structure | Relative Orientation |
|---|---|
| F | 170°/90° |

Comments




Downhole View
Fig. 2_069

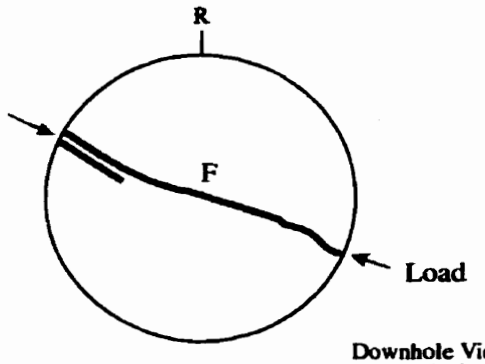
## Medium Grained Granite
Borehole 417-071-RT4

## Test Type (Brazilian)
Core Size: NQ-3, 45 mm

### Sample C

| Structure | Relative Orientation |
|---|---|
| F | 105°/90° |

Comments

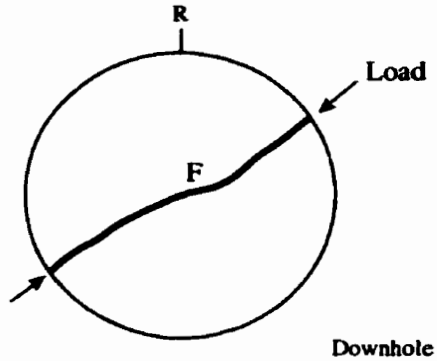



Downhole View
Fig. 2_070

### Sample D

| Structure | Relative Orientation |
|---|---|
| F | 065°/90° |

Comments




Downhole View
Fig. 2_067

## Medium Grained Granite
**Borehole 417-072-RT5**

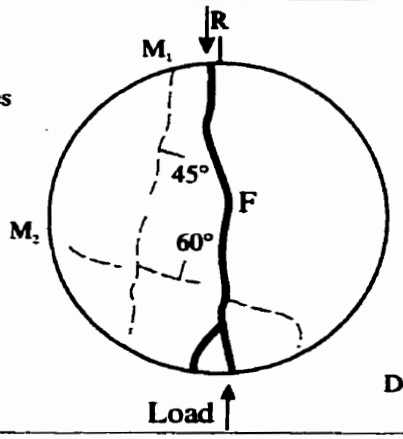
## Test Type (Brazilian)
**Core Size: NQ-3, 45 mm**

### Sample A

| Structure | Relative Orientation | |
|---|---|---|
| F | 000°/ 90° | |
| $M_1$ | 010°/45°E | Aligned mineral aggregates |
| $M_2$ | 110°/60°NE | |

Comments

Massive



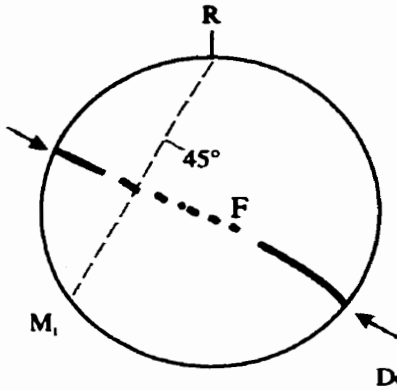

Downhole View
Fig. 2_062

### Sample B

| Structure | Relative Orientation |
|---|---|
| F | 120°/ 90° |
| $M_1$ | 025°/45°SE |

Comments

Nearly massive at this scale



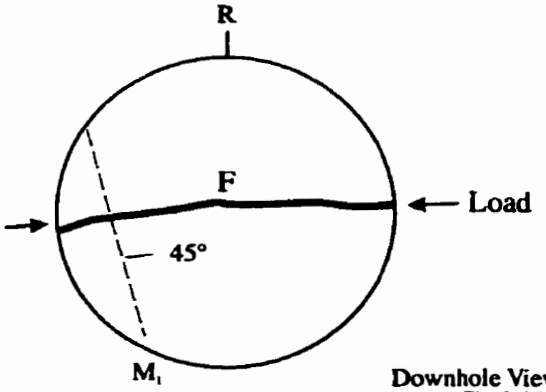

Downhole View
Fig. 2_063

### Sample C

| Structure | Relative Orientation |
|---|---|
| F | 088°/90° |
| $M_1$ | 000°/45°NE |

Comments

Nearly massive at this scale




Downhole View
Fig. 2_064

## Medium Grained Granite
Borehole 417-072-RT5

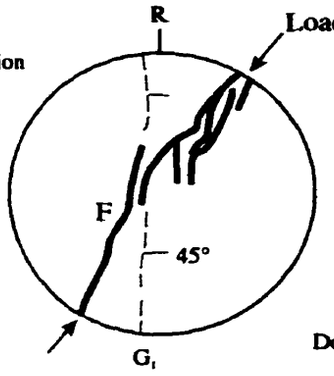
## Test Type (Brazilian)
Core Size: NQ-3, 45 mm

**Sample D**

| Structure | Relative Orientation |
|---|---|
| F | 035°/90° Many splays deviated by foliation |
| $G_t$ | 000°/45°E |

Comments

Nearly massive at this scale




Downhole View
Fig. 2_065

## Coarse Grained Granite
Borehole 418-051-RT2

## Test Type (Brazilian)
Core Size: NQ-3, 45 mm

Fig. 2_053

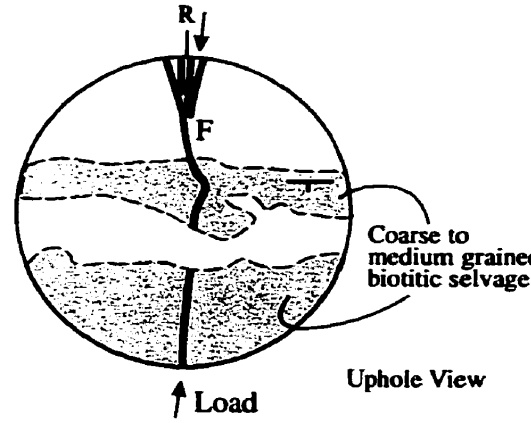
### Sample A

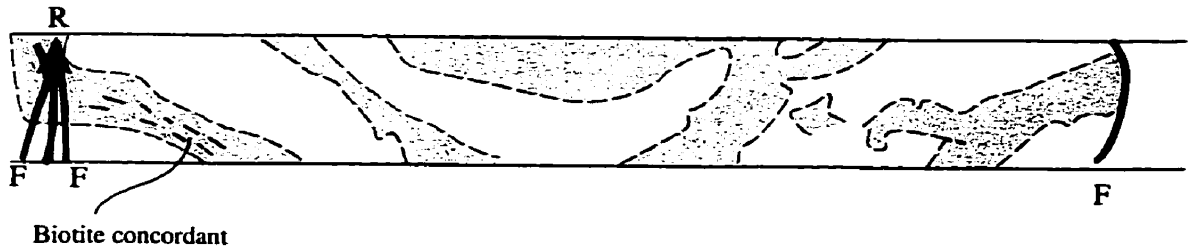
| Structure | Relative Orientation |
|---|---|
| F | 000°/90° |
| $G_1$ | 090°/33°S |
| $M_1$ | ~090°/35°S |

Comments

1. Note gap in fracture is in quartz/feldspar layer
2. 35° gneissosity and biotite

Loading angle to $G_1$ ~80°-90°




Perimeter View

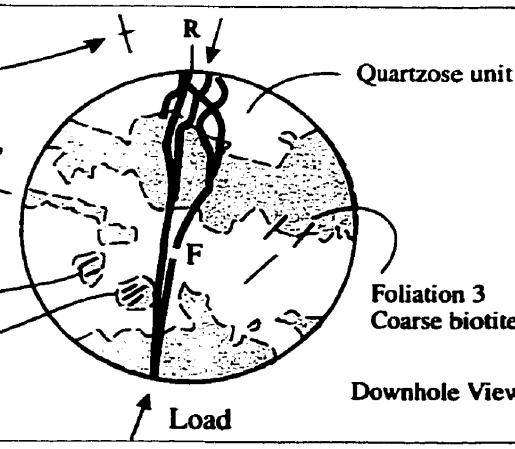
Structure Relative Orientation

Foliation 2: Coarse biotite in selvages guides fracture

35°

Comments

Vertical microcracks in quartz



## Coarse Grained Granite
Borehole 418-051-RT2

## Test Type (Brazilian)
Core Size: NQ-3, 45 mm

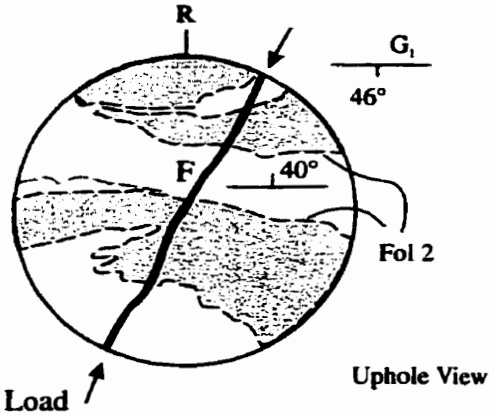
Fig. 2_054

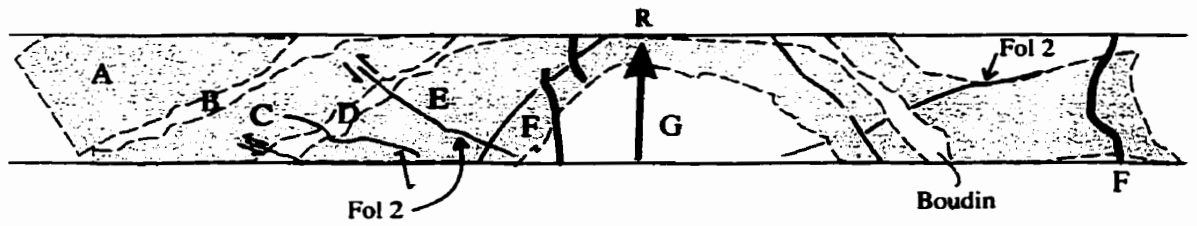
**Sample B**

Structure Relative Orientation

Comments (Perimeter View)

- Fine grained biotitic granite (A,C,E below)
- Leucocratic granite (B,D,G)
- Biotite selvage, (F)



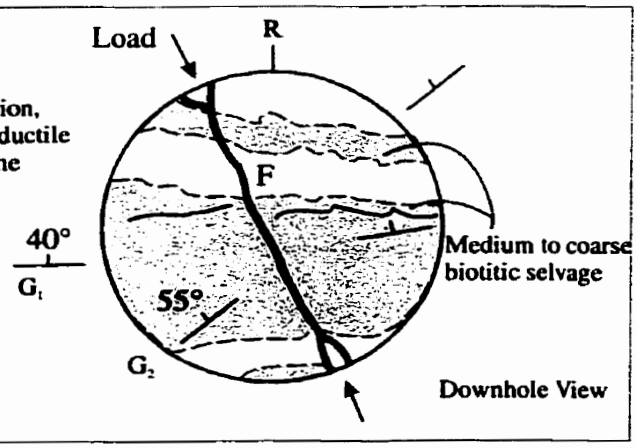




Perimeter View

| Structure | Relative Orientation | |
|---|---|---|
| $G_1$ | 180°/40°N | |
| $G_2$ | 180°/40°S | but opposite dip direction, resembles small semi-ductile faults, leaves leucosome boudined |
| F | 160°/90° | |

Comments

Fracture direction may coincide with intersection of the two foliations



## Coarse Grained Granite
**Borehole 418-051-RT2**

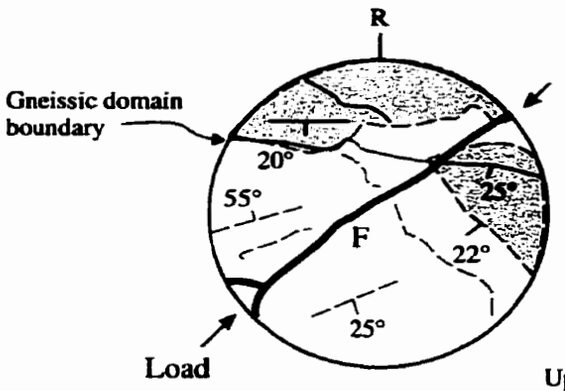
## Test Type (Brazilian)
**Core Size: NQ-3, 45 mm**

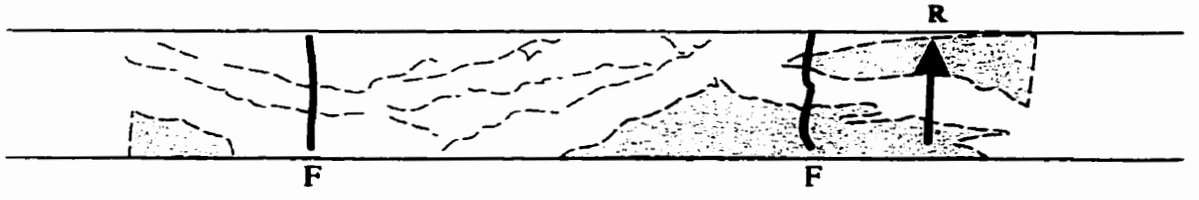
Fig. 2_055

**Sample C**

Structure | Relative Orientation

Comments



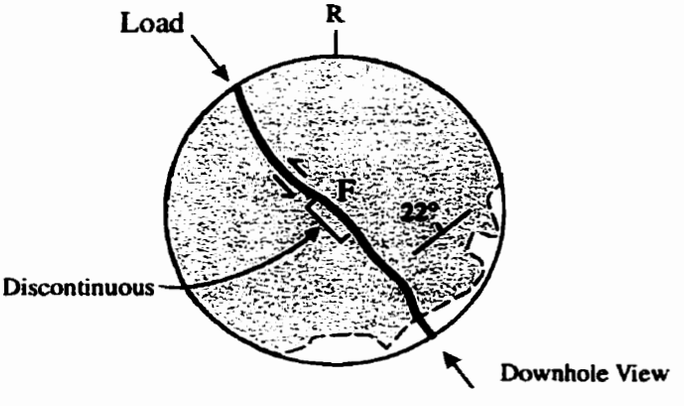




Perimeter View

| Structure | Relative Orientation |
|---|---|
| $G_t$ | 045°/25°N |
| F | 140°/90° |

Comments

Angles relative to loading



## Coarse Grained Granite
**Borehole 418-051-RT2**

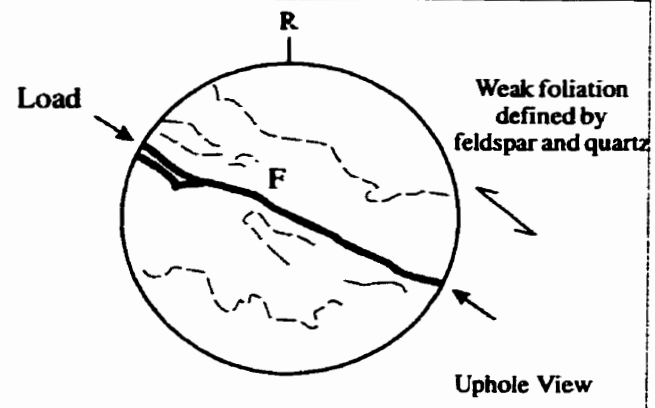
## Test Type (Brazilian)
**Core Size: NQ-3, 45 mm**

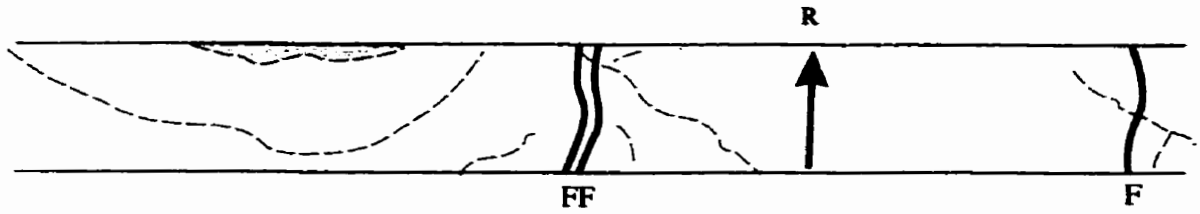
Fig. 2_056

**Sample D**

Structure Relative Orientation

Comments



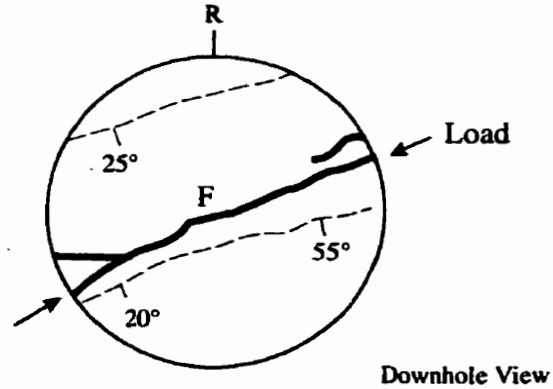




Perimeter View

| Structure | Relative Orientation |
|---|---|
| $G_1$ | 020° to 55° (1) and 025° (2) dips |

Comments

Angles relative to loading



## Coarse Grained Granite
Borehole 418-051-RT2

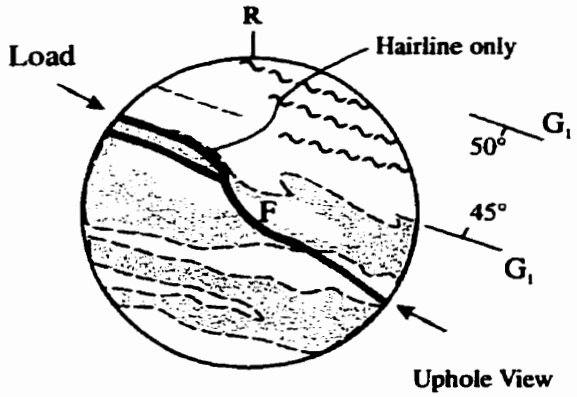
## Test Type (Brazilian)
Core Size: NQ-3, 45 mm

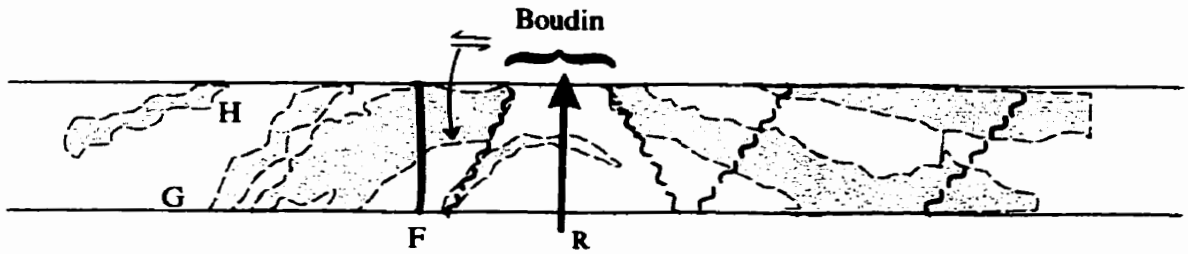
Fig. 2_057

**Sample F**

| Structure | Relative Orientation |
|---|---|
| F | 120°/ 90° |
| $G_1$ | 120°/45°NE |
| $G_2$ | 120°/50°SW |
| Load | 120°/00 |

Comments



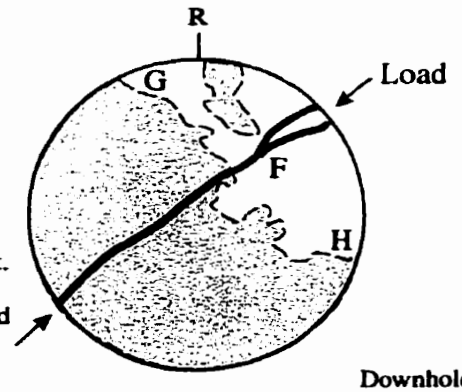




Perimeter View

Structure Relative Orientation

Comments

- Fracture partially follows banding and biotitic alignment.
- Biotite parallels compositional layering
- Cross-cutting spaced foliation cuts $G_1$ into poorly defined boudins foliation type.....(biotitic lenses)



## Coarse Grained Granite
**Borehole 418-051-RT2**

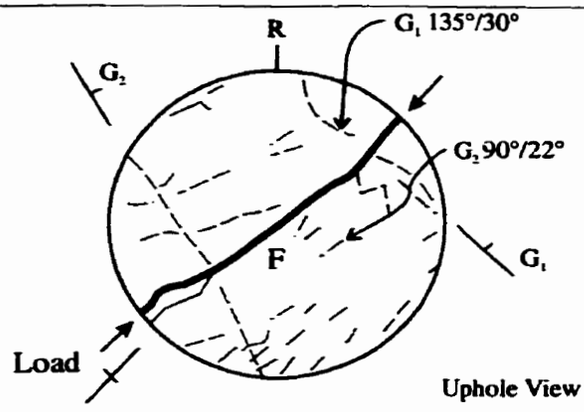
## Test Type (Brazilian)
**Core Size: NQ-3, 45 mm**

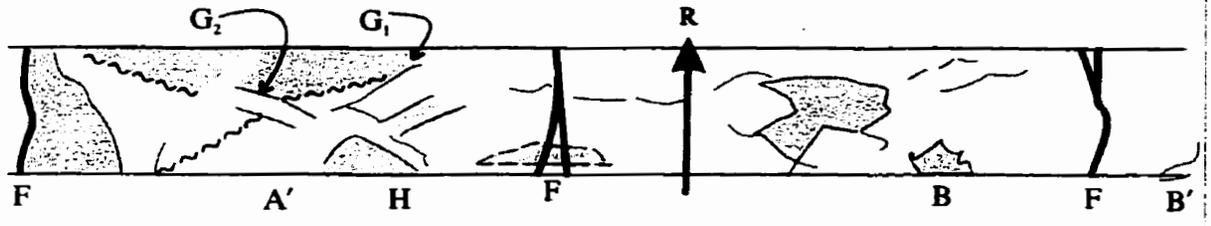
Fig. 2_058

**Sample G**

Structure | Relative Orientation

Comments

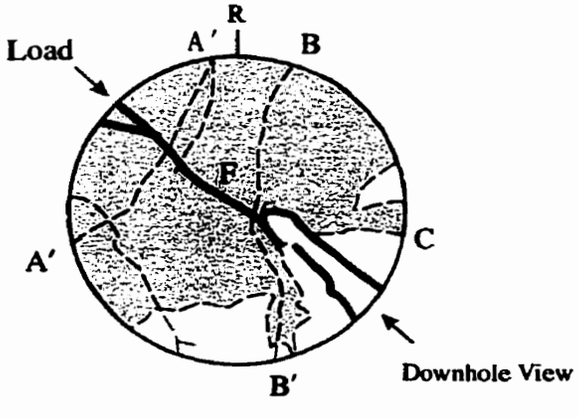



**Perimeter View**

| Structure | Relative Orientation |
|---|---|
| F | 130°/90° |
| $G_1$ | 120°/45°NE |
| $G_2$ | 120°/50°SW |
| Load | 130°/00° |

Comments



## Coarse Grained Granite
**Borehole 418-051-RT2**

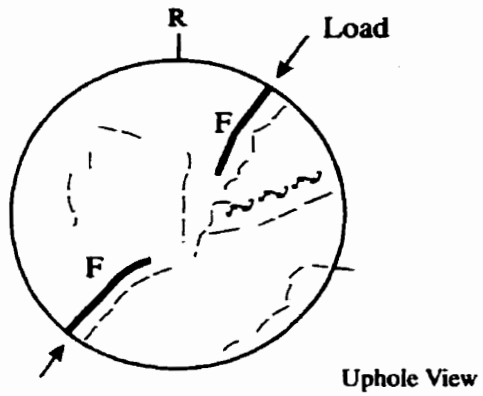
## Test Type (Brazilian)
**Core Size: NQ-3, 45 mm**

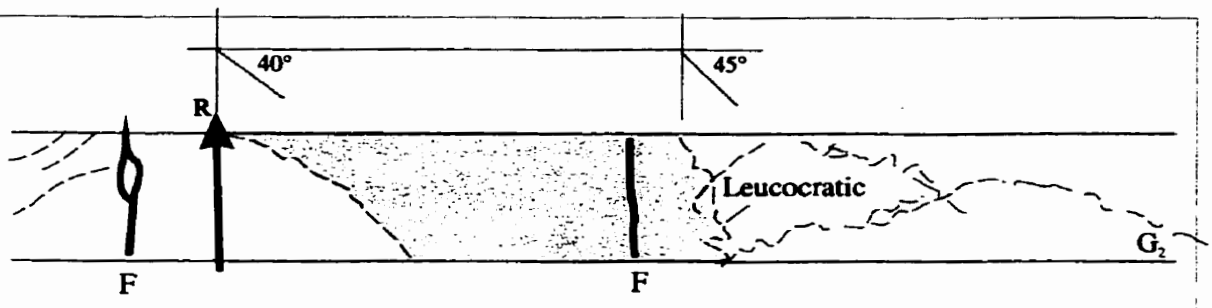
Fig. 2_059

**Sample H**

Structure | Relative Orientation

Comments



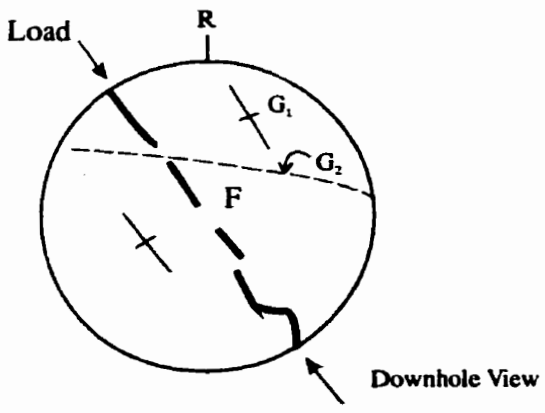




**Perimeter View**

| Structure | Relative Orientation |
|---|---|
| F | 145°/90° |
| $G_1$ | 135°/45°NE |
| $G_2$ | 100°/30°SW |
| Load | 145°/00 |

Comments

Homogenous textural domain from this view. Weakly developed foliations.



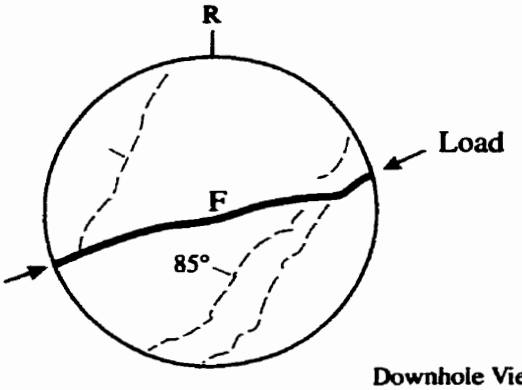
## Coarse Grained Granite
Borehole 418-051-RT3

## Test Type (Brazilian)
Core Size: NQ-3, 45 mm

### Sample A

| Structure | Relative Orientation |
|---|---|
| F | 080°/90° |
| $G_1$ | 035°/85°NW |

Comments



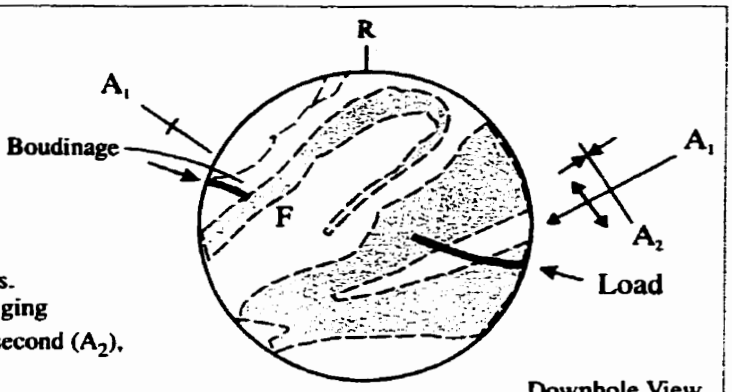
Downhole View
Fig. 2_047

### Sample B

| Structure | Relative Orientation |
|---|---|
| F | 100°/90° |
| $G_1$ | 050°/50°NW |
| | 090°/90°SE |
| $A_1$ | 070°/70°SW ? |
| $A_2$ | 135°/90° |

Comments

Sample shows at least two fold generations. One ($A_1$) tightly isoclinal and steeply plunging (in this relative orientation) and the later second ($A_2$), generation broad and open.



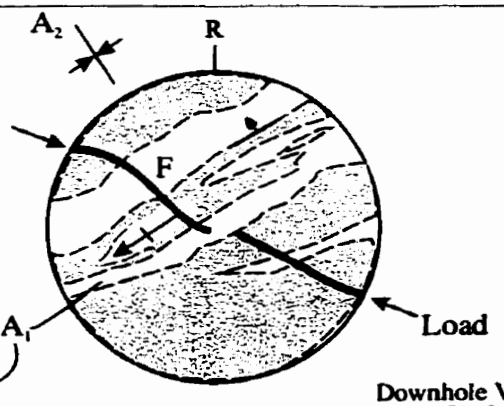
Downhole View
Fig. 2_048

### Sample C

| Structure | Relative Orientation | |
|---|---|---|
| F | 100°/90° | Fractures |
| $G_1$ | 045°/75°NW | Schlieric layering |
| $G_2$ | 065°/ 85°NW | Schlieric layering |

Comments

Xenolithic granite. Schlieric, fabric type 5



Downhole View
Fig. 2_049

## Coarse Grained Granite
**Borehole 418-051-RT3**

## Test Type (Brazilian)
**Core Size: NQ-3, 45 mm**

### Sample D

| Structure | Relative Orientation | |
|---|---|---|
| F | 130°/90° | |
| $G_1$ | 045°/55°NW | |
| $A_1$ | 045°/90° | Axial plane |
| $A_2$ | 130°/90° | Axial plane |

Comments

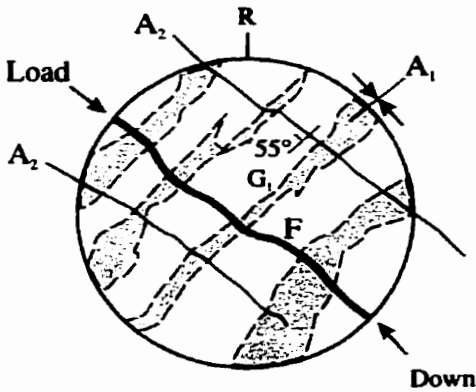


Fig. 2_050

### Sample E

| Structure | Relative Orientation |
|---|---|
| F | 150°/90° |
| $G_1$ | 045°/65°NW |

Comments

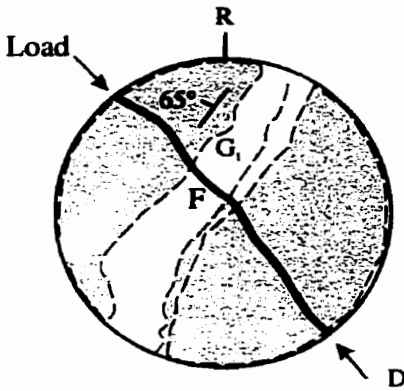


Fig. 2_051

### Sample F

| Structure | Relative Orientation |
|---|---|
| F | 040°/90° |
| $G_1$ | 045°/85°NW |

Comments

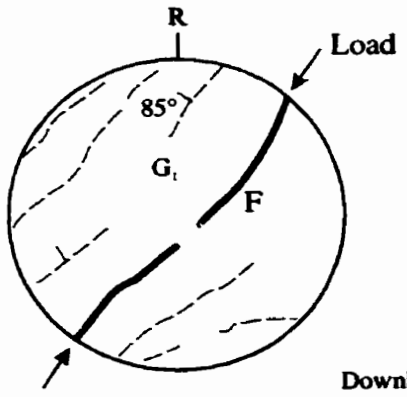


Fig. 2_052

## Coarse Grained Granite
Borehole 418-051-RT3

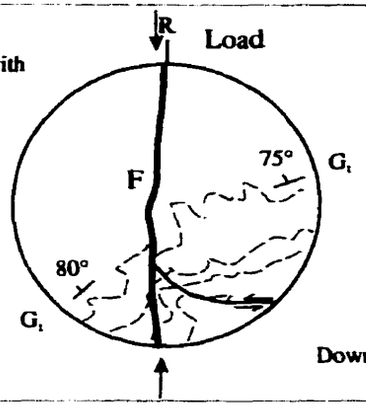
## Test Type (Brazilian)
Core Size: NQ-3, 45 mm

Sample G

| Structure | Relative Orientation | |
|---|---|---|
| F | 005°/90° | Folded on mm scale with crenulations and kinks |
| $G_1$ | 040° /80°N | |
| $G_t$ | 080° / 75°N | |
| A | 135°/?? | Axis |

Also folded normal to core axis

Comments




Downhole View
Fig. 2_060

## Coarse Grained Granite
**Borehole 418-051-RT4**

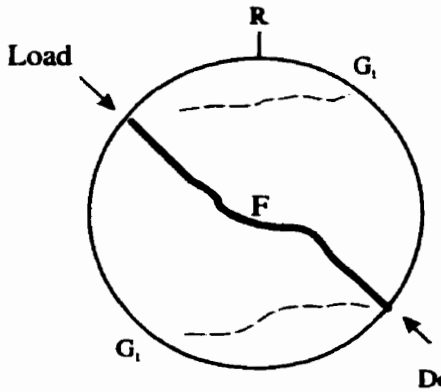
## Test Type (Brazilian)
**Core Size: NQ-3, 45 mm**

### Sample A

| Structure | Relative Orientation |
|---|---|
| F | 135°/90° |
| $G_1$ | 090°/90° |

Comments

Very coarse grained therefore boundaries somewhat irregular..

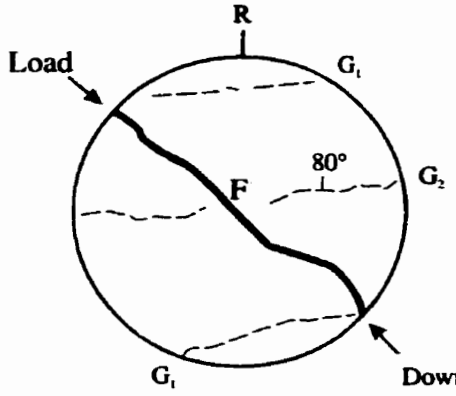



Downhole View
Fig. 2_038

### Sample B

| Structure | Relative Orientation | |
|---|---|---|
| F | 130°/90° | |
| $G_1$ | 080°/90° | |
| $G_2$ | 070°/80°N | Boudin Planes |

Comments

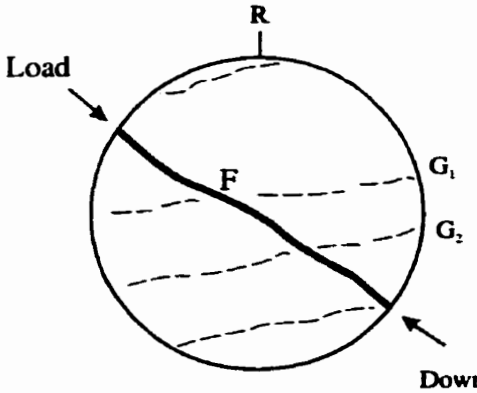



Downhole View
Fig. 2_039

### Sample C

| Structure | Relative Orientation | |
|---|---|---|
| F | 125°/90° | |
| $G_1$ | 085°/ 90° | Coarse Layering |
| $G_1$ | 085°/90° | Boudin Planes |
| $G_2$ | 075°/90° | Boudin Planes |

Comments




Downhole View
Fig. 2_040

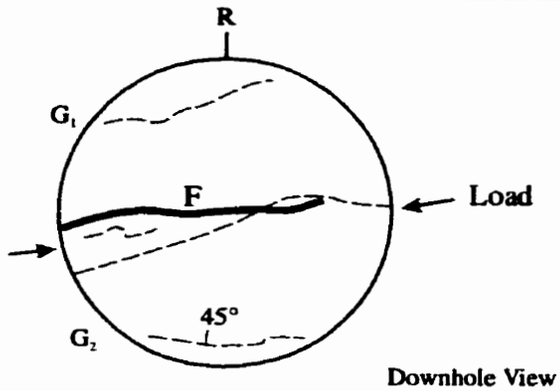
## Coarse Grained Granite
**Borehole 418-051-RT4**

## Test Type (Brazilian)
**Core Size: NQ-3, 45 mm**

### Sample D

| Structure | Relative Orientation |
|---|---|
| F | 080°/90° |
| $G_1$ | 090°/90° |
| $G_2$ | 080°/45°N |

Comments



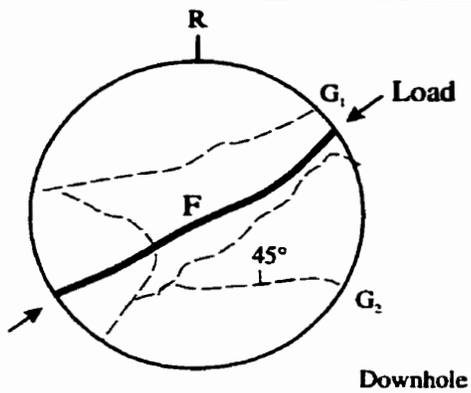

Downhole View
Fig. 2_041

### Sample E

| Structure | Relative Orientation |
|---|---|
| F | 055°/90° |
| $G_1$ | 070°/90° |
| $G_2$ | 050°/45°NW |

Comments

Boudin structure

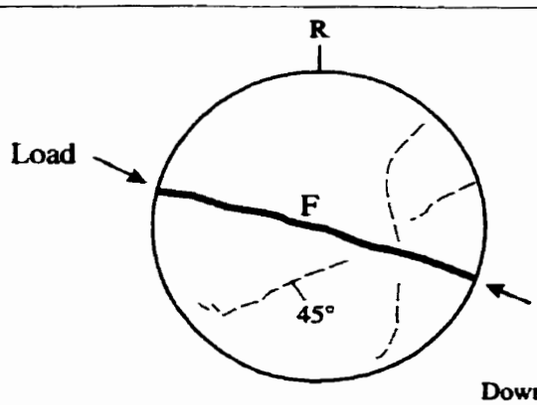



Downhole View
Fig. 2_042

### Sample F

| Structure | Relative Orientation |
|---|---|
| F | 105°/90° |
| $G_2$ | 060°/45°SE |

Comments




Downhole View
Fig. 2_037

## Coarse Grained Granite
Borehole 418-051-RT4

## Test Type (Brazilian)
Core Size: NQ-3, 45 mm

Sample G

| Structure | Relative Orientation |
|---|---|
| F | 130°/90° |
| $G_2$ | 075°/45°SE |

Comments

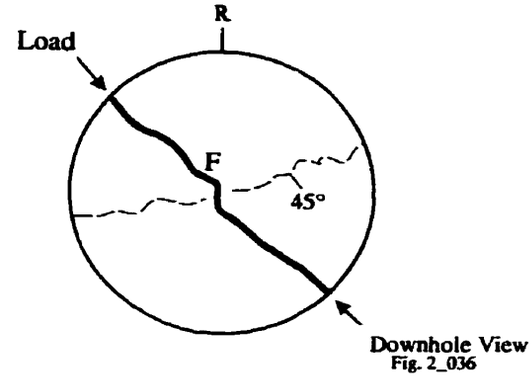


Fig. 2_036

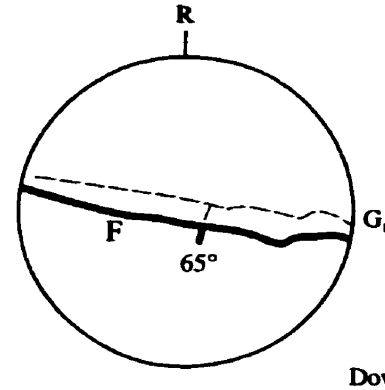
## Coarse Grained Granite
**Borehole 418-051-RT5**

## Test Type (Brazilian)
**Core Size: NQ-3, 45 mmm**

### Sample A

| Structure | Relative Orientation |
|---|---|
| F | 100°/90° |
| $G_1$ | 100°/65°S |

Comments

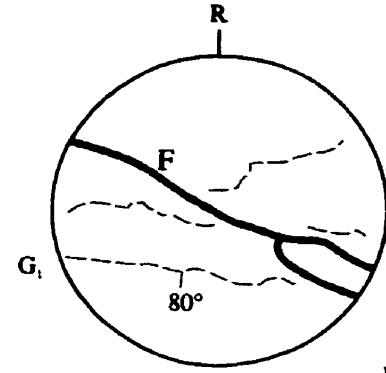



Downhole View
Fig. 2_043

### Sample B

| Structure | Relative Orientation |
|---|---|
| F | 120°/ 90° |
| $G_1$ | 090°/80°S |

Comments

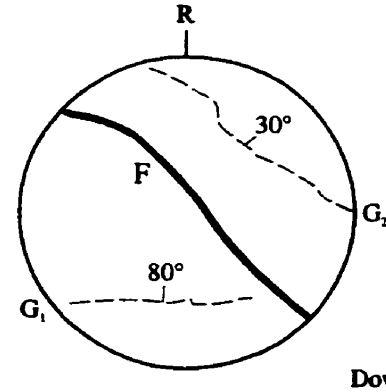



Downhole View
Fig. 2_044

### Sample C

| Structure | Relative Orientation |
|---|---|
| F | 135°/90° |
| $G_1$ | 090°/80°N |
| $G_2$ | 135°/30°NE |

Comments




Downhole View
Fig. 2_045

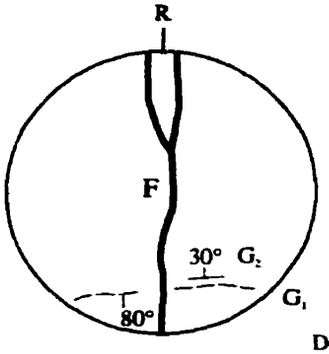
## Coarse Grained Granite
**Borehole 418-051-RT5**

## Test Type (Brazilian)
**Core Size: NQ-3, 45 mmm**

**Sample D**

| Structure | Relative Orientation |
|---|---|
| F | 000°/90° |
| $G_1$ | 085°/80°S |
| $G_2$ | 085°/30°N |

Comments




Downhole View
Fig. 2_046

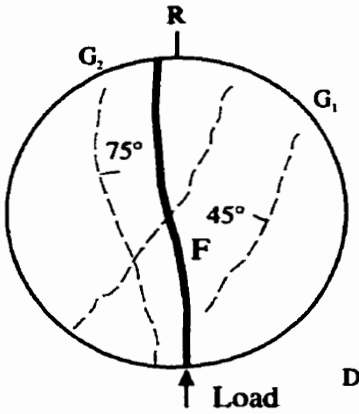
## Fine Grained Granite
Borehole 421-032-RT1

## Test Type (Brazilian)
Core Size: NQ-3, 45 mm

### Sample A

| Structure | Relative Orientation |
|---|---|
| F | 175°/90° |
| $G_1$ | 035°/45°W |
| $G_2$ | 175°/75°E |

Comments

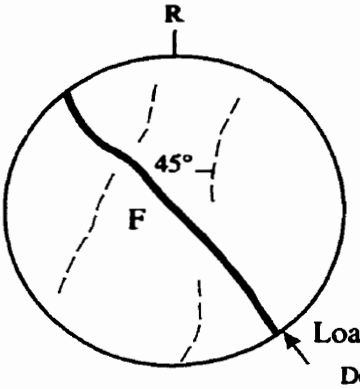



Downhole View
Fig. 2_005

### Sample B

| Structure | Relative Orientation |
|---|---|
| F | 135°/90° |
| G | 015°/45°W |

Comments

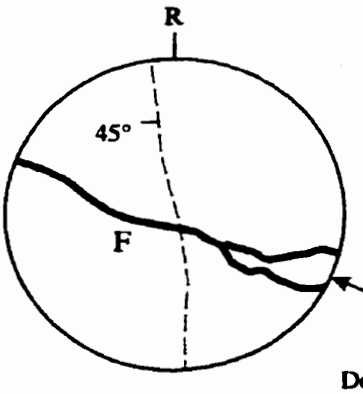



Downhole View
Fig. 2_006

### Sample C

| Structure | Relative Orientation |
|---|---|
| F | 115°/90° |
| $G_1$ | 175°/45°W |

Comments




Downhole View
Fig. 2_007

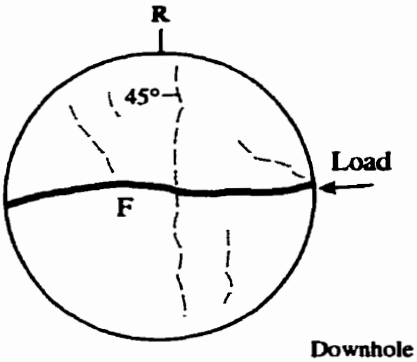
## Fine Grained Granite
**Borehole 421-032-RT1**

## Test Type (Brazilian)
**Core Size: NQ-3, 45 mm**

**Sample D**

| Structure | Relative Orientation |
|---|---|
| F | 090°/90° |
| G | 180°/45°W |

Comments




Downhole View
Fig. 2_008

## Fine Grained Granite
**Borehole 421-032-RT2**

## Test Type (Brazilian)
**Core Size: NQ-3, 45 mm**

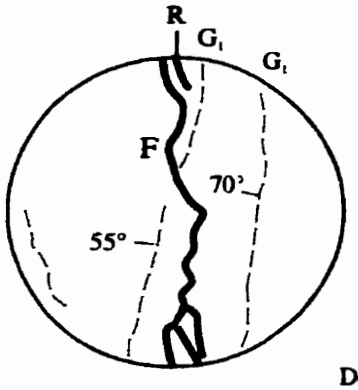
### Sample A

| Structure | Relative Orientation | |
|---|---|---|
| F | 000°/85°W<br>010°/85°W | } partially along banding |
| $G_1$ | 015°/70°W<br>015°/55°W | flow banding<br>mineral alignment within banding |

Comments

Loading nearly parallel to foliation

Failure at 7.0 MPa

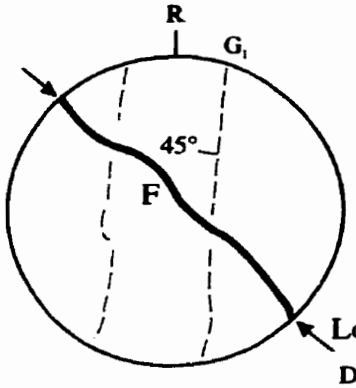



Downhole View
Fig. 2_014

### Sample B

| Structure | Relative Orientation | |
|---|---|---|
| F | 135°/90° | loading direction |
| $G_1$ | 010°/45°W | |

Comments

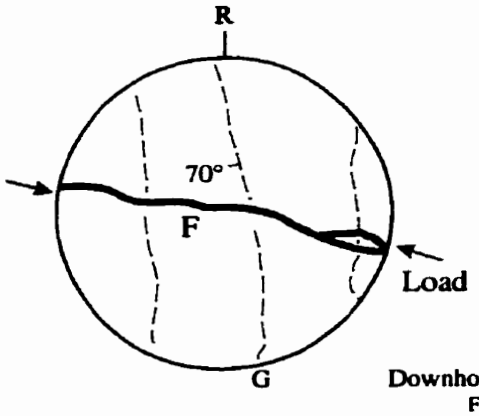



Downhole View
Fig. 2_016

### Sample C

| Structure | Relative Orientation |
|---|---|
| F | 105°/90° |
| G | 000°/70°W |

Comments




Downhole View
Fig. 2_017

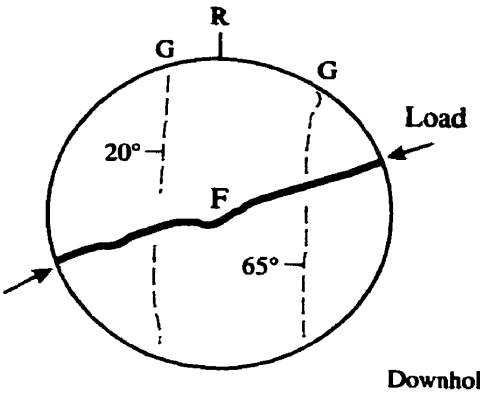
## Fine Grained Granite
**Borehole 421-032-RT2**

## Test Type (Brazilian)
**Core Size: NQ-3, 45 mm**

### Sample D

| Structure | Relative Orientation |
|---|---|
| F | 075°/90° |
| G | 005°/65°W |
| G | 005°/20°W |

Comments

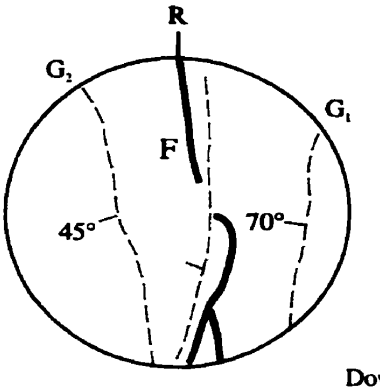



Downhole View
Fig. 2_018

### Sample E

| Structure | Relative Orientation | |
|---|---|---|
| F | 000°/70° | |
| G | 000°/70°W | banding |
| $G_1$ | 020°/70°W | banding |
| $G_2$ | 000°/45°W | mineral alignment |

Comments

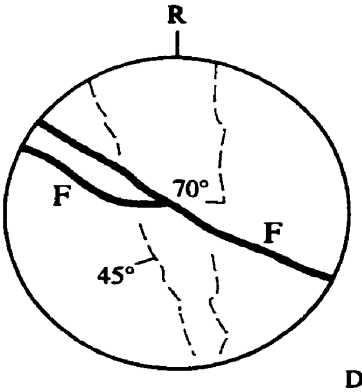



Downhole View
Fig. 2_015

### Sample F

| Structure | Relative Orientation | |
|---|---|---|
| F | 116°/90° | |
| $G_1$ | 175°/70°W | banding |
| $G_2$ | 165°/45°W | mineral alignment |

Comments




Downhole View
Fig. 2_013

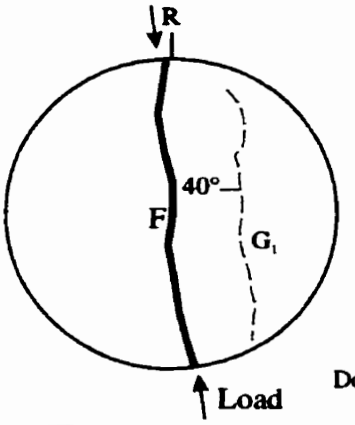
## Fine Grained Granite
Borehole 421-032-RT3

## Test Type (Brazilian)
Core Size: NQ-3, 45 mm

### Sample A

| Structure | Relative Orientation |
|---|---|
| F | 000°/90° |
| $G_1$ | 000°/40°W |

Comments

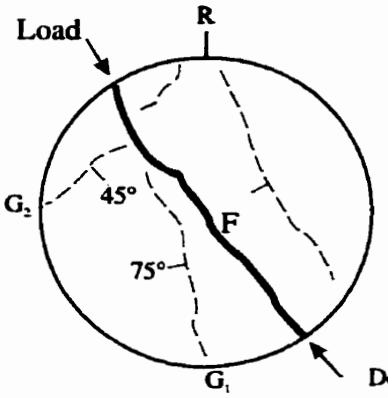



Downhole View
Fig. 2_024

### Sample B

| Structure | Relative Orientation |
|---|---|
| F | 150°/90° |
| $G_1$ | 150°/75°W |
| $G_2$ | 045°/45°SE |

Comments

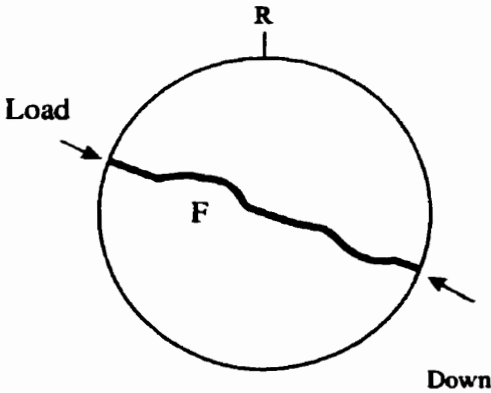



Downhole View
Fig. 2_025

### Sample C

| Structure | Relative Orientation |
|---|---|
| F | 120°/90° |

Comments




Downhole View
Fig. 2_026

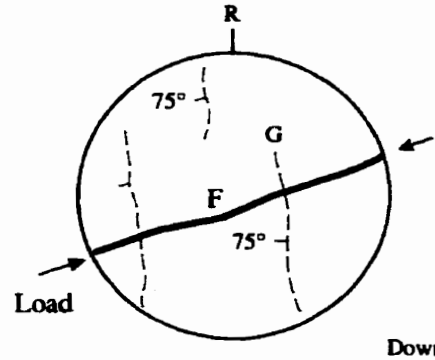
## Fine Grained Granite
Borehole 421-032-RT3

## Test Type (Brazilian)
Core Size: NQ-3, 45 mm

### Sample D

| Structure | Relative Orientation |
|---|---|
| F | 070°/90° |
| G | 175°/75°W |

Comments

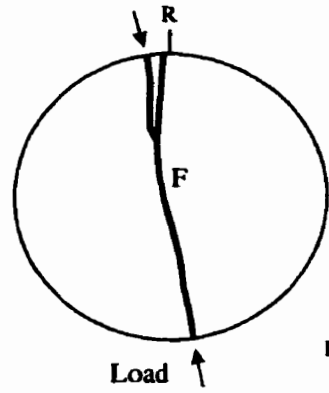



Downhole View
Fig. 2_027

### Sample E

| Structure | Relative Orientation |
|---|---|
| F | 175°/90° |

Comments




Downhole View
Fig. 2_023

## Fine Grained Granite
**Borehole 421-034-RT4**

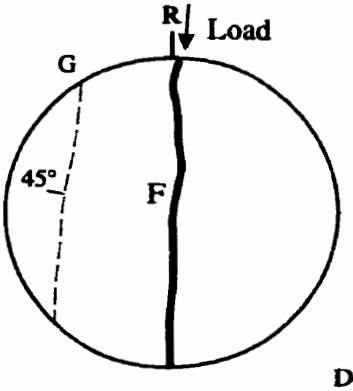
## Test Type (Brazilian)
**Core Size: NQ-3, 45 mm**

### Sample A

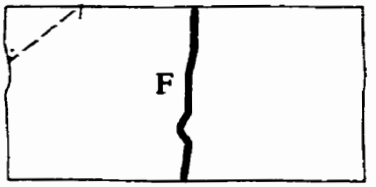
| Structure | Relative Orientation |
|---|---|
| F | 000°/90° |
| G | 010°/45°SW |

Comments

| | |
|---|---|
| Angle between strike of foliation and fracture | 05° |
| Dihedral angle between planes | 46° |




Downhole View



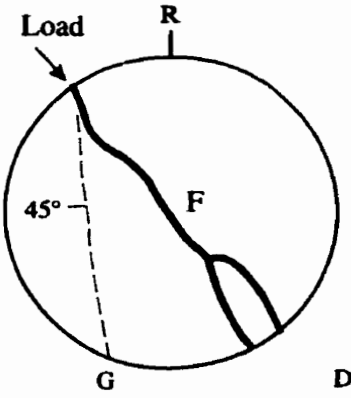

Section Normal to Fracture & Foliation

### Sample B

| Structure | Relative Orientation |
|---|---|
| F | 140°/90° |
| G | 170°/45°SW |

Comments

| | |
|---|---|
| Angle between strike of foliation and fracture | 30° |
| Dihedral angle between planes | 52° |




Downhole View
Fig. 2_034

## Fine Grained Granite
Borehole 421-034-RT4

## Test Type (Brazilian)
Core Size: NQ-3, 45 mm

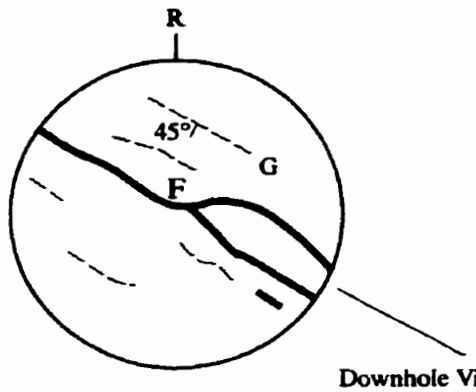
### Sample C

| Structure | Relative Orientation |
|---|---|
| F | 120°/90° |
| G | 170°/45°SW |
| Lineation | 120°/00° |

Comments

| | |
|---|---|
| Angle between strike of foliation and fracture | 50° |
| Dihedral angle between planes | 62° |

This sample is virtually massive although a non-penetrative lineation is observed. The foliation orientation is from Sample B.



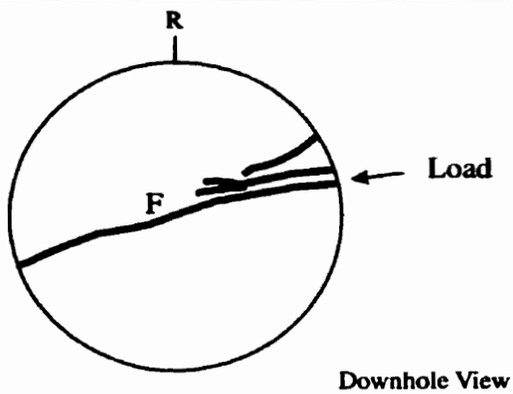

Downhole View
Fig. 2_035

### Sample D

| Structure | Relative Orientation |
|---|---|
| F | 070°/90° |
| G | 170°/45°SW |

Comments

| | |
|---|---|
| Angle between strike of foliation and fracture | ~80° |
| Dihedral angle between planes | 83° |



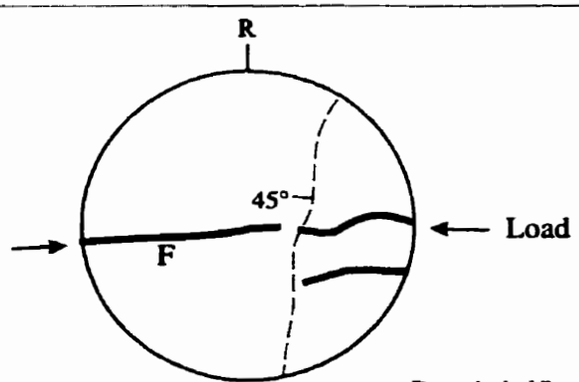

Downhole View
Fig. 2_035

### Sample E

| Structure | Relative Orientation |
|---|---|
| F | 085°/90° |
| G | 010°/45°W |

Comments

| | |
|---|---|
| Angle between strike of foliation and fracture | 75° |
| Dihedral angle between planes | 80° |




Downhole View
Fig. 2_033

## Fine Grained Granite
Borehole 421-034-RT4

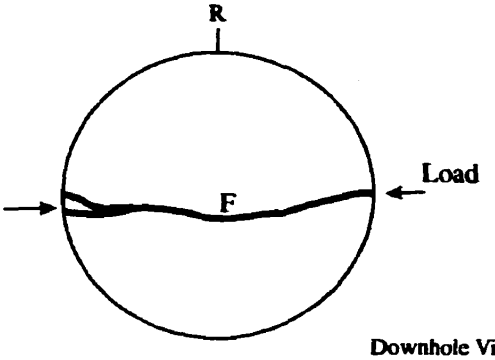
## Test Type (Brazilian)
Core Size: NQ-3, 45 mm

**Sample F**

| Structure | Relative Orientation |
|---|---|
| F | 085°/90° |
| G | 010°/45°W |

Comments

| | |
|---|---|
| Angle between strike of foliation and fracture | 75° |
| Dihedral angle between planes | 80° |




Downhole View
Fig. 2_033

# Appendix C

# UNIAXIAL COMPRESSION TESTS

This appendix presents the data obtained from the uniaxial compression testing conducted to define fabric-related anisotropy within the granite of the Lac Du Bonnet Batholith. The stress strain graphs and the original loading / strain data files are included. The stress strain diagrams show the axial, lateral and volumetric strains under accumulating load, and the associated data tables, listing the loads and corresponding strains as recorded by data acquisition system for each gauge. Volumetric strain was computed by using the following equation:

$$\varepsilon_v = \varepsilon_a + 2\varepsilon_c$$

where $\varepsilon_v$ = volumetric strain, $\varepsilon_a$ = axial strain, and $\varepsilon c$ = circumferential strain.

The inflection points on the stress-strain curves were determined manually and are listed in the tables in Chapter 9 of this report.

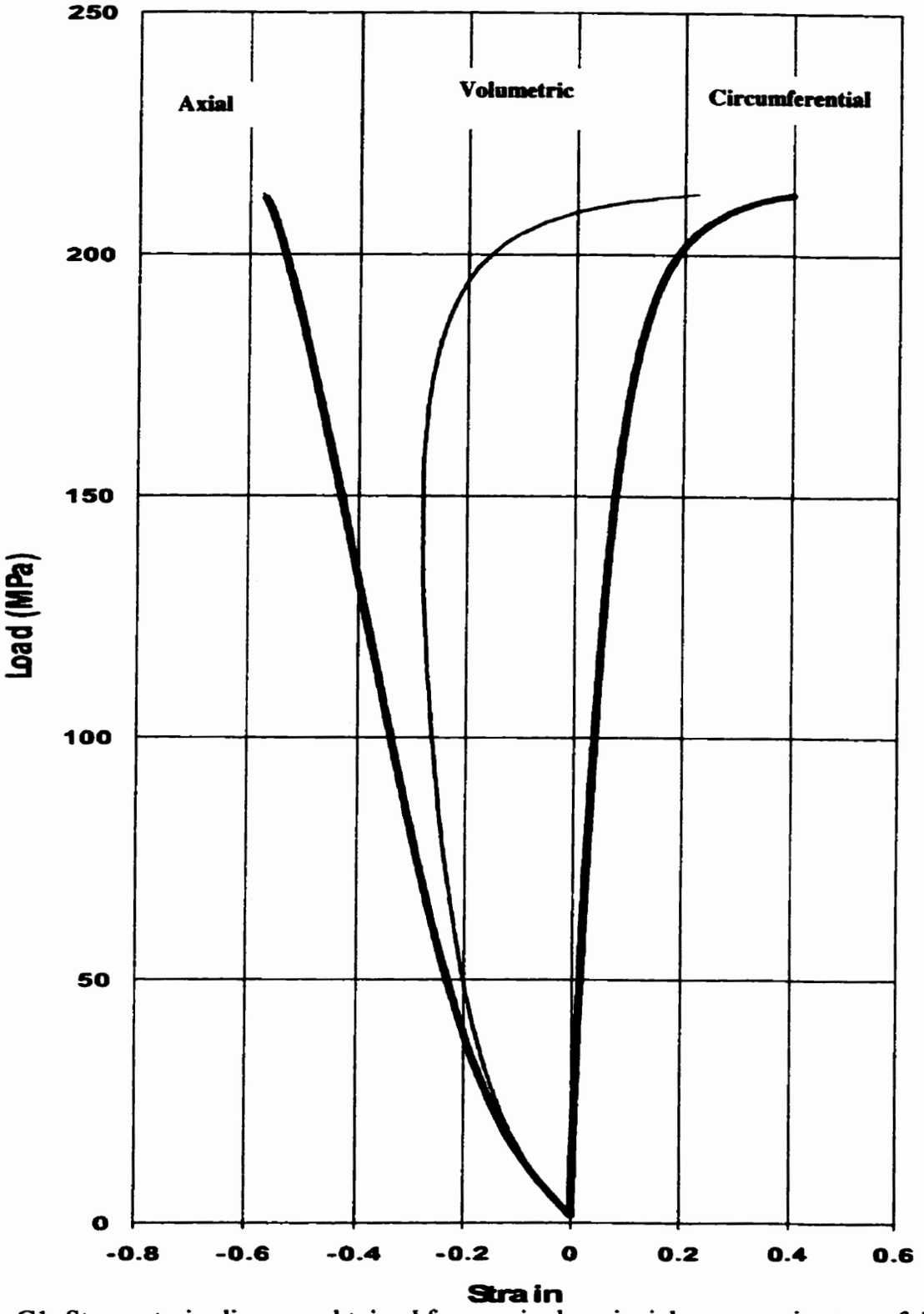



Figure C1. Stress-strain diagram obtained from a single uniaxial compression test of the fine-grained granite, borehole 421-032-RT1 (#1188).

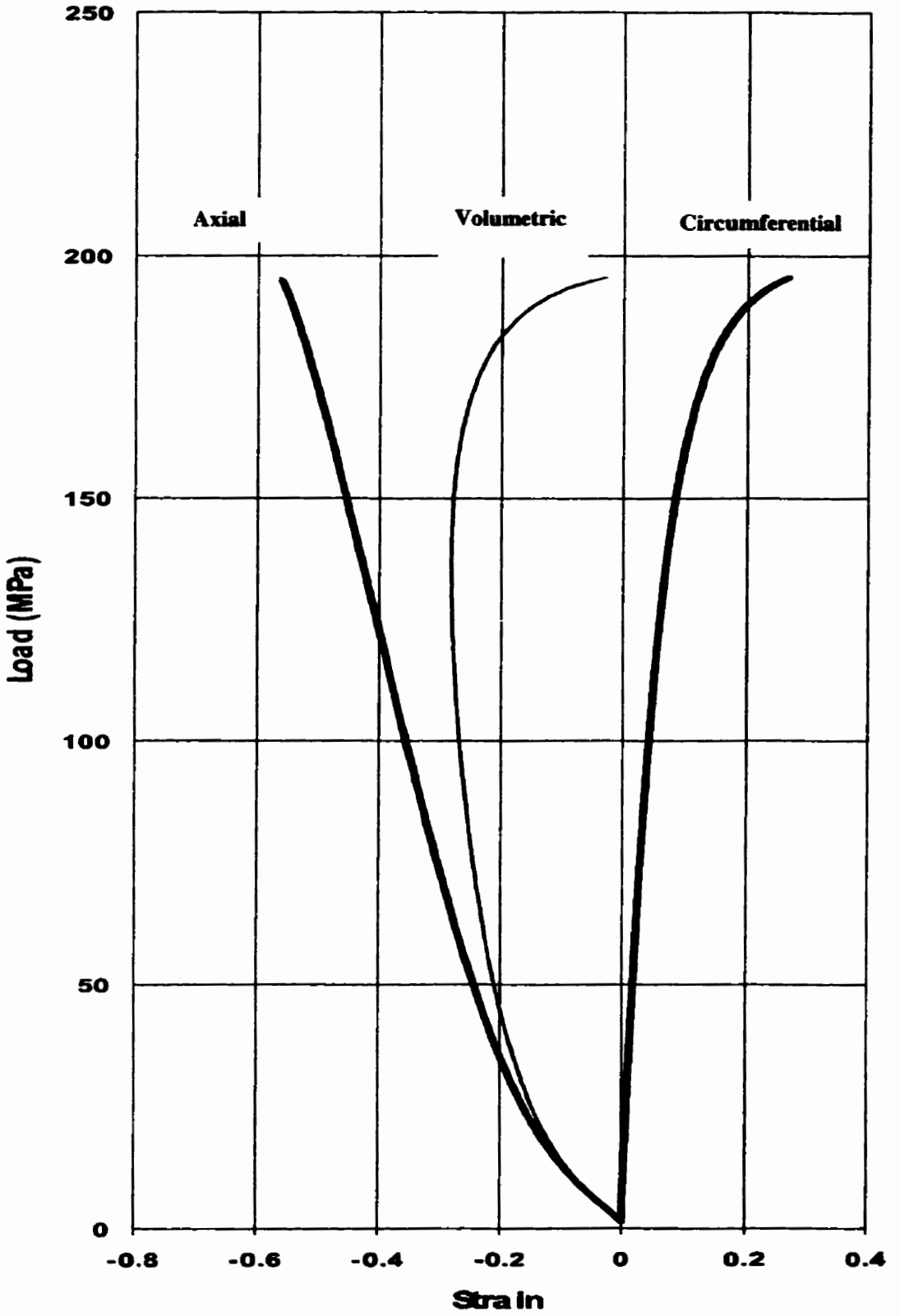



Figure C2. Stress-strain diagram obtained from a single uniaxial compression test of the fine-grained granite, borehole 421-032-RT2 (#1275).

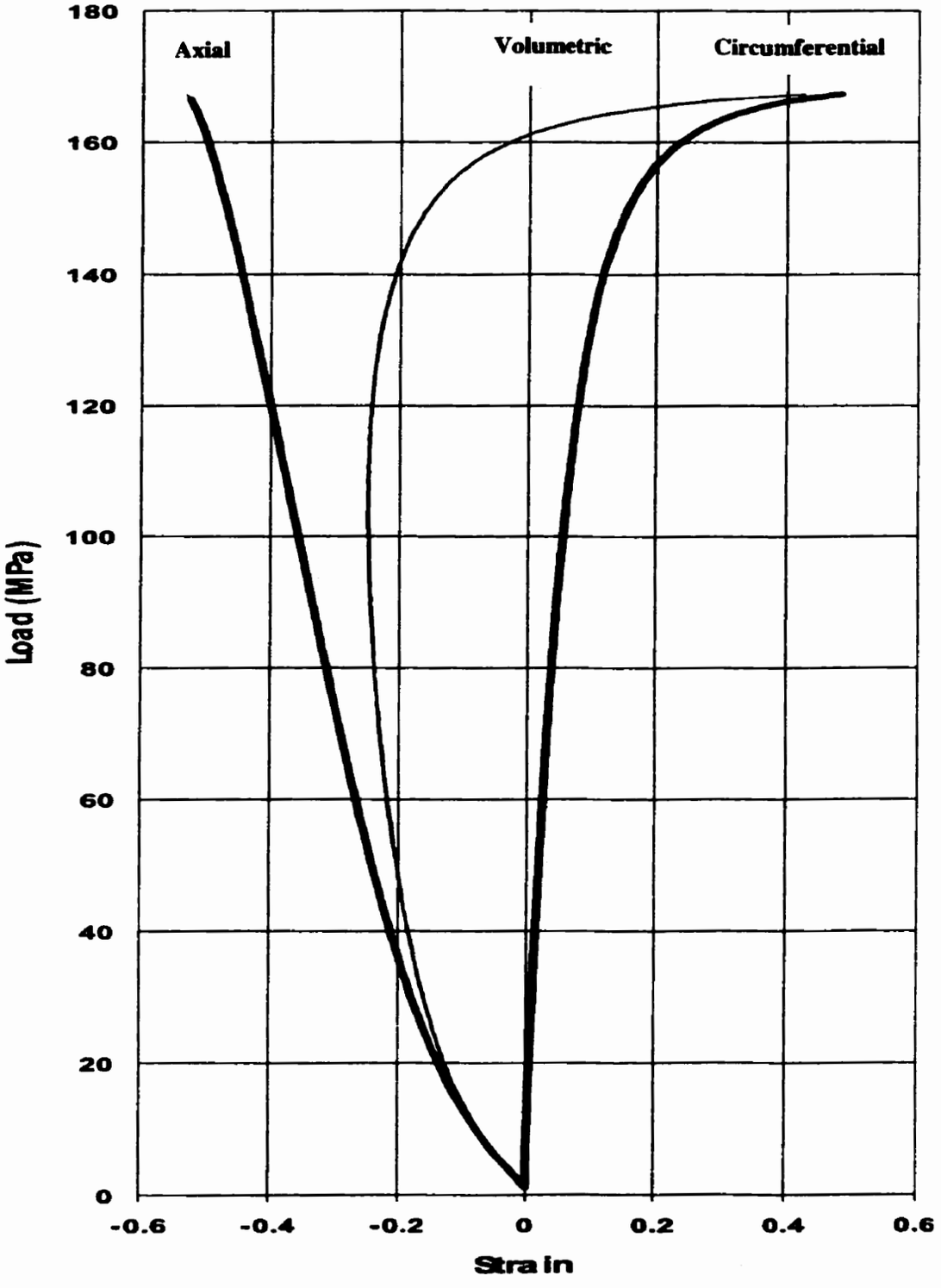



Figure C3. Stress-strain diagram obtained from a single uniaxial compression test of the fine-grained granite, borehole 421-032-RT3 (#1555).

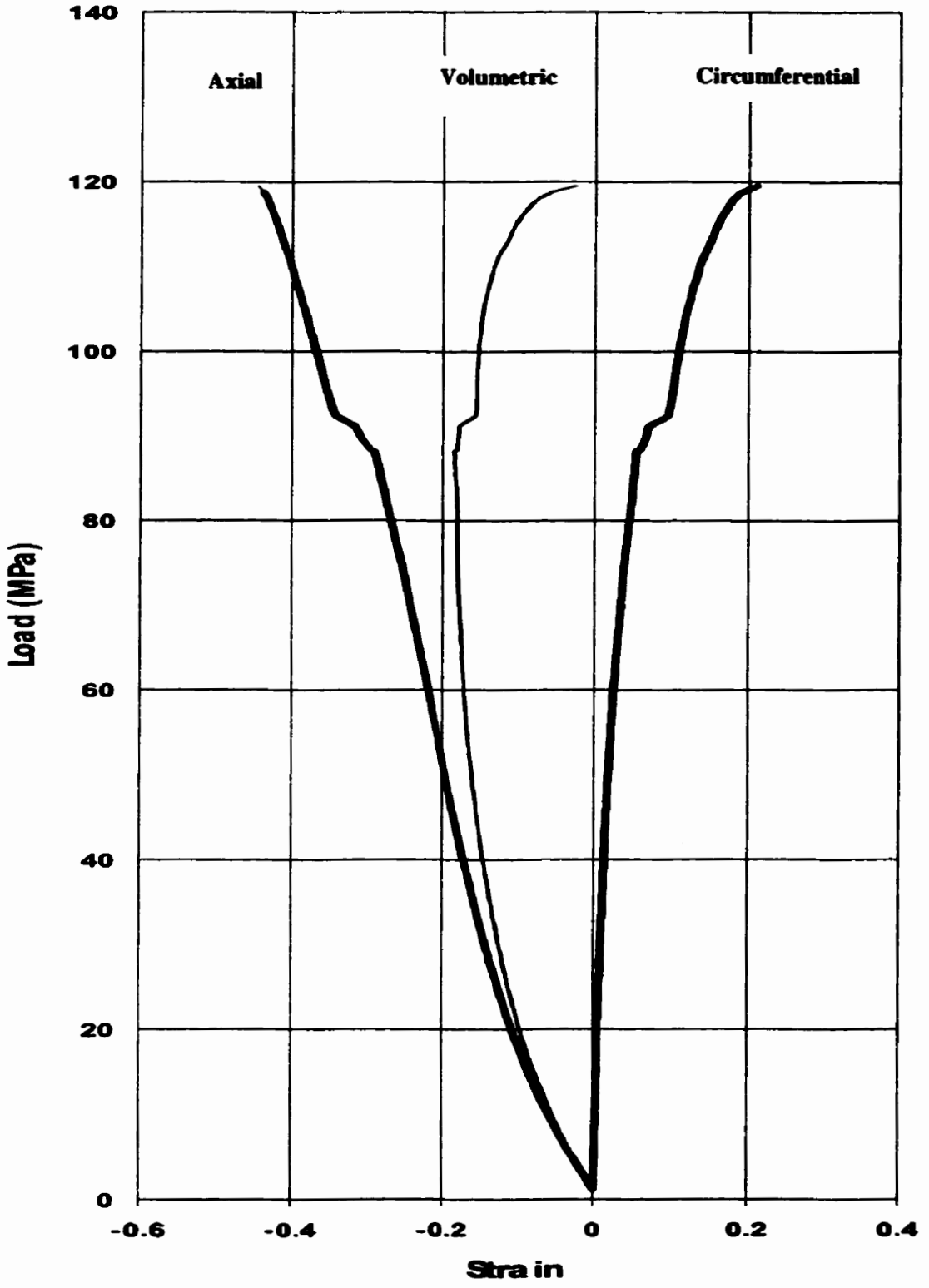



Figure C4. Stress-strain diagram obtained from a single uniaxial compression test of the fine-grained granite, borehole 421-032-RT4 (#1710).

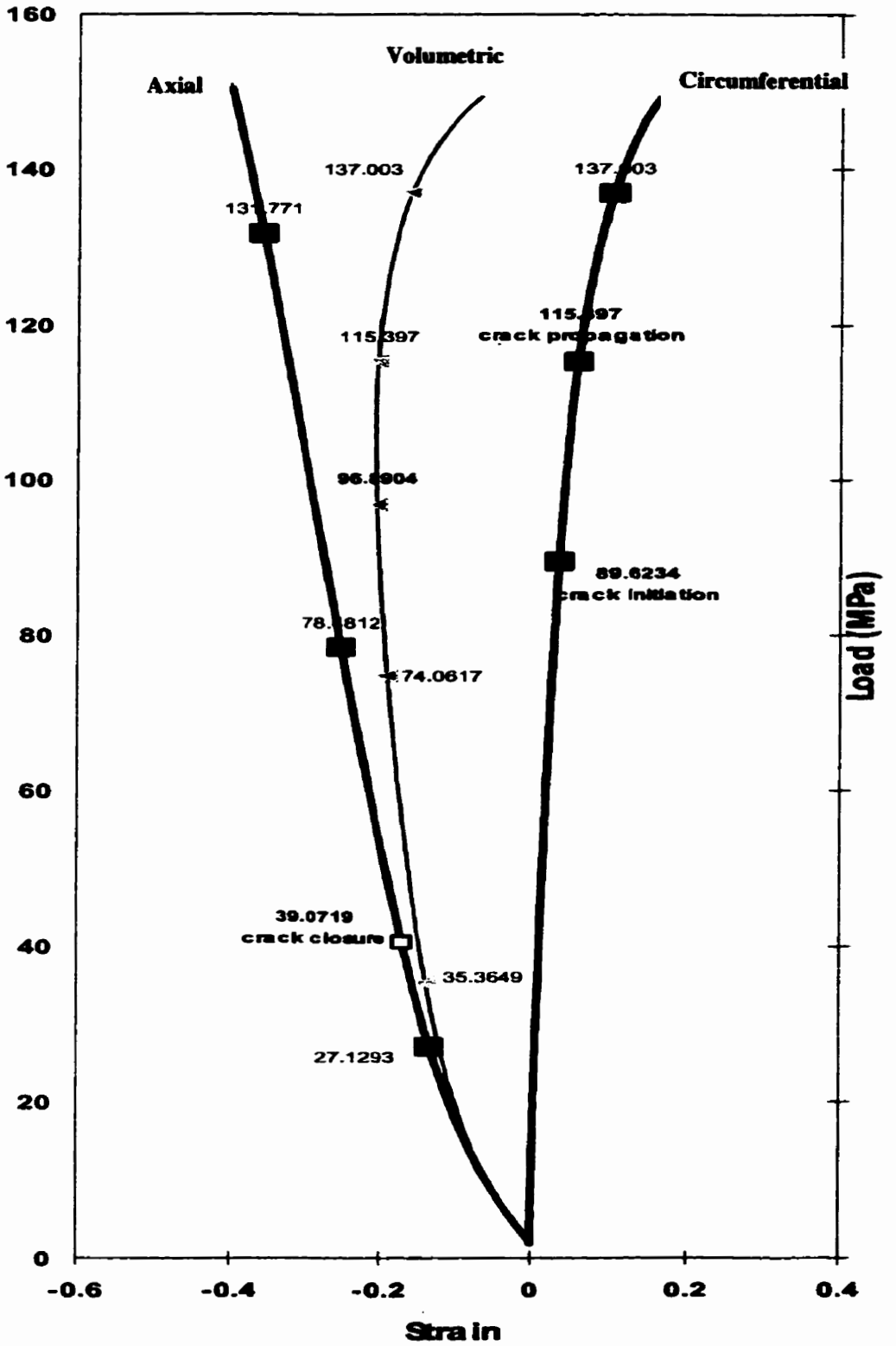



Figure C5. Stress-strain diagram obtained from a single uniaxial compression test of the medium-grained granite, borehole 417-071-RT4 (#1052).

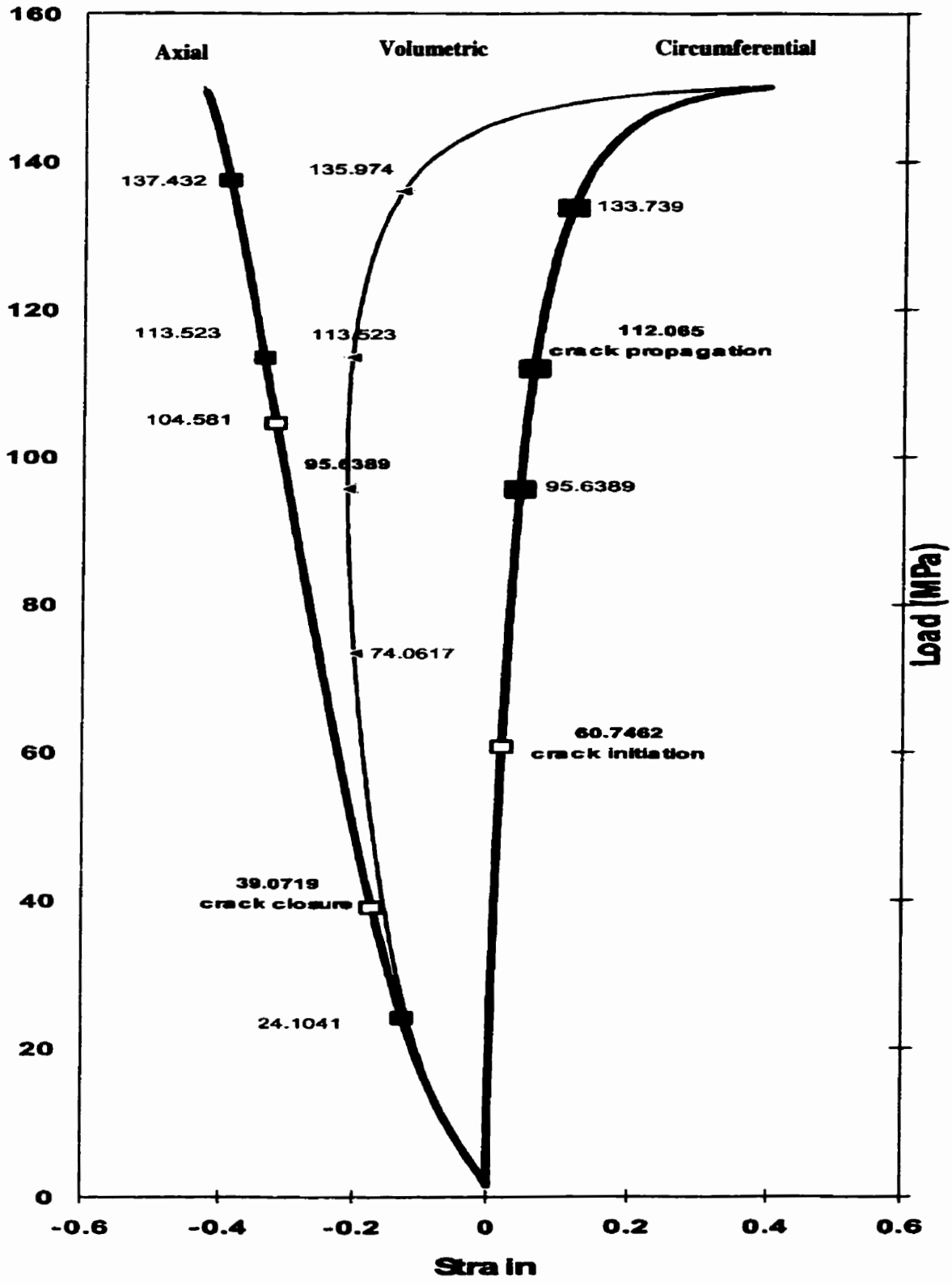



Figure C6. Stress-strain diagram obtained from a single uniaxial compression test of the medium-Grained granite, borehole 417-071-RT5 (#0831).

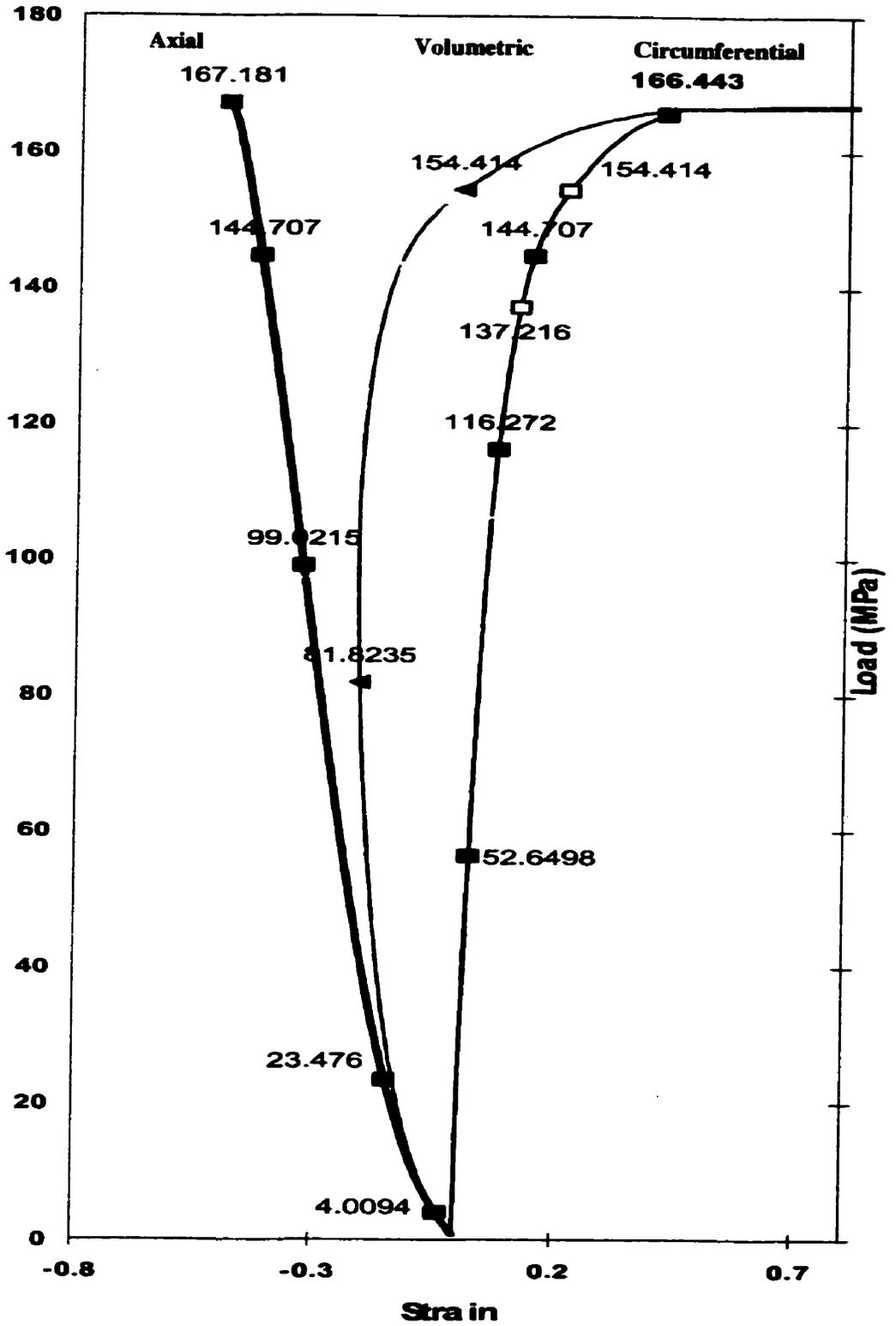



Figure C7. Stress-strain diagram obtained from a single uniaxial compression test of the medium grained granite, borehole 403-014-MB2 (#695).

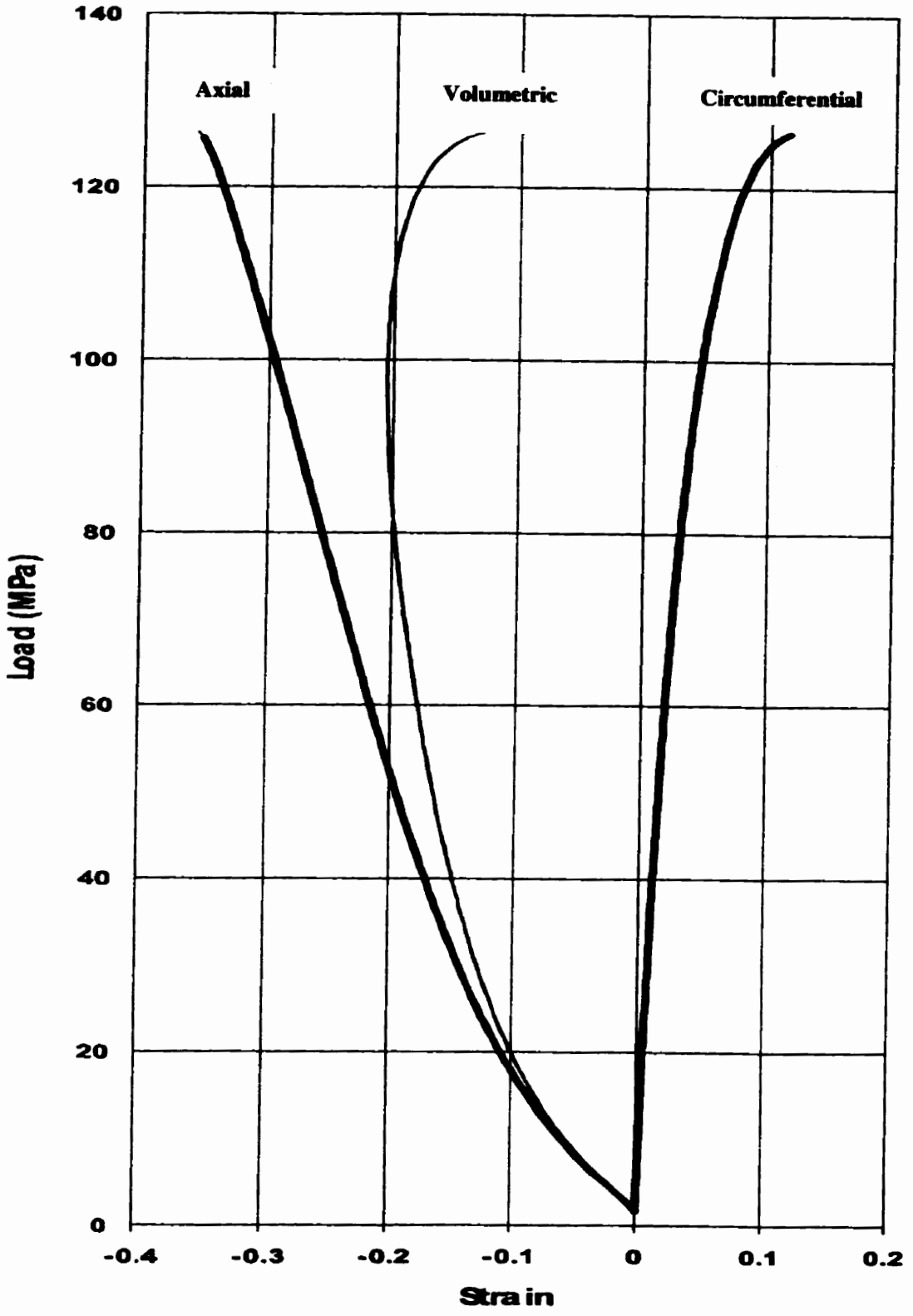



Figure C8. Stress-strain diagram obtained from a single uniaxial compression test of the coarse-grained granite, borehole 418-051-RT3 (#1335).

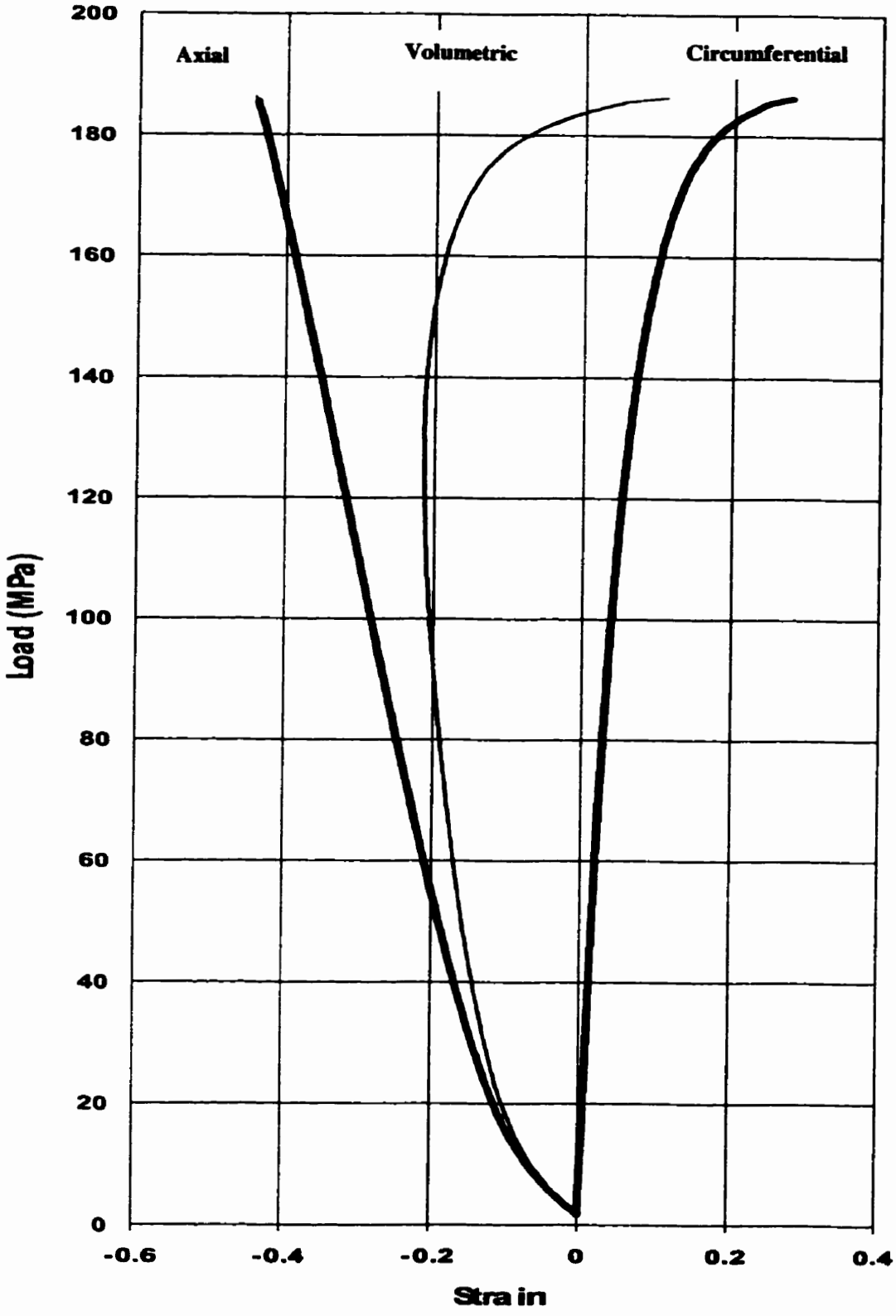



Figure C9. Stress-strain diagram obtained from a single uniaxial compression test of the coarse-grained granite, borehole 418-051-RT4 (#1342).

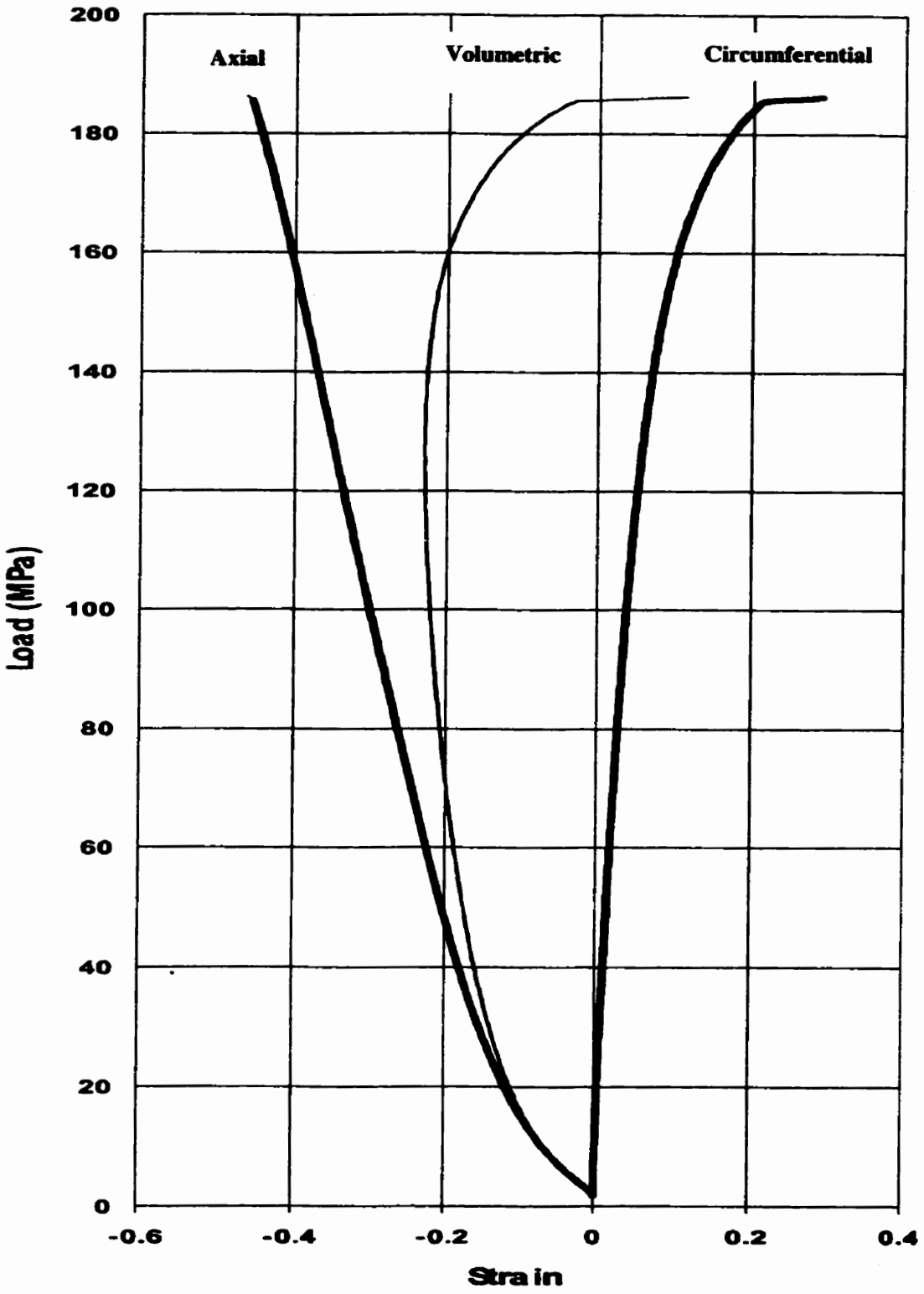



Figure C10. Stress-strain diagram obtained from a single uniaxial compression test of the coarse-grained granite, borehole 418-051-RT5 (#1114).

NQ

62637

UMI

MICROFILMED 2002

# INFORMATION TO USERS

This manuscript has been reproduced from the microfilm master. UMI films the text directly from the original or copy submitted. Thus, some thesis and dissertation copies are in typewriter face, while others may be from any type of computer printer.

**The quality of this reproduction is dependent upon the quality of the copy submitted.** Broken or indistinct print, colored or poor quality illustrations and photographs, print bleedthrough, substandard margins, and improper alignment can adversely affect reproduction.

In the unlikely event that the author did not send UMI a complete manuscript and there are missing pages, these will be noted. Also, if unauthorized copyright material had to be removed, a note will indicate the deletion.

Oversize materials (e.g., maps, drawings, charts) are reproduced by sectioning the original, beginning at the upper left-hand corner and continuing from left to right in equal sections with small overlaps.

Photographs included in the original manuscript have been reproduced xerographically in this copy. Higher quality 6" x 9" black and white photographic prints are available for any photographs or illustrations appearing in this copy for an additional charge. Contact UMI directly to order.

ProQuest Information and Learning
300 North Zeeb Road, Ann Arbor, MI 48106-1346 USA
800-521-0600

# Development of a One-Dimensional Electro-Thermophysical Model of the Snow Sea-Ice System: Arctic Climate Processes and Microwave Remote Sensing Applications

by

**John Michael Hanesiak**

A Thesis
Submitted to the Faculty of Graduate Studies
in partial fulfillment of the requirements
for the degree of

**Doctor of Philosophy**

Centre for Earth Observation Science
Department of Geography
University of Manitoba

Winnipeg, Manitoba, Canada, 2001

National Library of Canada
Acquisitions and Bibliographic Services
395 Wellington Street
Ottawa ON K1A 0N4
Canada

Bibliothèque nationale du Canada
Acquisitions et services bibliographiques
395, rue Wellington
Ottawa ON K1A 0N4
Canada

*Your file Votre référence*

*Our file Notre référence*

The author has granted a non-exclusive licence allowing the National Library of Canada to reproduce, loan, distribute or sell copies of this thesis in microform, paper or electronic formats.

The author retains ownership of the copyright in this thesis. Neither the thesis nor substantial extracts from it may be printed or otherwise reproduced without the author's permission.

L'auteur a accordé une licence non exclusive permettant à la Bibliothèque nationale du Canada de reproduire, prêter, distribuer ou vendre des copies de cette thèse sous la forme de microfiche/film, de reproduction sur papier ou sur format électronique.

L'auteur conserve la propriété du droit d'auteur qui protège cette thèse. Ni la thèse ni des extraits substantiels de celle-ci ne doivent être imprimés ou autrement reproduits sans son autorisation.

0-612-62637-7

Canada

**FACULTY OF GRADUATE STUDIES**
**FINAL ORAL EXAMINATION OF THE PH.D. THESIS**

**The undersigned certify that they have read, and recommend to the Faculty of Graduate Studies for acceptance, a Ph.D. thesis entitled:**

**"DEVELOPMENT OF A ONE-DIMENSIONAL ELECTRO-THERMOPHYSICAL MODEL OF THE SNOW SEA-ICE SYSTEM: ARCTIC CLIMATE PROCESSES AND MICROWAVE REMOTE SENSING APPLICATIONS"**

**BY**

**JOHN MICHAEL HANESIAK**

**In Partial fulfillment of the requirements for the Ph.D. Degree**

.....
Dr. David G. Barber, Advisor

.....

.....

.....

.....

**External Examiner**
**Dr. Claude Duguay**
**Université Laval**
**SainteFoy, Quebec, Canada**

Date of Oral Examination:.....February 12, 2001.....

The Student has satisfactorily completed and passed the Ph.D. Oral Examination.

.....
Advisor

.....

.....

.....

.....

.....
Chair of Ph.D. Oral

(The signature of the Chair does not necessarily signify that the Chair has read the thesis.)

THE UNIVERSITY OF MANITOBA

FACULTY OF GRADUATE STUDIES

*****

COPYRIGHT PERMISSION PAGE

# Development of a One-Dimensional Electro-Thermophysical Model of the Snow Sea-Ice System: Arctic Climate Processes and Microwave Remote Sensing Applications

by

John Michael Hanesiak

A Thesis/Practicum submitted to the Faculty of Graduate Studies of The University of Manitoba in partial fulfillment of the requirements of the degree of Doctor of Philosophy

John Michael Hanesiak © 2001

Permission has been granted to the Library of The University of Manitoba to lend or sell copies of this thesis/practicum, to the National Library of Canada to microfilm this thesis/practicum and to lend or sell copies of the film, and to Dissertations Abstracts International to publish an abstract of this thesis/practicum.

The author reserves other publication rights, and neither this thesis/practicum nor extensive extracts from it may be printed or otherwise reproduced without the author's written permission.

# ABSTRACT

Snow covered sea ice plays a crucial role in the earth's climate. This includes polar biology, local, regional and world weather and ocean circulations as well as indigenous people's way of life. Recent research has indicated significant climate change in the polar regions, especially the Canadian arctic. Polar climate processes are also among the most poorly misrepresented within global circulation models (GCMs). The goal of this thesis is to improve our understanding and capability to simulate arctic climate processes in a predictive sense. An electro-thermophysical relationship exists between the thermophysical characteristics (climate variables and processes) and electrical properties (dielectrics) that control microwave remote sensing of snow-covered first-year sea ice (FYI). This work explicitly links microwave dielectrics and a thermodynamic model of snow and sea ice by addressing four key issues. These includes: 1) ensure the existing one-dimensional sea ice models treat the surface energy balance (SEB) and snow/ice thermodynamics in the appropriate time scales we see occurring in field experiments, 2) ensure the snow/ice thermodynamics are not compromised by differences in environmental and spatial representation within components of the SEB, 3) ensure the snow layer is properly handled in the modeling environment, and 4) how we can make use of satellite microwave remote sensing data within the model environment. Results suggest that diurnal processes are critical and need to be accounted for in modeling snow-covered FYI, similar to time scales acting in microwave remote sensing signatures. The parameterization of incident short-wave ($K\downarrow$) and long-wave ($L\downarrow$) radiation contain seasonal and environmental biases in some arctic locations (i.e. near polynyas) and have been adjusted to account for these biases. The representation of surface albedo needs to be improved in sea ice models. Surface and airborne measurements of albedo over FYI in the melt season suggest four distinct cover types that dominate albedo that need to be included in process models and GCMs. The $K\downarrow$, $L\downarrow$, and albedo are critical in the SEB and therefore indirectly impact snow and ice microwave dielectrics through the electro-thermophysical relationship. The snow layer representation in current sea ice models is not sufficient for linkages to snow and ice microwave dielectrics. A detailed mass and energy balance 1-D snow model is coupled to a FYI model, a type of model that has not

existed until now. Results suggest the coupled version of the model improves simulations of FYI over annual cycles. The addition of salinity in snow thermal conductivity and specific heat improves ice temperatures during the melt season but causes over estimations of ice temperatures in the cold season. Addition of salinity in the mass balance of the model (in future implementations) will improve ice temperature simulation during the cold season. Output from the coupled snow sea-ice model provides the required input to microwave dielectric models of snow and sea ice to predict microwave penetration depths within the snow and sea ice (an Electro-Thermophysical model of the Snow Sea Ice System (ETSSIS)). Results suggest ETSSIS can accurately simulate microwave penetration depths in the cold dry snow season and wet snow season (funicular snow regime). Simulated penetration depths become too large in the pendular snow regime since liquid water is not generated soon enough within the snow pack in the spring season. The inclusion of salinity in the mass balance of ETSSIS will improve the simulation of penetration depths in the pendular snow regime in future implementations of the model.

# ACKNOWLEDGEMENTS

I am indebted to my supervisor Dr. David Barber who has broadened my interests and research skills. He has also become a good friend and one whom I admire for his dedication and insight. Thanks to all I have been affiliated with at the University of Manitoba (too many to mention all), especially Dave Mosscrop, soon to be Dr. John Yackel (now at U. of Calgary), John Iaacoza, and C.J. Mundy. Warm gratitude is extended to all personnel affiliated with the ice camps of C-ICE'97, C-ICE'00 and NOW'98, without their help, data collection toward this work would have been non-existent. I must extend sincere appreciation to Dr. Greg Flato (CCCMA) and Dr. Rachel Jordan (CRREL) who have expended much energy (and email) assisting me with model codes and expertise. My sincere appreciation to my committee members, Dr. Jim Gardner (U. Manitoba), Dr. Tim Papakyriakou (U. Manitoba), Mr. Richard Raddatz (MSC and U. Manitoba) and Dr. Claude Duguay (Laval U.) for taking the time to read my dissertation and providing valuable feedback toward its completion. I also acknowledge Dr. Ronald E. Stewart (MSC/CCRP) (my M.Sc. supervisor and Manitoba born) whom I greatly admire for his work, insight, tuning of my research skills and faith in my ability. Thanks goes out to all organizations who have assisted with funding, data and logistics: University of Manitoba, Canadian Ice Service (CIS), Polar Continental Shelf Project (PCSP) and the peoples of Resolute Bay, NSERC of Canada and Office of Naval Research (both under D. Barber), Northern Studies Training Program (NSTP), CRYSYS (Dr. Barry Goodison, PI), Meteorological Service of Canada (MSC). The contribution of NSERC, the Institute for Space and Terrestrial Science, and the Polar Continental Shelf Project towards the SIMMS program, directed by Dr. E. LeDrew and Dr. D. Barber, is acknowledged. Finally and most importantly, my wife (Teresa) and family (in Winnipeg, Toronto and Los Angeles) for encouragement, support and many hours that I have missed spending precious time with during this work. Party-on!

# DEDICATION

This thesis is dedicated to my loving wife (Teresa) who has endured the challenges (mentally, financially and emotionally) of my academic endeavors as well as my Mom (Nadia) and Dad (John) for providing the initial opportunity to pursue post-secondary studies. My sincere love and gratitude goes to them as well as my Sister (Bonnie), who have all provided undaunted and endless support, understanding and belief in me ... and of course, Sadie too (our dog), who sat beside me on those long afternoons of computer work!

# TABLE OF CONTENTS

**Abstract** ..... ii

**Acknowledgements** ..... iv

**Dedication** ..... v

**Table of Contents** ..... vi

**List of Tables** ..... xii

**List of Figures** ..... xiv

***Chapter 1: Introduction and Statement of Objectives*** ..... 1

**1.1 Motivation** ..... 1

**1.2 Thesis Objectives** ..... 11

**1.3 Thesis Structure** ..... 15

***Chapter 2: Background*** ..... 17

**2.1 First-Year Sea Ice and its Snow Cover Characteristics** ..... 17

**2.2 Thermodynamics of Snow and Sea Ice** ..... 26

2.2.1 Atmospheric Radiation Interaction with Snow and Ice ..... 29

*2.2.1.1 Incident Short-Wave Radiation* ..... 29

*2.2.1.2 Surface Reflection (Albedo)* ..... 30

*2.2.1.3 Surface Absorption* ..... 35

*2.2.1.4 Surface Transmission* ..... 35

*2.2.1.5 Long-Wave Radiation* ..... 36

2.2.2 Heat Flow Through Snow and Ice ..... 37

**2.3 Utility of Numerical Modeling for Snow Covered Sea Ice** ..... **43**

2.3.1 Historical Perspective ..... 43

2.3.2 Utility of Modeling ..... 44

2.3.3 Model Deficiencies ..... 55

**2.4 Utility of Microwave Remote Sensing for Snow Covered Sea Ice** ..... **58**

2.4.1 Background ..... 58

2.4.2 Microwave Remote Sensing of Sea Ice Through the Annual Cycle ..... 60

*2.4.2.1 Fall Freeze-Up* ..... 62

*2.4.2.2 Winter* ..... 65

*2.4.2.3 Early Melt* ..... 69

*2.4.2.4 Melt Onset* ..... 69

*2.4.2.5 Advanced Melt* ..... 70

**2.5 Summary** ..... **72**

***Chapter 3: Role of Diurnal Processes in the Seasonal Evolution of Sea Ice and its Snow Cover*** ..... **73**

**3.1 Introduction** ..... **73**

**3.2 Data and Methods** ..... **74**

3.2.1 Sea Ice Model ..... 74

3.2.2 Data ..... 76

**3.3 Results and Discussion** ..... **78**

3.3.1 Temporal Scaling Considerations ..... 78

*3.3.1.1 Hourly vs. Daily Forcing Comparisons* ..... 78

*3.3.1.2 Causes for Hourly vs. Daily Forcing Differences* ..... 86

3.3.2 Parameterized vs. In Situ Radiative Fluxes and Albedo ..... 92

3.3.3 Model Simulations Using Field Observation Forcing ..... 98

**3.4 Conclusions ..... 104**

3.4.1 Question 1 ..... 104

3.4.2 Question 2 ..... 105

3.4.3 Question 3 ..... 107

**3.5 Summary ..... 107**

***Chapter 4: Parameterization Schemes of Incident Radiation ..... 109***

**4.1 Introduction ..... 109**

**4.2 Data and Methods ..... 110**

4.2.1 Observational Data ..... 110

4.2.2 Radiative Flux Parameterizations ..... 117

*4.2.2.1 Short-Wave Clear Sky Flux* ..... 118

*4.2.2.2 Short-Wave All Sky Flux* ..... 119

*4.2.2.3 Long-Wave Clear Sky Flux* ..... 120

*4.2.2.4 Long-Wave All Sky Flux* ..... 120

4.2.3 Analysis Methods ..... 121

**4.3 Results ..... 123**

4.3.1 Parameterizations vs. Observation ..... 123

*4.3.1.1 Terrestrial and Fast-Ice Sites* ..... 123

*4.3.1.1.1 Incident Short-Wave Fluxes* ..... 123

*4.3.1.1.2 Incident Long-Wave Fluxes* ..... 127

*4.3.1.1.3 Seasonal Trends* ..... 132

*4.3.1.2 Ice Breaker Platform* ..... 136

*4.3.1.2.1 Incident Short-Wave Fluxes* ..... 137

*4.3.1.2.2 Incident Long-Wave Fluxes* ..... 139

**4.4 Conclusions** ..... **143**

4.4.1 Question 1 ..... 143

4.4.2 Question 2 ..... 145

**4.5 Summary** ..... **146**

***Chapter 5: Local and Regional Albedo Observations of Arctic First Year Sea Ice During Melt Ponding*** ..... **147**

**5.1 Introduction** ..... **147**

**5.2 Methods** ..... **148**

5.2.1 Site Description ..... 148

5.2.2 Surface Measurements of Albedo ..... 151

5.2.3 Airborne-Derived Albedo ..... 153

5.2.4 AVHRR-Derived Albedo ..... 158

5.2.5 Sea Ice Model Simulations ..... 160

**5.3 Results** ..... **163**

5.3.1 Surface Measurements of Albedo ..... 163

5.3.2 Upscaling Surface Albedo to Sub-Regional and Regional Scales ..... 173

5.3.3 Effect of Pond Fractions on Ice Thermodynamics ..... 181

**5.4 Conclusions** ..... **186**

5.4.1 Question 1 ..... 187

5.4.2 Question 2 ..... 187

5.4.3 Question 3 ..... 188

**5.5 Summary ..... 190**

***Chapter 6: Coupled 1-D Thermodynamic Snow Sea-Ice Model: Climate Processes ..... 192***

**6.1 Introduction ..... 192**

**6.2 Observational Data and Numerical Models ..... 194**

6.2.1 Observational Data ..... 194

6.2.2 Numerical Models ..... 196

**6.3 Methods ..... 199**

**6.4 Results ..... 200**

6.4.1 Seasonal Snow and Ice Evolution ..... 200

6.4.2 Thermodynamic Evolution ..... 204

**6.5 Conclusions ..... 209**

6.5.1 Question 1 ..... 209

6.5.2 Question 2 ..... 210

**6.6 Summary ..... 211**

***Chapter 7: Electro-Thermophysical Model of the Snow Sea-Ice System (ETSSIS) for Microwave Remote Sensing ..... 212***

**7.1 Introduction ..... 212**

**7.2 Field Data and Numerical Model ..... 215**

7.2.1 Observational Data ..... 215

7.2.2 Coupled 1-D Snow Sea-Ice Model ..... 216

7.2.3 Snow and Sea Ice Dielectric Modeling ..... 216

**7.3 Analysis Methods ..... 220**

**7.4 Results ..... 222**

7.4.1 Thermodynamic Evolution ..... 222

7.4.2 Physical Evolution ..... 230

**7.5 Microwave Dielectrics ..... 240**

7.5.1 Cold Dry Snow ..... 240

7.5.2 Pendular Snow Regime ..... 246

7.5.3 Funicular Snow Regime ..... 246

7.5.4 Dielectric Sensitivity to Measurement Errors ..... 247

**7.6 Conclusions ..... 248**

7.5.1 Question 1 ..... 249

7.5.2 Question 2 ..... 249

7.5.3 Question 3 ..... 250

**7.7 Summary ..... 251**

***Chapter 8: Conclusions and Future Research ..... 253***

**8.1 Conclusions ..... 253**

**8.2 Summary ..... 264**

**8.3 Future Research ..... 265**

***References ..... 272***

# LIST OF TABLES


2.1 Previous measurements of integrated albedo over various surfaces ..... 32

2.2 Variables obtainable from microwave remote sensing for sea ice processes ..... 61

3.1 Break-up dates and open water duration for various model runs ..... 88

3.2 Statistical comparisons between the modeled and observed net surface flux ($Q_{net}$) and surface temperatures ($T_{sfc}$) ..... 103

4.1 Parameterized clear-sky short-wave flux error ..... 123

4.2 Parameterized all-sky short-wave flux error ..... 125

4.3 Parameterized clear-sky long-wave flux error ..... 127

4.4 Parameterized clear-sky long-wave flux error after modifying emissivity ..... 127

4.5 Parameterized all-sky long-wave flux error ..... 129

4.6 Parameterized all-sky long-wave flux error when applying the new clear-sky schemes ..... 129

4.7 Parameterized all-sky long-wave flux error when applying the new clear-sky and new cloudy-sky schemes ..... 131

5.1 Surface type samples of spectral and broadband albedo conducted during C-ICE97 ..... 164

5.2 Comparison between aircraft and AVHRR-derived broadband albedo ..... 176

5.3 The "typical" or climatological fractional cover types and resulting total albedo over a 7 week period ..... 182

6.1 NC, C, and UPR $R^2$, mean bias error (MBE) and root-mean-squared error (RMSE) compared to observations of $Q_{net}$, $T_{sfc}$, and $T_{isfc}$ ..... 205

7.1 ETSSIS C, AVG, and UPR $R^2$, mean bias error (MBE) and root-mean-squared error (RMSE) compared to observations of $Q_{net}$, $T_{sfc}$ and $T_{isfc}$ ..... 227

7.2 ETSSIS C mean bias error (MBE) and root-mean-squared error (RMSE) snow density, grain sizes, water volume fraction, and snow salinity versus observations ..... 236

7.3 ETSSIS C, AVG, and UPR mean bias error (MBE) and root-mean-squared error (RMSE) permitivity (ε'), loss (ε") and penetration depth (δp) versus "observations" ..... 245

# LIST OF FIGURES


2.1 Stages of first-year sea ice growth, its thickness and external forcing ..... 18

2.2 First-year sea ice physical characteristics ..... 20

2.3 Typical salinity profiles for growing first-year sea ice ..... 21

2.4 Examples of two melt pond first-year sea ice surfaces measured with aerial video ..... 22

2.5 Typical new (faceted) and old (kinetic) snow grains and aggregates of partially melted grains (polycrystalline) ..... 25

2.6 The critical physical thermodynamic processes within snow covered first-year sea ice ..... 27

2.7 Two spectral albedo curves over snow covered sea ice ..... 33

2.8 Categories of snow covered first-year sea ice thermodynamic regimes between freeze-up to advanced melt periods ..... 42

2.9 Time series of observed and modeled ice and snow thickness at Alert, Nunavut between 1980-1990 ..... 46

2.10 Modeled interactions between external forcing and parameter values, and internal variables and fluxes for a one-dimensional thermodynamic sea ice model ..... 47

2.11 Effect of a 5-day period of snowfall (20 cm) on the open water duration of modeled first-year sea ice ..... 51

2.12 Modeled and observed temperature traces at various depths in the snow and sea ice ..... 54

2.13 Typical ERS and RadarSat backscatter ($\sigma^{o}$) at 5.3 GHz for thick First-Year and Multi-Year sea ice over the seasonal cycle ..... 62

2.14 Examples of arctic RadarSat imagery in Winter and Advanced Melt ..... 67

3.1 Resolute Bay, Nunavut and the first-year sea ice (FYI) field camps from SIMMS'92 and '93 ..... 77

3.2 Time series of observed and modeled snow and ice thickness between YD107, 1992 and YD170, 1993 ..... 81

3.3 Time series of observed and modeled snow thickness between YD120 and YD175, 1992 ..... 81

3.4 Time series of ambient air temperature and modeled surface temperatures between YD130 and YD175, 1992 ..... 83

3.5 Time series of modeled surface albedo between YD130 and YD175, 1992 ..... 84

3.6 Examples of modeled and observed vertical snow/ice temperature profiles ..... 86

3.7 Time series of modeled daily average K* difference between YD107 and YD175, 1992 ..... 90

3.8 Time series of modeled surface energy balance between YD130 and YD175, 1992 ..... 90

3.9 Time series of air temperature discrepancy between Resolute and the SIMMS'92 field site between YD107 and YD177, 1992 ..... 93

3.10 Modeled minus observed radiative fluxes ..... 94

3.11 Time series of parameterized and observed surface albedo ..... 97

3.12 Diurnal time series of observed and modeled surface albedo ..... 97

3.13 Comparison between observed and modeled snow depth between YD107-177, 1992 ..... 99

3.14 Comparison between observed and modeled net flux and surface temperature between YD107 and YD176, 1992 ..... 101

4.1 Geographical region of the NOW'98 project and locations of the terrestrial camp (Cape Herschel) and ice site in Rosse Bay ..... 111

4.2 Incident radiation instruments on the fast-ice and ice breaker ..... 114

4.3 The clear-sky short-wave flux error for two parameterizations ..... 124

4.4 The all-sky short-wave flux error for four combinations of parameterizations .. 126

4.5 The clear-sky long-wave flux error for the optimized emissivities ..... 128

4.6 The all-sky long-wave flux error using the optimized clear-sky emissivities in Figure 4.4 and optimized cloudy-sky emissivities ..... 130

4.7 The short-wave observed and Shine scheme seasonal clear-sky fluxes ..... 132

4.8 The short-wave observed and Shine/Shine scheme seasonal all-sky fluxes ..... 134

4.9 The long-wave observed, Maykut and Church, and Efimova schemes seasonal clear-sky fluxes ..... 135

4.10 The long-wave observed, MC/MC, and Efimova/MC schemes seasonal all-sky fluxes ..... 136

4.11 The short-wave observed and Shine scheme seasonal clear-sky fluxes for the ice breaker ..... 137

4.12 The short-wave observed and Shine/Shine scheme seasonal all-sky fluxes for the ice breaker ..... 138

4.13 The long-wave observed, Maykut and Church, and Efimova schemes seasonal clear-sky fluxes for the ice breaker ..... 139

4.14 The long-wave observed, MC/MC, and Efimova/MC schemes seasonal all-sky fluxes for the ice breaker ..... 140

5.1 Geographic locations of the CICE'97 land camp, ice camp and ice conditions within Wellington Channel and Lancaster Sound ..... 150

5.2 In-cabin helicopter photo of the albedometer mounted to the pontoon ..... 155

5.3 Geographic locations of the aircraft video surveys (linear transects) within Lancaster Sound and Wellington Channel ..... 156

5.4 Spectral albedo evolution of a deep snow pack undergoing rapid melt ..... 167

5.5 Directional hemispheric spectral albedo over various surfaces on first-year sea ice ..... 167

5.6 Spectral albedo of shallow melt ponds with and without a thin ice layer ..... 170

5.7 Broadband albedo over various surfaces on first-year sea ice ..... 170

5.8 Broadband albedo along a 1 - km ground transect ..... 172

5.9 Helicopter-measured broadband albedo at two altitudes ..... 172

5.10 Mean and standard deviation of fractional cover types along each aircraft transect ..... 175

5.11 Example of aircraft-derived broadband albedo along one transect ..... 175

5.12 NOAA-14 AVHRR-derived albedo on YD180 at 0909 UTC (0409 CDT) in Lancaster Sound and Wellington Channel ..... 178

5.13 Modeled and climatological broadband albedo over a seven week period ..... 183

5.14 Ice temperature gradient change (ITGC) versus pond fraction after four and seven days of model simulation ..... 183

5.15 Ice depth versus temperature between the ice-ocean interface and 25 cm from the top ice surface as a function of pond fraction ..... 186

6.1 Third order polynomials (AVG and UPR) of snow salinity versus snow depth ..... 196

6.2 Observed, UPR and NC snow and ice thickness evolution between SIMMS'92 and SIMMS'93 ..... 202

6.3 UPR and NC snow and ice thickness evolution for the freeze-up sensitivity study ..... 203

6.4 Observed, NC and UPR snow thickness evolution between YD107 and YD177, 1992 ..... 205

6.5 C snow surface temperature ($T_{sfc}$) versus observations (obs) between YD107 and YD177, 1992 ..... 206

6.6 NC, C, and UPR simulated ice surface temperature ($T_{isfc}$) versus observations (obs) between, a) YD107 to YD177 and b) YD107 to YD128 ..... 208

7.1a C, AVG, and UPR ETSSIS thermal conductivity ..... 224

7.1b C, AVG, and UPR ETSSIS specific heat ..... 225

7.2 C, AVG, and UPR ETSSIS temperatures versus observations ..... 226

7.3 ETSSIS temperature difference between C and UPR ..... 229

7.4 C, AVG, and UPR ETSSIS density versus observations ..... 233

7.5 C, AVG, and UPR ETSSIS snow grain sizes versus observations ..... 234

7.6 C, AVG, and UPR ETSSIS fractional water volume versus observations ..... 235

7.7 C, AVG, and UPR ETSSIS salinity versus observations ..... 237

7.8 C, AVG, and UPR ETSSIS brine volume versus "observations" ..... 239

7.9 C, AVG, and UPR ETSSIS dielectric permitivity versus "observations" ..... 242

7.10 C, AVG, and UPR ETSSIS dielectric loss versus "observations" ..... 243

7.11 C, AVG, and UPR ETSSIS microwave penetration depths at 5.3 GHz versus "observations" ..... 244

# CHAPTER 1: Introduction and Statement of Objectives

## 1.1 Motivation

Sea ice plays an important role in the earth's climate. It supplies freshwater to the North Atlantic via ice export, it moderates ocean-atmosphere heat, moisture, and momentum exchanges, and it provides a biological habitat for primary production up to mammals. Ice that is expelled from the polar regions and carried into mid-latitudes by wind and ocean currents eventually melts, supplying a fresh water source in the upper ocean. This creates varying strata within the ocean system that drives oceanic circulation. The sea ice provides a layer of insulation between the relatively warm ocean water and the cold atmosphere in winter, severely limiting the amount of heat and moisture transfer from the ocean to the atmosphere during this season. The ice also acts to limit the atmospheric contribution to ocean currents by suppressing wind drag directly on the ocean surface. Primary production in the arctic strongly depends on the existence of ice, and without it, aquatic life and land mammals would be directly affected through food chain dynamics.

We expect to see the first and largest impact of $CO_2$ induced global climate change within the polar regions. Various feedbacks occur within the ocean-sea ice-atmosphere system which tend to amplify climate processes (Moritz and Perovich, 1998). We expect that sea ice plays a central role in these processes due to its control on the exchange of energy and mass between the ocean and atmosphere. Current global circulation models[1] (GCMs) suggest that warming in the arctic will be 2-3 times larger

[1] GCMs are computer models that simulate the earth's climate.

than the global mean warming (Kattenberg et al., 1996). A primary process that controls this enhanced polar warming is the sea ice-albedo[2] feedback (Curry et al., 1995). This feedback is driven by increased arctic air temperatures (associated with a global warming scenario) that ablates more ice, in turn decreasing the aerial extent of the ice pack. The ocean surface becomes more exposed releasing more heat into the atmosphere, and with a much lower surface albedo (0.1) compared to the snow/ice surface (0.8), the system absorbs more solar energy. The increased heat release and solar absorption acts to further enhance atmospheric warming, thereby, ablating more ice, and so on. However, critical feedbacks such as the cloud-sea ice feedback and any precipitation changes to the system act in complicated ways to enhance and counteract other feedbacks (see for example, Barry et al., 1984). The feedback processes occur over a continuum of spatial and temporal scales and the interaction of these processes makes long range climate modeling a difficult task.

Over the past decade there has been a detectable change in both the atmosphere and the ocean of the arctic (Serreze et al. 2000) as well as northern hemisphere snow cover characteristics (Brown, 2000). Sea level pressure over the arctic has decreased dramatically, signifying an anomalous circulation within the atmosphere (Walsh et al. 1996). At the same time the Arctic Atlantic Layer (200-500m) has warmed by about 1.5°C (Carmack et al. 1995), which may be related to an observed reduction in the thermodynamic equilibrium thickness[3] of sea ice within the Arctic Basin from 3.1 m to 1.8 m (Rothrock et al., 1999). This decrease in ice thickness may not be the result of thermodynamic forcing alone but by dynamic forcing, a fact that negates ice thickness decreases solely caused by global

[2] albedo is defined as the ratio of energy reflected from a surface over energy incident upon the surface.

warming. Recent results show that sea ice concentrations are reducing at a rate of about 34,000 $km^2$ per year or 2.9% per decade (over the past 18 years) (Parkinson et al. 1999). Canadian GCMs are in agreement with these rates of decay and suggest a complete disappearance of the perennial ice pack (multi-year ice (MYI[4])) before 2100 (Boer et al. 2000). This implies that a greater amount of the arctic will become covered with seasonal sea ice (first-year ice (FYI) during the winter which has already been observed (Canadian Ice Service (CIS), pers comm). In fact, 1998 saw the least amount of MYI within the Canadian Arctic Archipelago since the CIS began records 20 years ago. Changes have been observed around the globe. Since 1945 the net mass of 80% of the glaciers in the arctic, Canadian and European mountains have significantly declined with an associated rise in sea level (8 cm between the years 1800 – 2000) (Greg Flato, pers. comm.). The North American snow pack has also seen a decline in its duration and southern extent over the past 20 years; Canadian GCMs are in agreement with these observations (Boer et al. 2000). However, it should be emphasized that it has yet to be proven that the changes being experienced have been caused by anthropogenic forcing and not due to natural climate variability. It is noteworthy however that the recent IPCC report concluded that the balance of evidence now suggests that there is a discernable influence of humans on the global climate system (IPCC, 1996).

With the importance of sea ice in the global climate and recent observations of significant changes in the polar regions, there is a need for improving our understanding of polar processes and further advancing our GCM capabilities. Polar processes are

---

[3] Thermodynamic equilibrium thickness is the mean pack ice thickness resulting from a balanced ocean-atmosphere climate system (i.e. atmospheric cooling is balanced by ocean warming).
[4] Formal definition of the sea ice types cited here can be found in MANICE (Manual of Standard Procedures for Observing and Reporting Ice Conditions; www.ec.gc.ca/manice/title_pg.htm)

among the most poorly understood and require greater attention (IPCC, 1996). Although the worlds GCMs generally agree that the arctic will warm as time goes on, they do not agree on the magnitude of these changes. Deficiencies arise from poorly understood processes and inadequate model physics/parameterizations[5]. Progress in climate modeling will depend on new data sets and their utility for improving and validating critical processes and parameterizations (IPCC, 1996). To illustrate a deficiency with current GCMs, the treatment of snow and sea ice is typically a single homogeneous slab with bulk physical properties. We know from field observations (Makshstas, 1991) and modeling (Maykut, 1978) that this treatment is not sufficient since the energy balance can be very different for various ice thicknesses and ice types (Papakyriakou, 1999). Snow has a very important role in the seasonal evolution of sea ice due to its high insulation and reflective properties, making its relationship with sea ice complex (Brown and Goodison, 1994; Barber et al., 1994; Papakyriakou, 1999). Snow warms the ice in winter because of its low thermal conductivity (high insulating properties) that inhibits ice growth. In the spring and early summer, the snow acts to enhance ice growth by reflecting most of the solar energy (inhibits internal heating) with its high albedo (Ebert and Curry, 1993). Thus, the timing of a heavy snowfall in a particular location will have different effects on the seasonal evolution of sea ice. The internal physical changes within the snow are also important for determining the net thermodynamic effect on sea ice through seasonal energy balance variations (Jordan et al., 1999).

[5] Parameterization is a technique to simulate a certain parameter through empirical formulations or combination of empirical and physical methods, as opposed to using a purely physical relationship.

One-dimensional[6] thermodynamic models of snow (Jordan et al., 1999) and sea ice (Flato and Brown, 1996; Ebert and Curry, 1993) are being used to investigate more detailed processes that control arctic climate processes. These models have much more detailed physical processes within the snow and sea ice than GCMs, and results from these investigations can provide better physical parameterizations for GCMs in the future. Most numerical simulations of multi-year and first-year sea ice using one-dimensional thermodynamic models have been conducted using daily average or monthly external forcing (see for example, Maykut and Untersteiner, 1971; Ebert and Curry, 1993; Flato and Brown, 1996). Studies such as those of Maykut and Untersteiner (1971) and Ebert and Curry (1993) focused on equilibrium simulations for multi-year sea ice using coarse monthly climatological forcing. Gabison (1987) used monthly averages of climatological data to compare a modeled annual first-year sea ice cycle to an observed climatological ice cycle for different Arctic locations. More recently, Flato and Brown (1996) used daily forcing simulations of land-fast first-year sea ice to illustrate the role of snowfall and air temperature in driving interannual variations in maximum ice thickness and the timing of snow/ice ablation. The importance of the snow layer on first-year sea ice has also been indicated by observations (Bown and Cote, 1992; Brown and Goodison, 1994; Barber et al., 1994). However, even current one-dimensional models have major deficiencies and work is ongoing to improve different components within them. Another limitation of these models is that they can not be directly applied to microwave remote sensing since the variables that control microwave dielectrics (defined in section 2.4) are either not

[6] One-dimensionality refers to thermodynamic processes operating in the vertical coordinate throughout the ocean-ice, snow-ice, snow-atmosphere interfaces and within the snow and ice layers.

simulated or not simulated in the proper time scales. Further discussion of this is provided throughout the thesis.

A particularly important time of year in the arctic is the spring and early summer (May and June) where the initiation of snow and ice melt toward complete melt can be fairly rapid (Ebert and Curry, 1993). In addition, the onset and progression of melt have large inter-annual and regional variations (Barry et al., 1989; Barber et al., 1998). The arctic spring and summer sees the largest solar irradiance which greatly enhances melt. The initiation, extent and rate of snow and ice melt are critical on the arctic sea ice balance and climate (Barry and Maslanik, 1989; Ebert and Curry, 1993). A controlling variable in the arctic energy balance is the short-wave[7] surface albedo. Central arctic albedo ranges from 0.8 (snow surface) in early spring before melt begins then declines to near 0.4 once melt ponds[8] and bare ice are present (see for example, Robinson et al., 1986; Lindsay and Rothrock, 1993, 1994). As the albedo declines, more short-wave energy is made available for melt, in turn decreasing the albedo further. Numerical models have major difficulty attempting to simulate the surface albedo with respect to its spatial and temporal variability, a characteristic that must be improved before we can adequately simulate seasonal changes in sea ice. More albedo observations are needed toward this goal.

The surface energy balance (SEB) (defined in Chapter 2), in which the albedo is a component, is crucial to sea ice (see for example, Zhang et al., 1996; Steffen and DeMaria, 1996) and its variations are critical for understanding the year to year differences in snow/ice decay. A key component of the surface energy balance is incident

[7] Short-wave refers to the portion of the electromagnetic spectrum between wavelengths 0.3 - 4 μm.
[8] Melt ponds (ponds of water) form on the ice surface from accumulated melted snow and surface ice melt.

atmospheric radiation (short-wave ($K\downarrow$) and long-wave[9] ($L\downarrow$)) upon the snow and ice surfaces that are major energy (heat) sources. The distinction between the wavelength ranges is that the short-wave has its origin from the Sun, and the long-wave is of terrestrial origin. Radiative fluxes are usually two orders of magnitude greater than the ocean heat flux at the ice underside and planetary boundary layer turbulent fluxes (see for example Ebert and Curry, 1993). These characteristics result in radiative fluxes dominating the overall sea ice thermodynamics. The diurnal SEB can affect snow grain growth, snow brine volume and snow wetness through thermodynamic processes which in turn all affect passive and active microwave remote sensing (see for example, Barber et al., 1994; Barber et al., 1995a). Changes in snow properties also strongly affect the surface albedo that feeds back into the SEB. In the absence of an adequate arctic radiation network, accurate model simulations of sea ice thermodynamics therefore require accurate incident radiation parameterizations.

Active[10] and passive[11] microwave remote sensing (in particular, active satellite Synthetic Aperture Radar (SAR) and passive satellite microwave radiometry) are proving to be valuable tools for investigating and monitoring arctic sea ice seasonal evolution. This portion of the spectrum provides information on the physical and electrical state of the surface with the added advantage of all-weather, day and night observations. The SAR scattering (or resulting returned signal) and SSM/I-measured emission is dominated by the combined geophysical and electrical conditions of the surface in the micro and macro scales. The evolution of the scattering and emission is related to the overall

[9] Long-wave refers to the portion of the electromagnetic spectrum between 3.5 – 50 μm.
[10] Active remote sensing is a device which sends out a signal to a target and then receives a returned signal from that target.
[11] Passive remote sensing is a device which measures an emitted signal from a target.

thermodynamics of the system and the control of temperature on both the physical and electrical properties of the volume (Barber et al., 1994). Recent work has shown that the temporal evolution of SAR scattering can be used to infer some physical (e.g., ice type, snow grain size, perhaps snow thickness, snow salinity, melt pond formation, and melt pond fraction) and SEB (e.g. short-wave flux, albedo, long-wave flux and thermal conductivity) attributes of the polar climate (see Barber et al., 1995a; Barber et al., 1998). Similarly, passive microwave sensors are valuable for monitoring sea ice extent and concentration throughout the polar regions (see for example, Weaver and Troisi, 1996; Weaver et al., 1987). The relationship between SAR and SSM/I signatures and the thermodynamic sea ice system brings about new ideas of how we can better understand the physical processes involved. Utilizing numerical techniques and remote observations in tandem can lead to improved ways in understanding and monitoring polar processes.

Considerable progress has been made in both the modeling and remote sensing communities dealing with sea ice processes, however, there are still major deficiencies with the models and research is still required to better understand the physical processes behind what we see in microwave satellite imagery.

The intent of my dissertation is to begin the process of linking thermodynamic models of snow and sea ice with satellite microwave remote sensing. A major impetus for combining remote sensing into sea ice numerical process models is to improve our understanding and monitoring of the arctic climate system where very few observations are available otherwise. Other research areas have also indicated the need for combining remote sensing with numerical models. Gower (1995) expressed that remote sensing products are essential to ocean research and "because of the major importance attached to

modeling (the earth's oceans), the WOCE (World Ocean Circulation Experiment) requires a major effort in assimilation of remote sensing data into the models". In the agricultural community, Brown (1995) has indicated that "remote sensing will be of great value for determining soil moisture over a wide area for initial soil moisture input to vegetation development models".

The integration of remote sensing and numerical sea ice models could lead to a new approach in looking at spatially and temporally complex, highly interrelated arctic cryospheric processes. The remote sensing data can also play a role in the initialization and validation of numerical sea ice models. Furthermore, remote sensing could become an integral component of the sea ice model creating a hybrid numerical model which explicitly considers spatial interrelationship of processes within the model environment. Remote sensing can also be used to replace some model parameterizations or constrain elements within the system. There are two main advantages of these approaches; 1) the combined use of remote sensing and modeling may be used to simulate as realistically as possible those variables that can not be derived from remote observations alone, such as the distribution of sea ice thickness, 2) using remote sensing in place of some model parameterizations, other processes in the model can be examined in more detail, such as, if surface albedo is well handled by remote observations, other variables/processes may be held accountable for "erroneous" model simulations. Another approach is through microwave forward and inverse scattering models that have advanced to a state that makes them valuable supplementary tools for understanding the remote observations (see reviews by Golden et al., 1998a,b) which in turn can be linked to thermodynamic models.

We are now in a position to begin the first steps of directly linking thermodynamic modeling with microwave remote sensing through the interactions of thermodynamics, dielectrics and the thermophysical properties of the snow/sea ice system. This is termed an 'electro-thermophysical' relationship and in what follows I explain how the thermodynamic and geophysical properties of the snow and sea ice control their microwave electrical and backscatter characteristics.

In the 'electro-thermophysical' framework, I strive to develop a model which couples knowledge of the scattering physics of the surface with a one-dimensional (1-D) thermodynamic model of the sea ice/snow system. The advantage of this approach is that the geophysical and electrical properties of the surface give rise to scattering and emission in the microwave portion of the spectrum; changes in the geophysical and or electrical state of the volume are driven by the thermodynamics across the interface. A model of these relationships would allow us to measure both the geophysical state of sea ice at any particular time and also evaluate the temporal evolution of this scattering over a period coinciding with the seasonal evolution of the surface. Since the SEB also evolves directly from the thermodynamic and physical state of the system, we also expect to be able to determine proxy measures of certain SEB state variables such as surface temperature, albedo, long-wave flux, surface atmospheric drag, etc. This information can also be useful in determining the timing of accretion and ablation and producing estimates of the mechanical strength of sea ice and/or timing of fast ice break-up.

My work includes a number of aspects dealing with snow covered FYI. This includes the critical processes that control the thermodynamics (growth and ablation) of snow covered smooth first-year sea ice (FYI) in a one-dimensional sense and how they

relate to microwave remote sensing. I limit my work to FYI not only to narrow my focus but to include a type of ice that is studied to a lesser degree than multi-year sea ice (MYI). In addition, at the current rate of arctic MYI depletion, there will be more FYI occupying the arctic oceans in the future. Numerical modeling and microwave remote sensing are also proving to be more useful tools for FYI. Smooth FYI (as opposed to rubbled and rafted FYI) is less complicated to study in terms of its flatter surface and more homogeneous physical structure, making it easier to simulate in numerical modeling and identifying relationships between ice thermodynamics and microwave remote sensing. Hence I also assume that the FYI is not a dynamic volume (i.e. the ice is not moving horizontally due to ocean currents or wind drag) which is typical of most FYI in the Canadian Arctic Archipelago (or fast-ice). This lends credence to the one-dimensionality assumption.

## 1.2 Thesis Objectives

The goal of my research is to conduct the initial steps necessary for combining a one-dimensional thermodynamic snow sea-ice model and microwave remote sensing that is related to arctic snow and sea ice processes. I eluded to several issues that need to be addressed and would advance the numerical model to a state where remote sensing could be better linked into the model environment. These include: 1) ensure the existing one-dimensional sea ice models treat the SEB and snow/ice thermodynamics in the appropriate time scales we see occurring in field experiments, 2) ensure the snow/ice thermodynamics are not compromised by differences in environmental and spatial representation within components of the SEB, 3) ensure that the snow layer is properly

handled in the modeling environment, and 4) how we can make use of satellite microwave remote sensing data with the model environment.

My specific research questions, which evolve directly from these issues are:

(1) Does the use of hourly and daily variable forcing result in a different simulation of the annual cycle of ice thermodynamics? If so, can the model parameterizations adequately represent specific processes over these shorter temporal scales, and is there a systematic bias in forcing the model with 'land based' versus 'on ice' measurements of input fluxes?

(2) How do selected parameterizations of $K\downarrow$ and $L\downarrow$ compare to *in situ* data and what are the characteristics of newer schemes that may be required?

(3) What are the broadband and spectral albedos over melting FYI as well as their local scale spatial variation (sub-km scale)? How do the surface albedos scale up to semi-regional (10's km) and regional (100's km) scales? What is the sensitivity of sea ice ablation to percent pond fraction and the associated spatial variability in surface albedo?

(4) Are there advantages of using a more sophisticated snow component within a coupled 1-D snow sea-ice model over a similar model that uses a single bulk property snow layer for annual cycles of snow-covered FYI?

(5) How can 1-D numerical models of snow-covered FYI be useful for microwave remote sensing applications?

In question (1) I examine whether hourly forcing (i.e. time step) results in a different simulation of ice thermodynamics than daily average forcing over an annual cycle using the sea ice model of Flato and Brown (1996). From field experiments, significant feedbacks operating between the atmosphere and the surface (e.g., albedo-snow grain growth; cloud cover-long-wave flux, etc.). The snow/ice properties do not scale linearly with radiative and turbulent fluxes in the SEB. I wish to examine this 'temporal scale' issue from both the perspective of parameterization of the model physics and integration of the thermodynamics over an annual cycle. In addition, I address the existing model parameterizations relative to *in situ* sea ice field observations made in the Canadian Archipelago between 1992 and 1993 during the Seasonal Ice Monitoring and Modeling Site (SIMMS) field programs (LeDrew and Barber, 1994). Specifically, I consider the model parameterizations of incident short-wave/long-wave radiation and albedo. This component will indicate how well the physics are handled within the existing parameterizations and help to ascertain where (if any) changes should be made when using shorter temporal scales. Finally, I am interested in differences in forcing which arise when using nearby land station versus *in situ* sea ice observations. Because regular annual observations do not exist over the ice, arctic meteorological observations from land stations are typically used to force sea ice models. However, local climate differences between on-ice and land are significant. This addresses the differences in ice simulations between using land-based forcing and on-ice forcing.

Question (2) deals with environmental and spatial differences in the incident radiative fluxes. That is, how accurate are the current incident radiation parameterizations for different arctic environments? For this analysis, I use unique *in situ* observations

taken in an arctic polynya[12] over terrestrial, fast-ice, marginal ice zone and open water environments during the 1998 International North Water (NOW) Polynya Project (Barber et al., 2001).

In question (3), I address the issue of albedo measurements (broadband and spectral) over a variety of FYI surfaces (including melt ponds) due to the dearth of information currently available. I then scale up the surface albedo measurements to regional scales (100's km) using aircraft aerial video and satellite data to show melt pond spatial variability and its albedo over FYI. Then I examine how the thermodynamic state (and mechanical strength) of the sea ice is affected by variations in fractional pond cover. In effect I am asking the question: "How important is it that the magnitude and spatial pattern of surface albedo (as examined above) is correctly parameterized within the annual cycle of sea ice ablation?" Measurements were made during C-ICE97 (Collaborative - Interdisciplinary Cryospheric Experiment 1997) conducted in Wellington Channel and Lancaster Sound, Nunavut.

Question (4) investigates model simulation differences in snow and ice physical and thermal evolution between a multiple layered coupled snow sea-ice model and a similar model that uses only a single bulk property snow layer. The one-dimensional mass and energy balance snow model (SNTHERM; Jordan et al., 1999) and sea ice model (Flato and Brown, 1996) are coupled at the snow-ice interface where fluxes of heat and radiation are transferred. Both models are coded in standard FORTRAN. The Simulation comparisons were conducted over a full annual cycle of FYI between 1992-93 as well as a specific shorter time period in 1992 to compare the models against measured

[12] A polynya is a body of water that remains primarily ice-free during the arctic winter months.

field data. This allowed a very detailed investigation into whether using a more sophisticated model is warranted for snow-covered FYI annual cycles.

The final question (5) illustrates the direct application of the coupled snow sea-ice model toward microwave remote sensing interpretation. The coupled model is linked to a microwave dielectric model; the amalgamation of the models produces a type of Electro-Thermophysical model of the Snow Sea-Ice System (ETSSIS). ETSSIS is directly compared to measured physical, thermodynamic and microwave dielectric variables to show how well and when the model reproduces them.

## 1.3 Thesis Structure

I begin by providing a discussion of the critical processes that control the thermodynamics within the FYI and snow surfaces over a range of temporal and spatial scales (Chapter 2 – section 2.1 and 2.2). Chapter 2 (section 2.3) also discusses how we can use numerical modeling to understand the critical processes, what processes are handled in these models, and what some of the deficiencies are. I limit the discussion primarily to one-dimensional thermodynamic models. Chapter 2 (section 2.4) outlines the utility of microwave remote sensing toward FYI thermodynamics. In particular, what processes we can observe in microwave imagery over the course of an annual cycle of FYI. Chapter 2 provides a thorough background to following chapters that address the five objectives in as many chapters. Research question 1 is discussed in Chapter 3 with the results being previously peer reviewed and published (see Hanesiak et al., 1999). Question 2 appears in Chapter 4 with the results being previously peer reviewed and

published (see Hanesiak et al., 2001a). Question 3 is addressed in Chapter 5 with the results being previously peer reviewed and published (see Hanesiak et al., 2001c). Chapter 6 contains results from question 4. Finally, question 5 results appear in Chapter 7 and in peer review for publication during the final version of this dissertation. The thesis concludes with a summary of results and research that should be conducted in the future (Chapter 8).

The usefulness of this work will be to advance our understanding and monitoring abilities of arctic snow covered sea ice processes. This will enable the climate modeling community to utilize new insights in future GCM implementation strategies. It can also be used toward operational sea ice applications when arctic ice decay becomes important for ocean navigation. Lastly, ETSSIS can be used to assist in the interpretation of satellite microwave remote sensing, a type of model that is unique for this purpose.

# CHAPTER 2: Background

## 2.1 First-Year Sea Ice and its Snow Cover Characteristics

This section describes the critical thermodynamic processes within snow covered first-year sea ice over a variety of temporal and spatial scales. Sea ice has several stages of growth and development, classified by the World Meteorological Organization (WMO) and described in detail by the Manual of Standard Procedures for Observing and Reporting Ice Conditions (MANICE) (Figure 2.1). I pay particular attention to first-year sea ice (FYI) and begin with a brief description of the physical mechanisms of ice growth and snow evolution. An in depth discussion of snow and ice thermodynamics follows.

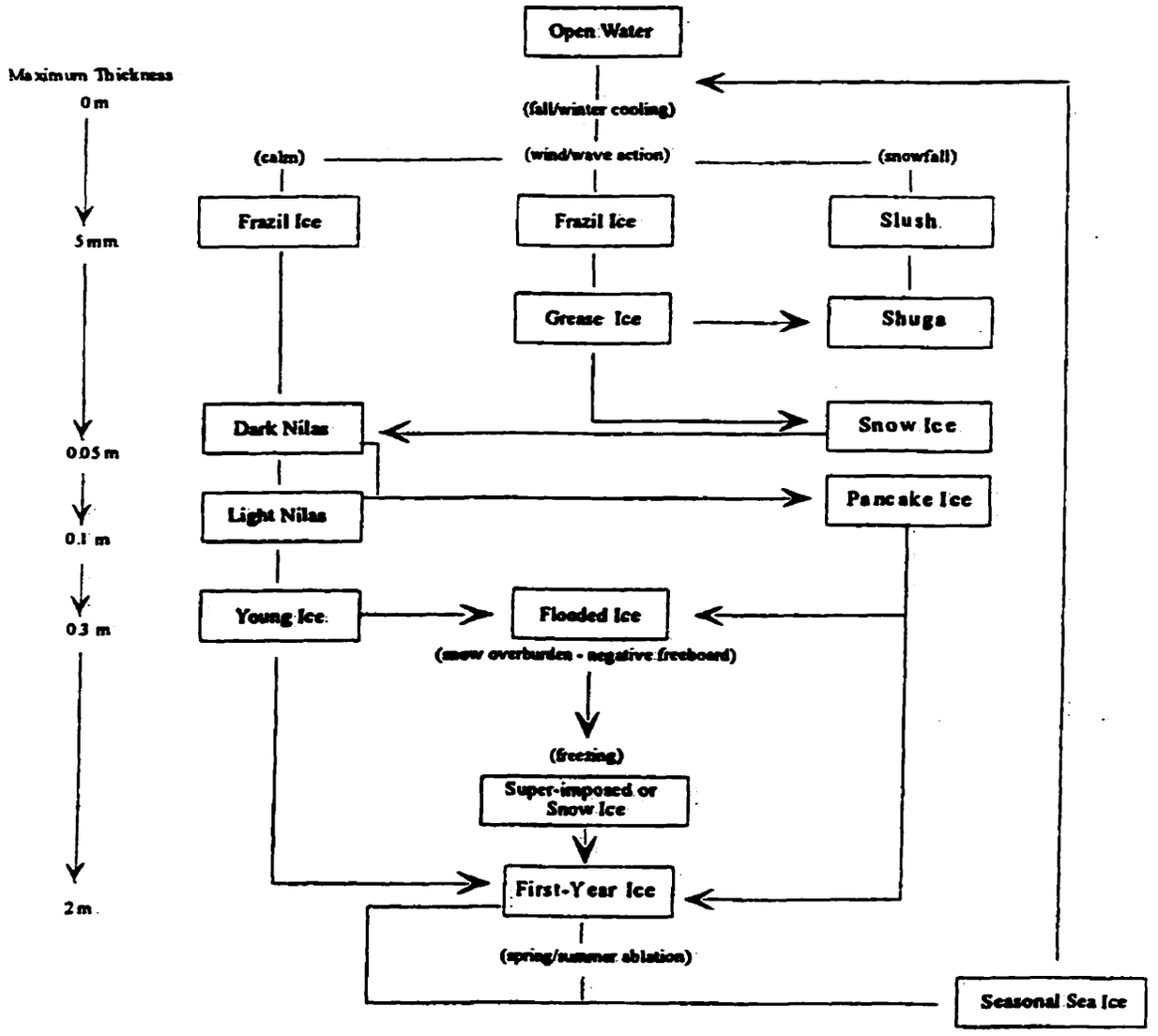



*Figure 2.1: Stages of first-year ice growth, its thickness and external forcing (in brackets). Modified from Papakyriakou (1999).*

New ice begins as frazil or slush (Figure 2.1) in the fall period during freeze-up (in temperatures ≤ -1.8°C) where heat is lost from the ocean to the atmosphere typically under windy conditions. Sea water has an average salinity of 32 ppt (parts per thousand), depressing its freezing temperature to below 0°C. As ice growth continues, grease ice and shuga (Figure 2.1) are formed (mixture of ice crystals and sea water), and continues to grow into more consolidated ice forms (nilas and pancake ice) over time. Conduction of

heat continues from the bottom to the top of the ice forming thicker young ice (0.3 m; Figure 2.1). As young ice grows, vertical brine inclusions (mixture of salt and water) form between ice columns that are near 1 mm in width but vary in size according to growth rate (Figure 2.2) (Lofgren and Weeks, 1969). Salt that is trapped in the ice structure is either in liquid brine form or solid depending on the ice temperature (Lock, 1990). Ice growth continues as a columnar structure to typical FYI thicknesses (30 – 200 cm). FYI is usually made up of 5-25% frazil ice (top layer) and 75-95% columnar ice (Figure 2.2) (see for example, Weeks and Gow, 1980; Tucker et al., 1987). Typical FYI densities are 880-910 kg $m^{-3}$ with the variation primarily caused by air and salt content (Fukusako, 1990). Density increases with increased salinity (if bubble content remains constant) and decreases with increased porosity (Fukusako, 1990). This makes the average bulk ice thermal conductivity near 2.1 W $m^{-1}$ $C^{-1}$. All of the ice forms vary widely over large spatial areas depending on proximity to land, other ice formations, atmospheric wind and thermal conditions. This creates a myriad of ice conditions and ice thickness once full consolidation takes place. Dynamic processes (ocean currents, wind and ice contraction/expansion) also create various degrees of rough ice conditions (ridged, rafted and rubbled ice) with pieces of ice extending upward from 0.5 – 3 m (not including MYI floes). However, I am specifically interested in smooth FYI.

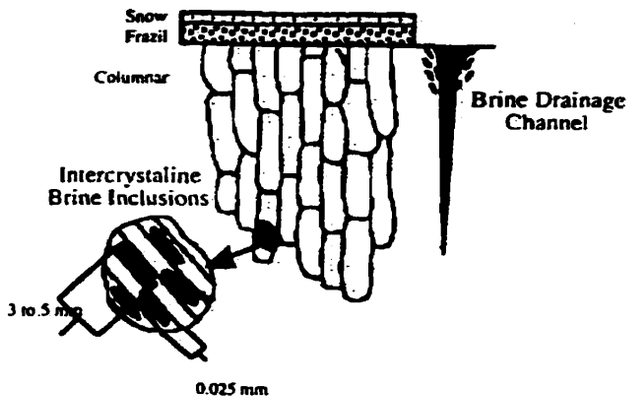



*Figure 2.2: First-year sea ice physical characteristics showing the frazil and columnar positions relative to intercrystalline brine pockets and macroscale brine drainage channels. Adapted from Vant et al. (1978).*

The brine inclusions are very important for heat conduction and latent heating within the sea ice (shown later). Brine is forced upward and downward within the ice layer as it grows. Brine and salt impurities that are forced upward out of the ice surface during initial ice growth form frost flowers (deformed tree-like structures that extend vertically on the ice surface between 2 – 6 cm). This results in the bottom snow layer (once snow has fallen) containing brine as well (see for example, Barber et al. 1995a) significantly affecting snow thermal properties (Papakyriakou, 1999). However, brine drainage is primarily of gravitational origin moving downward along the brine drainage channels (Bennington, 1963) and is ongoing over time as the ice grows. Brine pockets interconnect forming the tree-like brine channel within the ice from cm's to tens of cm's long (see Figure 2.2) (Lake and Lewis, 1970). The resulting FYI salinity profile over time appears as C-shaped (Figure 2.3). The rejection of brine out of the bottom of the sea ice has global ocean circulation ramifications. The more dense sea water at the ice underside creates an unstable ocean layer that generates the thermohaline ocean circulation and is a

fundamental process by forcing and maintaining global and regional ocean circulations (Aagaard et al., 1981).

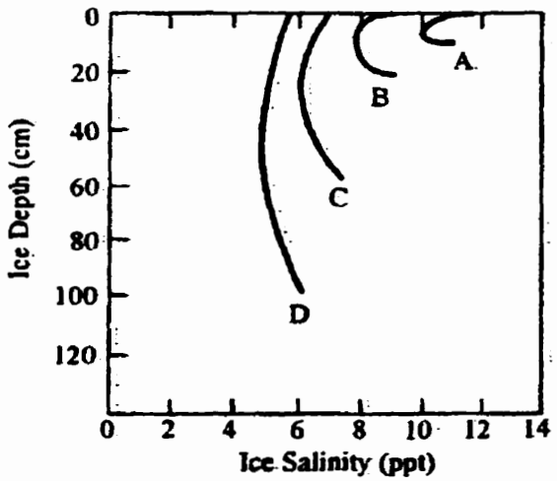



*Figure 2.3: Typical salinity profiles for growing first-year sea ice. Curves A-D represent ice thickness between 10 to 100 cm. Adapted from Maykut (1985).*

During the melt season (April – August), the available liquid water from the snow pack and melted top layer of ice forms melt ponds on the ice surface (Figure 2.4). Aerial melt pond fractions vary dramatically with time and location over the course of the melt season. Melt ponds reduce the surface albedo significantly with maximum pond aerial fractions near 75% on smooth FYI (see for example, Barber and Yackel, 1999). Pond fractions tend to decline to values between 35-50% once drainage takes place due to seal holes and cracks in the ice (Barber and Yackel, 1999). FYI typically breaks-up in summer due to dynamic effects and rarely completely ablates in the same location where it originally formed.

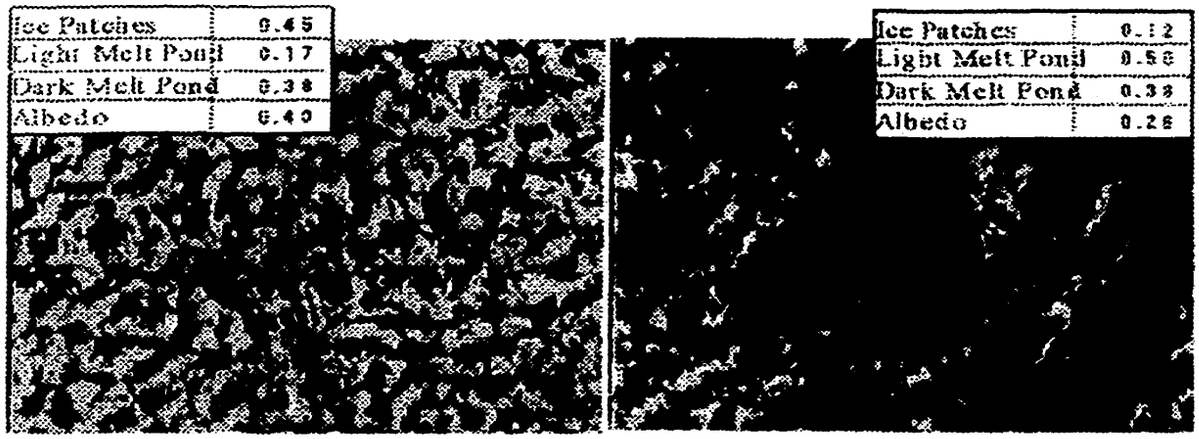



*Figure 2.4: Examples of two melt pond first-year sea ice surfaces measured with aerial video (adapted from Barber and Yackel, 1999). Values in the tables suggest the average fractional areas of ice/snow patches, light colored ponds, dark colored ponds and aerial averaged albedo for both surfaces.*

Snow becomes an important physical presence on sea ice once ice has formed. Snow is a porous mixture of ice grains of either single crystals or clusters of crystals (Alford, 1974) and has average bulk densities between 50 – 400 kg $m^{-3}$. This makes the average bulk snow thermal conductivity considerably lower than FYI (0.4 W $m^{-1}$ $C^{-1}$). The pores between ice crystals are filled with air, water vapor and liquid water if the snow is wet and/or melting (Colbeck, 1982). The snow layer and its physical properties are highly variable in space and time, controlled by various factors. Immediately after deposition, the snow transforms into a discrete layered medium through metamorphic[13] processes, overlying snow weight (overburden), freeze-thaw cycles, and wind transport (see for example Jordan et al., 1999). Density typically increases with depth provided there are no appreciable temperature gradients in the snow (Colbeck, 1982; Sturm, 1991). Most of what we know about snow and its internal physical changes is largely based on terrestrial surfaces (e.g. Alford, 1974; Anderson, 1976; Colbeck, 1982). Snow on sea ice

is different than terrestrial snow due to salinity and oceanic heating conducted upward within the ice (see for example, Drinkwater and Crocker, 1988; Barber et al., 1994). From field observations, we typically see a vertical salinity gradient in the snow between 10-17 ppt at the bottom-most snow layer and tapers off non-linearly to 0 ppt by 10-25 cm from the snow-ice interface. This reduces the snow thermal conductivity in the brine-laden regions by at least 100% in most cases (Papakyriakou, 1999). With such a low thermal conductivity, snow significantly suppresses (enhances) ice growth in the Fall/Winter (Spring) by insulating the heat transfer between the ocean-ice (atmosphere) and the atmosphere (ice) (see for example, Flato and Brown, 1996).

Metamorphism is a critical process that snow undergoes as it ages and evolves. There are two types of snow metamorphism, dry and wet, depending on temperature, density and liquid water profiles in the considered snow layer (Colbeck, 1973, 1980, 1983). Snow metamorphism drives snow cover evolution and affects snow density, porosity, strength, albedo, and thermal conductivity (Brun et al., 1992). Once a fresh snowfall occurs over sea ice, it rapidly undergoes metamorphic changes under a dry snow scenario, unless rain or warm air temperatures complicate the process.

Dry snow metamorphism implies that no liquid water is present in the snow pack, temperatures are sub-freezing, and the snow is in solid state in equilibrium with vapor. The equilibrium crystal shape (that is highly dependent on the environmental conditions) dominates the grain shape in the absence of strong temperature gradients. The growth rate of snow grains is limited by vapor diffusion that is driven by the vapor pressure gradient (Alford, 1974). The vapor pressure gradient is controlled by temperature, radius of

[13] Metamorphic processes are changes in snow morphology that take place as functions of temperature and pressure.

curvature and size of each snow grain. For example, warmer snow pack areas near the ice surface in winter can hold more vapor than colder areas at the snow surface. This produces a vertical vapor pressure gradient in which vapor travels upward through the snow pack. In addition, smaller and more curved snow grains contain higher vapor pressures than larger and less rounded grains (Colbeck, 1987). This implies the larger snow grains tend to grow at the expense of smaller grains (Colbeck, 1973).

There are two main types of dry snow metamorphism, equitemperature (ET) and temperature gradient (TG) processes. ET is defined by a small or no temperature gradient in the snow (0 to ~7°C $m^{-1}$) usually seen in spring, whereas TG has very high temperature gradients (15 – 70°C $m^{-1}$) typical of a winter snow pack (see for example, Colbeck, 1982). In dry snow ET metamorphism, snow grains reduce their surface free energy tending toward their stable state that is controlled by their radius of curvature and the surface to volume ratio. Snow grains naturally reduce their surface to volume ratio by expelling excess vapor. Both processes (curvature and surface to volume reduction) act to produce more rounded grains and increase snow density (Alford, 1974). Rounded crystals are typically small between less than 0.1 mm to 1.0 mm.

In dry snow TG metamorphism, the rate of vapor transport is very fast (kinetic growth) that builds angular, highly faceted grains that bond very poorly (Colbeck, 1982) (Figure 2.5). This process increases the crystal size but decreases the crystal number density (Colbeck, 1987, 1991) resulting in weak snow strengths and decreases in density. It should be mentioned that convective processes (air movement) within the snow pack also complicates the grain growth process and is largely unexplored, however, it is believed to be a non-negligible factor (Colbeck, 1989). Convective processes are difficult

to measure and are therefore ignored in detailed snow models. A wide range of faceted crystal types are found in the snow cover, including complex scrolls, cup-shaped crystals, plates, sheaths, and needles (Bradley et al., 1977). Faceted crystals can grow up to 3-4 cm in length (see Figure 2.5) and are commonly found in the hoar snow layer. The hoar layer is defined by larger faceted snow grains near the base of the snow pack where warmer temperatures and high temperature gradients exist, or within the snow volume where two discontinuous density layers exist. The bottom layer of the snow is typically termed the basal layer. Another type of faceted crystal known as surface hoar can develop at the snow-atmosphere interface. These crystals form by condensation from the atmosphere that is supersaturated compared to the snow surface. This can occur on cold clear arctic nights.

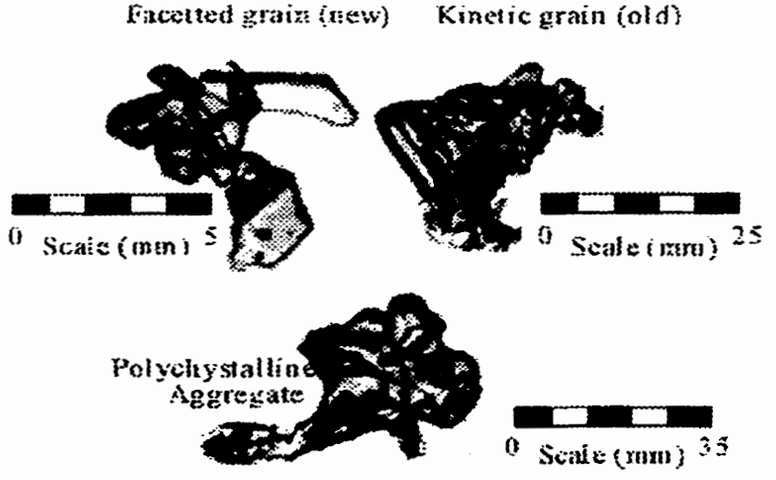



*Figure 2.5: Typical new (faceted) and old (kinetic) snow grains and aggregates of partially melted grains (polycrystalline) that result from the metamorphic processes described in the text (Adapted from Barber et al., 1999).*

Wet snow metamorphism is distinguished by low and high liquid water contents, typically termed the pendular and funicular regimes, respectively. In the pendular regime (1-8% liquid by volume), air makes continuous channels through the snow while liquid water travels between the crystals and air (Colbeck, 1982). Grain growth rates are accelerated due to liquid water with the small grains preferentially destroyed and larger grains become more rounded and well bonded (polycrystalline aggregate in Figure 2.5) (Colbeck, 1982). Snow strength is relatively high in the pendular regime since no melting occurs at the grain contacts and is further increased if the snow pack undergoes re-freezing (firnification). Conduction of heat within snow is increased if re-freezing takes place. In the funicular regime (8-15% liquid by volume), water occupies the pore spaces with little air and drains freely toward the ice surface (Colbeck et al., 1990). Grains are surrounded by water and have difficulty adhering, unlike the pendular regime. The snow pack shrinks in depth due to snow mass loss and also loses its overall mechanical strength as liquid water creates vertical channels within the snow. The smaller particles rapidly disappear once reaching a critical size (Colbeck, 1973). Increases in mean grain size in "wet" conditions can occur from about 0.3 to 0.8 mm in six days, and from 0.21 to 1.78 mm in nearly 40 days (Colbeck, 1982, 1986). In the funicular regime, since the snow is saline, brine is flushed toward the ice surface reducing the snow's layered salinity profile, forming slush at the base of the snow pack (called basal ice if refrozen).

## 2.2 Thermodynamics of Snow and Sea Ice

The atmosphere is the primary force behind the growth and decay of FYI with the surface energy balance (SEB) dominating. Ocean heat fluxes also contribute to ice

evolution while ice is present to counteract and balance atmospheric forcing (to produce an equilibrium ice thickness in MYI). However, the ocean heat flux is assumed to be somewhat constant and relatively small (2-30 W $m^{-2}$) compared to atmospheric forcing for the type of ice under study here (Shirasawa and Ingram, 1997; Flato and Brown, 1996). A graphical depiction of the SEB and the thermodynamic variables involved in the snow and ice layers outlines the major components (Figure 2.6).

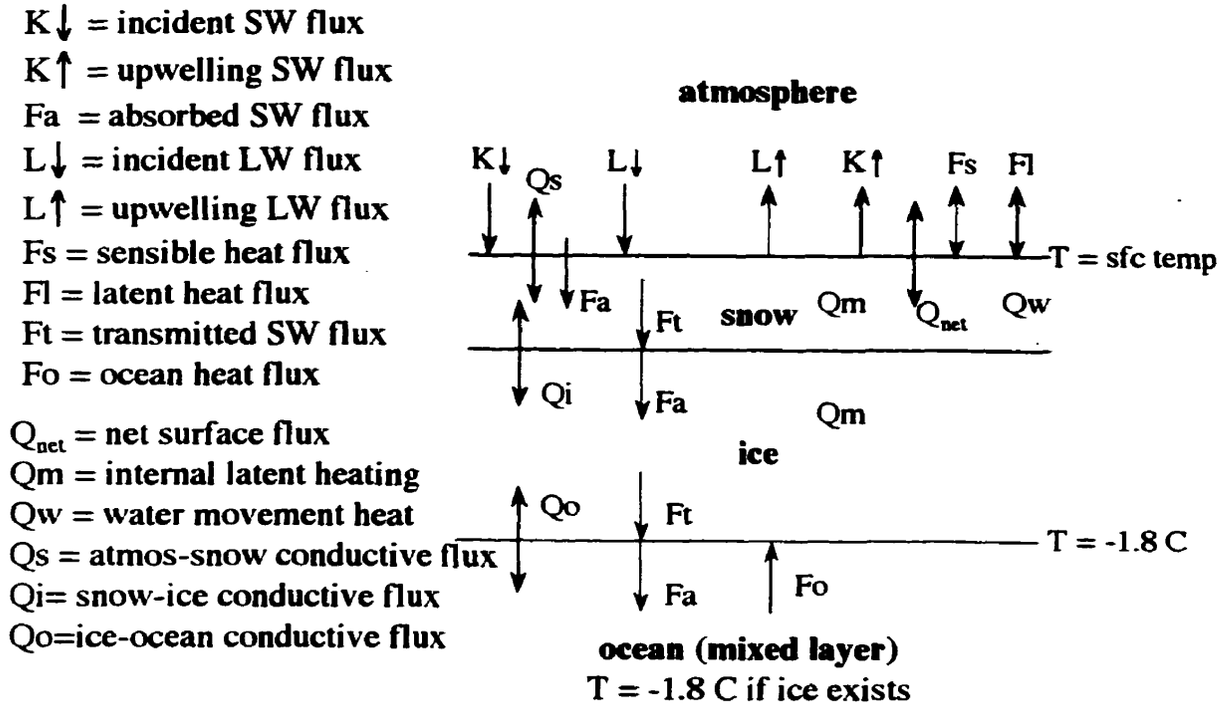



*Figure 2.6: Depiction of the critical physical thermodynamic processes within snow covered first-year sea ice including the surface energy balance (SEB), internal snow/ice processes and conductive fluxes across each interface.*

The corresponding SEB equation appropriate for modeling snow over sea ice (Jordan et al., 1999) is (positive from the atmosphere toward the surface),

$$Q_{net} = K\downarrow - K\uparrow + L\downarrow - L\uparrow + F_s + F_l + F_{prec} \quad (2.1)$$

where, $Q_{net}$ is the net atmospheric heating of the snow or ice surface, $K\downarrow$, $K\uparrow$, $L\downarrow$, $L\uparrow$ are the incident and reflected short-wave radiation, incident and up-welling long-wave radiation, respectively. $F_s$, $F_l$ and $F_{prec}$ are the sensible and latent heat fluxes and heat conducted by precipitation, respectively. The surface short-wave albedo is defined as the ratio of reflected solar radiation over the incident solar radiation,

$$\alpha = K\uparrow / K\downarrow \tag{2.2}$$

The energy within the snow volume can be estimated by (Papakyriakou, 1999),

$$dQ_s + Q_i - Q_s + Fa + Q_{ms} + Q_w = 0 \tag{2.3}$$

where, $dQ_s$ is the net atmospheric heating ($Q_{net} + F_s + F_l$), $Q_i$ is the conductive heat flux at the ice surface, $Q_s$ is the conductive heat flux at the snow surface, Fa is the absorbed solar radiation, $Q_{ms}$ is the phase transition energy, and $Q_w$ is heat energy associated with water flow . Similarly, the energy contained within the ice volume is,

$$Q_o - Q_i + F_{ai} + Q_{mi} + F_o = 0 \tag{2.4}$$

where, $Q_o$ is the conductive heat flux at the ice-ocean surface, $F_{ai}$ and $Q_{mi}$ are the absorbed solar radiation (in ice) and internal latent heating of the sea ice layer, respectively, and $F_o$ is the heat flux from the ocean mixed layer toward the ice underside. I now continue with individual critical physical processes that control snow and sea ice thermodynamics, beginning with atmospheric forcing down to the ice underside.

### *2.2.1 Atmospheric Radiation Interaction with Snow and Ice*

The solar short-wave (SW) portion of the spectrum occupies wavelengths between 0.15 – 3.0 μm. This is subdivided into visible wavelengths (0.36 – 0.75 μm), near-infrared (0.7 – 1.3 μm), and mid-infrared (1.3 to 3.0 μm). The solar component of the arctic energy balance is seasonal due to the Earth's obliquity. The maximum solar zenith angle[14] in the arctic during the summer solstice is limited to about 47 degrees because of the earth's obliquity. High latitude locations (>66.5 degrees) are therefore in a state of darkness for up to six months of the year. Only when the sun rises above the horizon for longer periods of time through the spring does the short-wave radiation become increasingly dominant. This dominance continues through the spring and summer months when, at high latitudes, the sun does not drop below the horizon at all.

Long-wave radiation (3.0 – 100 μm; thermal infrared) is of terrestrial origin, with incident long-wave (L↓) emitted from the atmosphere all of the time. Up-welling long-wave radiation (L↑) is emitted from the surface and is also a continual annual process.

#### *2.2.1.1 Incident Short-Wave Radiation*

Incident SW (K↓) responds to seasonal, latitudinal, diurnal and atmospheric variations that significantly alter the arctic environment. The K↓ variations are highly variable due to rapid changes in cloud cover, water vapor, aerosols, and the relatively slower changes in surface features. The K↓ parameter is most sensitive to aerosol optical

depth and surface albedo under clear skies and cloud optical depth, cloud liquid water content and surface albedo under cloudy skies (Key et al., 1996; Leontyeva and Stamnes, 1993). $K\downarrow$ is a key component in the SEB and accurate measurements and/or model parameterizations are required to numerically simulate seasonal snow and ice evolution. More will be said about $K\downarrow$ in following sections.

### *2.2.1.2 Surface Reflection (Albedo)*

Surface albedo is defined in equation (2.2). Several factors affect albedo, including wavelength, solar zenith angle, surface characteristics, and clouds (Ebert and Curry, 1993). In sea ice climatology, the albedo of the surface (with snow and/or melt ponds on it) is typically measured in the optical region of the spectrum over wavelengths between 0.3 – 3 μm. However, the most responsive region within this spectrum to changes in the sea ice surface is the visible (VIS) and part of the near infra red (NIR) region (0.3 – 1.1 μm). Surface features such as a new snow cover, old snow, wet snow, saturated snow, and melt ponds, determine the overall albedo magnitude over first-year sea ice. Fresh snow acts as a highly reflective surface in the VIS and NIR whereas, older snow absorbs more of the VIS and NIR energy. The direct albedo of snow can show considerable daily variation (Petzold, 1977). This is due to several factors. For instance, a snow pack will slowly settle under its own weight, changing the size, shape, and spacing of crystals at the surface. Dust particles from the atmosphere will also accumulate on the snow pack darkening the surface and lowering the albedo (Warren and Wiscombe, 1981).

[14] Solar zenith angle is the angle between the local normal to the Earth's surface, and a line between a point

Wind is a significant factor since it can redistribute and compact snow. A 1 cm layer of freshly accumulated snow will result in a new albedo that is completely independent of that for the previous surface (Petzold, 1977). Finally, a change from powdery to granular snow occurs at the surface of a melting snow pack. This means that the albedo of a melting snow pack is going to change more rapidly than an accumulating snow pack (Petzold, 1977). When the snow melts in spring, melt ponds develop on the sea ice surface and significantly decreases the overall surface albedo. The formation of ponds and the associated drop in albedo causes further rapid sea ice ablation through the ice-albedo-feedback process (Curry et al., 1995). The extra SW radiation absorbed by the melt ponds is used to increase their depth and the volume of brine within the ice (Maykut, 1982).

Clouds and solar zenith angle also affect the surface albedo. Maximum total cloud cover occurs in the summer (up to 90%), minimum coverage occurs in the winter (40-50%), and the transitions between maximum and minimum values occur over short time periods in the spring and fall (Curry and Ebert, 1992). Clouds influence $K\downarrow$ and hence surface albedo in three ways: 1) clouds reflect a portion of the radiation that they intercept back into space, that is, they have their own albedo, 2) clouds selectively absord SW radiation at wavelengths greater than 0.7 microns, and 3) cloud cover controls the relative proportions of direct and diffuse radiation. Since the albedo of snow is greater for visible light than for near infrared radiation, albedo generally increases as cloud cover and cloud opacity increase (Petzold, 1977). Furthermore, the albedo of surfaces that do not scatter radiation equally in all directions (non-Lambertian reflectors) will increase as

---

on the Earth's surface and the sun.

the diffuse component increases (Curry and Ebert, 1992). Solar zenith angle effects show an 8-20% decrease in clear sky albedo as the zenith angle decreases from 80 - 45 degrees (Petzold, 1977).

Sea ice albedo is a critical climatological parameter controlling short-wave responses in the SEB, and has been a major focus in several modeling and observational studies (see for example, Shine and Henderson-Sellers, 1985; Ebert and Curry, 1993; Robinson et al., 1986; Ross and Walsh, 1987; De Abreu et al., 1994; Grenfell and Maykut, 1977; Grenfell and Perovich, 1984). Perovich (1996) summarized many of the scarce bulk (broadband[15]) and spectral[16] albedo measurements collected *in situ* for various surfaces over different kinds of sea ice. Broadband albedos range from 0.21 for melt ponds to 0.87 for dry new snow over first year ice (Table 2.1).

*Table 2.1: Previous measurements of integrated albedo over various surfaces (data adapted from Perovich, 1996).*

| Surface Type | Integrated Albedo |
|---|---|
| Open Water | 0.05 |
| Old Pond | 0.15 |
| Ponds (1st Year) | 0.21 |
| Mature Pond | 0.29 |
| Melting Ice (blue) | 0.33 |
| Refrozen Pond | 0.40 |
| Bare First-Year Ice | 0.52 |
| Melting White Ice | 0.56-0.68 |
| Frozen White Ice | 0.70 |
| Melting Snow | 0.77 |
| Wind Packed Snow | 0.81 |
| New Snow | 0.87 |

[15] Broadband albedo refers to the integrated spectral albedo across the optical wavelengths.
[16] Spectral albedo refers to the reflectance of energy at a particular wavelength in the optical spectrum.

Spectral snow albedos are typically highest in the VIS and drop off in the NIR (Figure 2.7). The magnitude and shape of the spectral albedo curves are strongly dictated by the snow grain size and amount of liquid water in the snow (Wiscombe and Warren, 1981). Few albedo measurements exist from the Canadian Archipelago compared to the Arctic Basin and Antarctic. De Abreu et al. (1995) measured first-year sea ice (FYI) spectral albedo in the Canadian Archipelago during the melt season showing the gradual decline in albedo is attributed to the evolving geophysical structure of the snow/ice volumes, increasing the contrast between the VIS and NIR spectral albedo.

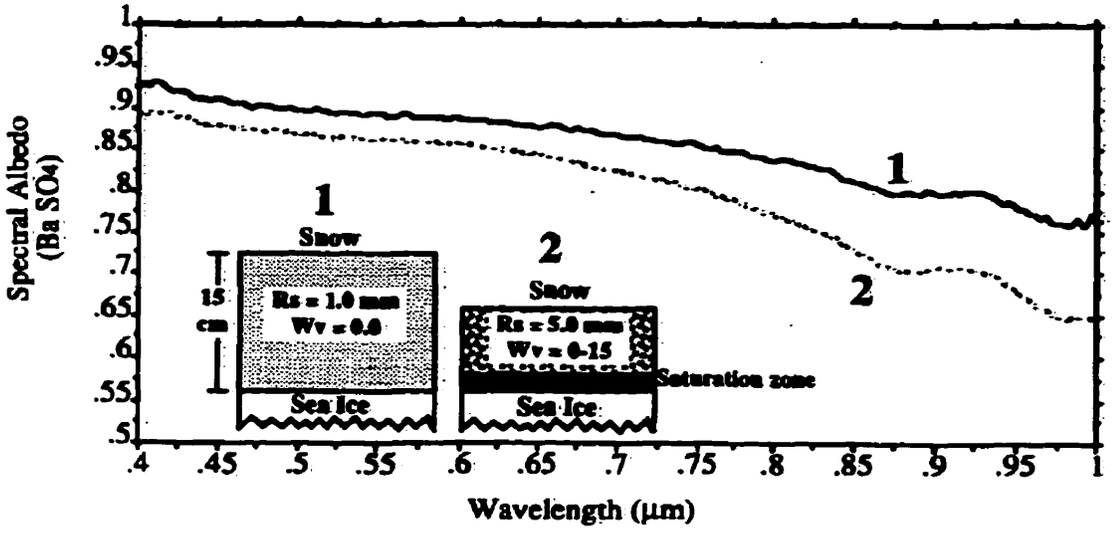



*Figure 2.7: Two spectral albedo curves over snow covered sea ice. One (1) over dry snow, and the other (2) over freely draining snow. The curves are representative with respect to a barium sulphate reference panel. Site (2) ranges from 0 - 15% water by volume from the top of the snow pack to the bottom. (Adapted from Barber, 1993).*

Very little work has been done to assess the spatial variability of albedo, mostly devoted to broad satellite-scale variations. Langleben (1971) suggested a linear decrease in albedo as a function of the degree of puddling in the Beaufort Sea, ranging between 0.6 (0% pond cover) to 0.3 (70% pond cover). More recently, Barber and Yackel (1999) used

video-derived albedo estimates near Resolute, Nunavut to show its spatial variability over linear transects during spring melt. Albedos ranged between 0.3 and 0.5 depending on its proximity to land, ice type and ice roughness (Barber and Yackel, 1999). Sea ice surface albedo derived from satellite sensors in the VIS/NIR wavelengths can obtain regional pixel-scale (~1 km) albedo estimates. Robinson et al. (1986) and Robinson et al. (1992) used DMSP (Defense Meteorological Satellite Program) data to produce realistic albedo maps of the Arctic Basin that was a good indicator of inter-annual variations in the mass and energy budget of the arctic. Robinson et al. (1986) calculated a net areal albedo of 0.53 for the central arctic, corroborating earlier estimates. More recently, De Abreu et al. (1994, 1996) compared *in situ* albedo to AVHRR (Advanced Very High Resolution Radiometer) albedo over FYI showing that the two agreed favorably (within ±5%) when accounting for atmospheric attenuation, viewing geometry, and sensor spectral response. Lindsay and Rothrock (1993, 1994) found mean monthly Arctic Basin albedos to range from 0.76 in April to 0.47 in August, 1989. More observations of surface albedo are needed over FYI, in particular, local albedos of specific surface features (different snow types, melt ponds, etc.) and how these scale up to regional aerial albedo. In addition, little is known about the sensitivity of FYI ablation and break-up potential to the seasonal progression of fractional melt pond cover and their associated albedo changes from year to year. This is important for understanding the partitioning of energy and how we can temporally and spatially represent albedo in numerical modeling studies.

### *2.2.1.3 Surface Absorption*

Absorbed energy is used to heat the medium and create internal phase changes, if the melting temperature is reached or can be re-emitted as long-wave radiation (see below). The absorbed SW radiation for a given level (z) in the snow (or ice) is given by (adapted from Papakyriakou, 1999),

$$K_z^* = \int_{\lambda_1}^{\lambda_2} (K\downarrow_{z,\lambda} - K\uparrow_{z,\lambda}) d\lambda \tag{2.5}$$

where

$$K\uparrow_{(z,\lambda)} = \alpha_{(s,\lambda)} K\downarrow_{(o,\lambda)} \exp(-\kappa_\lambda z) \tag{2.6}$$

and $K\downarrow_{(o,\lambda)}$ is the incident snow surface short-wave radiation, $\kappa_\lambda$ and $\alpha_{(s,\lambda)}$ are the spectral extinction coefficient ($m^{-1}$) and spectral albedo, respectively, and z is the vertical distance.

### *2.2.1.4 Surface Transmission*

The remaining SW energy that is not reflected or absorbed by the snow layer is transmitted into the ice. This available ice energy acts as internal heating of the ice of which a certain portion is also transmitted into the top ocean (mixed) layer. Beer's law dictates the amount of energy that is transmitted through the ice, as in snow above,

$$K\downarrow_{(z,\lambda)} = K\downarrow_{(o,\lambda)} \exp(-\kappa_\lambda z) \tag{2.7}$$

where, $\kappa_\lambda$ is the spectral extinction coefficient ($m^{-1}$). Maximum transmission of SW energy through the snow and ice occurs between the 0.45 to 0.55 μm region of the spectrum (Maykut and Grenfell, 1975) with most absorption in the infrared. Typical bulk extinction coefficients for dry snow range between 37 to 97 $m^{-1}$ and 23 to 58 $m^{-1}$ for wet snow (Fukami et al., 1985), depending on the snow density.

### *2.2.1.5 Long-Wave Radiation*

All matter above absolute zero (-273.15°C or 0 K) emits energy as a function of its temperature, governed by the Stefan-Boltzman law,

$$L = \varepsilon \sigma T^4 \tag{2.8}$$

where, L is the emitted long-wave radiation (W $m^{-2}$), ε is the long-wave emissivity, σ is the Stefan-Boltzman constant, and T is the medium's radiative temperature (K). The snow emissivity is near 0.99, whereas sea ice and water are typically 0.97 (Maykut, 1986). During the months of little or no short-wave input from the sun, the incident long-wave (LW) flux ($L\downarrow$) is dominant in the arctic, although the LW flux remains important even in the spring. There are essentially three sources of LW radiation: 1) the atmosphere, 2) clouds, and 3) the surface. The $L\uparrow$ is primarily controlled by the snow or ice surface temperature and its associated emissivity. As with the SW flux, clouds play an important role in the LW flux because they are able to absorb and re-emit a portion of the LW radiation they receive from the surface and surrounding atmosphere. Overall, they serve to reduce the amount of radiation lost to space by 28-55 W $m^{-2}$, although a large

portion of L↑ radiation is transmitted through clouds less than 320 m thick (Herman, 1986). The L↓ is most sensitive to precipitable water and aerosol optical depth under clear skies and cloud base height and cloud optical depth for cloudy skies (Key et al., 1996). The clear-sky L↓ flux is important during the polar winter. This is because temperatures in the troposphere dictate that condensates be in a predominantly crystaline state even under clear skies (ice crystals). The ice crystals are presumed to emit long-wave energy between 10-40 W $m^{-2}$ greater than modeled winter clear-sky values (Curry and Ebert, 1992). L↓ is a key component in the SEB throughout the annual cycle and accurate measurements and/or model parameterizations are required to numerically simulate seasonal snow and ice evolution.

### *2.2.2 Heat Flow Through Snow and Ice*

The temperature field in snow and ice is described by equations of heat and mass transfer with terms corresponding to internal melting, absorption of short-wave solar radiation, and evaporation/condensation. Water flow within the snow is also usually considered. Thermal processes in the snow and ice are different since snow has a much different physical make-up where compaction and internal evaporation/condensation affect the snow only. In addition, salt (brine) impurities in the snow and ice play a crucial role in their thermal behavior (Maykut and Untersteiner, 1971).

Wind transport of snow (blowing snow) is also a key factor for determining the overall snow depth over a particular point (Pomeroy et al., 1997) which affects the thermal and metamorphic processes in the snow layer over time (Jordan et al., 1999). In

fact, blowing and drifting snow create a spatial pattern of "snow dunes" on FYI that may be related to the spring melt pond spatial distribution. Snow distribution and depth on FYI is also critical for seasonal ice thermodynamics (see for example, Flato and Brown, 1996). The snow model (SNTHERM) by Jordan et al. (1999) allows one to tune the fractional stress gradient ($\tau_s^{-1}\ \partial\tau_s / \partial x$ ; where $\tau_s$ is the snow surface stress) to recreate actual snow depths (positive gradients remove snow, negative gradients accumulate snow). This parameter can be held constant or be allowed to vary in time.

The one-dimensional equation for the conservation of energy (of snow or ice) for the various constituents (i = ice, l = liquid, v = water vapor) is (from Jordan et al., 1999)

$$\frac{\partial}{\partial t}\int \rho_t h_t dz = -\sum_{k=i,l,v}\sum_{s} h_k \mathbf{J}_k \bullet \mathbf{n} + \sum_{s} k_t \nabla T \bullet \mathbf{n} + \sum_{s} \mathbf{R}_s \bullet \mathbf{n} \qquad (2.9)$$

where t is time, $\rho_t$ is the overall snow density (= $\rho_l + \rho_i$ ; bulk density of liquid and ice, respectively, discounting the gas), z is the vertical position relative to the interface between the snow and the sea ice, $J_k$ is the flux of water constituent k (positive upward), the subscript *t* refers to the total medium (snow or sea ice), T is temperature (in kelvins), $k_t$ is the thermal conductivity, and $R_s$ is the net solar radiation (positive downward). The integral is taken over a control volume of thickness $\Delta z$; and the summation $\Sigma_s$ is over the top and bottom surfaces, where $\mathbf{n}$ is the unit vector normal to the surface. The specific enthalpy $h_k$ of a water constituent is (from Jordan et al., 1999)

$$h_k = \int_{273.15}^{T} c_k(T)dT + L_k \qquad (2.10)$$

where $c_k$ is the specific heat of the constituent at constant pressure, $L_t$ and $L_v$ are the latent heats of freezing and sublimation, and $h_t = \Sigma_{k=i,l,v} h_k$. The first term on the right hand side of (2.9) is the heat flux due to water flow and vapor diffusion (disallowed in sea ice), the second term on the right side is the conduction, the last term is the sub-surface absorption of solar radiation. The term on the left hand side represents the change in stored heat.

The parameters $c_k$, $L_t$, $L_v$, $k_t$, brine content/volume ($v_b$), and phase transition temperature for snow and sea ice depend strongly on its salinity and temperature. Due to the wide variety of empirical formulae for these parameters in snow and sea ice, they will not be described in detail. The main point is that brine tends to retard any heating or cooling forcing in the snow or sea ice.

Snow metamorphic and deposition-removal processes also need to be included as indirect thermodynamics factors since these processes affect snow thermodynamics. I only briefly discuss the relevant processes here. Metamorphism and deposition-removal of the snow pack directly affects snow density (at various degrees and levels in the snow), making the snow medium a layered structure. Hence the thermodynamics are affected, and in turn feed back into the metamorphic processes. Grain size and crystal shape also affect the snow density where smaller crystals and simpler shapes pack most efficiently, leading to denser snow (Jordan et al., 1999). Four primary processes act in the deformation rate of the snow pack, namely (after Jordan et al., 1999),

$$\frac{\partial}{\partial z}\left(\frac{\partial z}{\partial t}\right) = \left[\frac{\partial}{\partial z}\frac{\partial z}{\partial t}\right]_{metamorphism} + \left[\frac{\partial}{\partial z}\frac{\partial z}{\partial t}\right]_{overburden} + \left[\frac{\partial}{\partial z}\frac{\partial z}{\partial t}\right]_{melt} + \left[\frac{\partial}{\partial z}\frac{\partial z}{\partial t}\right]_{wind} \qquad (2.11)$$

where, the term on the left side is the resulting deformation rate, the next term describes the metamorphic processes (wind packing, density changes, erosion and accumulation), followed by the overburden (weight of snow causing lower layers to compress), the melting term (when melt (freezing) causes the snow pack to shrink (remain the same) in depth), and finally an extra wind term (snow deposition-removal by the wind; i.e. blowing/drifting snow).

The surface temperature gradient is determined by turbulent sensible heat exchange between the snow and air, long-wave and short-wave radiation from the atmosphere and re-radiation from the snow surface, and by latent heat of snow evaporation (or condensation of water from the air) upon the snow surface. The melting of sea ice at its surface produces low-salinity brine and causes a drop in sea ice salinity (Ono and Krass, 1993). Water from snow melt leads to brine dilution and raises the freezing temperature, possibly causing further ice formation depending on the temperature. This process is important for ice generation during the Spring season (Ono and Krass, 1993).

The growth or decrease of the ice thickness at the ice-ocean interface is determined by heat-mass balance on the ice bottom (i.e. conductive flux in the ice close to the ice-ocean boundary and the turbulent heat flux from the ocean) and penetrating short-wave to the ice-ocean interface. This process is represented by (as in Flato and Brown, 1996),

$$\frac{\partial h_i}{\partial t}\Big|_{z=h} = \left( k_i \frac{\partial T}{\partial z}\Big|_{z=h} - \int_h^{\infty} \mathrm{K}\downarrow I_o(1-\alpha)\kappa e^{-\kappa z} dz \right) \frac{1}{L_{fi}} \qquad (2.12)$$

where, $h_i$ is the ice thickness, $k_i$ is the ice thermal conductivity, $I_o$ is the fraction of short-wave radiation that penetrates the surface, $\alpha$ is the surface albedo, $\kappa$ is the bulk extinction coefficient for short-wave radiation (i.e. asymptotic decay of short-wave energy into the ocean), and $L_{fi}$ is the heat of fusion of sea ice. The integral represents the short-wave energy available to the ocean mixed layer that is assumed to be absorbed there to keep the mixed layer temperature at the freezing point.

Figure 2.8 shows the simplified resulting seasonal progression of temperature profiles within the snow and sea ice layers. Note the SW component with maximum daily average radiation (~400 W $m^{-2}$) at melt onset, where a significant amount of energy is available to the surface at this period. A snow layer may or may not be present during Freeze-Up, depending on the timing of snowfall. A snow layer during Freeze-Up significantly retards the ice growth rate (see for example, Flato and Brown, 1996). The snow temperature gradient in winter is greater than the ice due to thermal conductivity (and thermal diffusivity[17]) differences between the two mediums. During Early Melt, the temperature gradient in the snow reflects the diurnal nature of atmospheric heating and cooling, producing temperature waves in the snow over the course of a day. It also signals the beginning of snow metamorphism (Livingstone et al., 1987). The ice volume begins to show signals of warming. At Melt Onset, the snow volume contains liquid water at all times with the pedular and funicular regimes active at this stage; snow grain growth is accelerated (Colbeck, 1982) and albedo decreases significantly; conductive fluxes become larger within the snow; diurnal temperature variations are still present; and the ice continues to warm considerably. By Advanced Melt, the snow pack shrinks (if

present at all) with a near-isothermal temperature profile, liquid water is always present thereby decreasing the albedo further; bulk salinity in the snow and top ice layer decreases (Tucker et al., 1987); melt ponds develop on the ice surface and brine drainage in the ice is fully active; the ice volume becomes near-isothermal which creates ideal break-up conditions (due to a marked decrease in ice strength (Barber et al., 1998)).

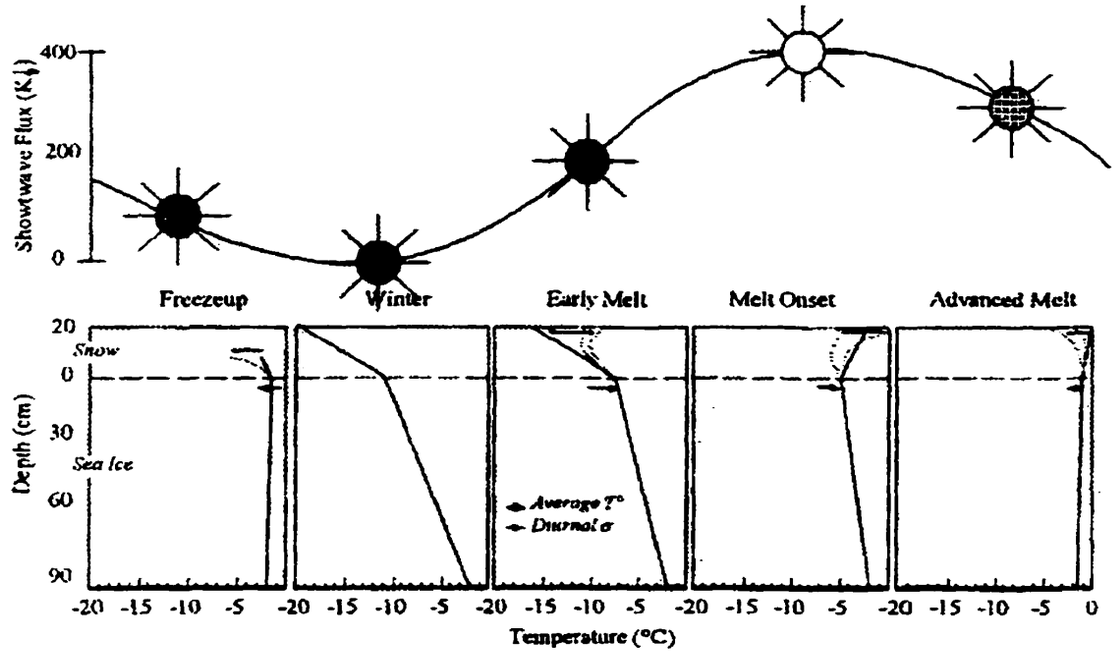



*Figure 2.8: Categories of snow covered first-year sea ice thermodynamic regimes between freeze-up to advanced melt periods. Profiles are typical from 8 years of in situ data. Daily average solar energy (W $m^{-2}$) available for each regime is also shown. (Adapted from Barber, 1997).*

[17] Thermal diffusivity (khs) is the rate at which the surface induced temperature wave can travel within the medium (snow or ice); and related to thermal conductivity (ks) and heat capacity (C) by khs = ks / C.

## 2.3 Utility of Numerical Modeling for Snow Covered Sea Ice

This section is devoted to illustrating the utility of numerical modeling for advancing our understanding of snow covered sea ice climate processes. I begin with a brief historical perspective of sea ice modeling and continue with research results pertinent to this work that demonstrate the use of models to study snow and sea ice in a one-dimensional sense.

### *2.3.1 Historical Perspective*

Before the advent of higher-speed computers (prior to the late 1960's), empirical relationships for ice growth and snow/ice thermodynamics were developed but offered little physical insight into the problems. Budyko (1966) developed one of the first physically representative thermodynamic models of ice through numerical techniques, however a major problem was the specification of surface temperature that ignored heat conduction effects on surface ablation and surface temperature. Maykut and Untersteiner (1971) (MU) offered a more complex thermodynamic model that considered atmospheric and oceanic fluxes to dictate the thermal regime of sea ice and snow in discrete layers. A general approach to estimating internal ice temperatures and thickness was to use heat conduction equations similar to section 2.2 that are still used in modern snow and sea ice models. The MU model included the effects of internal heating from solar radiation penetration, internal heat storage in brine pockets, and variation of specific heat and thermal conductivity of the ice/snow with temperature and salinity. However, the model

was limited by numerical techniques (required larger time steps and layer spacing), course albedo parameterization, poor turbulent flux parameterizations and could not grow ice from open water. More recently, Ebert and Curry (1993) (EC) developed a thermodynamic model as an offshoot of MU but with much more detail in many physical processes. These included a revised turbulent flux parameterization, improved albedo parameterization, greater flexibility in numerical technique that improved the overall thermodynamics, and improved atmospheric radiative forcing parameterizations. Flato and Brown (1996) (FB) considered a thermodynamic model for FYI (land-fast ice) that is similar to EC, with a different albedo parameterization and slightly different atmospheric forcing parameterizations, mainly for the purpose of inter-annual simulations. Recently, a model for snow layer thermodynamics and mass balance over sea ice has been developed by Jordan et al. (1999). The Jordan model is state of the art in that it treats most critical processes in a physical way with more detail paid to heat conduction specific to snow (layered mixture of dry air, ice, liquid water, and water vapor), phase changes, water flow, snow densification, grain growth, and effects from blowing/drifting snow.

### *2.3.2 Utility of Modeling*

Results to date from numerical modeling studies of snow and sea ice have proven to be useful for a number of reasons. A key advantage of models is futuristic insight, i.e. so we can determine what the conditions may be like in the future. A second advantage is the output they produce. Models can generate a time series of information from any particular variable that is simulated by the model, for example surface albedo or snow/ice

surface temperature. A third advantage is conducting sensitivity studies. Sensitivity studies are an attractive component of models where we can perturb a single variable (or multiple variables) to investigate any changes in the simulation. For example, a control run simulation (initial simulation with known input and forcing) would be generated, then a second simulation would be produced by altering a variable (e.g. albedo). A comparison can be made between the control simulation and the altered albedo simulation to observe any major differences in the output parameters (such as snow/ice thickness over time). Sensitivity studies are useful for identifying: 1) variables and processes that control snow/ice evolution, 2) feedback mechanisms, and 3) which variables may be more "active" than others. This also allows us to investigate processes that would not otherwise be determined by field observations since we can not (in most cases) alter the physical characteristics of the forcing variables[18].

Current sea ice models simulate inter-annual and climatological (inter-yearly) sea ice conditions reasonably well (Figure 2.9) (Flato and Brown, 1996; Ebert and Curry, 1993), although major deficiencies are still present.

[18] Forcing variables are contained in the SEB (i.e. meteorological) and ocean mixed layer beneath the ice (ocean heat flux). They are the primary parameters that drive any processes acting on the snow and ice.

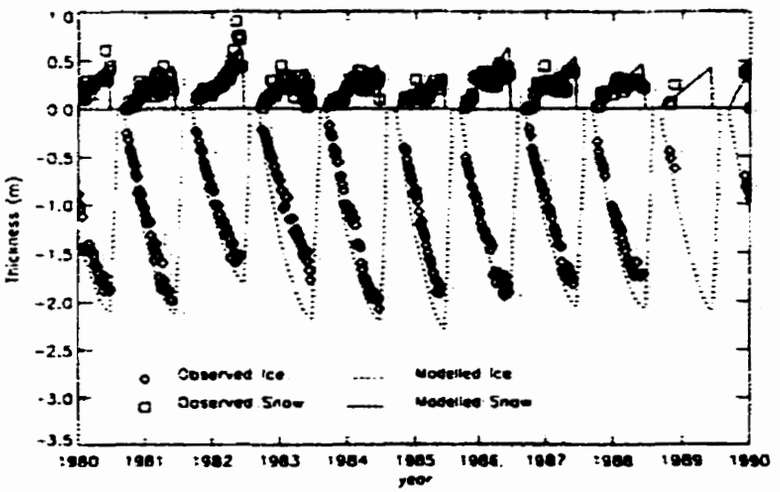



*Figure 2.9: Time series of observed and modeled ice and snow thickness at Alert, Nunavut between 1980-1990. Ice thickness appears as negative values. (Adapted from Flato and Brown, 1996).*

Data presented in Figure 2.9 stems from the FB model where they simulated the inter-annual variability of land-fast FYI between 1955 – 1990 in two different arctic locations (only Alert, Nunavut between 1980-90 is shown). The model simulation was generated using *daily-averaged* meteorological forcing (air temperature, relative humidity, wind, total cloud amount, and snowfall) from Alert spanning the years of simulation. Figure 2.9 shows what model output can generate and illustrate its utility for studying sea ice processes. The complexity of the EC model physics and processes is illustrated in Figure 2.10 (adapted from the sea ice model of Ebert and Curry, 1993).

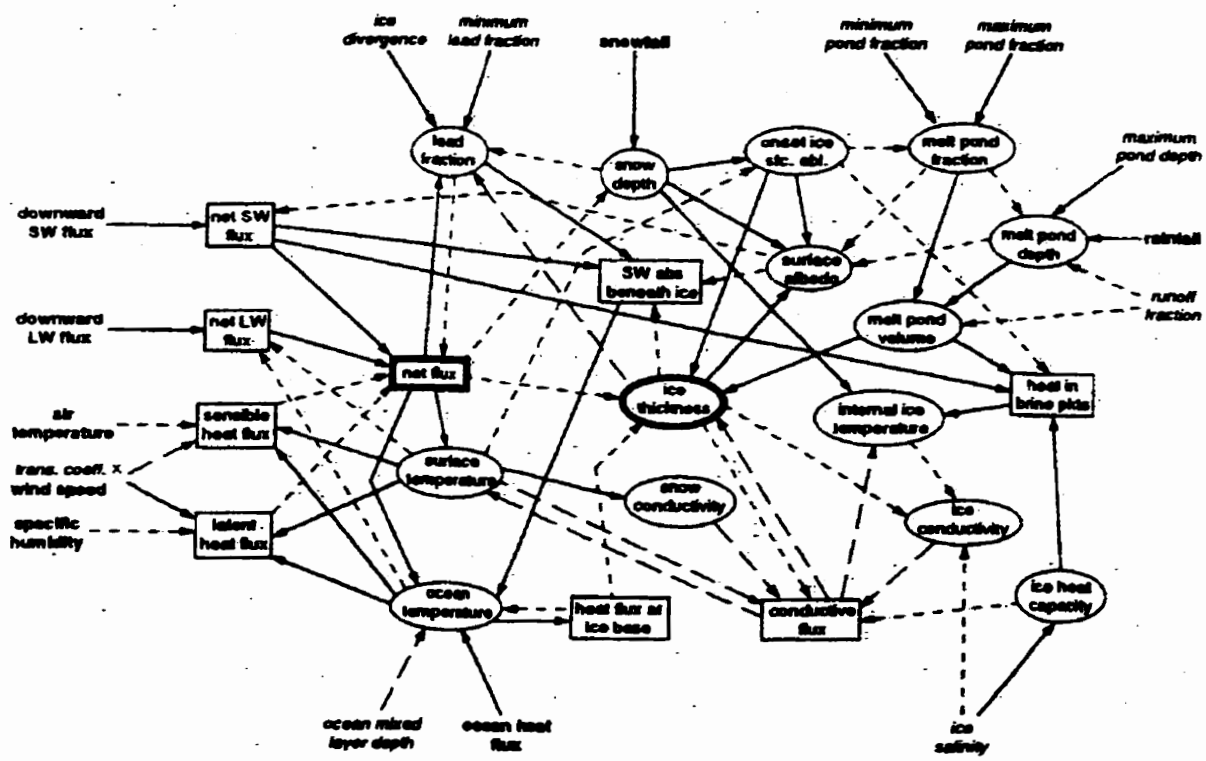



*Figure 2.10: Modeled interactions (see text) between external forcing (bold) and parameter values (italic), and internal variables (ovals) and fluxes (boxes) for the Ebert and Curry (1993) one-dimensional thermodynamic sea ice model. (Adapted from Ebert and Curry, 1993).*

Figure 2.10 shows the interactions and feedbacks relating the model variables and fluxes (ovals and boxes, respectively) to each other and to the specified values of the parameters and external forcing (italics and bold, respectively around the outside of the figure). A solid arrow indicates that a positive change in the first variable has a direct positive impact on the second variable (positive interaction), a short dashed arrow indicates that a positive change in the first variable impacts negatively on the second variable (negative interaction). Long dashed arrows represent positive or negative interactions, depending on the season. These interactions can be amplified or dampened by positive or negative feedback loops (Ebert and Curry, 1993). An even number of negative interactions between a series of interactions results in a positive feedback loop

where the initial variable becomes more positive because of the feedback, and so on. An odd number of negative interactions results in a net decline of the original variable (negative feedback loop). As an example, the surface albedo feedback indicates an increase in the SW or LW flux (or air temperature) acts to increase the net flux on the ice, warming the ice surface (Ebert and Curry, 1993). A warmer surface accelerates melt onset and melt rate, lowering the surface albedo and increasing the short-wave flux. Ice ablation is directly increased as a result of greater input energy (Ebert and Curry, 1993). This example clearly illustrates the utility of numerical models in identifying primary sea ice processes through sensitivity studies.

The degree of ice thickness changes to air temperature changes can be examined through similar sensitivity studies. Flato and Brown (1996) showed a 5°C increase (decrease) in air temperature would result in a decrease (increase) in land-fast ice thickness of 30 cm. Other studies (Parkinson and Kellogg, 1979; Semtner, 1987; Ebert and Curry, 1993) indicate temperature perturbations between 2-4°C would have great impacts on ice characteristics and cause the ice pack to disappear by the late melt season. It is very difficult to determine why the results are so different between the models, however, subtle differences between the formulations of ice thermodynamics, numerical methods and atmospheric fields may be enough to account for the discrepancies (Holland et al., 1993).

Mentioned earlier, the surface albedo is of primary interest for sea ice processes and modeling its seasonal evolution (parameterizing it) is very important for accurate sea ice simulations. Numerical models must parameterize the albedo since a complete physical model is not yet available, and a purely physical albedo model would be

computationally expensive. Modeling studies have shown that multiple albedos must be parameterized to properly account for dry snow as opposed to wet snow as well as bare, melting ice (Shine and Henderson-Sellers, 1985). The observed albedo for bare, melting ice ranges between 0.5 and 0.58 and this range can affect ice thickness by a factor of 2. This range in albedo also determined whether the ice was seasonal or multiyear ice (Shine and Henderson-Sellers, 1985). Flato and Brown (1996) also stressed that fixing the melting ice albedo to 0.55 and not accounting for intermediate melting ice albedos produced unrealistic multiyear ice scenarios in some land-fast arctic ice regions. Shine and Henderson-Sellers (1985) altered the dry snow albedo between 0.8 and 0.75 in which the difference in equilibrium ice thickness were 1 m and the date of total snow melt occurred one week earlier for the lower albedo. The wet snow albedo was altered between 0.7 and 0.65 where the difference in ice thickness at equilibrium was 1 m and total snow melt occurred 4 days earlier with the lower albedo.

Ebert and Curry (1993) included a more complex treatment for melting snow and ice where they explicitly considered melt pond effects on surface albedo. They found that the melt water runoff fraction (Fr) significantly altered the ice thickness. If no melt ponds were not allowed to exist (Fr=1), the equilibrium ice thickness was much too large (5.0 m compared to an average observed thickness of 3.0 m). If Fr < 0.5, the ice would completely disappear for a few weeks in the summer and have a mean annual thickness of 0.8 m. It is evident that the parameterization of albedo can be very complex and important for the short-wave and long-wave radiative processes. It is also evident that slight variations in albedo parameterizations between the different models can lead to differences in model output in many situations.

Another important part of the SEB are the incident radiative fluxes ($K\downarrow$ and $L\downarrow$), mentioned earlier. Incident radiation is primarily controlled by the atmosphere (discussed earlier) and snow and sea ice models parameterize these fluxes since more sophisticated methods are computationally inefficient (Key et al., 1996). Parameterizing these fluxes is tricky due to their relatively quick response to variables that control their variation (minutes to hours). With current radiation schemes, the equilibrium ice thickness of a one-dimensional thermodynamic sea ice model varies by 4 m for a ± 5% change in $K\downarrow$ and 12 m for a similar change in $L\downarrow$ (Ebert and Curry, 1993). The ice pack could completely disappear in summer for an increase in $L\downarrow$ greater than 2% (Ebert and Curry, 1993). Equilibrium ice thickness variations are greater for $L\downarrow$ since it acts all of the time compared to $K\downarrow$ that acts only during solar illumination. Recent examinations of arctic incident radiative flux sea ice model parameterizations have shown that the best existing parameterizations (Shine, 1984) can estimate *daily average* $K\downarrow$ to within 2% of the mean with an RMSE (root-mean-squared error) of 4% for clear skies and 21% for cloudy skies. The $L\downarrow$ parameterizations performed better, with an accuracy of about 1% of the mean and RMSE of 6% for both clear and cloudy skies (Efimova, 1961 for clear skies; Jacobs, 1978 for cloudy skies) (Key et al., 1996). The $L\downarrow$ parameterizations (clear and cloudy skies) of Maykut and Church (1973) were very close in accuracy to Efimova (1961) and Jacobs (1978).

The snow cover has a profound effect on ice growth and ablation, shown by modeling studies (and observations; see section 2.2). During the winter season, Ebert and Curry (1993) found that the snow layer decreased the ice thickness due to its insolating effects and allowing ablation at the ice underside to dominate the ice growth process. The

inter-annual variability in land-fast ice thickness is dominated by snowfall variations (Flato and Brown, 1996). An increase in snowfall rate caused the maximum ice thickness to decline until about double the present day value (1.7 mm $day^{-1}$) where the ice thickness would increase thereafter. This is due to a sufficient snow cover that submerges the ice into the water forming "slush ice" at the surface, increasing the ice thickness. The timing of a snowfall dictates the seasonal progression of FYI thickness where a large snowfall during initial ice growth (in the Fall) will retard ice growth, whereas as a large snowfall in Spring will enhance ice growth (Figure 2.11; G. Flato and R. Brown, pers. comm). The insulating snow in the Fall dominates (decreases) ice growth with little or no solar component at this time of year, and the high albedo snow surface in spring dominates (increases) ice growth by delaying ice melt.

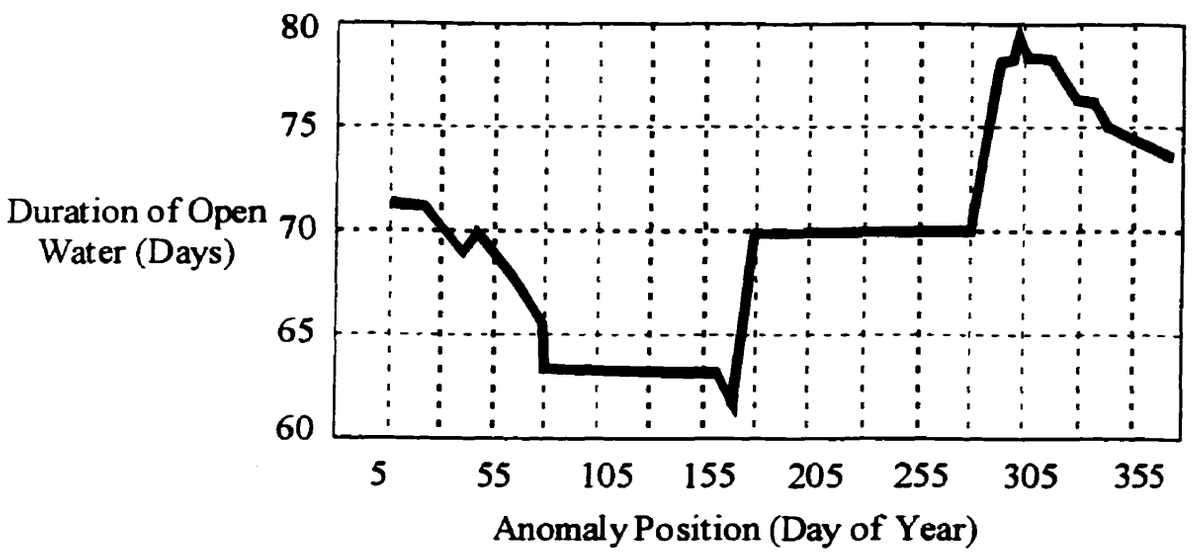



*Figure 2.11: Effect of a 5-day period of snowfall (20 cm) on the open water duration of modeled first-year sea ice. The 5-day period was moved along, throughout the annual cycle (anomaly position) to show the impact of the timing of new snowfall on open water duration. (G. Flato and R. Brown, pers. com.).*

Maykut and Untersteiner (1971) and Semtner (1976) both showed that a snowfall rate of 120 cm $yr^{-1}$ increased the ice thickness without bound for two 1-D thermodynamic sea ice models. For land-fast ice scenarios, this unbounded ice thickness problem from large snowfall rates (extended high surface albedo) did not occur since the warm spring and summer air temperatures melted the snow which alleviated the strong snow albedo influence (Flato and Brown, 1996). Holland et al. (1993) found that increasing the snow fall rate to 200 cm $yr^{-1}$ increased the average ice thickness by about 0.5 m in which case did not produce unstable ice characteristics for their dynamic-thermodynamic[19] sea ice model. It was concluded that the dynamic-thermodynamic model was less sensitive to snow parameter changes than a thermodynamic model (Holland et al., 1993). The reason for this was that there is a negative feedback between the dynamics and thermodynamics of the sea ice model (see Owens and Lemke, 1990) where a thermodynamic model would not capture these feedbacks. The dynamic-thermodynamic model example was included into the discussion since it further illustrates how different models (and their sophistication) can produce a variety of results.

Several models have been developed that contain detailed physics/parameterizations of physical processes and energy and mass exchange in the snow layer [e.g. Brun et al., 1989; Bader and Weilenmann, 1992; Loth et al., 1993; Gruell and Konzelmann, 1994; Lynch-Steiglitz, 1994; Jordan et al., (1999)]. These models simulate snow cover properties like density, grain size, liquid water (and others mentioned earlier) over time. However, most have applications to terrestrial surfaces and not over sea ice, except that of Jordan et al. (1999). Figure 2.12 illustrates the

performance of SNTHERM (snow model of Jordan et al., 1999) over a 6 month period. The model simulates the top snow layer temperature (z = +50 in Figure 2.12) reasonably well, but has greater difficulty in the spring period. This may be the result of snow depth not being accurate (too deep in model) during a storm passage in late February (Jordan et al., 1999). In Spring and Summer, the short-wave radiation is the main term in the SEB. Short-wave energy that was absorbed by the snow created a sub-surface temperature maximum similar to observations (Jordan et al., 1999). In Winter, the net long-wave balance is the main term in the SEB. The snow and ice cool in response to long-wave losses, however the sensible heat from the air to the surface mitigates the losses and nearly mirrors the emitted long-wave flux (Jordan et al., 1999).

[19] A dynamic-thermodynamic sea ice model includes two or three dimensional components taking into consideration ocean currents, wind and other dynamical effects.

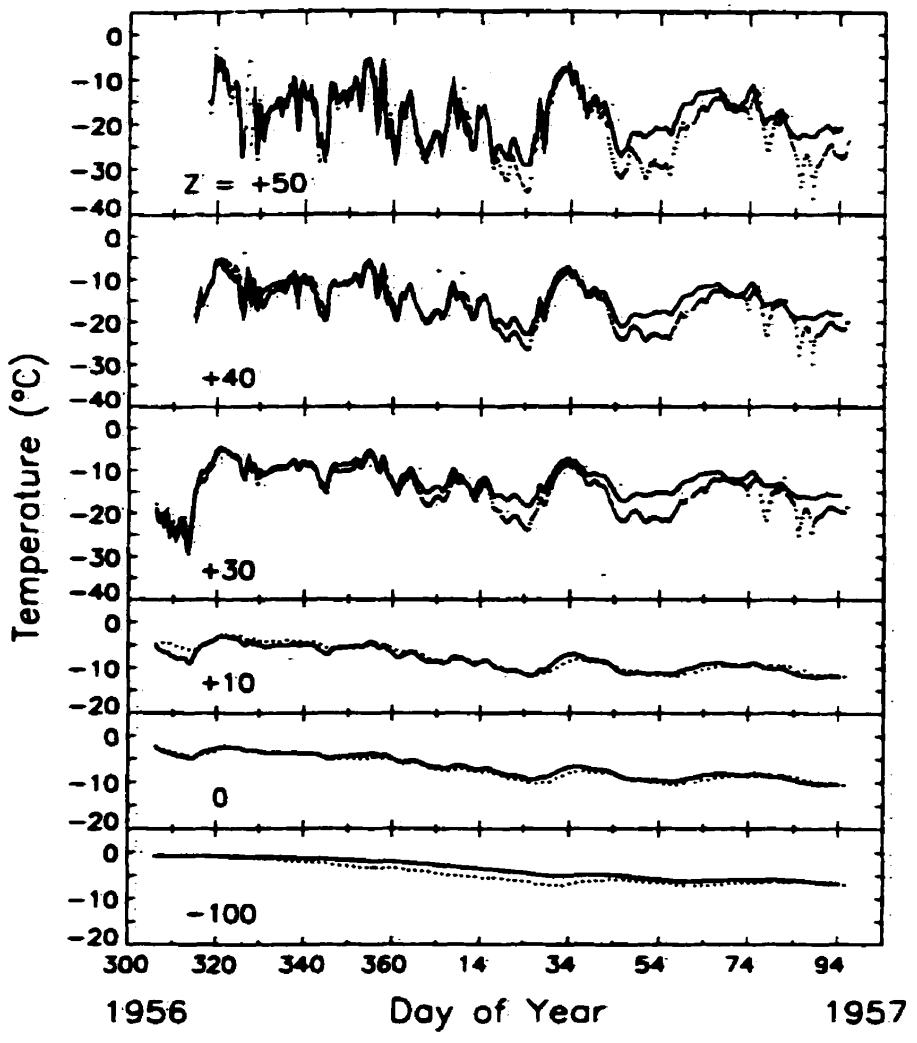



*Figure 2.12: Modeled (line) and observed (dots) temperature traces at various depths in the snow and sea ice between November 1, 1956 to April 4, 1957. The snow-ice interface is at 0 cm. (Adapted from Jordan et al., 1999).*

Jordan et al. (1999) suggests that SNTHERM has the third-generation snow parameterization for sea ice models. SNTHERM models more in-snow processes than EC and FB who treat the snow layer as a single slab with bulk homogeneous properties. SNTHERM has an order of magnitude greater vertical resolution and runs with an hourly time step[20], instead of a daily time step in EC and FB. This is important since we know (through field observations) that the snow layer and its processes act over smaller spatial

and temporal scales than the current sea ice models allow. This ultimately affects the sea ice that responds to snow layer thermodynamics. We also know that the SEB acts over diurnal (hourly and shorter) time scales that fits into the proper scaling of the snow model. An attractive approach is to have a detailed snow model and sea ice model coupled together that both operate over diurnal time scales.

### *2.3.3 Model Deficiencies*

Current snow and ice models simulate many of the critical processes, however there are deficiencies that require further investigation, primarily in the SEB but also involving the thermal properties. Deficiencies outlined here are not exhaustive but are considered of primary interest.

An assumption that has been made by the modeling community is that the SEB can be simulated sufficiently over daily time scales for sea ice (daily time steps). From field work we see processes acting over much shorter time scales (and various spatial scales) ranging from hourly or less. These include surface radiative fluxes, micro and macro-scale processes in snow, sub-daily or daily melt pond evolution, and weekly and monthly ice thickness variations. All of these are tied together through feedback loops (see Figure 2.10) and illustrates the problem of scaling, temporally and spatially.

A technique that is used to force sea ice models is to use land-based meteorological data since measurements on the ice are typically not available (such as the study by FB). Conditions (either atmospheric or surface properties) spatially vary quite

[20] Time step refers to the spacing of time between consecutive iterations of the model (i.e. if the time step is

widely over the arctic and the assumption being made here may not be valid in some cases. For example, incident radiation varies over relatively short time scales and locations. The application of current incident radiation parameterizations may not be valid for all arctic regions due to factors above and is the reason why there are many types of parameterizations (see for example, Shine (1984); Bennett (1982); Efimova (1961); Maykut and Church (1973). In particular, the meteorological conditions, as we see in field observations, over arctic terrestrial, sea ice and open water regions may be quite different.

The surface albedo is another parameter not handled well in current snow and sea ice models. It also changes dramatically in time and space. Changes in micro and macro-scale snow properties (grain size and water content) alters the snow albedo over short periods and the spatial distribution (and thickness) of snow dictates the spatial representation of albedo. Currently, there are no physically-based snow albedo parameterizations under all conditions. Once melt ponds form, their spatial pattern and evolution significantly alter the local and aerial averaged albedo. Capturing these characteristics in an albedo parameterization (and being able to simplify them) is important and more work is required to do this.

A major deficiency with the sea ice models is the lack of detail in the snow layer. Accurate representation of snow over sea ice has been emphasized here as well as other studies (see for example, Brown and Goodison, 1994; Brown and Cote, 1992). A first step toward this is a coupled snow sea-ice model that contains enough detail in both mediums for future investigations of snow and sea ice processes.

---

1 h, the external forcing parameters must be provided at hourly intervals)

Other model deficiencies include:

1) Snow grain evolution is only parameterized and current schemes require validation during the melt season. More physically-based schemes are required.
2) Salinity effects on snow thermal conductivity and specific heat are not modeled (Jordan et al., 1999; Papakyriakou, 1999) nor is it accounted for in snow mass balance (i.e. brine generation in sub-freezing conditions).
3) Turbulent transfer parameterizations in the SEB with regards to vapor transport over cold snow surfaces need more attention (Papakyriakou, 1999).
4) Ice models may over-predict ice growth in Fall since they assume a constant (and too large) snow density during this period (Jordan et al., 1999).
5) Snow models do not account for ponding of water *within* the snow (Jordan et al., 1999).
6) Snow models create excessive basal ice due to the lack of water filtration across the snow-ice interface (Jordan et al., 1999).
7) More work is needed for drifting-blowing snow on erosion-accumulation of the snow pack (Pomeroy et al., 1997; Jordan et al., 1999).

A drawback to SNTHERM and ice models such as FB is they are too detailed for use in large-scale climate modeling (long integrations into the future). This is due to limited computing power and length of time to produce GCM-type simulations, even now with simplified physics in today's GCMs. However, detailed models are useful for the purpose of investigating small scale processes (temporally and their spatial

representations) and used to derive relationships that can be scaled-up to climatic time and spatial scales. We can then incorporate these scaled-up relationships into current GCMs.

## 2.4 Utility of Microwave Remote Sensing for Snow Covered Sea Ice

### *2.4.1 Background*

This section is devoted to illustrating the utility of microwave remote sensing for advancing our understanding of snow covered sea ice climate processes. In particular, I outline research results that demonstrate the use of microwave remote sensing to study snow and sea ice.

Remote sensing of snow covered sea ice has created a diversification of how we use electromagnetic radiation (EMR) interactions to study sea ice processes and its spatial and temporal evolution. Much attention is being paid to linking reflectance, scattering and emission of EMR to physical variables which in turn are coupled through ocean-surface and surface-atmosphere processes to the linkages which control sea ice evolution. Remote sensing technologies are particularly appropriate for studies that focus on the temporal and spatial aspects of processes. Remote sensing measurements can also provide us with a broad area of fine spatial resolution data that can otherwise not be obtained with sparse surface instrumentation.

In the microwave region, both active and passive techniques are widely used for remote sensing of sea ice due to dampened cloud and precipitation effects. Scattering, thermal emission and to a lesser extent, rotational processes (polarization) are all important in the microwave interaction mechanisms with the surface (for example,

Barber et al., 1994; Drinkwater, 1989; Winebrenner et al., 1994). Scattering plays the most important role in microwave remote sensing which is a complex function of the dielectric properties[21], surface roughness and volume inhomogeneities. The relative scattering cross section[22] ($\sigma^o$) can change over time and space. The spatial variability is mostly a function of the geophysical properties that contribute to the volume dielectrics or the surface roughness of the snow/ice (Barber et al., 1995b; Onstott, 1992). The temporal changes are controlled by the dielectric mismatch across the air-snow and snow-ice interfaces. The key variables involved are the relative phases of water (ice, liquid and vapor), snow grain size, brine volume and surface roughness. In addition, scattering can be broken down into two key components; surface scattering and volume scattering. If there is a strong dielectric mismatch at a particular interface then the surface scattering will dominate.

Snow cover plays an important role in the microwave backscatter characteristics. Snow is essentially transparent in the Winter season but as the water content in the snow pack increases (due to increased Spring-time air temperatures), the penetration depth of the microwaves into the snow decreases (Drinkwater, 1989). As the penetration depth decreases, the average scattering from a snow covered ice surface contains an increasing contribution from the snow volume and snow surface geometry. The relative contribution of snow surface roughness to the backscatter of microwave energy increases as the free water content of the snow increases (Barber et al., 1991).

---

[21] Dielectric properties define the electrical conductivity of a material (such as snow or ice) relative to the wavelength and polarization of the electromagnetic energy (see Vant et al, 1974; Ulaby et al, 1986).
[22] The relative scattering cross section is an average of the scattering contributions (surface and volume) made by a medium (or mediums including their interfaces) (see for example Barber, 1993).

### *2.4.2 Microwave Remote Sensing of Sea Ice Through the Annual Cycle*

Microwave remote sensing of sea ice has been available since the late seventies in which significant progress has been made since then in the amount of information that is obtainable. However, not all techniques that are used to extract geophysical variables from remote sensing are applicable throughout an annual cycle of sea ice due to many factors (see for example Drinkwater, 1989; Barber et al., 1995b). The spatial and temporal dynamic and physical changes in which sea ice undergoes throughout an annual cycle can be compartmentalized into 5 different stages, similar to Figure 2.8 in section 2.2. These include fall freeze up, winter, early melt, melt onset and advanced melt periods (after Livingstone et al., 1987). The following section will follow these 5 sea ice stages. This section includes a description of the current geophysical variables that are extractable from remote sensing technology during the different stages of sea ice growth and decay.

Table 2.2 summarizes the types of geophysical variables obtainable from microwave remote sensing of sea ice along with the types of sensors, frequencies/wavelengths, spatial resolution and temporal coverage of the remote measurements. Sample references of each technique are also included in the table. Each subsection will provide a discussion of the remote sensing techniques used to infer the variables in Table 2.2 as they are related to the different stages of sea ice growth/decay.

*Table 2.2: Variables obtainable from microwave remote sensing for sea ice processes.*
* Passive microwave satellite processing is done at the National Snow and Ice Data Center.
** Ice typing is done operationally at the Canadian Ice Service.

| Variable | Satellite/Sensor/ Band | Wavelength/ Frequency | Resolution | Coverage | Sample Reference |
|---|---|---|---|---|---|
| Ice Mass Balance: | | | | | |
| Extent | ESMR (p. μwave) | 19 GHz | 25 km | | |
| | SMMR (p. μwave) | 8,21,22,31,37 GHz | 55 km | 2 per day | |
| | SMM/I (p. μwave) | 8,21,22,31,37, GHz | 12 - 25 km | daily | Weaver & Troisi, 1996 |
| Surface melt; Thickness | SMMR (p. μwave) | 8,21,22,31,37 GHz | 55 km | 2 per day | Weaver et al., 1987 |
| | SMM/I (p. μwave) | 8,21,22,31,37, GHz | 12 - 25 km | daily | Johannessen et al., 1996 |
| | ERS-1/2 (a. μwave) | 5 cm (C-band), | 30 - 100 m | 1-2 per day | Weaver & Troisi, 1996 |
| | Radarsat (a. μwave) | 23 cm (L-band) | | | Fily & Rothrock, 1986 |
| | | | | | Barber et al., 1991 |
| | | | | | Markus & Cavalieri, 1996 |
| | | | | | Shuchman et al., 1996 |
| | | | | | Winebrenner, 1996 |
| | | | | | Malinas & Shuchman, 1994 |
| Type; Concentration | SMM/I (p. μwave) | 8,21,22,31,37, GHz | 25 km | daily | Weaver & Troisi, 1996 |
| | SMMR (p. μwave) | 8,21,22,31,37 GHz | 25 km | 2 per day | NSIDC * |
| | Radarsat (a. μwave) | 5 cm (C-band) | 30 - 100 m | 1-2 per day | Can. Ice Service ** |
| Motion/Velocity | ERS-1/2 (a. μwave) | | | | |
| | Radarsat (a. μwave) | 5 cm (C-band) | 30 - 100 m | 1-2 per day | Fowler et al., 1994 |
| | | | | | Wilson et al., 2001 |
| Surface SW Flux: | | | | | |
| Surface albedo; Melt ponds | ERS-1/2 (a. μwave) | 3, 5 cm (C-band) | 30, 100 m | 1 - 2 per day | Barber & LeDrew, 1994 |
| | Radarsat (a. μwave) | | | | Weaver & Troisi, 1996 |
| | | | | | Yackel and Barber, 2000 |
| Snow Thickness | ERS-1/2 (a. μwave) | 3, 5 cm (C-band) | 30, 100 m | 1 – 2 per day | Barber and Nghiem, 1999 |
| Net SW flux | ERS-1/2 (a. μwave) | 3, 5 cm (C-band) | 30, 100 m | 1 - 2 per day | Barber et al., 1994 |
| Transmitted PAR | ERS-1/2 (a. μwave) | 3, 5 cm (C-band) | 30, 100 m | 1 - 2 per day | Barber & LeDrew, 1994 |
| Surface LW Flux: | | | | | |
| Surface temperature | SMM/I (p. μwave) | 8,21,22,31,37, GHz | 25 km | Daily | Weaver & Troisi, 1996 |
| | | | | | Comiso, 1986 |
| | | | | | St. Germain & Cavalieri, 1996 |
| | ERS-1/2 (a. μwave) | 3, 5 cm (C-band) | 30, 100 m | 1 – 2 per day | Barber et al., 1994 |
| LW flux | ERS-1/2 (a. μwave) | 3, 5 cm (C-band) | 30, 100 m | 1 - 2 per day | Barber et al., 1995a |
| Net All Wave Flux: | ERS-1/2 (a. μwave) | 5 cm (C-band) | 30, 100 m | 1 - 2 per day | Barber et al., 1995a |
| Sensible Heat Flux: | | | | | |
| Floe size stats | Radarsat (a. μwave) | 5 cm (C-band) | 100 km | Weekly | Weaver & Troisi, 1996 |
| Atmos. drag | ERS-1/2 (a. μwave) | 23 cm (L-band) | 10 -100 m | 1 - 2 per day | Burns, 1990 |

### *2.4.2.1 Fall Freeze Up*

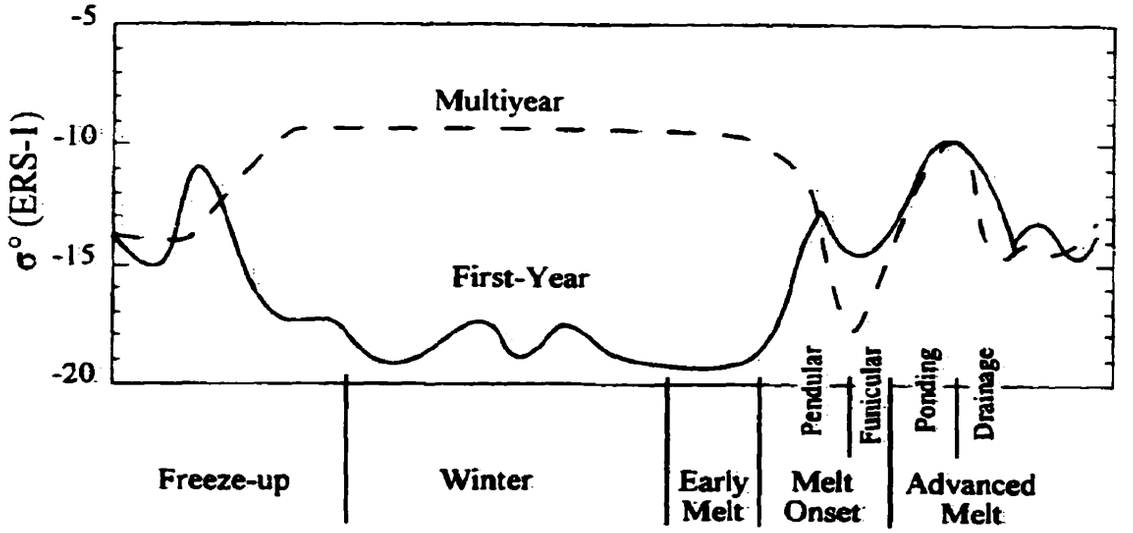



*Figure 2.13: Typical ERS and RadarSat backscatter ($\sigma^o$) at 5.3 GHz for thick First-Year and Multi-Year sea ice over the seasonal cycle. (adapted from Barber et al., 1999).*

Active microwave remote sensing of sea ice varies according to ice type (Figure 2.13). Note the difference between MYI and FYI. This is due to differences in surface and volume physical characteristics of the two types of ice. I limit the discussion to FYI only. The scattering is variable over the ocean surface as a function of wind speed and its high dielectric constant. As ice thickens, there is a decrease in scattering (Figure 2.13). Snow deposition further reduces scattering. The detection of first Fall freeze-up and tracking of its evolution using active microwave remote sensing has not been extensively exploited. However, some studies using SAR technology (satellite and space shuttle) in relation to leads have indicated that a distinction between thin ice regions and open water may be detected (Shuchman et al., 1996) as well as its thickness (Winebrenner, 1996). Shuchman et al. (1996) used ERS-1 SAR data to show that $\sigma^o$ decreases (< -15 dB) within thin ice regions compared to higher $\sigma^o$ values (> -7 dB) in open water along arctic leads. This trend may be applicable to the Fall freeze-up period in general since similar

physical processes occur. Shuchman et al. (1996) pointed out that the higher radar returns may also be caused by frost flowers on a thin ice surface as opposed to open water. Winebrenner (1996) indicated that thin ice (< 70 cm) thickness may be extracted from L-band (23 cm) SAR imagery using co-polar ratios and phases in backscatter, however, extensive testing and validation of the technique is required due to the limited number of ground observations in the study.

Burns (1990) showed SAR (23 cm wavelengths) may be useful in the marginal ice zone for inferring the atmospheric drag coefficient. SAR data in the 23 cm wavelength bands can reveal large scale topography (ridges, floes edges) in which high backscatter returns occur from small, broken and heavily deformed floes and rubble areas (Burns, 1990). In addition, the atmospheric drag coefficient increases with increased ice concentration and ice floe deformation in neutrally stratified atmospheric conditions. Atmospheric drag decreases in more atmospherically stable regimes in which surface roughness has a smaller influence. With the help of SAR data to determine the ice characteristics and a simple atmospheric boundary layer model, estimates of the atmospheric drag coefficient could be obtained in the marginal ice zone under neutral atmospheric stability (Burns, 1990).

Passive microwave sensors have been shown to decipher between open water, thin ice (0 - 9 cm) and thicker first year ice. Microwave surface observations indicated that a sharp rise in surface temperature during the first 1-2 cm of growth coincided to a decrease in brightness temperature at higher frequencies (≥ 37 GHz) (Wensnahan et al., 1993a). They attributed this increase in surface temperature to the upward transport of warm brine from the interior of the ice and resulted in the formation of a salinity

enhanced surface layer. Wensnahan et al. (1993b) later applied a principle component analysis technique to decipher between very thin (0 - 4 cm) and thin ice (4 - 9 cm) on microwave surface observations of sea ice. They pointed out that these thin ice ranges would be very difficult to detect by satellites due to spatial resolution and periodic enhanced speeds at which the ice can grow at these thicknesses. However, satellite-based microwave sensors have been shown to be good proxy indicators to thin ice regions outlined below.

Satellite passive microwave techniques at 85 GHz and 37 GHz using the SMM/I sensor have been shown to be able to distinguish between open water, thin ice and first year ice regions (Markus and Cavalieri, 1996; Grenfell, 1996). Two techniques [NASA team ice concentration algorithm (Cavalieri, 1994) and the Polynya Signature Simulation Method (PSSM) (Markus and Burns, 1995)] to distinguish thin ice from open water were tested against aircraft measurements in which good agreement was found (Markus and Cavalieri, 1996). Both techniques provided accurate open water/thin ice discrimination, however, the NASA team algorithm was better suited for larger scale analysis (25 km or greater) whereas, the PSSM was appropriate for spatial scales near 6 km. It was also pointed out that both techniques are only valid for seasonal ice zones and not as of yet applicable to the entire arctic.

A recent study by St. Germain and Cavalieri (1996) have linked SSM/I passive microwave radiances with surface temperatures in the arctic seasonal ice zone. Comparisons with other satellite brightness temperatures were good for several case studies in the seasonal ice zone, however, further research is required to refine the technique (St. Germain and Cavalieri, 1996).

Markus and Cavalieri (1996) also correlated SMM/I ice type/thickness (open water, thin ice and first year) recognition (using the PSSM and NASA algorithms) with other satellite brightness temperatures. They showed that as ice thickness increased, the brightness temperatures decreased with maximum temperature differences between the two techniques (PSSM and NASA) near 2 K. They also pointed out that the large temperature difference between the open water (275 K) and thin ice (268 K) in their case magnifies the importance of the first few centimeters of ice growth on the latent and sensible heat fluxes in that region.

### *2.4.2.2 Winter*

Within first year ice, microwave scattering appears to oscillate according to changes in the oceanic and atmospheric heat fluxes but still provides a stable surface in which remote sensing can be useful (Barber et al., 1994). It has been shown (Barber and Nghiem, 1999) that this oscillation is driven by atmospheric forcing of the snow-ice interface temperature. An increase in the brine volume of the basal snow layer is induced by temperature waves that travel to the base of the snow pack. This increase in brine volume together with large kinetic snow grains create a sufficiently large volume scattering term to increase the total scattering above that of the ice surface (i.e. these oscillations are only found over smooth thick FYI). These thermodynamically driven oscillations can be found for surfaces below about –19 dB (Barber and Thomas, 1998).

Because of the stable nature of microwave scattering in winter over first year ice, several geophysical variables can be extracted from active and passive remote sensing

during this period (see Table 2.2). The separation between the scattering of different ice types during Winter (see Figure 2.13) also allows one to decipher smooth from rougher ice (Figure 2.14). In the active regime, the short-wave (0.3 - 3.0 μm) surface albedo and transmitted Photosynthetically Active Radiation (PAR) (0.4 - 0.7 μm) for a snow cover can be derived from ERS-1 SAR imagery at 5.3 GHz and 9.25 GHz (Barber and LeDrew, 1994). However, measurement of snow thickness may be required to make remote PAR observations a useful quantity. Although recently, a relationship between $\sigma^o$ and the thermodynamics of the snow cover have shown that snow thickness may be extractable from SAR, but breaks down once water becomes present in the snow pack (Barber and Nghiem, 1999). Snow thickness determination over smooth FYI is in its infancy but results are promising. Barber and LeDrew (1994) developed a statistical relationship (quadratic) between $\sigma^o$ and albedo and $\sigma^o$ and transmitted PAR where as $\sigma^o$ increased, albedo decreased and transmitted PAR increased. The primary physical variables causing the relationship between $\sigma^o$ and albedo were the increased water volume of the snow pack and a larger contribution to snow volume scattering from the increase in the snow crystal radius (Barber and LeDrew, 1994). At very low water volumes (< 1% water by volume) in the snow pack, transmitted PAR to the snow/ice surface was almost non-existent. Once the water volume increased above 5%, transmitted PAR increased logarithmically with $\sigma^o$ due to the increase in snow grain size and increased number density of the water inclusions (Barber and LeDrew, 1994).

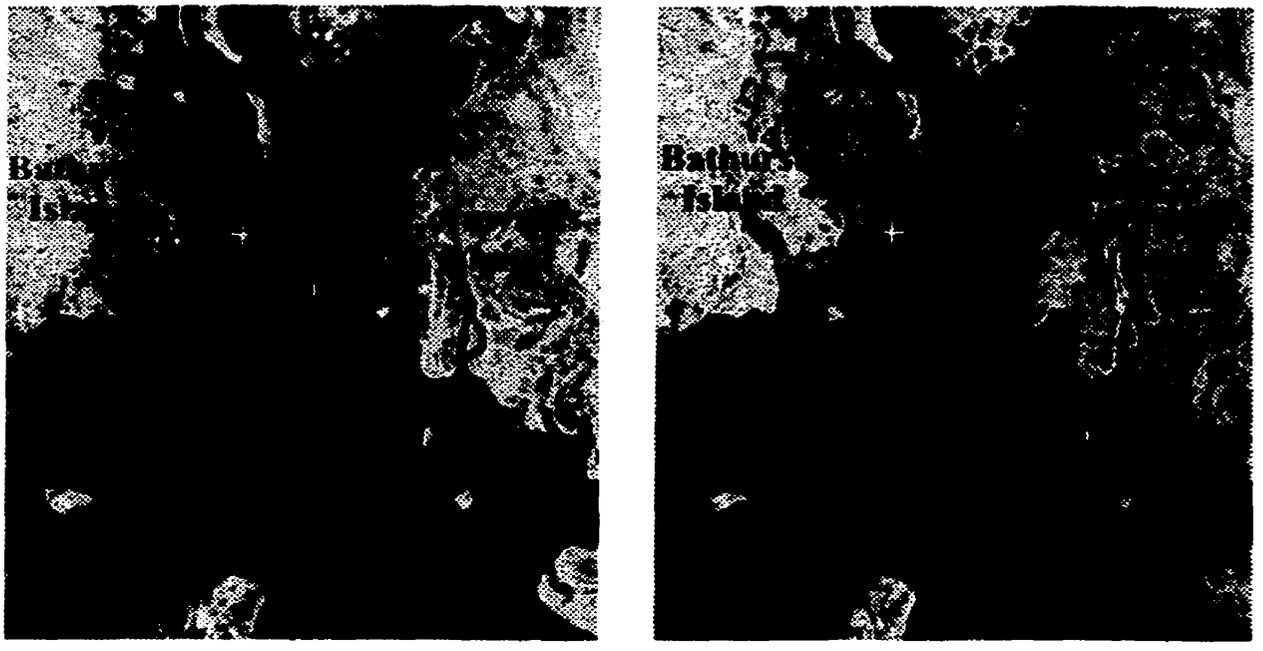


*Figure 2.14: RadarSat imagery on January 31, 1999 (Winter; left) and June 27, 1999 (Advanced Melt; right) in McDougall Sound, Nunavut. Dark shades in Winter are smooth first-year ice, whiter shades are rougher (ridged and rubbled) first-year ice. During Advanced Melt, detail is lost due to water in liquid phase dominating the scene (see text). The "+" signifies smooth first-year sea ice. Due north is toward the top of the images.*

Barber et al. (1994) showed there is a statistical relationship between the seasonal evolution of $\sigma^o$ with surface temperature ($T_{sfc}$) and the net short-wave flux (K*) over first year snow covered sea ice. This is due to both variables ($T_{sfc}$ and K*) either causing or covarying with the change in the dielectric properties responsible for the change in $\sigma^o$. An inverse relationship existed between both $T_{sfc}$ and K* relative to $\sigma^o$, and the influence of the surface temperature was about twice that of the net short-wave flux in explaining the variation in $\sigma^o$. Furthermore, Barber et al. (1995a) determined that $T_{sfc}$ and the long-wave flux (L*) could be considered statistically indistinguishable between first-year and multiyear sea ice, for conditions experienced during SIMMS'93 (Seasonal Sea Ice Monitoring and Modeling Site, 1993). The K* and net all-wave flux (Q*) were however

statistically distinguishable between the MYI and FYI sites. They attributed this result to the fact that $T_{sfc}$ and L* were, to a large degree, determined by characteristics of the atmospheric boundary layer and that K* and Q* were largely prescribed by the snow metamorphic state. Because the snow is distributed differently over multiyear versus first-year ice types, differences occurred in the seasonal evolution of these fluxes.

Maslanik and Maybee (1994) and Fowler et al. (1994) derived ice motions from interpolated ice displacements from ERS-1 data. The observed ice motion vectors were then used to initialize and validate a two-dimensional sea ice model. More recently, Wilson et al. (2001) is applying RadarSat imagery and operational software (TRACKER – used at the Canadian Ice Service) to derive ice motion vectors over time. The technique is has been validated using ice buoys from the 1998 International North Water (NOW) Polynya Project with good success.

Malinas and Shuchman (1994) developed a transfer function to convert SAR $\sigma^{o}$ values to ice thicknesses using aircraft SAR and upward looking sonar data. The general observation was that an increase in ice thickness corresponded to an increase in $\sigma^{o}$ for multi-year and first year ice types. However, the technique requires further verification and testing.

Satellite passive microwave remote sensing sensors provide another valuable source for winter sea ice monitoring and have been widely used for many studies (see for example Preller et al., 1992; Maslanik and Maybee, 1994). Algorithms for multi-channel microwave data can extract sea ice concentration for any given pixel, from which ice extent, ice area, water area and overall ice concentration can be derived for broader regions (see for example, Johannessen et al., 1996). Individual pixel concentrations can

be derived from SSM/I radiances in the frequencies at 19, 22, 37 and 85 GHz in conjunction with an algorithm such as NORSEX (Norwegian Remote Sensing Experiment; Svendsen et al., 1983) or the PSSM or NASA team algorithms (see Fall Freeze-Up section). Time series of ice extent, ice area and overall ice concentration can then be derived from individual-pixel ice concentrations of consecutive SMM/I imagery.

### *2.4.2.3 Early Melt*

Within first-year sea ice, microwave scattering appears to be dominated by a combination of basal layer volume scattering and ice surface scattering (see Figure 2.13). During this period we would expect to find a significant difference in solar noon versus solar midnight observations as small amounts of water in liquid phase would contribute both to grain growth and the elevated temperatures would significantly increase the brine volume of the snow basal layer. The overall magnitude of $\sigma^{\circ}$ will be dependent on the ice surface micro-scale roughness and its surface brine volume. All of the geophysical variables extracted from microwave remote sensing (active and passive) in Table 2.2 are transferable to the Early Melt season.

### *2.4.2.4 Melt Onset*

In first-year sea ice, melt onset is denoted by a rapid increase in $\sigma^{\circ}$ (see Figure 2.13). There are 2 mechanisms which are likely candidates for the observed increase. At relatively low water volumes (1 to 3 percent) the large wet snow grains in the basal

layer may contribute a significant volume scattering term to $\sigma^{\circ}$. As the water in liquid phase continues to increase (but is maintained within the pendular regime) it is likely that the snow surface may also contribute a surface scattering term to $\sigma^{\circ}$ (Drinkwater, 1989; Livingstone and Drinkwater, 1991; Barber and LeDrew, 1994). In a case study done by Barber et al. (1995a) it was found that the snow volume scattering term was the dominant mechanism. A distinct dip in $\sigma^{\circ}$ at the FYI site corresponded with the transition from the pendular and funicular regimes. This transition marks the reduction of brine within the basal layer to near zero, an increase in the water in liquid phase at the base of the snow cover and a reduction of water in liquid phase in the top parts of the snow volume (as the surface begins to drain). These processes could lead to a reduction of both the volume scattering and snow surface scattering hypothesized to dominate the pendular regime conditions.

All of the geophysical variables extracted from microwave remote sensing (active and passive) in Table 2.2 may be transferable to the Melt Onset season. However, the high variability in the spatial patterns of most geophysical parameters within the icescape may produce large differences between pixels within the satellite imagery depending on the spatial resolution of the remote sensor.

### *2.4.2.5 Advanced Melt*

In first-year sea ice, advanced melt is denoted by a rapid increase in $\sigma^{\circ}$ (see Figure 2.13). As surface water forms in the melt ponds there is an increase in the discontinuity at the air/water interface. We can expect a penetration depth on the order of

1 cm at 5.3 GHz (Ulaby et al. 1986). An increase in scattering occurs if the melt pond surfaces are wind roughened. The separation in scattering between ice types is also decreased (Figure 2.13) making detail in satellite SAR imagery "washed-out" due to water in liquid phase dominating during this period (see Figure 2.14). Once the ice surface begins to drain there is a pronounced decrease in the $\sigma^{\circ}$. This period coincides with a reduction in the aerial extent of the melt ponds. The exact mechanisms for first-year ice scattering during the advanced melt period are also largely unknown.

Due to the high variability in the icescape during the Advanced Melt period, it is not known whether most of the geophysical variables that are extracted from remote sensing (appearing in Table 2.2) are valid. Most studies do not include this stage of sea ice evolution within the temporal scales of the techniques. This is mostly due to the limited number of surface observations for validation during this period in which dangerous ice conditions exist. However, Barber and Yackel (1999) and Yackel and Barber (2000) showed that SAR imagery may be able to infer surface albedo and melt pond fractions using ERS-1 or RADARSAT during advanced melt but only applied to first year land-fast sea ice in windy (2-3 m $s^{-1}$) conditions. They also indicated that the microwave scattering could be used to obtain an unambiguous measure of the onset of melt, by an increase in scattering over first year ice. Furthermore, the transition between pendular and funicular regimes may be detectable in first year scattering (Barber and Yackel, 1999).

## 2.5 Summary

In summary, the critical characteristics of snow-covered FYI including its physical and thermal attributes have been discussed in sections 2.1 and 2.2. These attributes occur over a wide range of spatial scales (microscale to synoptic) and temporal scales (seconds to years). I have also included a discussion of the current types of one-dimensional numerical models that are available to simulate the physical and thermal attributes of snow-covered FYI over various spatial and temporal scales (section 2.3). The discussion illustrated the utility of one-dimensional models for improving our understanding of arctic snow and sea ice processes, their advantages and deficiencies. The following section (2.4) illustrated the utility of satellite microwave remote sensing for monitoring and understanding physical processes within FYI over the various seasons. Chapter 2 has thus provided background material to issues that are addressed in my dissertation that primarily deal with using *in situ* data for advancing one-dimensional sea ice models to a state where they can be useful for application to microwave remote sensing. The first chapter (Chapter 3) deals with the issue of ensuring the one-dimensional sea ice models properly handle the temporal scales we see acting in field observations. Previous field work and microwave remote sensing observations suggest that diurnal time scales are important over seasonal snow-covered sea ice. In Chapter 3, I investigate the ability of a one-dimensional thermodynamic sea ice model to simulate annual cycles of FYI and its various physical and thermal processes over diurnal time scales, an investigation that has previously never been conducted.

# CHAPTER 3: Role of Diurnal Processes in the Seasonal Evolution of Sea Ice and its Snow Cover

## 3.1 Introduction

In this chapter, I examine whether there is a need to consider diurnal time scale processes in a snow-covered FYI 1-D numerical modeling environment (using Flato and Brown, 1996) and whether the existing model appropriately handles these time scales in several critical parameterizations. Stated in Chapters 1 and 2, it is important that snow and sea ice models operate at surface energy balance time scales we see in field observations (typically hourly). It is also important to ensure the model parameterizations are able to work in these time scales and they are not compromised by spatial differences. An investigation of this type has previously never been done since it has been assumed that diurnal processes would produce similar modeling results as longer time scales. The results from this chapter have been published in the peer reviewed literature in Hanesiak et al. (1999). The primary objectives of this chapter are to examine the following research questions:

1) Does the use of hourly and daily varying forcing data result in a different simulation of the annual cycle of ice thermodynamics?
2) Do the parameterizations used in the model adequately represent specific processes over shorter temporal scales and are they valid when compared with 'on ice' measurements?
3) Is there a systematic bias in forcing the model with 'land based' versus 'on ice' measurements of input fluxes when using shorter temporal scales?

## 3.2 Data and Methods

### *3.2.1 Sea Ice Model*

A thorough description of the one-dimensional thermodynamic sea ice model used in this work can be found in Flato and Brown (1996) where many of the physical processes discussed in chapter 2 are implemented. The model was designed to run with daily average forcing but was adapted to use hourly forcing and time step. This required modifications to the down-welling short-wave parameterization to account for diurnal variations in solar heating. Since a more detailed investigation involving the parameterized and observed down-welling short-wave (K↓) and long-wave (L↓) and surface albedo are conducted (sections 3.3.2 and 3.3.3), brief descriptions of how the model parameterizes these variables is given here. Chapter 4 looks at these parameterizations and tests their capabilities in greater detail.

L↓ is computed using the parameterization of Maykut and Church (1973) for both clear-sky and all-sky fluxes and the K↓ parameterization of Shine (1984) is used for both clear-sky and all-sky fluxes. Both parameterizations were found to be among the best for Arctic climates (Key et al., 1996) and will be further tested in chapter 4. Both the L↓ and K↓ parameterizations use a single total cloud fraction for computing all-sky fluxes. The K↓ parameterization explicitly considers changes in albedo, relative humidity and cloud fraction however I employ a constant cloud optical depth (0.7 after Curry and Ebert, 1992 and used also in Flato and Brown, 1996) that may cause negative biases in this work under overcast conditions (see section 3.3.2). Due to the non-linearity of cloud fraction and relative humidity in the K↓ parameterization, the daily average K↓ computed over hourly intervals may be different than the single daily value K↓ computed using a

daily-average cloud fraction and relative humidity. This holds even under clear skies with a constant surface albedo due to relative humidity changes. The impact of this in combination with other processes is shown in section 3.3.1.

The surface albedo parameterization is dictated by surface type (ice, snow, open water), surface temperature (melting or subfreezing), and ice thickness. The latter acts as a proxy for time during the melt season that attempts to capture melt pond evolution (Flato and Brown, 1996). The parameterization does not explicitly treat melt ponds nor does it partition the albedo into discrete wavelengths as in Ebert and Curry (1993). The dry snow albedo is set to 0.75 if the snow depth is $\geq$0.1 m and decreases according to snow depth less than 0.1 m and the cold ice albedo of Maykut (1982). When snow is melting, its albedo decreases to 0.65. The melting ice albedo parameterization is adapted from Heron and Woo (1994). The full formulation of the model's albedo parameterization can be found in Flato and Brown (1996).

The snow layer in the model is treated as a single layer with bulk homogeneous physical properties (ie: density = 330 kg $m^{-3}$ when dry). When new snow is deposited on the surface, the albedo and density do not change, however, if the mean snow layer temperature is at or above freezing, its density increases to 450 kg $m^{-3}$. This serves as a proxy for the presence of liquid water and snow densification due to increased liquid water. The model does not physically consider variations in density over time, snow grain processes, liquid water content or brine inclusions.

The number of layers within the ice is arbitrary, however, fifty vertical layers are used here (one for snow and 49 for ice) to resolve diurnal thermodynamic fluctuations in the ice. A time step of 1 h is used for the diurnal cycle simulations and one day for the daily-average forcing simulations. The model is initialized with *in situ* data discussed in Section 3.2.2.

### *3.2.2 Data*

The field observations were collected in the spring/summer periods during the SIMMS (Seasonal Ice Monitoring and Modeling Site) experiments between 1992-1993 near Resolute, Nunavut. The 1992 field season extended from Year Day (YD) 107 (April 18) to YD 177 (June 27) and the 1993 field season ran between YD 117 (April 27) and YD 170 (June 19). Figure 3.1 shows the geographic locations of Resolute and the first-year ice (FYI) field observations for both SIMMS experiments. The 1992-93 sea ice cycle was selected to perform the numerical model simulations since 1992 was atypically cool and 1993 was atypically warm. In 1992, a prolonged cool period extended into the late spring delaying the onset of melt. The 1993 spring field season was significantly warmer with much earlier melt than in 1992.

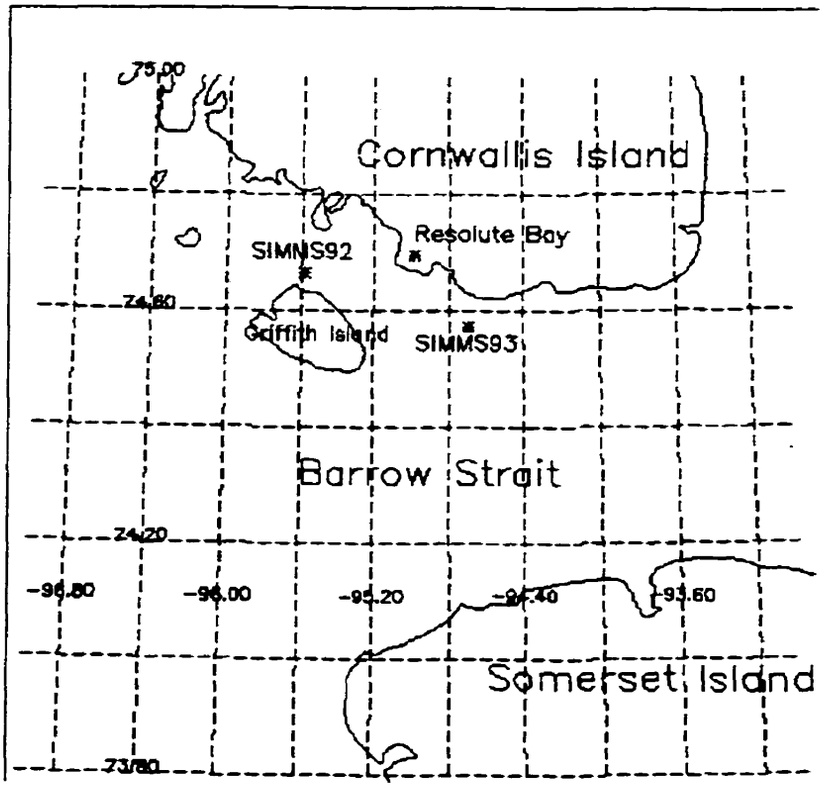



*Figure 3.1: Geographical region of Resolute Bay, Nunavut and the first-year sea ice (FYI) field camps from SIMMS'92 and '93.*

The field observations pertinent to this work include surface energy balance parameters such as incident and reflected short-wave radiation ($K\downarrow$, $K\uparrow$), net short-wave ($K^* = K\downarrow - K\uparrow$) down-welling and up-welling long-wave radiation ($L\downarrow$, $L\uparrow$), net long-wave ($L^* = L\downarrow - L\uparrow$) and net radiation ($Q^* = K^* + L^*$). The observations also include vertical temperature profiles within the snow/ice volumes and snow and ice thickness. A description of the instrumentation and measurement methods can be found in Barber et al. (1994; 1995a). The surface energy balance parameters and snow/ice temperature profiles were logged at 15 or 30 minute intervals depending on the field season. Instantaneous on-the-hour data were used for this work since this corresponds to the shortest time step/forcing used in the numerical model simulations. Snow and ice thickness was measured periodically throughout each field season.

Other surface observations include hourly standard meteorological observations taken at Resolute by the Meteorological Service of Canada (MSC). The pertinent data include air temperature ($T_a$), relative humidity (RH), wind speed (u), precipitation (sfall), and total cloud fraction (amount) (c).

## 3.3 Results and Discussion

### *3.3.1 Temporal Scaling Considerations*

#### *3.3.1.1 Hourly vs. Daily Forcing Comparisons*

The sea ice model (Flato and Brown, 1996) was forced using hourly data and daily average data consisting of land-based MSC Resolute meteorological observations: $T_a$, RH, u, sfall, and c between YD 107, 1992 and YD 170, 1993. For simplicity, the hourly-forced simulation and daily average-forced simulation will be referred to as hourly simulation and daily simulation, respectively. The ice thickness obtained from both models agrees quite well with observations over the 1992 spring period except toward the end of the season (Figure 3.2). More discussion of this is given in section 3.3.3. Ice thickness measurement variability was about ± 0.1 m during the SIMMS field experiment. The 1992 break-up (BU) and freeze-up (FU) dates for the hourly (BU=YD189, FU=YD297) and daily (BU=YD202, FU=YD289) simulations indicates a 21 day longer open water duration for the hourly simulation. Care should be taken when interpreting the BU dates in particular since the model does not account for dynamic sea ice processes (i.e., wind and ocean currents). The modeled BU is defined by the complete disappearance of the sea ice whereas in reality, wind and oceanic forcing may induce BU prior to complete ablation in most Arctic regions. The Canadian Ice Service (CIS) charts

indicated the 1992 BU may have occurred between YD190 and YD197 and FU occurred between YD288 and YD295 within the surrounding Resolute Bay region. This suggests that the model-predicted BU and FU dates above are within the uncertainty of the observed dates. It again should be mentioned that it is not the intent to compare the predictive capabilities (explicitly using BU and FU) between the two forcing simulations, but to examine in a diagnostic sense, the differences between using hourly as opposed to daily average forcing.

Overall, the obvious differences between the hourly and daily simulations in Figure 3.2 are: 1) the open water duration is much longer (by 21 days) in the hourly run with break-up (freeze-up) occurring sooner (later); and 2) the snow ablates earlier but more gradually in the hourly simulation. The snow thickness differences between the field site measurements and model (using Resolute snowfall amounts) will be discussed later.

The 1993 maximum ice thickness was reached on an earlier date in the hourly simulation (2.19 m on YD149) compared to the daily simulation (2.13 m on YD155) (see Figure 3.2). This rather modest difference is related to the differences in freeze-up and break-up. Delayed melt in the daily run during the 1993 spring allowed the maximum ice thickness to be reached later in the year. In addition, the daily simulation ice growth rate is less than the hourly simulation due to a quicker FU date allowing an overall greater accumulation of snow thereby stunting ice growth.

A closer look at the modeled snow thickness during the 1992 spring period (YD120-YD175 in Figure 3.3) reveals that Year Days 133, 134, 139, 142, 149, and 154 onward were days with snow ablation in the hourly simulation, with complete snow ablation by YD165. In contrast, the daily run did not produce snow ablation until YD169 onward, with an abrupt disappearance of the snow pack (over 5 days). The observed snow thickness is much different than the modeled snow thickness due to: 1)

redistribution effects (blowing/drifting snow); 2) differences in precipitation amounts between the field site and Resolute; 3) sfall measurement error at Resolute; and 4) the assumption of a constant dry snow density in the model (330 kg $m^{-3}$).

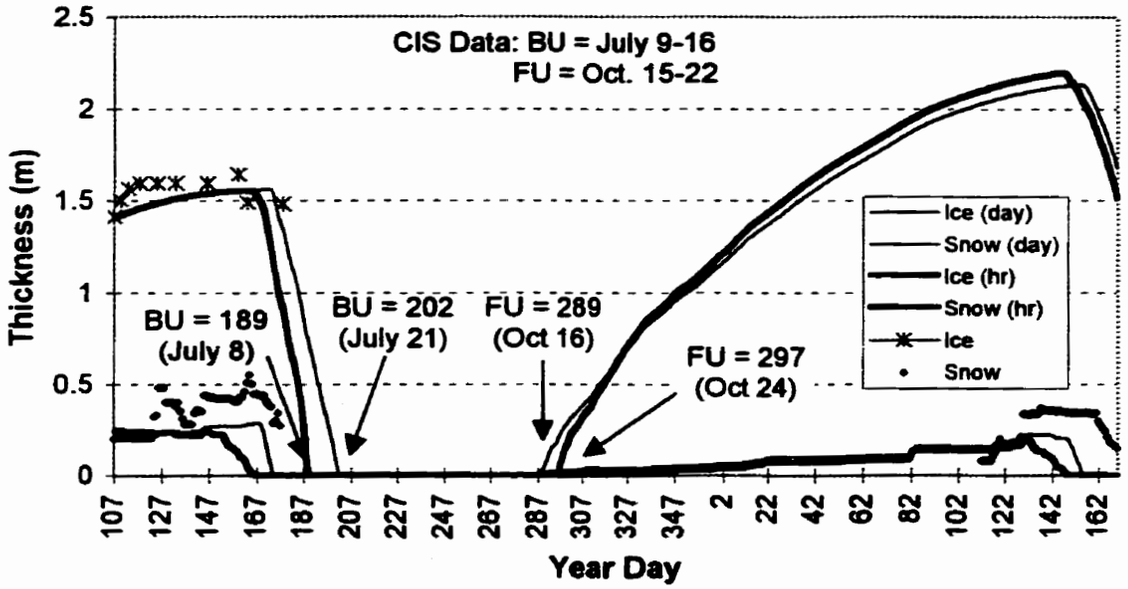



*Figure 3.2: Time series of observed snow (dots) and ice (*) thickness, hourly forcing modeled snow/ice thickness (thick solid lines), and daily forcing modeled snow-ice thickness (thin solid line) from YD107 (April 18), 1992 to YD170 (June 19), 1993. Units are in m. Canadian Ice Service (CIS) estimated freeze-up and break-up dates are also indicated near the top of the figure.*

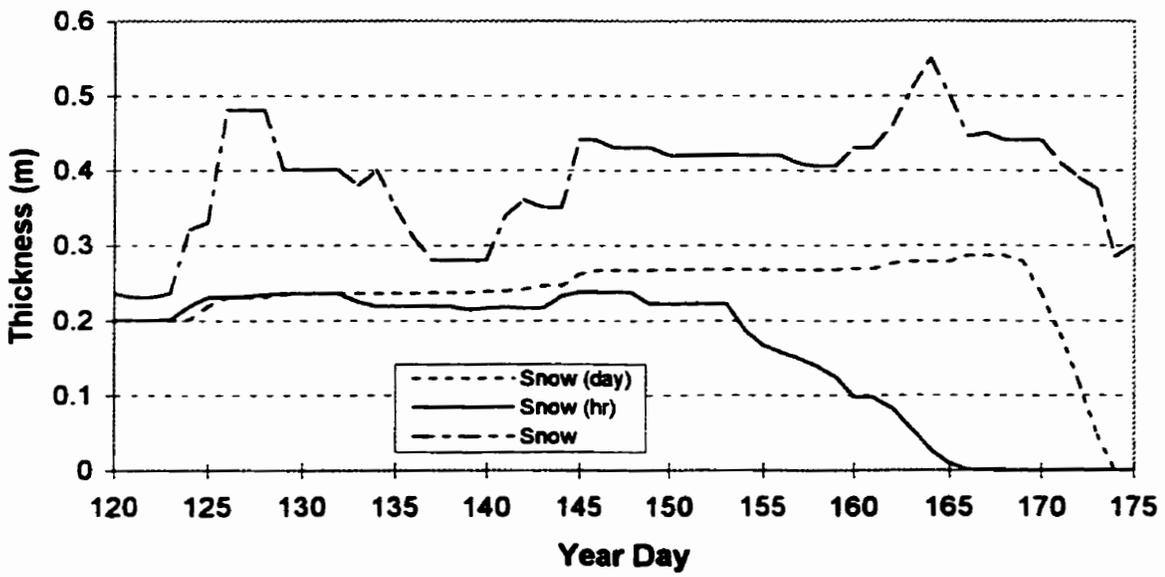



*Figure 3.3: Time series of observed snow thickness (upper-most line), hourly forcing modeled snow thickness (solid line), and daily forcing modeled snow thickness (dashed line) from YD120 (May 1), 1992 to YD175 (June 25), 1992.*

Sporadic diurnal snow ablation (some days with snow melt and some with no melt) and gradual decrease in snow thickness were associated with melting surface temperatures during certain hours, even though air temperatures were well below freezing (Figure 3.4a). However, the air temperature and surface temperature differences in Figure 3.4a are exaggerated in some cases indicated by the observed surface temperatures (see Figure 3.14c). More discussion of this is provided in section 3.3.3. The intent of Figure 3.4a is to show the associated changes in snow characteristics with surface temperature for the hourly simulation and compare them with the daily simulation that did not reach melting snow conditions until YD169 (Figure 3.4b). The gradual decline in snow thickness over the spring period (hourly run) is attributable to the freeze-thaw diurnal cycles of surface temperature as well as the addition of new snow during that period. These diurnal processes are typical of our SIMMS observations.

The diurnal freeze-thaw cycle is more obvious when looking at the modeled surface albedo over the 1992 spring period (Figure 3.5a). The surface albedo drops to the wet snow value (0.65 in the hourly simulation) when surface temperatures reach the melting point and returns to the dry snow value (0.75 in the model) once re-frozen. The albedo does not drop below the dry snow value in the daily run until YD169 (Figure 3.5b). It should also be noted that the diurnal albedo decreases take place during peak to medium solar exposure hours on days associated with diurnal melt. The impact of the diurnal albedo cycle is to increase the short-wave absorption and hence the amount of energy available for snow melt. More discussion of this is provided in the following section.

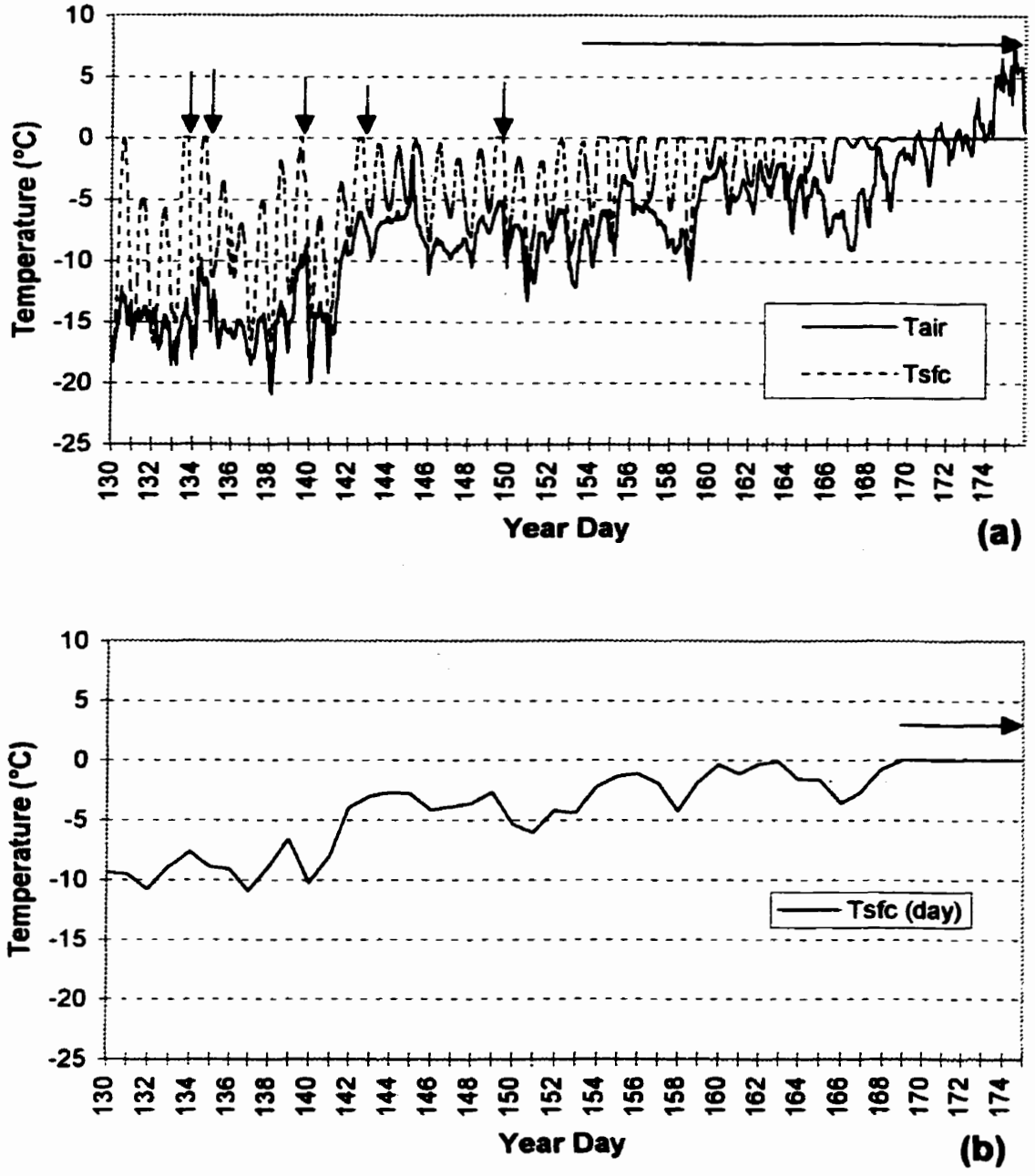



*Figure 3.4: Time series of a) hourly forcing modeled surface temperature and ambient air temperature (°C) and b) daily forcing modeled surface temperature (°C) from YD130 (May 11), 1992 to YD175 (June 25), 1992. Arrows indicate days with surface melt.*

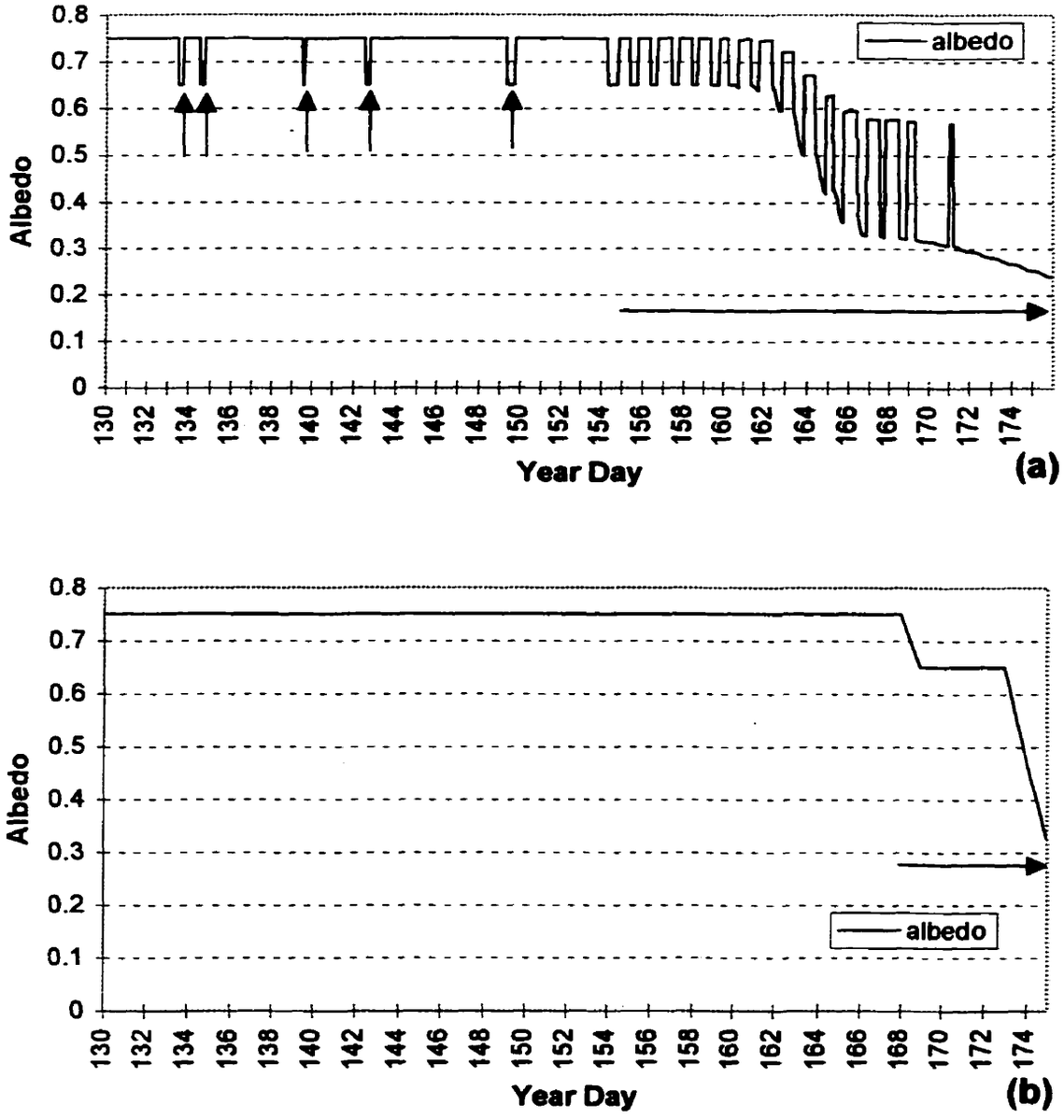



*Figure 3.5: Time series of a) hourly forcing modeled surface albedo and, b) daily forcing modeled surface albedo from YD130 (May 11), 1992 to YD175 (June 25), 1992. Arrows indicate days with surface melt.*

Other diurnal features that are typically observed are vertical temperature waves that propagate through the snow and ice layers due to the cyclic heating/cooling of the snow surface (Figure 3.6). These temperature changes alter the overall dielectric properties of the snow/ice making this phenomenon important for active and passive microwave satellite signatures of sea ice. The modeled vertical ice temperature profile (Figure 3.6) does exhibit these temperature waves in the early to mid spring season [however it is very difficult to see in Figure 3.6 since the amplitudes of these waves (<0.1°C) are significantly less than those observed in the field (0.4-0.5°C maximum)]. The diurnal temperature waves appear only when the snow cover is very thin (< 2 cm) late in the spring season (Figure 3.6). The dampened temperature waves in the early to mid spring season may be the result of the sea ice model's single slab representation of the snow layer and possibly deviations between actual and modeled snow/ice thermal diffusivities and heat capacities. The single-slab snow representation may hinder important thermodynamic processes by not accounting for density changes, snow grain processes, liquid water content and brine inclusions that affect the thermal conductivity of the snow. Formal testing of these speculations are beyond the scope of this work but are addressed again in later chapters.

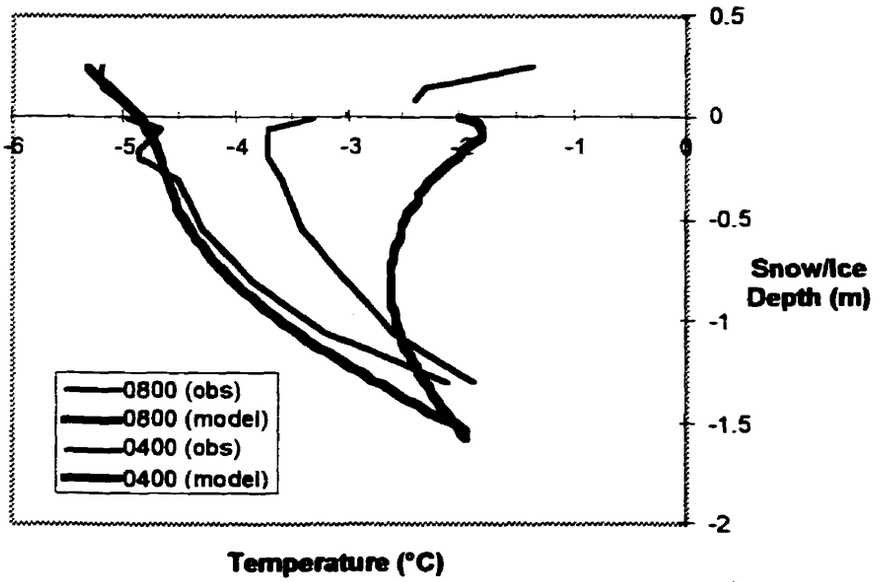



*Figure 3.6: Examples of modeled (thick solid line) and observed (thin solid line) vertical snow/ice temperature profiles during the mid spring season at 0800 solar time on YD148 (colder profiles) and late spring season at 0400 solar time on YD165, 1992 (warmer profiles). Note that the snow array temperature measurements were only available to the +25 cm snow depth (with respect to the snow/ice interface), where actual snow depths were ~ 45 cm on YD148 and YD165. Modeled snow depths were near 25 cm on YD148 and 0 cm on YD165. The large discrepancies on YD165 are due to differences in modeled snow/ice conditions compared to observations (see text).*

### *3.3.1.2 Causes for Hourly vs. Daily Forcing Differences*

In this section I conduct sensitivity tests with the hourly and daily simulations to investigate the processes responsible for the differences in the previous section. From the results above, it is difficult to conclude immediately whether the open water duration difference is due mostly to the difference in snow ablation rates or other processes linked to the forcing differences. In other words, what processes (in the snow or ice) respond differently to shorter temporal scales in the forcing?

In order to investigate the snow ablation effects on the simulation differences, hourly and daily runs were conducted with no snow cover. In this case, the hourly

simulation open water duration is only 3 days longer (with BU occurring 1 day sooner) than the daily model run compared to 21 open water days (BU occurring 13 days sooner) when snow is included. It is clear then that processes involving the snow cover are responsible for the bulk of the break-up date and open water duration differences between the daily and hourly simulations. The question then becomes - what are the controlling processes?

A noticeable difference between the daily and hourly simulation time series of K* (net short-wave flux) is that the hourly run produces a greater daily average K* of 20.1 W $m^{-2}$ (or 1400 W $m^{-2}$ accumulated) over the 1992 spring period (Figure 3.7). The hourly run produces an even greater daily average K* of 30 W $m^{-2}$ (>1400 W $m^{-2}$ accumulated) when considering only the time period from melt onset and beyond. The daily average K* from the hourly simulation was computed by summing the hourly K* values and dividing by the total numbers of hours in the day. The single K* value computed from the daily simulation was then subtracted from the hourly simulation average to plot Figure 3.7. Once again, although the optical depth is held constant (0.7) and the albedo may be the same over the entire day, the daily average K* for the hourly simulation may still be different than the daily simulation K*, as discussed in section 3.2.1. The combination of incident short-wave and absorbed short-wave radiation are therefore quite different between the daily and hourly simulations.

A major difference between the hourly and daily runs is the distribution of down-welling short-wave radiation (K↓) throughout the day. The hourly simulation K↓ is distributed about solar noon, whereas the daily run K↓ distribution is a constant equal to the daily average K↓. This can have significant effects on snow melt initiation and duration. To illustrate this, a sensitivity run (TEST 1) was conducted on the hourly simulation with its hourly K↓ set to its daily average value for each day. In this case, the

BU date was delayed to YD193 and the open water duration shortened by 5 days compared to the original hourly simulation (Table 3.1). This occurred due to a reduced number of days having snow melt as well as a shortening of snow melt duration for each day. This suggests that the diurnal cycle of K↓ has an influence on snow melt initiation, but its neglect accounts for only about a quarter of the hourly/daily simulation difference.

*Table 3.1: The break-up dates and open water duration (days) for the daily model, hourly model, and sensitivity simulations TEST 1, TEST 2, TEST 3, TEST 4 and TEST 5 (see text).*

| | Daily | Hourly | TEST 1 | TEST 2 | TEST 3 | TEST 4 |
|---|---|---|---|---|---|---|
| BU Date | JD202 | JD189 | JD193 | JD195 | JD200 | JD191 |
| OW Duration | 87 days | 108 days | 103 days | 98 days | 91 days | 105 days |

As mentioned previously, the diurnal decreases in albedo (when melt occurred) took place during peak to medium solar exposure hours in the hourly simulation. This implies that an even greater amount of short-wave energy is absorbed by the snow during these hours due to the positive albedo feedback. To investigate the impact of this effect, the albedo was not allowed to decrease to its wet snow value (0.65) during melt events in the hourly simulation (TEST 2). In this case, the BU date was delayed to YD195 and the open water duration shortened to 98 days compared to the original hourly simulation (Table 3.1). This occurred due to a reduction of days having snow melt and therefore a shortening of snow melt duration. TEST 2 indicates that about half of the difference between the hourly/daily simulations can be attributed to the addition of short-wave energy during melt events due to the positive snowmelt albedo feedback.

To investigate the diurnal K↓ distribution and albedo feedback impacts simultaneously, a sensitivity run (TEST 3) was conducted on the hourly simulation combining TEST 1 and TEST 2 (ie: daily-average K↓ and no effect of melt on albedo). In

this scenario, the BU date was delayed to YD200 and the open water duration shortened by 17 days compared to the original hourly simulation (Table 3.1). The combined effects of increased short-wave absorption via albedo feedback and the $K\downarrow$ diurnal distribution therefore explains 10 out of the 12 BU days and 17 out of the 21 open water days difference between the daily/hourly simulations. The remaining difference must be attributable to some other component of the surface energy balance.

Analysis of the surface energy balance reveals that days associated with diurnal surface melt corresponded to days with very low latent and sensible heat fluxes (mostly due to calm winds) at peak daylight hours (Figure 3.8). The premise here is that the timing of very low latent/sensible heat loss in the hourly model was capable of increasing the likelihood of melt during peak to medium solar exposure. In contrast, the daily run uses a daily average wind speed that would in most cases never be zero, thus reducing the likelihood of melt for that day. To investigate this further, an hourly run (TEST 4) was conducted with wind speeds increased to 1.0 m $s^{-1}$ when wind speeds were actually zero for any particular hour. The 1 m $s^{-1}$ was approximately the daily average wind speed for only those days with melt occurring. In TEST 4, the BU date was delayed to YD191 and the open water duration shortened by 3 days compared to the original hourly simulation (Table 3.1). Increasing the wind (latent/sensible heat loss) in this fashion reduced the duration of melt events and also completely eliminated some melt days compared to the original hourly simulation, in turn prolonging BU.

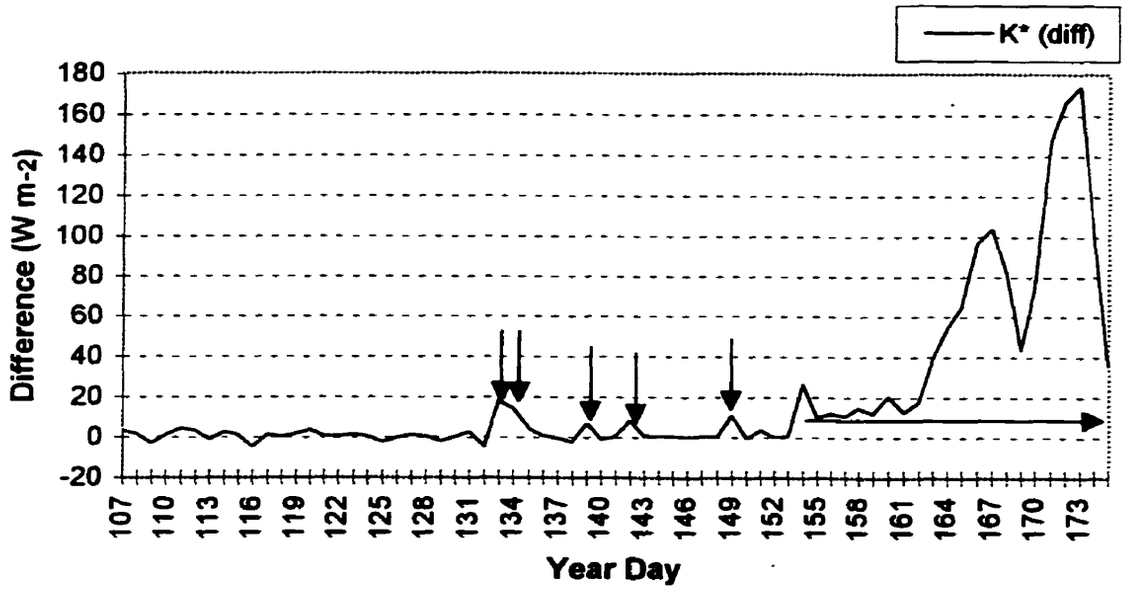



*Figure 3.7: Time series of daily average K* difference (hourly model derived minus daily model) between YD107 (April 17) and YD175 (June 25), 1992. Positive difference: imply greater daily average K* in the hourly model. Arrows signify days with diurnal melt in the hourly model simulation.*

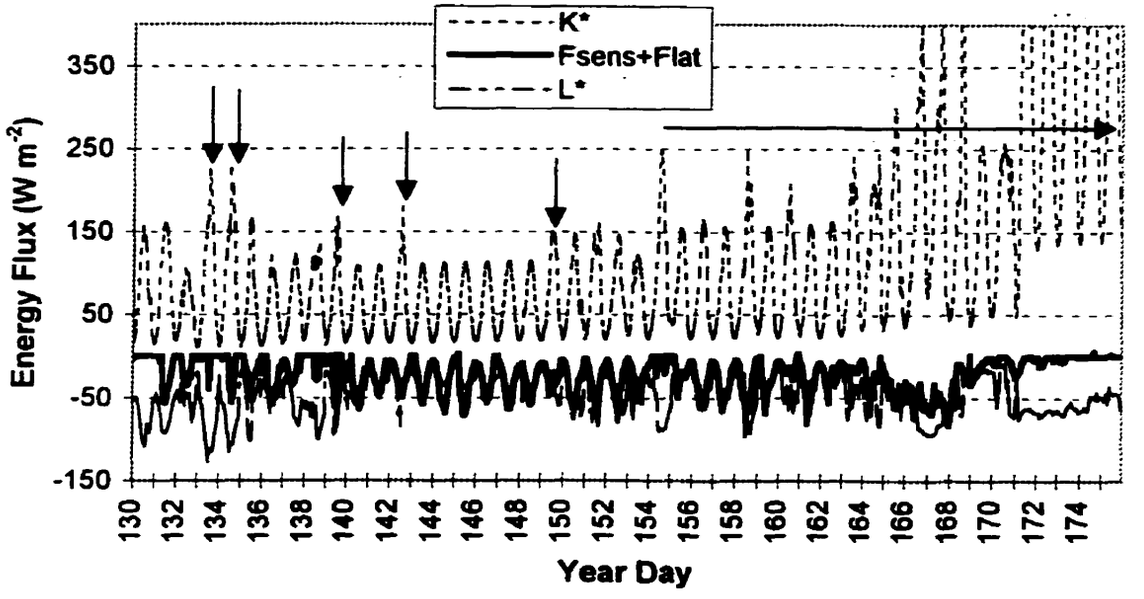



*Figure 3.8: Time series of hourly forcing modeled surface energy balance between YD130 (May 11), 1992 to YD175 (June 25), 1992. Shown are the net short-wave flux (K*), net long-wave flux (L*) and the sum total of latent and sensible heat fluxes (Fsens + Flat). Units are in W m $^{-2}$ . Positive values represent surface energy gain. Arrows indicate days with diurnal surface melt.*

It is useful to quantify the combined effects of all processes above on the daily/hourly simulation differences (ie: TEST 3 and TEST 4 simultaneously). In producing an hourly simulation when combining TEST 3 and 4 (TEST 5), the BU date and open water duration become very similar to the original daily simulation (Table 3.1). TEST 5 therefore shows that the bulk of the daily/hourly simulation differences can be attributed to: 1) the diurnal distribution of K↓ which acts to enhance snow melt events at peak to medium solar exposure hours; 2) the timing of very low latent/sensible heat loss that occurs during peak to medium solar exposure hours, further increasing the likelihood of a melt event; and most importantly 3) the timing of a melt event (during peak to medium solar exposure) which enhances the absorption of short-wave energy into the snow pack by reducing the surface albedo.

In further interest of model forcing, another model run was generated using forcing data by linearly interpolating the daily average values to hourly intervals which were then used to force the hourly simulations. This is typically done in modeling studies over larger time scales where monthly data may be interpolated to daily values for daily forcing or even shorter time steps (see for example, Maykut and Untersteiner (1971); Ebert and Curry (1993). The daily average incident short-wave radiation was also interpolated to hourly intervals, similar to the other meteorological forcing. In this case, the BU and open water duration dates were exactly midway (BU = YD196; open water duration = 99 days) between the standard hourly simulation and daily simulation. This is primarily due to the enhanced short-wave responses of the model that act on the snow layer over hourly time scales (as shown by TEST 3). This is illustrated further when the normal diurnal K↓ distribution is used instead of the hourly interpolated daily average K↓ forcing, where the simulated BU and open water duration becomes closer to the original hourly simulation (BU = YD192; open water duration = 103 days).

### *3.3.2 Parameterized vs. In Situ Radiative Fluxes and Albedo*

In this section I compare parameterizations such as incident short-wave (K↓) and long-wave (L↓) radiation and surface albedo ($\alpha$) used in the Flato and Brown (1996) model with *in situ* SIMMS'92/93 field data. The purpose is to assess whether the parameterizations used in the model adequately represent specific processes over shorter temporal scales and whether they accurately reflect 'on ice' measurements. This would then estimate the potential error in modeled ice and snow thermodynamics due to errors in forcing. It should be emphasized that differences between modeled and *in situ* data should be expected since the SIMMS'92/93 field sites were roughly 30 km from Resolute leading to differences in atmospheric forcing. In addition, some model parameterizations are intended to reflect area-averaged conditions and so may not represent the point conditions measured at the SIMMS field sites (eg. albedo). Unfortunately, no cloud observations were available during the SIMMS experiments, thus cloud variation comparisons (and the associated changes in K↓ and L↓) with Resolute were not possible. These factors create a rather crude comparison but can nevertheless illustrate the spatial and temporal scale characteristics between parameterized and *in situ* albedo/radiative fluxes.

As an example of atmospheric forcing differences between Resolute and the field sites, Figure 3.9 shows the time series of air temperature differences between Resolute and SIMMS'92 observations (Resolute - SIMMS). A mean bias of +0.5°C in Figure 3.9 indicates that land-based temperatures (Resolute) can be somewhat larger compared to on-ice conditions (SIMMS). Major differences occur during the morning hours (overnight lows) where land-based temperatures can be significantly warmer (~ 6-8°C) than over the ice.

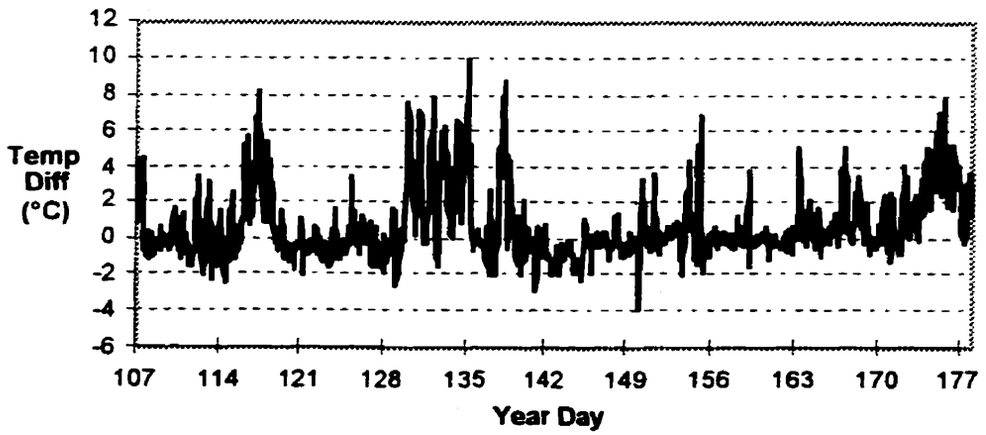



*Figure 3.9: Time series of air temperature discrepancy (Resolute minus SIMMS) between Resolute and the SIMMS'92 field site between YD107 and YD177, 1992.*

The hourly time series of K↓ and L↓ error (modeled - observed) over the spring period of 1992 are shown in Figures 3.10a-c. The K↓ error time series (in Figure 3.10a) only includes the first 9 days of the simulation/field experiment (for illustrative purposes) since it is difficult to visualize the entire time series. Low sun angle (< 25°) data were excluded from the analysis due to instrument limitations. The coefficient of determination between the modeled and observed K↓ was $R^2 = 0.91$ with the error varying between nearly zero and ± 250 W $m^{-2}$ over the entire SIMMS'92 field period. The K↓ mean error was -23 W $m^{-2}$ with a standard deviation of 89 W $m^{-2}$. The K↓ all sky mean error found here is larger than the mean error found by Key et al. (1996) (-3.2 W $m^{-2}$ for the Shine (1984) scheme) who compared daily average observed and parameterized fluxes for two Arctic locations. Differences in local cloud conditions between the SIMMS field sites and Resolute may be a significant factor.

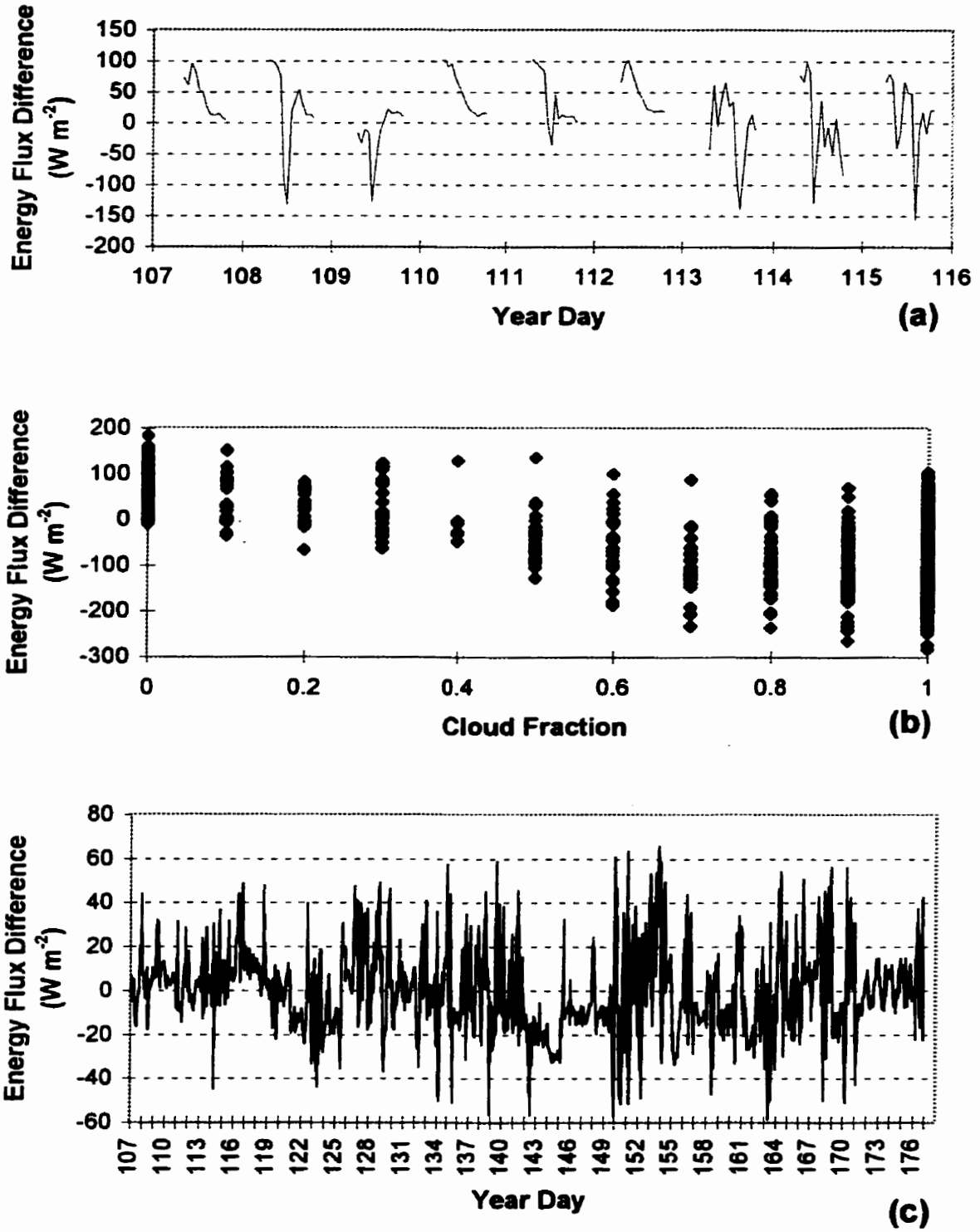



*Figure 3.10: Modeled minus observed (a) incident short-wave error between YD107-116, 1992, (b) incident short-wave error versus total cloud fraction, and (c) incident long-wave error between YD107-177, 1992. Low sun angle (< 25°) data were excluded from the short-wave analysis due to measurement limitations.*

The K↓ parameterization tended to overestimate clear sky K↓ by roughly 40-50 W $m^{-2}$ on average and K↓ differences varied widely during complete overcast conditions (+125 W $m^{-2}$ to -275 W $m^{-2}$). This can be visualized in Figure 3.10b where cloud-free conditions produced more positive biases (mean of +40 W $m^{-2}$) in K↓, and when complete overcast conditions occurred the discrepancy between modeled and observed values grew much larger with a more negative bias (-50 W $m^{-2}$). The main differences between modeled and observed K↓ is due to location/cloud differences between the SIMMS field sites and Resolute as well as cloud optical depth (held constant at 0.7) and tropospheric aerosol inhomogeneities pointed out by Key et al. (1996). Key et al. (1996) showed that the Shine (1984) clear sky and all sky K↓ parameterizations did not have sun angle biases but the all sky parameterizations had greater errors as the solar zenith angle (SZA) decreased. The 55-70° SZA errors for K↓ all sky point estimates were shown to range between ± 150 W $m^{-2}$ (Key et al., 1996) which may be a significant contributor to the errors observed here. Larger discrepancies between modeled and observed K↓ values are expected to occur during total overcast situations due to the factors above, indicated in Figure 3.10b.

The time series of modeled-observed differences for L↓ (Figure 3.10c) are not as extreme as K↓ and no biases were found when comparing cloud-free conditions and total overcast conditions (not shown). However, the discrepancies in L↓ (modeled-observed) grew larger (+60 W $m^{-2}$ to -40 W $m^{-2}$) during complete overcast events compared to clear skies (+20 W $m^{-2}$ to -40 W $m^{-2}$) as expected due to similar factors as those for K↓. The mean L↓ error was -1 W $m^{-2}$ with a standard deviation of 18.8 W $m^{-2}$ and an $R^2 = 0.91$. The mean error found here is smaller than the all sky L↓ value (-9.7 W $m^{-2}$) reported by Key et al. (1996) for the same L↓ parameterization of Maykut and Church (1973).

The time series of surface albedo shows that the dry/new snow values used in the

model (0.75) are lower than observed (0.8-0.9) (Figure 3.11). The modeled albedo is intended to represent an area average that may differ from local conditions. Typical Arctic basin dry/new snow albedo observations have been shown to range between 0.8-0.98 (Radionov et al., 1997). The mean dry/new snow albedo for SIMMS'92 was 0.83 and 0.82 for SIMMS'93 (excluding the low sun angle data). The larger observed albedo compared to the modeled albedo may be significant since less short-wave radiation would be absorbed by the snow layer and may lead to less days with surface melt (see section 3.3.3).

In addition, there were several days when diurnal changes in albedo were observed during SIMMS'92/93 similar to what is portrayed in the model simulation. However, the model simulated more frequent diurnal melt events. Definite diurnal albedo patterns associated with snow surface freeze-thaw cycles were observed after YD168, 1992 where a mean albedo decrease of 0.07 ± 0.03 took place during snow melt events (Figure 3.12). Other Arctic observations indicate maximum diurnal albedo changes on the order of 0.11 (Radionov et al., 1997). The model assumes a 10% decline in albedo during snow surface melt over an area average. Observations suggest the associated drop in modeled albedo due to diurnal snow melt may be slightly excessive in most cases. This could also lead to enhanced snow melt and more frequent melt events in the model simulation. Further discussion will be provided in section 3.3.3.

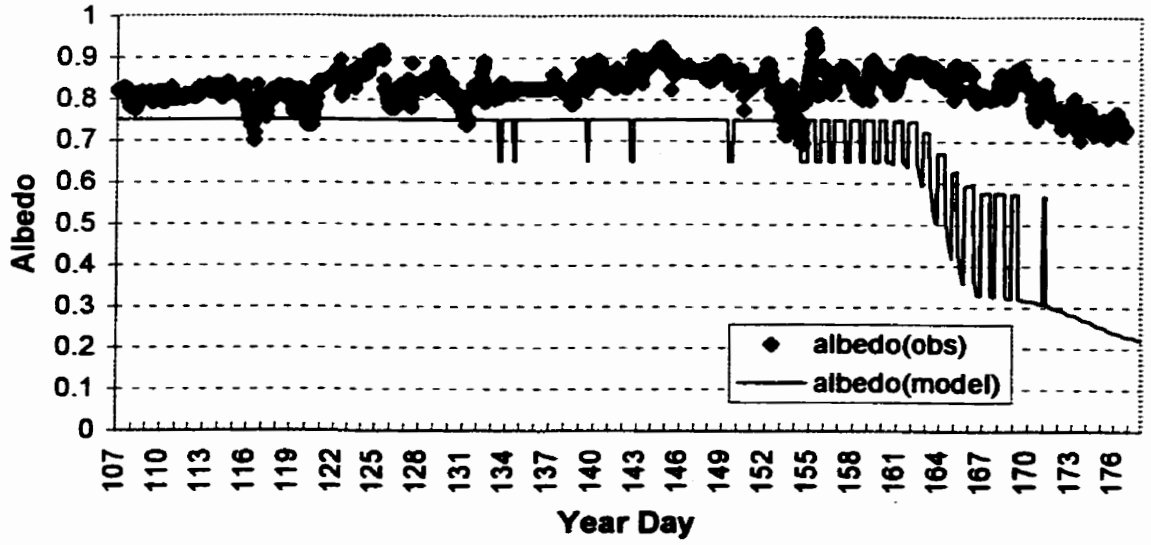



*Figure 3.11: Time series of parameterized (solid line) and observed (*) surface albedo between YD107 and YD177, 1992.*

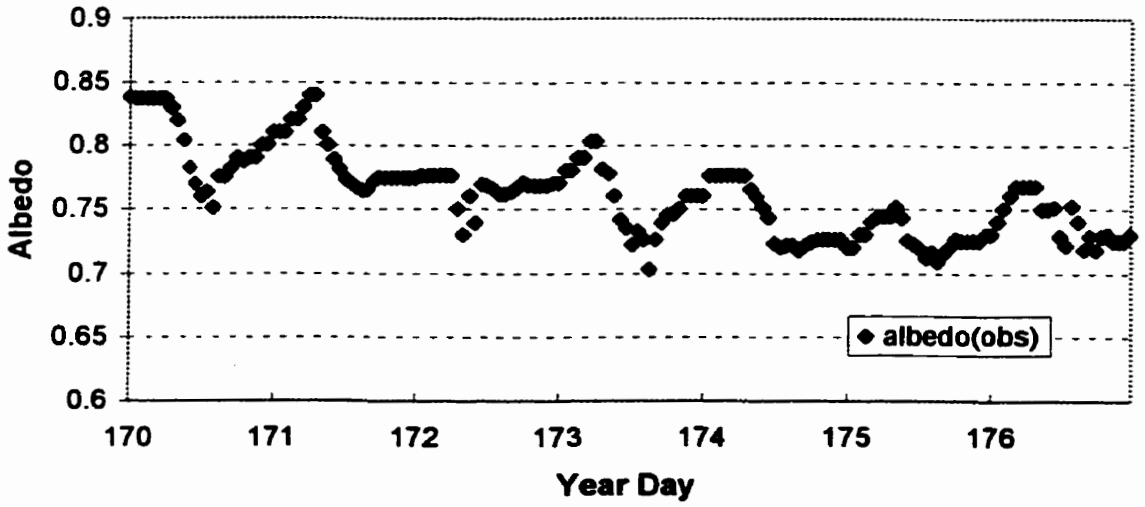



*Figure 3.12: Diurnal time series of SIMMS'92 observed surface albedo between YD170-177.*

### *3.3.3 Model Simulations Using Field Observation Forcing*

The sea ice model was run with SIMMS'92 field data ($K\downarrow$, $L\downarrow$, $T_a$, wind speed and surface albedo) to investigate differences between using on-ice forcing as opposed to land-based forcing during the spring of 1992. The simulation was conducted for days 107-176, 1992 since full day field data did not exist beyond YD176. Direct comparisons of modeled surface temperatures ($T_{sfc}$), $Q_{net}$ and snow/ice ablation rates were made between the standard hourly simulation (CONT) (section 3.3.1) and the SIMMS field data forcing simulation (SFS) over the 1992 spring period. Since relative humidity was not recorded at the SIMMS'92 field site, Resolute observations were used in the SFS simulation. The SIMMS'92 snow thickness measurements were not used to estimate actual snowfall amounts since it was impossible to exclude redistribution effects (blowing/drifting snow) from the data. Cloud fractions were not necessary since $K\downarrow$ and $L\downarrow$ field data were used instead of the $K\downarrow$ and $L\downarrow$ parameterizations. Resolute wind data was used between YD107-117 since usable SIMMS'92 wind speed data were only available beginning on YD118.

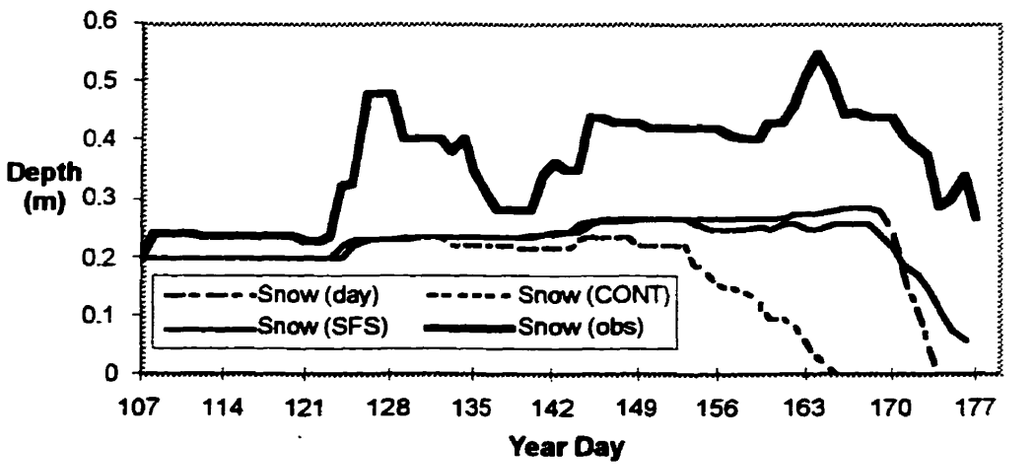



*Figure 3.13: Comparison between observed (thick solid line), CONT (lower line), SFS (solid line), and daily forcing (dash-dot line) snow depth between YD107-177, 1992.*

Using the SIMMS field data as forcing and replacement of the K↓, L↓ and albedo parameterizations produces very different spring period simulations compared to the control simulation (Figures 3.13 and 3.14). The snow layer does not disappear in the SFS run (minimum of 0.07 m on YD176) compared to complete ablation in CONT by YD165 (Figure 3.13). The SFS snow ablation is similar to the daily forcing model simulation but SFS has a much more realistic ablation curve compared to observations beyond YD165 (Figure 3.13). SIMMS researchers did not observe complete snow ablation and advanced melt (ponding) conditions in 1992 (p. comm, T. Papakyriakou). This suggests that the SFS simulation more accurately reproduced the regional conditions since complete snow ablation and advanced melt were not generated in the SFS run, whereas, CONT produced complete snow ablation and advanced ice melt stages.

The SFS run produced diurnal melt events similar to CONT, however, the onset of melt was much later (YD154 in SFS, YD133 in CONT) causing the number of melt events to be significantly reduced in the SFS simulation between YD107-177 (see Figure

3.13). Surface melt took place in the SFS simulation on YD154, 155, 160, 162, 163, and 168 onward. The later melt onset is most likely due to the overall increase in surface albedo (SFS mean $\alpha$ = 0.828 ± 0.5, CONT mean $\alpha$ = 0.63 ± 0.1) and wind speeds (SFS mean u = 4.1 ± 2.4 m $s^{-1}$, CONT mean u = 1.4 ± 1.0 m $s^{-1}$; the means were calculated between YD107-176, 1992). This suggests that the stronger winds in SFS compared to CONT play a significant role in the melt onset dates. It is interesting to note that the mean K↓ and L↓ for SFS (272.6 W $m^{-2}$, 232.4 W $m^{-2}$) and CONT (272.3 W $m^{-2}$, 231.4 W $m^{-2}$) were very similar but the SFS standard deviations ($\sigma$(K↓) = 195.2 W $m^{-2}$, $\sigma$(L↓) = 45.3 W $m^{-2}$) were slightly larger than CONT ($\sigma$(K↓) = 176.4 W $m^{-2}$, $\sigma$(L↓) = 38.3 W $m^{-2}$). This may indicate that although the mean radiative forcing was similar for both simulations, the different variations in radiative forcing may account for some of the differences between the simulations. Differences in air temperature between SFS and CONT is expected to have a minor influence on the snow layer simulation since $T_a$ was quite similar in both simulations (SFS $T_a$ mean = -11.7 ± 8.3°C, CONT $T_a$ mean = -12.0 ± 8.4°C).

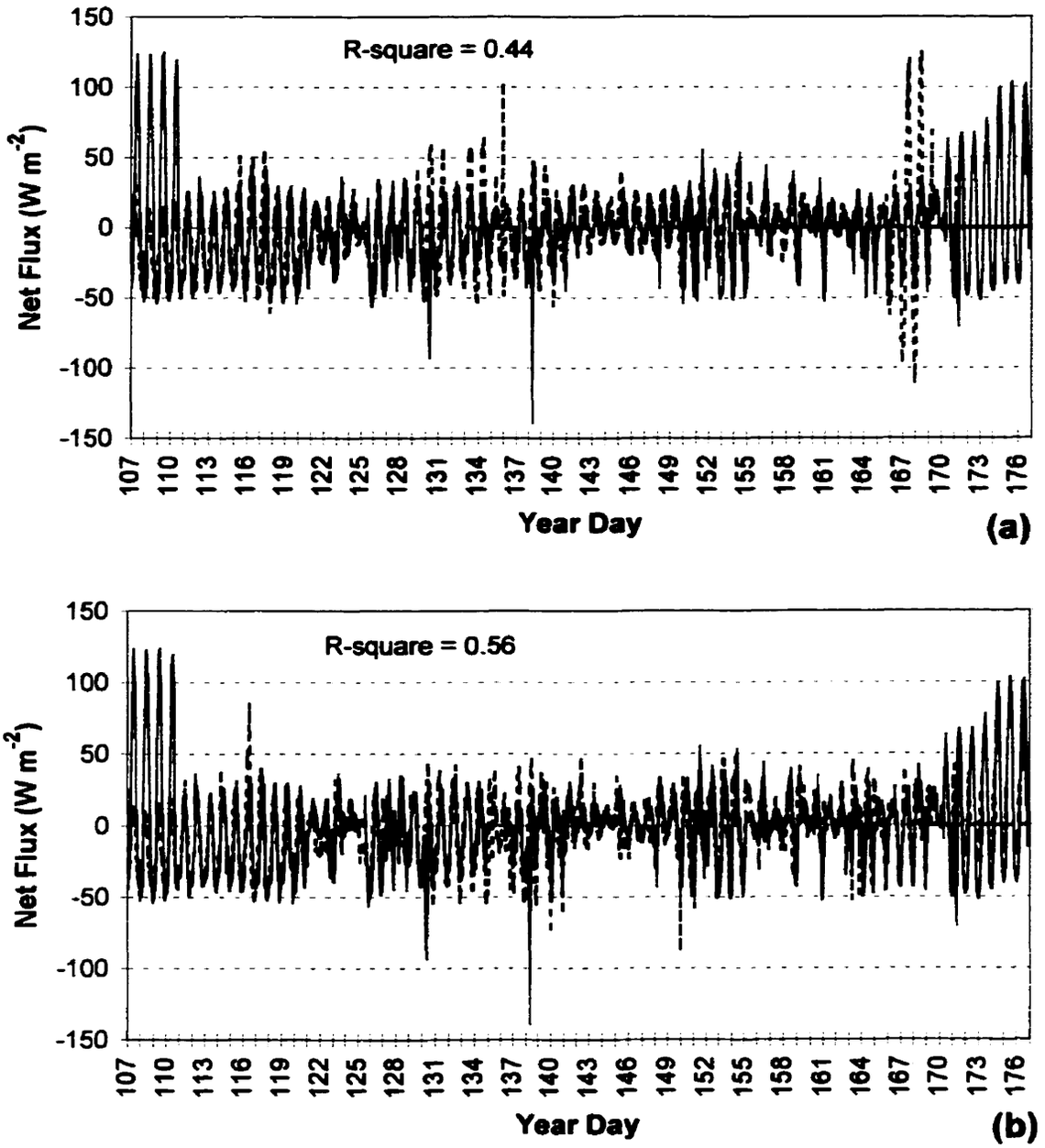



*Figure 3.14a,b: Comparisons of Qnet between (a) CONT (black line) and SIMMS'92 observations (red line), (b) SFS (black line) and SIMMS'92 observations (red line).*

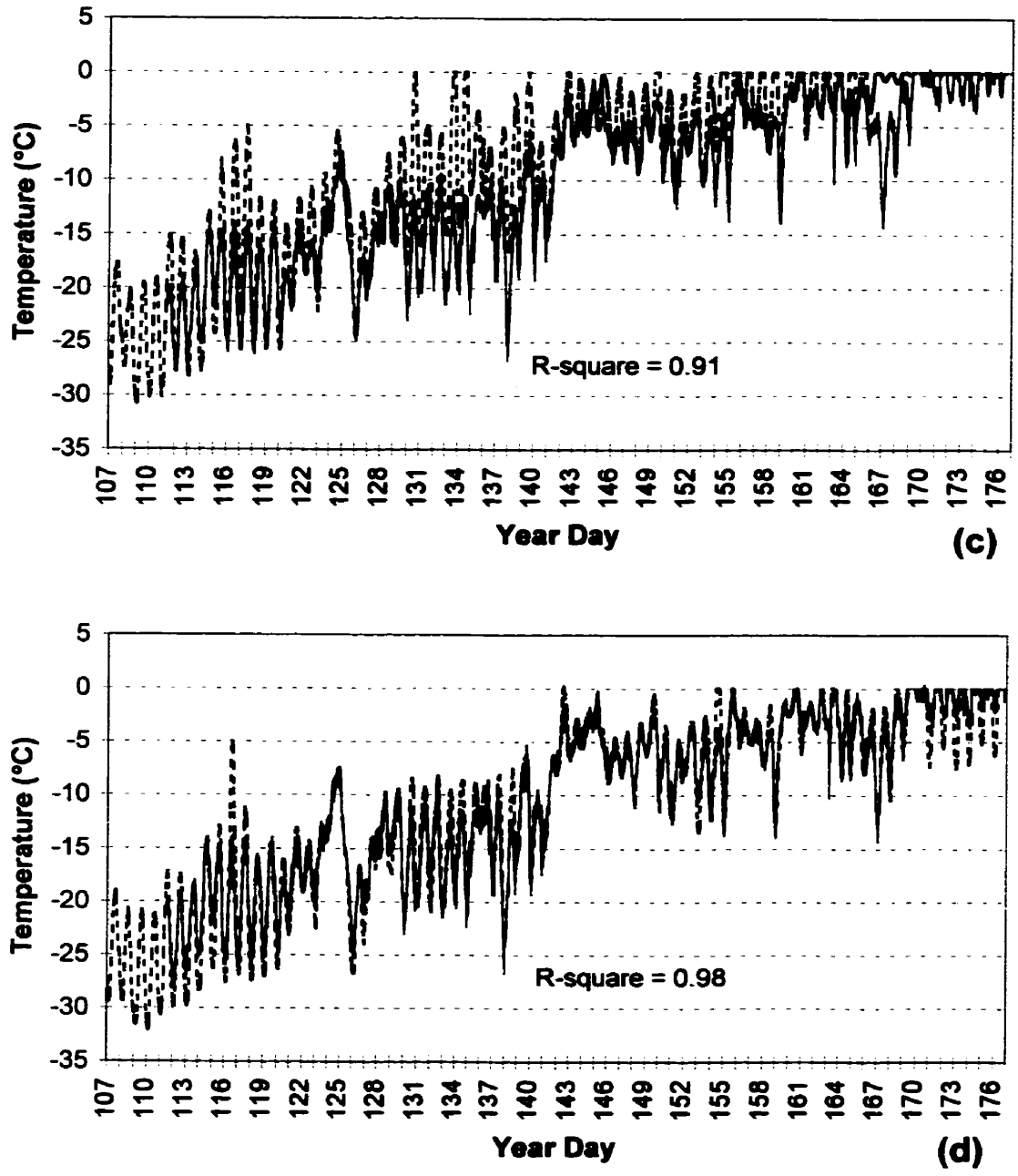



*Figure 3.14c,d: Comparisons of Tsfc between (c) CONT (black line) and SIMMS'92 observations (red line), (d) SFS (black line) and SIMMS'92 observations (red line).*

The SFS simulation produced improved estimates of $Q_{net}$ and $T_{sfc}$ compared to CONT, determined by direct comparison with SIMMS'92 measured values (Figure 3.14a-d). The $Q_{net}$ $R^2$ for SFS is larger than the CONT simulation and the $Q_{net}$ mean error (modeled-observed) and standard deviation (error) for SFS was slightly smaller than the CONT simulation (Table 3.2). YD107-110 were excluded from the statistics due to $Q_{net}$ measurement errors on those days.

*Table 3.2: Statistical comparisons ($R^2$, mean error (modeled - observed) and standard deviations) between the modeled (CONT and SFS) and observed net surface flux ($Q_{net}$) and surface temperatures ($T_{sfc}$). CONT is the standard hourly model simulation using 'land-based' forcing (from Section 3.3.1) and SFS is the 'on-ice' forcing simulation using SIMMS'92 field data.*

| | Qnet CONT/obs | Qnet SFS/obs | Tsfc CONT/obs | Tsfc SFS/obs |
|---|---|---|---|---|
| R-Square | 0.44 | 0.56 | 0.91 | 0.98 |
| Mean Error | -2.5 | -2 | 2.5 | -0.08 |
| St. Dev. | 30 | 25 | 3.1 | 1.7 |

The improvement in simulation using the field observations compared to using Resolute forcing and model K↓, L↓ and albedo parameterizations is also manifested when comparing the SFS and CONT $T_{sfc}$ time series (Figure 3.14c,d). $T_{sfc}$ was estimated using the measured L↑ with an emissivity of 0.99 (similar to the model prescribed value) since direct measurements of surface temperatures were not available. The $T_{sfc}$ $R^2$ for SFS is noticeably larger than the CONT simulation and the $T_{sfc}$ mean error (modeled-observed) and standard deviation (error) for SFS was significantly smaller than the CONT simulation (Table 3.2). YD107-110 were excluded from the statistics due to unavailable L↑ measurements on those days. The SFS simulation reproduced all of the days that had diurnal melting in the observations (Figure 3.14d). Observed L↑ fluxes indicated that surface melt may have occurred on YD155, 160, 162, 163 and 169 onward. The SFS

simulation produced diurnal melt on all of these days as well as YD154 and 168. It is not known why there is discrepancy on these days. The statistics and Figures 3.14c,d suggest that the SFS simulation provides a better representation of the measured $T_{sfc}$ than CONT over the spring of 1992.

## 3.4 Conclusions

The purpose of this chapter is to illustrate whether diurnal time scales are important to consider in modeling seasonal snow-covered FYI and whether the model parameterizations appropriately handle these time scales. This is important not only for resolving these time scales and processes in the models for arctic climate studies, but for advancing the model to a state that will be useful for microwave remote sensing applications. This chapter was also concerned with spatial differences inherent in the model environment through its forcing data. Typically, sea ice models use "land-based" forcing data to drive the simulations, however, difficulties may arise due to differences between "on-ice" meteorological conditions and "land-based" conditions. Three research questions were designed to address the above issues.

### *3.4.1 Question 1*

Direct comparisons between a one-dimensional thermodynamic sea ice model using hourly forcing data and daily average forcing data showed significantly different results in terms of break-up dates, open water duration and snow ablation. Break-up occurred 13 days sooner, open water duration was lengthened by 21 days and snow ablation was initiated sooner but progressed gradually and more realistically (compared to

field observations) using hourly forcing. However, the hourly simulation produced much earlier melt onset and created advanced melt stages in 1992 which were not observed by SIMMS'92 researchers. Sensitivity tests showed that 12 of the 13 break-up day and 18 of the 20 open water day differences between the daily/hourly simulations is due to snow cover evolution differences within the simulations. Analysis has showed that short-wave exchange dominates these snow processes, as observed in the field. The diurnal distribution of down-welling short-wave energy is responsible for about 33% of the snow cover evolution (and hence the daily/hourly simulation) differences by acting to enhance snow melt events at peak to medium solar exposure hours. The diurnal variations in latent/sensible heat fluxes (mostly due to wind speeds) can explain about 17% of the daily/hourly model differences where some days have very low latent/sensible heat loss during peak to medium solar loading hours. The timing of these low latent/sensible heat loss events create a greater likelihood of a melt event to take place. In contrast, the daily average wind speed is rarely zero, thus reducing the likelihood of melt for that day. The dominant process, which leads to 50% of the snow cover evolution differences, is enhanced diurnal short-wave absorption due to albedo decreases during diurnal melt events occurring at peak to medium solar exposure hours (ie: a positive albedo feedback). It can therefore be concluded that non-linearities in the surface energy balance play a critical role in the melting of the snow layer which in turn affect the seasonal evolution (especially the spring melt period) of first-year sea ice.

### *3.4.2 Question 2*

Comparisons between observed and model-parameterized $K\downarrow$, $L\downarrow$ and albedo showed the model parameterizations represented observations fairly well in some cases

and not so well in others. The Shine (1984) parameterization for K↓ had a positive error of roughly 40 W $m^{-2}$ on average during clear conditions and a slight negative bias of -50 W $m^{-2}$ under complete overcast conditions. The mean K↓ error (modeled-observed) was -23 W $m^{-2}$ with a standard deviation of 89 W $m^{-2}$ over the SIMMS'92/93 field seasons. The mean error here is larger than that reported by Key et al. (1996) (-3.2 W $m^{-2}$). Larger K↓ discrepancies occurred under cloudy conditions compared to clear skies due to differences between SIMMS field sites and Resolute, NWT, (spatial/location limitations) as well as cloud optical depth (held constant in the model) and tropospheric aerosol inhomogeneities (Key et al., 1996). Errors can also be largely attributable to solar zenith angle variations as suggested by Key et al., 1996.

The L↓ parameterization of Maykut and Church (1973) had smaller errors and no trends were found when comparing cloud-free conditions to clear skies. The mean L↓ error was -1 W $m^{-2}$ with a standard deviation of 18.8 W $m^{-2}$. The mean error is smaller than that of Key et al. (1996) (-9.7 W $m^{-2}$) and shows this L↓ parameterization performs well even at hourly intervals. However, the low L↓ mean error and standard deviation may be the result of a cancellation of errors due to cloud amount bias between Resolute and the SIMMS field site (spatial limitations).

The time series of modeled and observed surface albedo showed that the model may underestimate the dry/new snow albedo by roughly 0.05 to 0.15. The sea ice model assumes a dry/new snow albedo of 0.75 but regionally representative SIMMS'92/93 albedos were between 0.8-0.9 in the early spring periods. The modeled albedo decrease for wet snow (0.1) was slightly higher than observed (0.07 ± 0.03) and so the model's snow melt albedo feedback may be slightly too strong.

### *3.4.3 Question 3*

When the sea ice model was forced directly with SIMMS'92 field data ($K\downarrow$, $L\downarrow$, air temperature, wind speed and albedo) it produced better representations of the modeled snow/ice ablation evolution as well as $Q_{net}$ and surface temperature over the 1992 spring period. This case reproduced observed diurnal melt periods faithfully whereas the control run produced many more and earlier melt events. Also, complete snow ablation and advanced melt were not generated in this case, in agreement with observations. The overall improvement when forced with SIMMS observations was primarily due to greater turbulent fluxes (larger wind speeds), higher surface albedo and different radiative forcing compared to using land-based forcing (Resolute) and modeled $K\downarrow$, $L\downarrow$ and albedo parameterizations. I also note, based on eight (1990- 1997) consecutive years of the SIMMS and follow on C-ICE (Collaborative-Interdisciplinary Cryospheric Experiment) experiments that the differences between surface energy fluxes observed on-ice and those estimated from nearby land stations do not appear to be systematic. This means that it may be difficult to establish a correction term to account for various flux terms required in thermodynamic modeling of sea ice.

## 3.5 Summary

In summary, results suggest diurnal processes are important to consider and the parameterizations handle these time scales well in some cases and not so well in others.

This is an important step toward understanding the necessary model forcing time scales required for linking these types of models to microwave remote sensing in later chapters. The results also illustrate the time scales necessary to resolve when attempting to simulate annual cycles of snow-covered FYI. It is also important to consider the environmental conditions on the ice surface as opposed to land for monitoring sea ice evolution. That is, if *in situ* data are not available on the ice, remote observations from satellites may provide insight to environmental conditions there. The following chapter (Chapter 4) further investigates and supplements the analysis in this chapter in terms of the incident short-wave and long-wave model parameterizations. Incident radiation is a critical process within the SEB of FYI and a more thorough analysis of their parameterization schemes is warranted given the results thus far.

# CHAPTER 4: Parameterization Schemes of Incident Radiation

## 4.1 Introduction

Thermodynamic sea ice models require precise parameterization of incident short-wave and long-wave fluxes over a variety of time and space scales as illustrated in Chapter 3. Chapters 1 and 2 outlined previous investigations of these fluxes, difficulties while trying to parameterize them and their importance to the snow/sea ice SEB. This chapter builds on results in chapter 3 by investigating in more detail sea ice model parameterization schemes of incident radiation. Coincident *in situ* radiation data comparing different arctic environments and spatial variations have previously never been available. Data presented here were collected in a marine polynya environment including terrestrial, shore fast-ice, marginal ice zone and open water regions during the International North Water Polynya (NOW) Project conducted between March – July 1998 (Barber et al., 2001). The two research question of this chapter are:

1) Are there any seasonal or environmental/spatial biases within selected $K\downarrow$ and $L\downarrow$ parameterizations compared against *in situ* field measurements?
2) Can we improve the parameterization representations for the different environments where possible?

Results from this investigation have previously appeared in the peer reviewed literature (Hanesiak et al., 2001a).

## 4.2 Data and Methods

### *4.2.1 Observational Data*

Data collected were from the International North Water Polynya (NOW) Project during an intensive field campaign between March to the end of July 1998 (Barber et al., 2001). Three primary data collection platforms included a terrestrial camp (Cape Herschel) and fast-ice site (Rosse Bay) on the east coast of Ellesmere Island, Nunavut and the Canadian Coast Guard ice breaker Pierre Radisson in the NOW region (Figure 4.1).

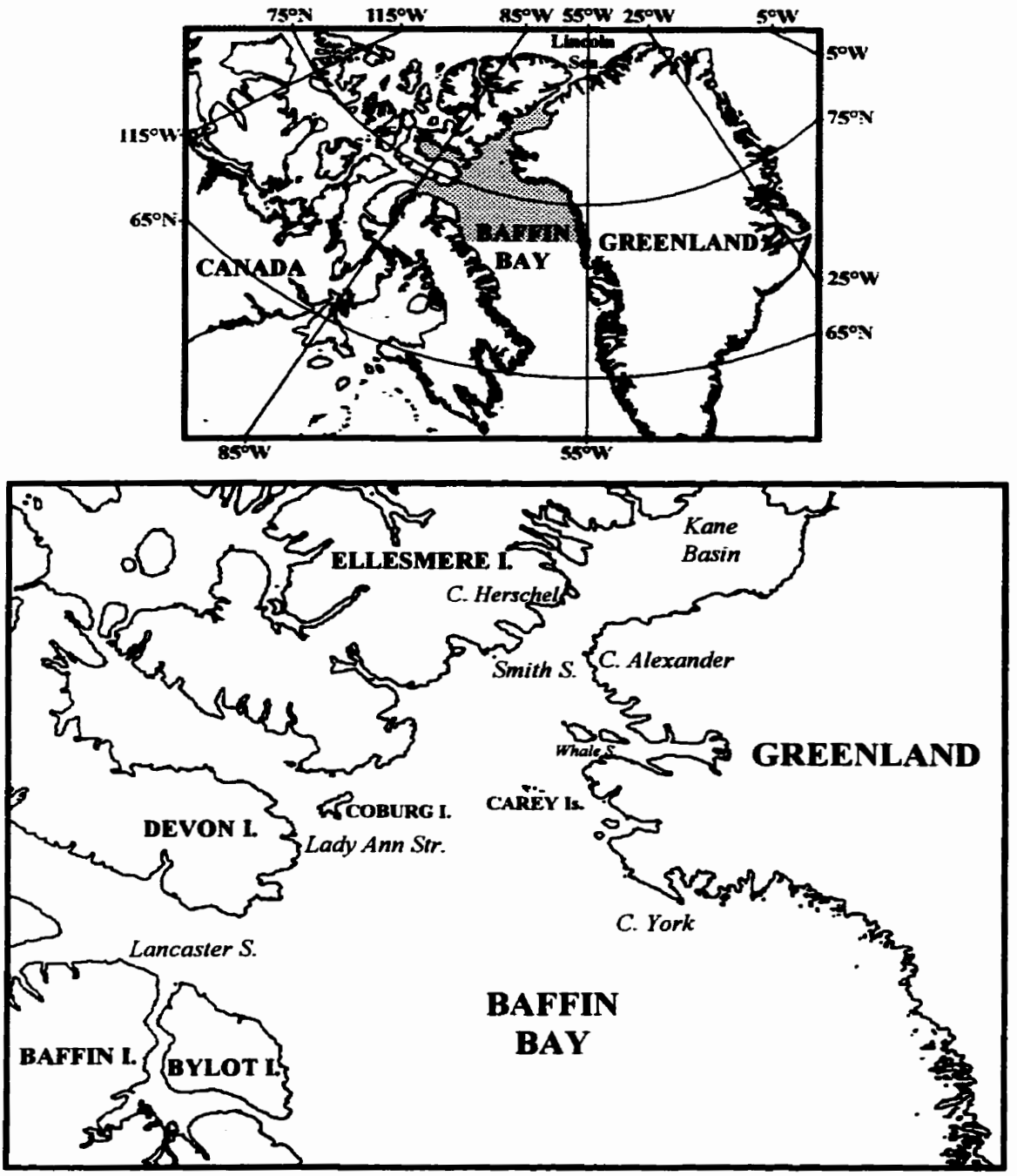



*Figure 4.1a: Geographical region of the NOW'98 project (shaded region in top image). Smith Sound, Kane Basin and Cape Herschel appear at the north end of NOW (bottom image).*

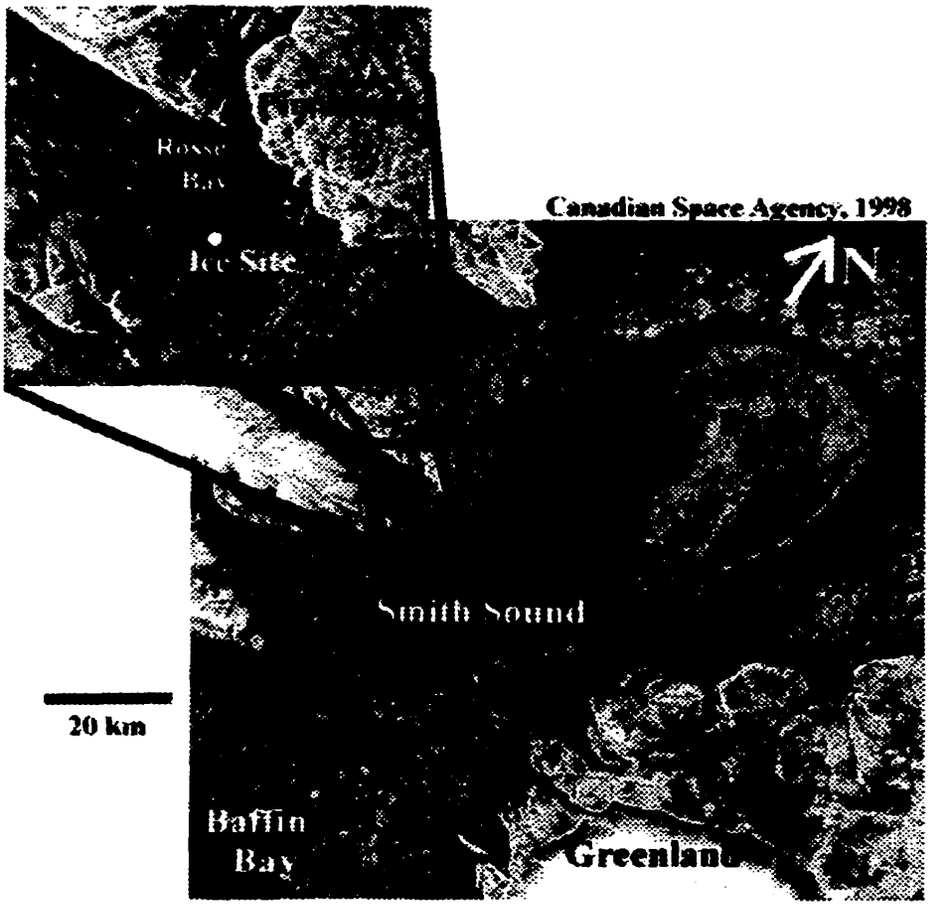



*Figure 4.1b: Locations of the terrestrial camp (Cape Herschel) and ice site in Rosse Bay during the NOW'98 project. The image is a synthetic aperture radar (SAR) image from Radarsat-1. Dark shades represent smooth first-year ice in Rosse Bay and open water mixed with ice in Smith Sound and Kane Basin.*

Cape Herschel is located 74.67° W, 78.65° N and the fast-ice site was 3.5 km due north of the terrestrial camp. The fast-ice regime was typified by smooth first year sea ice (Figure 4.1b; dark shades in Rosse Bay) with a snow cover ranging between 5 and 15 cm up to complete snow melt by early June. The terrestrial site was 100 m AMSL, with a full snow cover and gradual sloping terrain toward Rosse Bay. The fast-ice edge at Cape Herschel was initially 3-4 km off shore and decreased to less than 1 km by early June. Measurements of $K\downarrow$ and $L\downarrow$ were available from instruments that were mounted on the foredeck of the Canadian Coast Guard ice-breaker, *Pierre Radisson* from transects within NOW (Figure 4.2). Radiation flux densities were measured over several ice types and under a variety of atmospheric conditions. Data were measured every second and stored as one-minute averages. The terrestrial and ice camps were operational between Year-Day (YD) 89 (March 30) and YD155 (June 4) and the ice-breaker between approximately YD99 (April 9) to YD203 (July 22). This resulted in 601 h and 822 h of useful terrestrial site $K\downarrow$ and $L\downarrow$ data, respectively, 635 h and 915 h at the fast-ice site and 659 h and 988 h on the ice-breaker. There were less short-wave hours due to night-time darkness in the early spring and data quality control (see below).

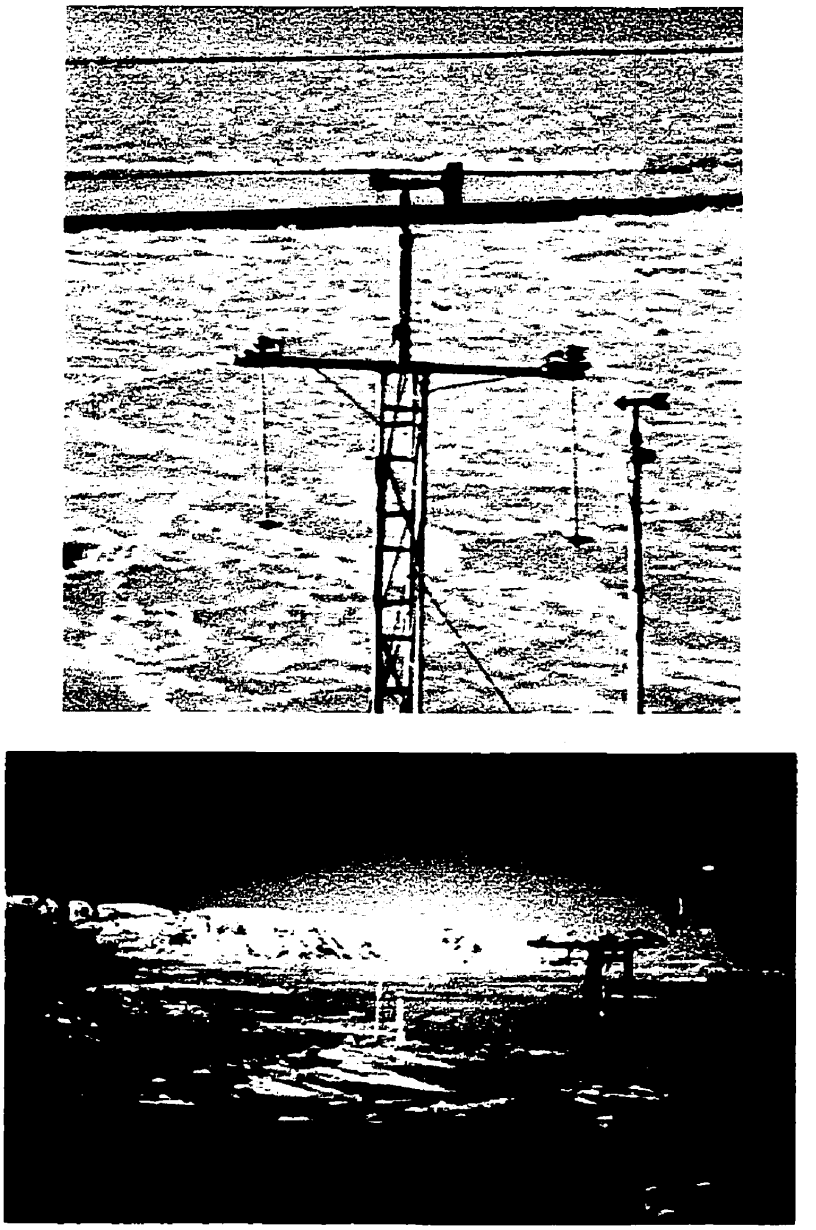


*Figure 4.2: **Incident radiation instrumentation on the ice breaker Pierre Radisson (top picture) and the fast-ice site near Cape Herschel (centre right in the bottom picture). Both photos were taken looking in a northerly direction.***

Incident short-wave and long-wave radiation were sampled every 5 s by an Eppley pyranometer (model PSP) and pyrgeometer (model PIR) which was installed at both the Cape Herschell and Rosse Bay sites (Figure 4.2). The radiometers were mounted 1.2 m above the surface and recorded 15-minute radiation averages. Incident short-wave and long-wave radiation were measured using Eppley 8-48 "Black and White" Pyranometers on the ice-breaker, and pyrgeometers (model PIR) mounted on gimbles 6 m above the ship deck to ensure proper leveling in the mean. Measurement accuracy is estimated at $\pm 10$ W $m^{-2}$ and $\pm 2.5\%$ for the pyrgeometers (Philipona et al., 1995) and pyranometers (Latimer, 1978), respectively. The spectral response of the PSP and 8-48 pyranometers are determined by the glass dome over the sensors, which has uniform transmission for radiation in the wavelength range 0.285 to 2.8 μm. The spectral response of the PIR, determined by an interference filter inside a silicon dome, is from 3.5 to 50μm. A down-facing pyranometer (Eppley, model PSP) was installed off of a scaffold-type tower 2.55 m above the fast ice surface along with a down-facing pyranometer (Eppley 8-48) extending 3 m out from the bow of the *Pierre Radisson* to sample K↑. These radiation instruments are designed to measure the critical wavelengths of the solar and infra-red spectrums that control radiative fluxes. Supplementary hourly data included air temperature ($T_a$) (error approximately $\pm 0.1$°C), relative humidity (RH) (error ± 5%), total cloud fraction (c) (error ± 10%), and surface albedo (± 0.03 albedo units) used as input for the radiative flux parameterizations (discussed below). The $T_a$ and RH were measured 2 m above the surface at Cape Herschel using a relative humidity probe (CSI

207F) and 15 m above the water line of the ship. Surface albedo was estimated as the ratio of K↑ to K↓ at the Rosse Bay site and from the *Pierre Radisson.* Short-wave albedo estimates were continuous and stable at Rosse Bay, but intermittent on the ice-breaker due to rime build-up on the sensor. Total cloud fractions were determined from human hourly observations at the terrestrial/fast-ice sites and an all-sky time lapse video camera on the ship. The all-sky camera instrument is an all-weather camera looking downward on a hemispheric mirror that produced a 180° view of the celestial dome. Hourly averages of total cloud fraction were determined from analyses of the video images sub-sampled at 10-minute intervals. A discussion of the cloud conditions is provided by Hanafin and Minnett (2001). Incident radiation was measured over a temperature range between –30°C and +3°C from late winter to early summer at the terrestrial and fast-ice sites and –8°C to +12°C on the ice-breaker. As a result, there was little overlap of temperature and vapor pressure between the fast-ice and ice breaker measurements (i.e. the two environmental regimes were quite different).

Data were stratified into clear-sky (0 to 1 tenth coverage) and all-sky (2-10 tenths coverage) conditions, solar zenith angle (Z), and location (terrestrial, fast-ice, ice breaker). Seasonal variations according to air temperature were also investigated for parameterization seasonal biases. Clear-sky data were combined into 0 to 1 tenth sky cover since there was no difference between the mean short-wave and long-wave radiative fluxes associated with them. This resulted from hourly averaging and the short time the solar disk was obscured by cumulus clouds. The Z is used as an independent variable since some parameterizations have shown to contain Z biases (Key et al. 1996). In

addition, all K↓ data with a Z > 75° were omitted due to measurement error at these angles, especially albedo (K↑ fluxes). Location stratification was performed for parameterization performance between the terrestrial, fast-ice and full marine environments. The fast-ice and terrestrial data sets are more accurate due to their stable measurement platforms and less prone to the harsh marine environment; they show smaller errors between the observed and estimated fluxes (see below).

### *4.2.2 Radiative Flux Parameterizations*

Key et. al. (1996) discussed several simple K↓ and L↓ schemes that performed well in the Arctic that are still used in one-dimensional thermodynamic sea ice models and two-dimensional dynamic-thermodynamic models (see for example, Flato and Brown, 1996; Maslanik et al., 1995; Ebert and Curry, 1993). Only those parameterizations that were selected to outperform others (by Key et al.) were used in this work. The K↓ parameterizations used include Bennett (1982) and Shine (1984) for clear skies and Jacobs (1978) and Shine (1984) for cloudy skies. The L↓ parameterizations include Ohmura (1981), Efimova (1961) and Maykut and Church (1973) for clear skies and Jacobs (1978) and Maykut and Church (1973) for cloudy skies. The schemes are attractive for sea ice modeling due to their simplicity and computational ease and perform fairly well over daily and hourly averages in most cases (Key et al., 1996; Hanesiak et al., 1999). However, their performance can be affected by observational errors in input parameters

and unknown total column atmospheric conditions. I refrain from detailed discussions of the various parameterizations in light of Key et al.'s comprehensive review, and simply outline their empirical formulations.

The parameterizations have been categorized into clear sky and all-sky fluxes with similar conventions as Key et al. (1996), for consistency. The down-welling short-wave flux ($K\downarrow$), down-welling long-wave flux ($L\downarrow$) and solar constant (S = 1356) are in W $m^{-2}$; solar zenith angle (termed Z above) in degrees, near surface air temperature ($T_a$) in Kelvin, near surface vapor pressure (ea) in hPa, cloud fraction (c in tenths), cloud optical depth ($\tau$), and surface albedo ($\alpha$). A more detailed discussion of each parameterization can be found in Key et al., (1996). A brief description of each parameterization is outlined.

### *4.2.2.1 Short-Wave Clear Sky Flux*

Key et al. (1996) indicated that the best clear sky $K\downarrow$ ($K\downarrow_{clr}$) parameterization in their study was that of Shine (1984) with Bennett (1982) performing reasonably well over daily averages. The Shine (1984) parameterization attempts to account for the near surface vapor pressure explicitly and was tested with a radiative transfer model. In comparison the scheme of Bennett (1982) is less comprehensive and is primarily intended for the estimation of mean monthly values.

The form of Bennett (1982) for $K\downarrow_{clr}$ is

$$K\downarrow_{clr} = 0.72\ S\ \cos(Z) \qquad (4.1)$$

The form of Shine (1984) for $K\downarrow_{clr}$ is

$$K\downarrow_{clr} = (S \cos^2(Z)) / [1.2 \cos(Z) + (1.0 + \cos(Z))x10^{-3} ea + 0.0455] \quad (4.2)$$

### *4.2.2.2 Short-Wave All-Sky Flux*

To parameterize cloud effects on the all-sky $K\downarrow$ flux, two approaches were found to perform reasonably well. One is to simply multiply the clear sky flux by a function of cloud fraction as in Jacobs (1978). Another more sophisticated approach is to account for surface changes (or albedo changes) affecting multiple reflections between the surface and cloud as well as cloud properties (or optical depth/thickness) as in Shine (1984) that performs well over shorter time periods (Key et al., 1996). I assume a constant cloud optical thickness (7.0) according to Ebert and Curry (1993) and input the computed hourly surface albedo from the Rosse Bay and ice-breaker platforms.

The form of Jacobs (1978) under all-sky conditions is

$$K\downarrow_{all} = K\downarrow_{clr} (1 - 0.33\ c) \quad (4.3)$$

The form of Shine (1984) under all-sky conditions is

$$K\downarrow_{cld} = (53.5 + 1274.5 \cos(Z)) \cos^{0.5}(Z) / [1 + 0.139\ (1 - 0.9345\ \alpha)\tau \quad (4.4a)$$

$$K\downarrow_{all} = (1 - c)\ K\downarrow_{clr} + c\ K\downarrow_{cld} \quad (4.4b)$$

### *4.2.2.3 Long-Wave Clear Sky Flux*

Down-welling clear-sky long-wave radiation ($L\downarrow_{clr}$) is typically a function of near surface temperature and vapor pressure. Three commonly used parameterizations differ in their treatment of the atmospheric emittance. Ohmura (1981) considers the atmospheric emittance as a function of near surface air temperature. Maykut and Church (1973), on the other hand, assign a constant value to the atmosphere emittance, and Efimova (1961) make the term a function of near surface vapor pressure.

The form of Ohmura (1981) for $L\downarrow_{clr}$ is

$$L\downarrow_{clr} = \sigma\,(8.733\text{x}10^{-3}\,T_a^{0.788})\,T_a^4 \tag{4.5}$$

The form of Maykut and Church (1973) for $L\downarrow_{clr}$ is

$$L\downarrow_{clr} = 0.7855\,\sigma\,T_a^4 \tag{4.6}$$

The form of Efimova (1961) for $L\downarrow_{clr}$ is

$$L\downarrow_{clr} = \sigma\,(0.746 + 0.0066\,ea)\,T_a^4 \tag{4.7}$$

### *4.2.2.4 Long-Wave All Sky Flux*

An increase in $L\downarrow$ associated with clouds is a function of cloud fraction in expressions by Jacobs (1978) and Maykut and Church (1973). Jacobs (1978) developed a simple linear relationship between cloud emittance, cloud fraction and the clear-sky flux,

while the relationship of Maykut and Church (1973) is in power form. Cloud emissivity also differs between the two formulations.

The form of Jacobs (1978) $L\downarrow_{all}$ is

$$L\downarrow_{all} = L\downarrow_{clr}(1 + 0.26\ c) \tag{4.8}$$

The form of Maykut and Church (1973) $L\downarrow_{all}$ is

$$L\downarrow_{all} = L\downarrow_{clr}(1 + 0.22\ c^{2.75}) \tag{4.9}$$

### *4.2.3 Analysis Methods*

The $K\downarrow_{clr}$ and $L\downarrow_{clr}$ schemes were individually tested. The all-sky $K\downarrow$ and $L\downarrow$ parameterizations require a clear-sky component, hence different combinations of the various schemes were performed. The $K\downarrow$ all-sky combinations included using: 1) the clear-sky of Shine with the cloudy sky correction of Shine, 2) the clear-sky of Bennett with the cloudy sky correction of Jacobs, 3) the clear-sky of Shine with the cloudy sky correction of Jacobs, and 4) the clear-sky of Bennett with the cloudy sky correction of Shine. The $L\downarrow$ all-sky combinations included using: 1) the clear-sky of MC with the cloudy sky correction of MC, 2) the clear-sky of Ohmura with the cloudy sky correction of Jacobs, 3) the clear-sky of Efimova with the cloudy sky correction of Jacobs, 4) the clear-sky of MC with the cloudy sky correction of Jacobs, and 5) the clear-sky of Efimova with the cloudy sky correction of MC.

The radiation and meteorological data were converted to hourly averages (centred on the hour). The meteorological data from each observing platform was used as input for

the parameterizations. The difference between the modeled and observed flux (difference parameter) was used to measure the performance of each scheme. Performance was quantified using statistical indices on the difference parameter such as simple linear regression tests for biases, mean bias error (MBE) and the root-mean-square error RMSE. The linear slope of regression lines fit along the difference parameter were statistically tested with an associated confidence interval to infer significant biases (slope significantly different from the mean) as functions of solar zenith angle (Z), season and the type of environment. Residual plots revealed that a linear model is appropriate for each regression line along the difference parameter.

For the seasonal bias analysis, the observed and estimated radiation were stratified into temperature regimes. The measured and estimated fluxes were binned according to the following temperature categories: (1) $T_a \leq -20°C$, (2) $-20°C < T_a \leq -14°C$, (3) $-14°C < T_a \leq -9°C$, (4) $-9°C < T_a \leq -2°C$, and (5) $T_a > -2°C$. The categories correspond to the main "seasonal" transitory atmospheric and sea ice conditions for the NOW region (Hanesiak, 1999) and allowing enough statistically viable data points for the comparisons in each temperature regime. The results of this analysis illustrates the seasonal bias inherent in the parameterizations and does not necessarily suggest that one scheme is superior since the simulated fluxes were within measurement errors in some cases.

## 4.3. Results

### *4.3.1 Parameterizations vs. Observations*

#### *4.3.1.1 Terrestrial and Fast-Ice Sites*

The results from the analysis using data from the terrestrial and fast ice site were similar (generally within 2%). Consequently the following sections describes the performance of the models from the fast-ice site.

##### *4.3.1.1.1 Incident Short-wave Fluxes*

The results are in agreement with Key et al., (1996) showing that the Shine scheme is preferable to Bennett's equation (Figure 4.3 and Table 4.1). The Bennett scheme contains a negative Z bias with decreasing Z (Figure 4.3) within a 99% confidence interval. The Shine scheme performs well for the fast-ice data set with only 1.1 W $m^{-2}$ mean error (5.0 W $m^{-2}$ at the terrestrial site). The parameterization statistics in Table 4.1 are slightly better than Key et al. (1996). The cancellation of errors between negative and positive values may contribute to the low mean errors.

*Table 4.1: Parameterized clear-sky short-wave flux error (estimated flux minus the measured fast-ice flux). Bracketed values are the corresponding ice-breaker results. The mean, mean error and RMSE are in W $m^{-2}$. The number of observations is also shown.*

| No. obs = 145(104) | Mean | $R^2$ | Mean Error | RMSE |
|---|---|---|---|---|
| Observed | 391.9(419.1) | | | |
| Shine | 393.1(442.9) | 0.99(0.91) | 1.1(23.9) | 15.9(48.8) |
| Bennett | 373.1(426.3) | 0.98(0.91) | -18.8(7.3) | 21.6(47.8) |

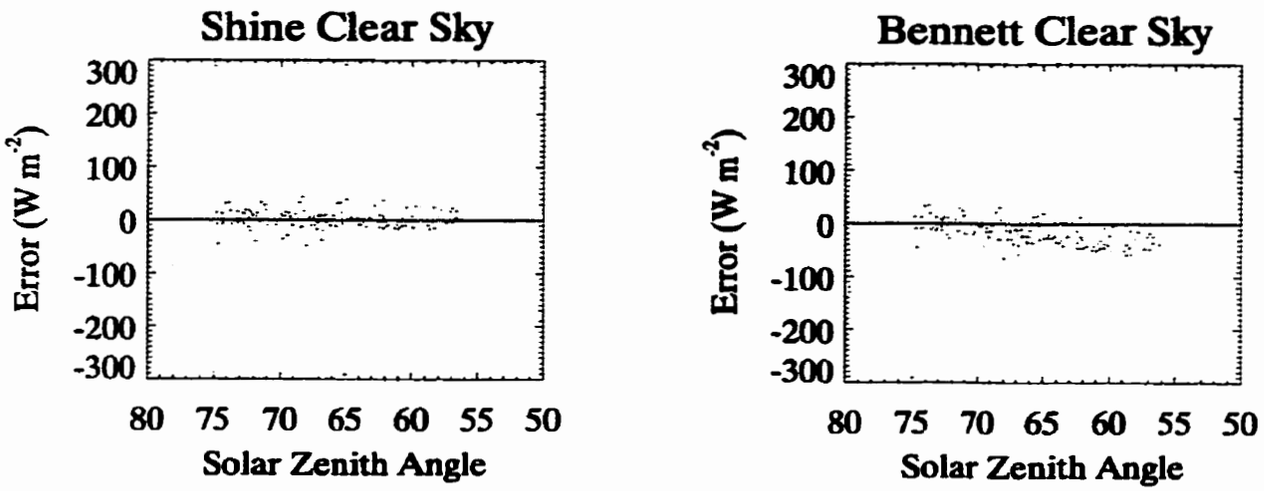



*Figure 4.3: The clear-sky short-wave flux error for two parameterizations (modeled flux minus the observed flux at Rosse Bay).*

For the K↓ all-sky combinations the results again support Key et al. (1996). The error statistics (Table 4.2) and error plots (Figure 4.4) show the best parameterization combination is Shine/Shine with a mean error of –2.1 W $m^{-2}$ (3.1 W $m^{-2}$ at the terrestrial site). The Jacobs cloudy-sky scheme depletes too much radiation when combined with the Shine and Bennett clear sky routines, shown by the highly negative mean errors.

*Table 4.2: Parameterized all-sky short-wave flux error (estimated flux minus the measured fast-ice flux). Bracketed values are the corresponding ice-breaker results. The mean, mean error and RMSE are in W $m^{-2}$. The number of observations is also shown.*

| No. obs = 456(555) | Mean | $R^2$ | Mean Error | RMSE |
|---|---|---|---|---|
| Observed | 322.4(250.7) | | | |
| Shine/Shine | 320.3(332.4) | 0.83(0.76) | -2.1(81.7) | 59.6(90.1) |
| Bennett/Jacobs | 289.9(351.0) | 0.85(0.71) | -32.4(100.3) | 60.6(95.9) |
| Shine/Jacobs | 304.0(370.1) | 0.85(0.71) | -18.4(119.4) | 57.9(96.5) |
| Bennett/Shine | 315.1(322.5) | 0.82(0.75) | -7.4(71.8) | 61.3(90.8) |

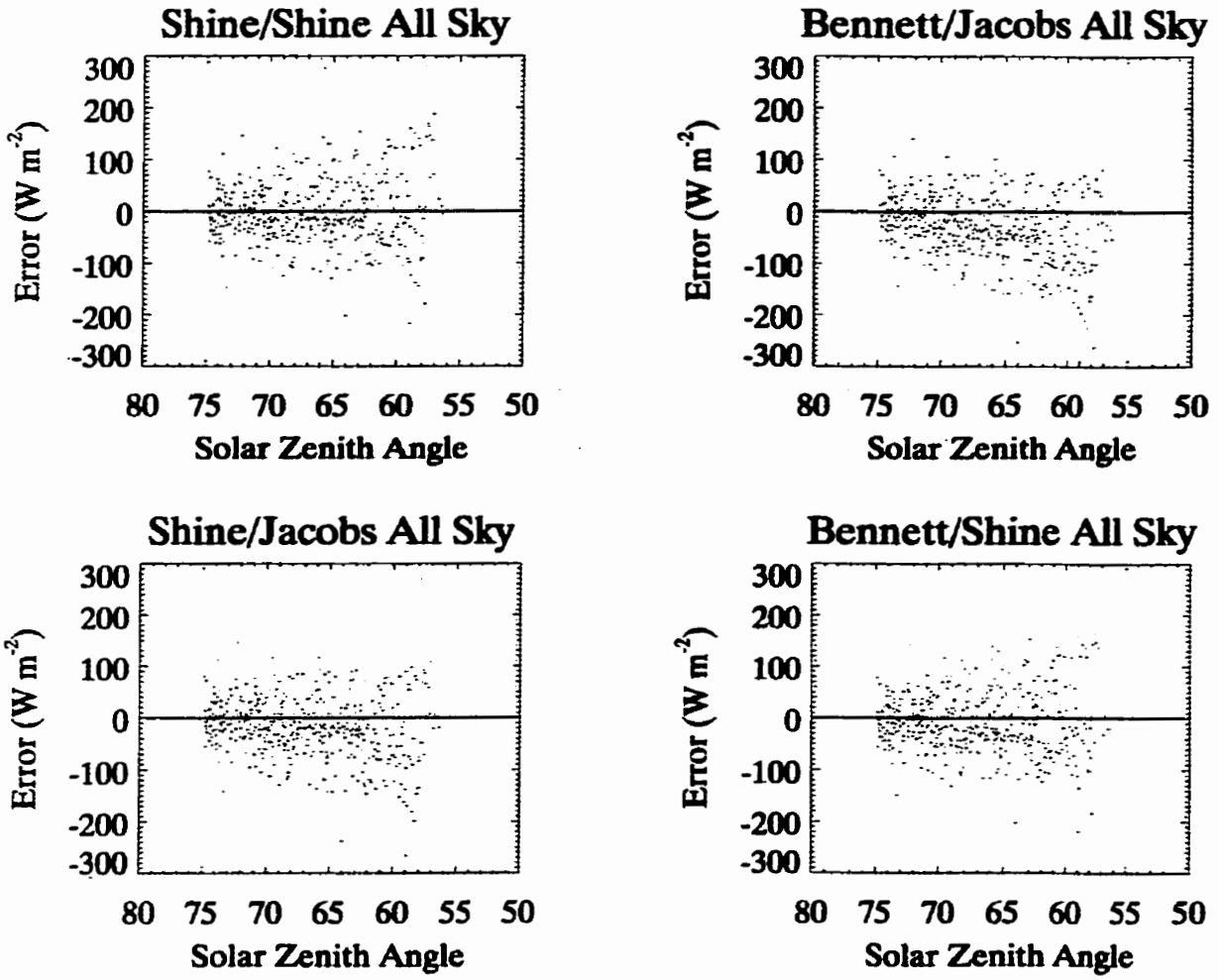



*Figure 4.4: The all-sky short-wave flux error for four combinations of parameterizations (modeled flux minus the observed flux at Rosse Bay). The first name refers to the clear-sky parameterization used for that combination from Figure 4.3.*

### *4.3.1.1.2 Incident Long-wave Fluxes*

The parameterization of Ohmura performs better than both the Maykut-Church (MC) and Efimova's schemes (Table 4.3). Key et al., (1996) found the Efimova routine to perform best. The emittance of long-wave radiation is too high in both Maykut and Church's and Efimova's schemes for incident fluxes less than about 240 W $m^{-2}$. A similar trend is shown by Key for the MC model. If the clear-sky atmospheric emissivity is adjusted by roughly the same amount in MC (from 0.7855 to 0.729) and Efimova (0.746 to 0.7; the first coefficient independent of vapor pressure) during periods of low emittance ($L\downarrow < 240$ W $m^{-2}$) the results from both parameterizations improves (Table 4.4 and Figure 4.5). The emissivities for each scheme was optimized to fit this data by minimizing the mean bias error (MBE) for each scheme.

*Table 4.3: Parameterized clear-sky long-wave flux error (estimated flux minus the measured fast-ice flux). Bracketed values are the corresponding ice-breaker results. The mean, mean error and RMSE are in W $m^{-2}$. The number of observations is also shown.*

| No. obs = 200(131) | Mean | $R^2$ | Mean Error | RMSE |
|---|---|---|---|---|
| Observed | 197.5(235.4) | | | |
| MC | 209.8(243.9) | 0.97(0.80) | 12.2(8.5) | 7.3(17.8) |
| Ohmura | 192.5(230.6) | 0.97(0.80) | -4.9(-4.8) | 6.6(17.3) |
| Efimova | 207.9(247.9) | 0.98(0.80) | 10.4(12.5) | 6.1(17.0) |

*Table 4.4: Parameterized clear-sky long-wave flux error (estimated minus measured fast-ice flux) when altering the emissivity in MC (from 0.7855 to 0.729) and Efimova (from 0.746 to 0.70). Bracketed values are corresponding ice-breaker results. The mean, mean error and RMSE are in W $m^{-2}$. The number of observations is also shown.*

| No. obs = 200(131) | Mean | $R^2$ | Mean Error | RMSE |
|---|---|---|---|---|
| Observed | 197.5(235.4) | | | |
| MC | 197.6(238.0) | 0.98(0.81) | 0.1(2.7) | 5.2(16.7) |
| Efimova | 197.6(242.9) | 0.98(0.80) | 0.1(7.5) | 5.2(17.8) |

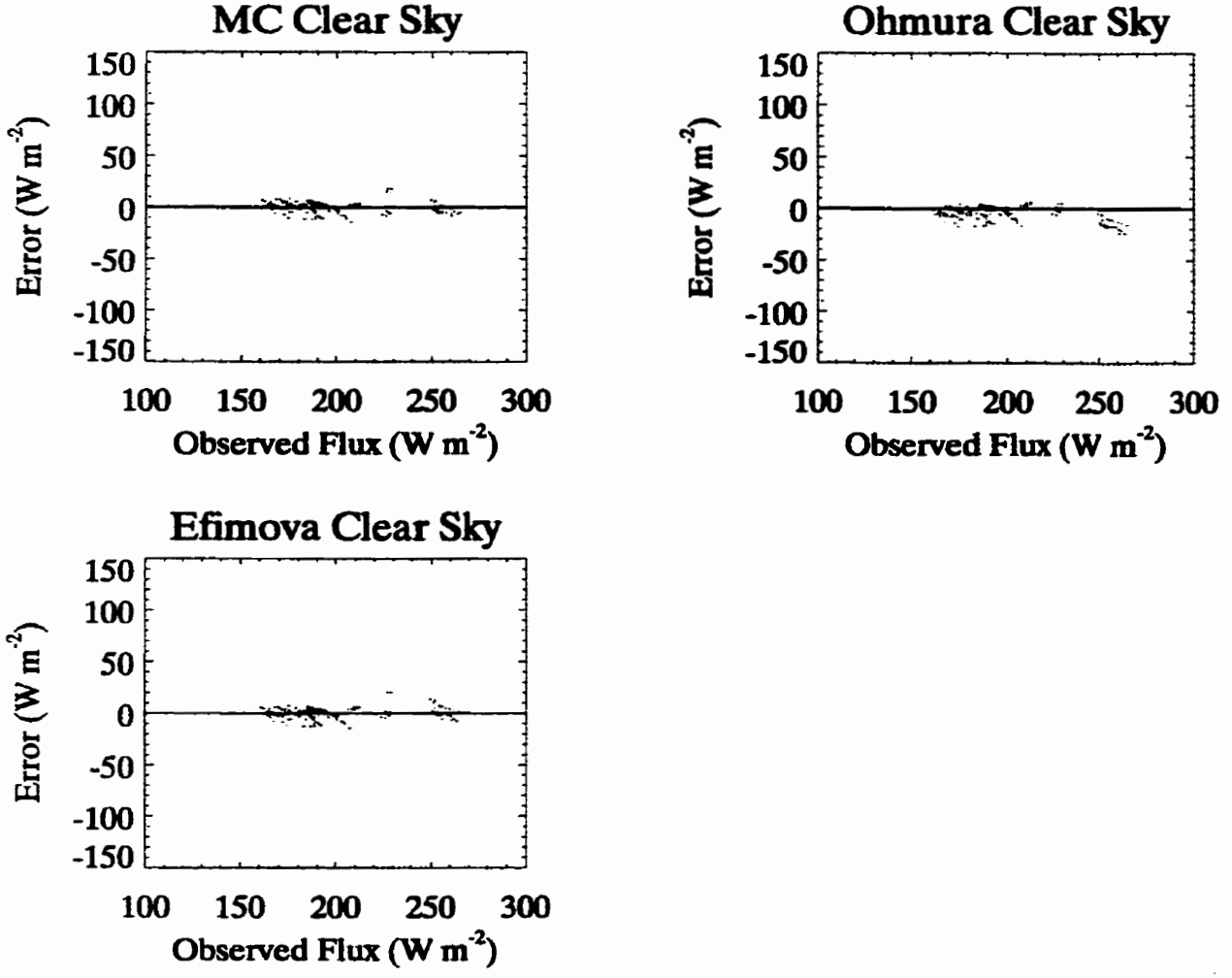



*Figure 4.5: The clear-sky long-wave flux error (modeled flux minus the observed flux) for the optimized emissivities in MC and Efimova to fit the Rosse Bay data (see text).*

For the L↓ all-sky combinations, Efimova/MC had both the smallest MBE and RMSE of the combinations examined (Table 4.5). Key suggested the Efimova/Jacobs schemes perform best. If we consider the modifications to the clear-sky emission characteristics (above), the Efimova/Jacobs and the MC/Jacobs combination perform better than the Efimova/MC model (Table 4.6), with Efimova/Jacobs performing the best. The MC cloudy-sky correction for long-wave radiation consistently underestimates fluxes in cloudy skies.

*Table 4.5: Parameterized all-sky long-wave flux error (estimated flux minus the measured fast-ice flux). Bracketed values are the corresponding ice-breaker results. The mean, mean error and RMSE are in W $m^{-2}$. The number of observations is also shown.*

| No. obs = 622(857) | Mean | $R^2$ | Mean Error | RMSE |
|---|---|---|---|---|
| Observed | 231.6(277.8) | | | |
| MC/MC | 238.0(262.8) | 0.92(0.73) | 6.4(-15.1) | 14.9(19.3) |
| Ohmura/Jacobs | 226.1(260.5) | 0.89(0.78) | -5.5(-17.3) | 16.9(17.6) |
| Efimova/Jacobs | 243.7(280.1) | 0.90(0.79) | 12.1(2.2) | 16.3(17.4) |
| MC/Jacobs | 251.8(282.6) | 0.90(0.77) | 20.1(4.8) | 16.2(18.1) |
| Efimova/MC | 230.4(260.4) | 0.92(0.75) | -1.2(-17.4) | 14.9(18.4) |

*Table 4.6: Parameterized all-sky long-wave flux error (estimated flux minus the measured fast-ice flux) when applying the new clear-sky formulations of Maykut-Church and Efimova from Table 4. The mean, mean error and RMSE are in W $m^{-2}$. The number of observations is also shown.*

| No. obs = 622 | Mean | $R^2$ | Mean Error | RMSE |
|---|---|---|---|---|
| Observed | 231.6 | | | |
| MC/MC | 221.8 | 0.92 | -9.8 | 15.4 |
| Ohmura/Jacobs | 226.1 | 0.89 | -5.5 | 16.9 |
| Efimova/Jacobs | 229.6 | 0.89 | -2.0 | 16.8 |
| MC/Jacobs | 234.6 | 0.89 | 2.9 | 16.9 |
| Efimova/MC | 217.1 | 0.92 | -14.5 | 15.2 |

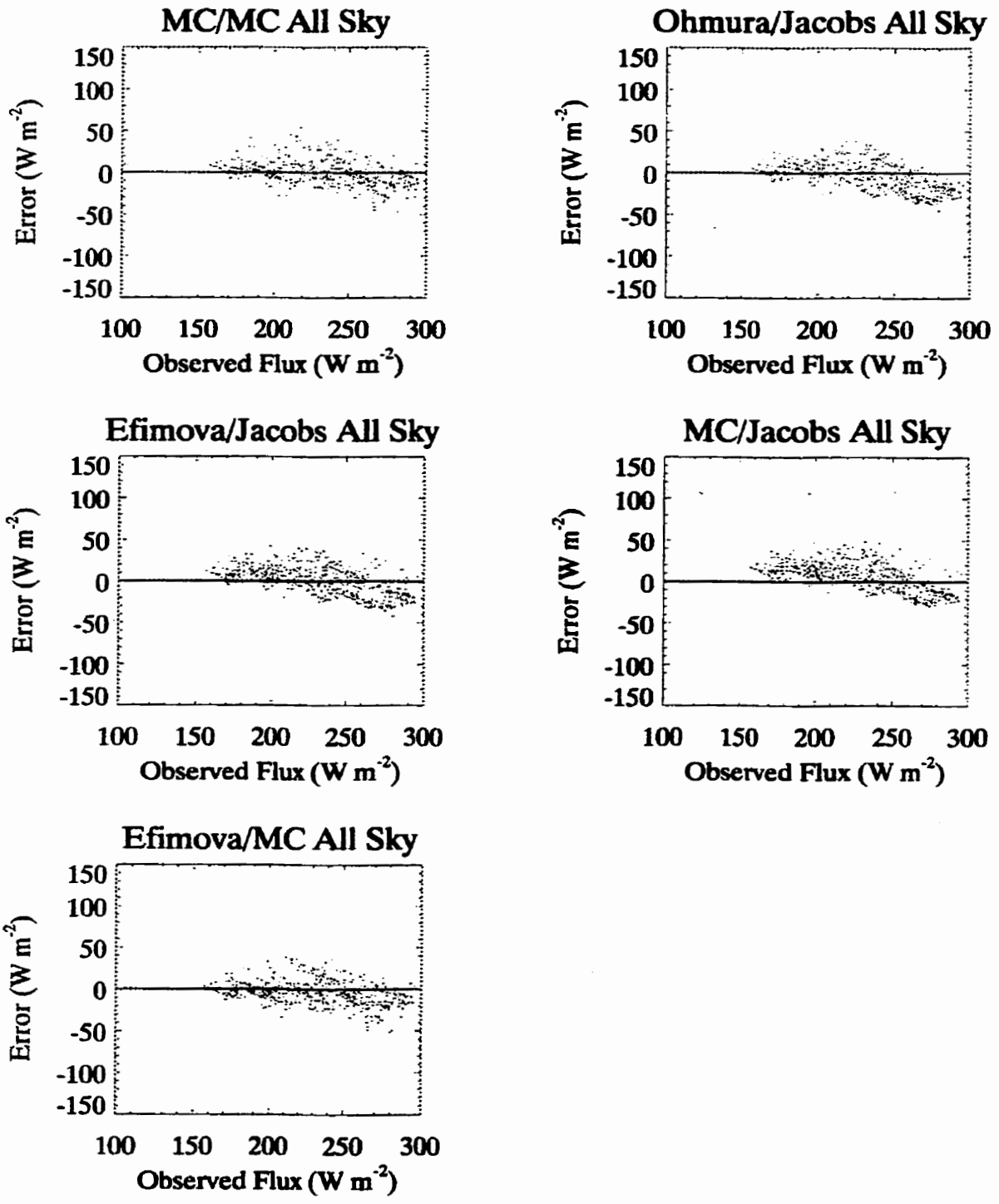



*Figure 4.6: All-sky long-wave flux error for five combinations of parameterizations (modeled flux minus the observed flux at Rosse Bay). The first name refers to the clear-sky parameterization used for that combination using the optimized clear-sky emissivities in Figure 4.5 and optimized cloudy-sky emissivities of MC and Jacobs to fit the Rosse Bay data (see text).*

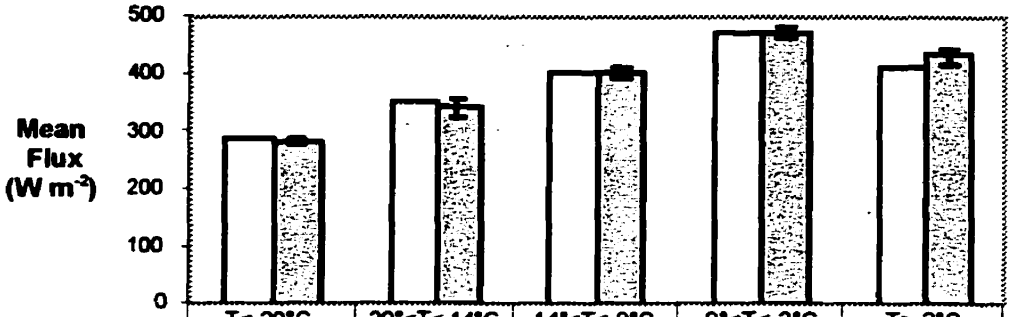
A noticeable trait in the data and the results of Key et al. (1996) is a consistent negative slope (within a 99% confidence interval) in the $L\downarrow_{cld}$ error as the observed flux in Figure 4.5 increases (ie: the estimated flux underestimates real fluxes as the magnitudes of the fluxes increase). This suggests the cloudy sky emissivity is too small and needs to be increased. If the cloudy sky emissivity is increased in MC (from 0.2232 to 0.32) and Jacobs (from 0.26 to 0.275) the results are again different (Table 4.7 and Figure 4.6). The cloudy-sky emissivity of MC and Jacobs were optimized as before by minimizing the MBE for each scheme. After optimizing the cloud-sky emissivities, all of the parameterizations perform well based on the error indices (Table 4.7), however, the modified MC cloudy-sky formulation relaxes the negative slope (with 95% confidence) in the $L\downarrow_{cld}$ error (Figure 4.6) with increasing flux. This is due to the MC cloudy-sky exponential dependence in equation (4.9). The Jacobs cloudy-sky formulation neglects this exponential factor. Very similar results (within a few percent) were obtained using the terrestrial data.

*Table 4.7: Parameterized all-sky long-wave flux error (estimated flux minus the measured fast-ice flux) when applying the new clear-sky formulations of Maykut-Church and Efimova from Table 4.4 and the new cloudy-sky formulation of Maykut-Church and Jacobs (see text). Bracketed values are the corresponding ice-breaker results. The mean, mean error and RMSE are in W m$^{-2}$. The number of observations is also shown.*

| No. obs = 622(857) | Mean | $R^2$ | Mean Error | RMSE |
|---|---|---|---|---|
| Observed | 231.6(277.8) | | | |
| MC/MC | 231.9(265.1) | 0.92(0.76) | 0.3(-12.7) | 13.9(20.3) |
| Ohmura/Jacobs | 228.1(262.6) | 0.89(0.78) | -3.5(-15.3) | 16.6(17.7) |
| Efimova/Jacobs | 231.7(276.5) | 0.90(0.79) | 0.1(-1.3) | 16.4(18.3) |
| MC/Jacobs | 236.7(278.8) | 0.90(0.80) | 5.0(0.9) | 16.5(17.5) |
| Efimova/MC | 227.1(262.9) | 0.93(0.75) | -4.6(-14.9) | 13.8(21.0) |

### *4.3.1.1.3 Seasonal Trends*

The superior parameterizations above and those modified to fit the NOW region were evaluated for seasonal bias using methods described in section 4.2.3. The Shine $K\downarrow_{clr}$ scheme has a positive seasonal bias (confidence interval of 99%) by slightly underestimating fluxes in the early part of the season and overestimating fluxes in the latter part of the season (Figure 4.7).




| | T<-20°C | -20°<T<-14°C | -14°<T<-9°C | -9°<T<-2°C | T>-2°C |
|---|---|---|---|---|---|
| obs | 287.5 | 350.7 | 402.2 | 473.7 | 411.7 |
| Shine | 282.4 | 339.1 | 402.4 | 471.8 | 432.1 |

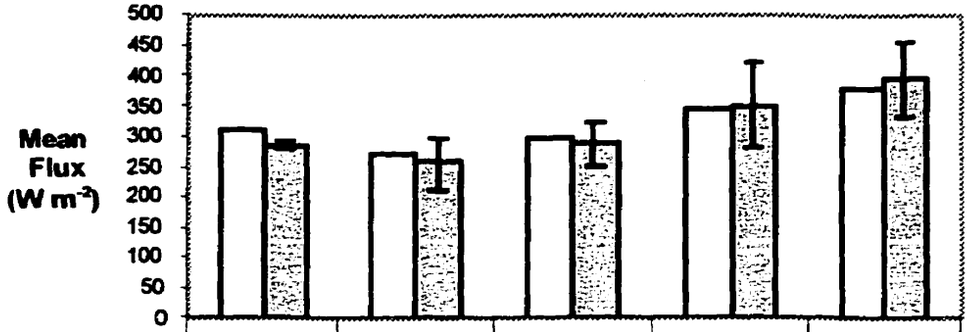
*Figure 4.7: The short-wave observed and Shine scheme seasonal clear-sky fluxes (W $m^{-2}$). The RMSE error bars are also shown. The coldest temperature category contained 13 data points, 34 in the next warmest, 48 in the next warmest, 20 in the next warmest and 32 in the warmest category.*

The overestimation in the late season could be, in part, due to the model's inability to represent the increase in atmospheric optical depth associated with rising temperatures and water vapor loading (see also, Leontyeva and Stamnes, 1993; Blanchet and List, 1983). Integrated total column water vapor using radiosondes launched from the ice breaker increased from 2 kg $m^{-2}$ in mid-April to 5.5 kg $m^{-2}$ by early May then 9 kg $m^{-2}$ in early June and 11 kg $m^{-2}$ by early July. Surface vapor pressures also increased over the

warm season but a de-coupling between the surface and lower troposphere (if it occurred) may affect the results. This is especially important over fast-ice in the cold season under inversion conditions where the near-surface temperature and humidity may not represent total column characteristics. Hanafin and Minnett (2001) built upon results here to further improve the Shine scheme over the polynya in the warm season by increasing the coefficient of the water vapor term. Aerosol effects are unknown since no measurements were available, however, these effects would lead to the most serious errors in the parameterizations (Shine, 1984); the radiative effects of the aerosols being strongly dependent on the atmospheric humidity (e.g. Hänel, 1976) and once again total column variations (Bergin et al., 2000) that are largely unknown in the arctic. Radiative transfer models perform well in mid-latitudes (Jing and Cess, 1998) but have not been tested to the same degree in arctic environments.

The Shine K↓ all-sky scheme (Figure 4.8) excessively depletes radiation in the early and middle part of the season and not enough in the late season (within a 99% confidence interval). This is likely because of optically thinner clouds during the colder part of the season than in the warmer season. Recall that the Shine model has an inherent clear-sky bias that affects the all-sky calculation. If we decrease the cloud optical depth (from its assumed constant value of 7.0) to 1.0 for $T_a \leq -20°C$, the all-sky Shine/Shine mean error drops to the Shine clear-sky mean error for the same temperature category. Using the same procedure for $-20°C < T_a \leq -14°C$, the cloud optical depth becomes 5.5 and stays near 7.0 for warmer temperatures. This suggests an increasing cloud optical

depth would improve the K↓ all-sky estimates as the warm season approaches. The results are consistent with Curry and Ebert (1992) and Leontyeva and Stamnes (1993) who suggest a cloud optical depth near 7.0 for the warm season. I assume the presence of mid and upper cloud has no seasonal trend when obscured by low cloud during ground observations. Other radiative transfer issues such as solar zenith angle, atmospheric gas variations, 3-D cloud variations, and averaging procedures also affect the results (see for example, Li et al., 1993; Evans, 1998; Arking et al., 1992; Bergin et al., 2000; Barker and Davies, 1992) but can not be realistically accounted for here. These issues also apply to the ice breaker results (section 4.3.1.2).



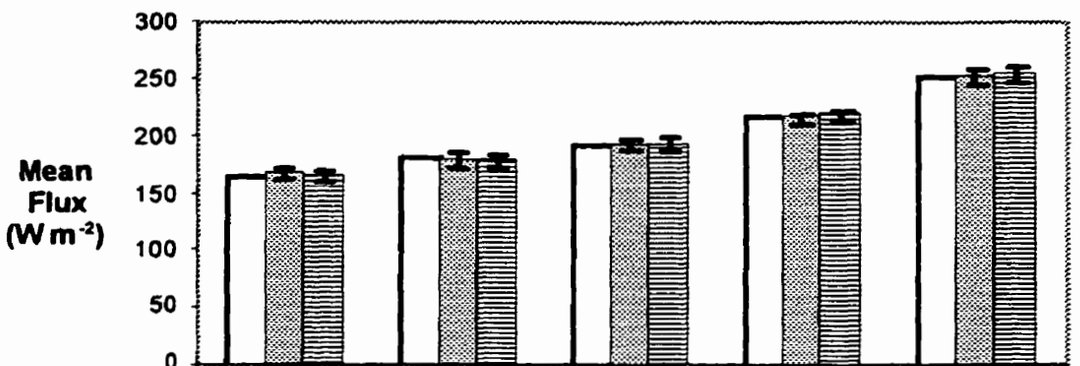

| | T<-20°C | -20°<T<-14°C | -14°<T<-9°C | -9°<T<-2°C | T>-2°C |
|---|---|---|---|---|---|
| obs | 311.4 | 273.6 | 300.2 | 344.3 | 379.1 |
| Shine/Shine | 286.2 | 257.4 | 288.5 | 351.2 | 397.9 |

*Figure 4.8: The short-wave observed and Shine/Shine scheme seasonal all-sky fluxes (W $m^{-2}$). The RMSE error bars are also shown. The coldest temperature category contained 11 data points, 61 in the next warmest, 156 in the next warmest, 235 in the next warmest and 25 in the warmest category.*

The $L\downarrow_{clr}$ seasonal trends of MC (modified) and Efimova (modified) are depicted in Figure 4.9. The MC parameterization does not contain a seasonal bias (with 99% confidence). The scheme of Efimova contains a slight negative bias (with 99% confidence) early in the season and positive bias late in the season. The Efimova emissivity was

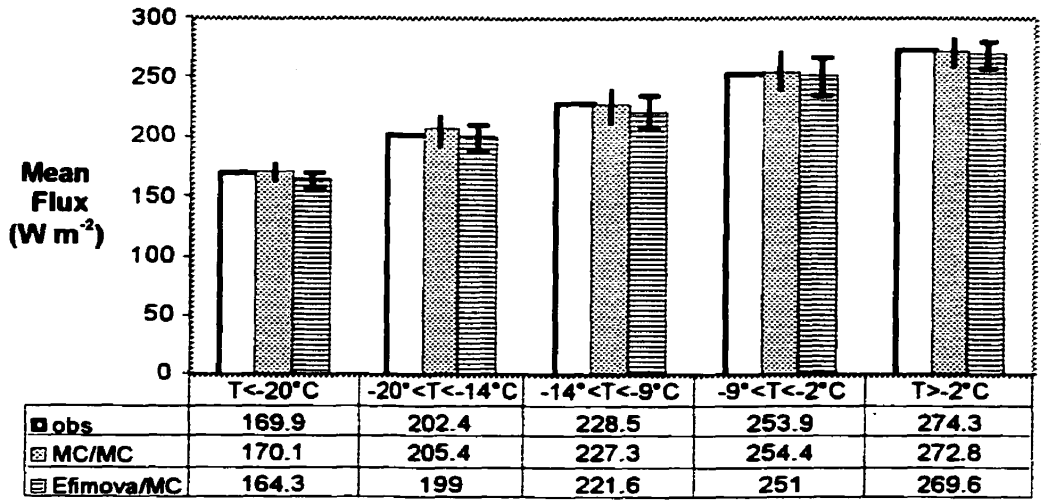
adjusted as in MC, but the vapor pressure dependence in Efimova may introduce other complications.

The L↓ all-sky seasonal trends of MC/MC (modified) and Efimova/MC (modified) are depicted in Figure 4.10. The MC/MC combination does not contain a seasonal bias (99% confidence). The Efimova/MC combination also did not have a seasonal bias (99% confidence) but contains a consistent negative error similar to what was shown earlier in Table 4.7. It should be emphasized that this analysis is intended to illustrate seasonal biases only and not compare each parameterization since their fluxes are very similar in magnitude (i.e. one scheme may not necessarily be more accurate than the other).




| | T<-20°C | -20°<T<-14°C | -14°<T<-9°C | -9°<T<-2°C | T>-2°C |
|---|---|---|---|---|---|
| obs | 166.1 | 180.3 | 193.4 | 216.8 | 252.2 |
| MC | 166.9 | 179.4 | 194 | 216.8 | 252.5 |
| Efimova | 165.3 | 178.5 | 194.1 | 219 | 255.4 |

*Figure 4.9: The long-wave observed (clear bar), MC (middle bar) and Efimova schemes seasonal clear-sky fluxes (W m$^{-2}$). The RMSE error bars are also shown. The coldest temperature category contained 29 data points, 61 in the next warmest, 73 in the next warmest, 21 in the next warmest and 34 in the warmest category.*




| | T<-20°C | -20°<T<-14°C | -14°<T<-9°C | -9°<T<-2°C | T>-2°C |
|---|---|---|---|---|---|
| obs | 169.9 | 202.4 | 228.5 | 253.9 | 274.3 |
| MC/MC | 170.1 | 205.4 | 227.3 | 254.4 | 272.8 |
| Efimova/MC | 164.3 | 199 | 221.6 | 251 | 269.6 |

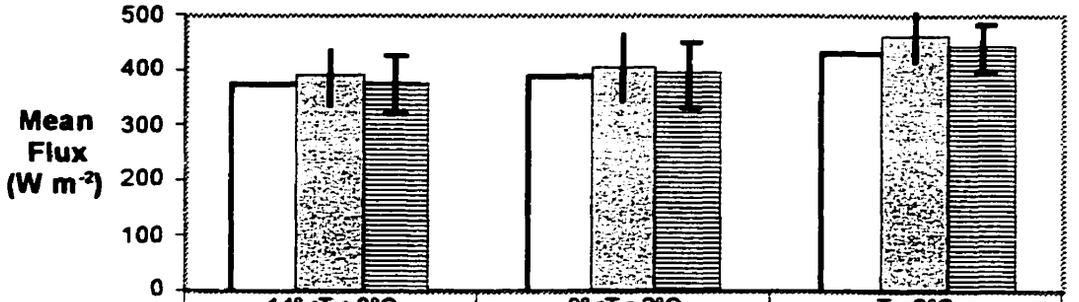
*Figure 4.10: The long-wave observed (clear bar), MC/MC (middle bar) and Efimova/MC schemes seasonal all-sky fluxes (W $m^{-2}$). The RMSE error bars are also shown. The coldest temperature category contained 34 data points, 139 in the next warmest, 239 in the next warmest, 257 in the next warmest and 28 in the warmest category.*

### *4.3.1.2 Ice Breaker Platform*

A similar analysis to the last section was conducted on the ice breaker K↓ and L↓. All of the different parameterizations were used for the analysis, including the original and modified long-wave schemes. The same seasonal bias analysis was used except polynya air temperatures were marine-like (greater than –8°C).

### *4.3.1.2.1 Incident Short-wave Fluxes*

The $K\downarrow_{clr}$ results show the Shine and Bennett schemes perform worst within the polynya relative to the fast-ice and terrestrial sites (Table 4.1). Note that all ice-breaker data in the Tables appear in brackets. The schemes overestimate fluxes with RMSE 2-3 times larger than the fast-ice results. Both the Shine and Bennett schemes contained a more pronounced positive seasonal bias (with 99% confidence) (Figure 4.11). The general overestimation of clear-sky fluxes is likely due to greater total column water vapor than the schemes were originally intended for, especially for the marine environment. Hanafin and Minnett (2001) have corrected for this in the Shine scheme.

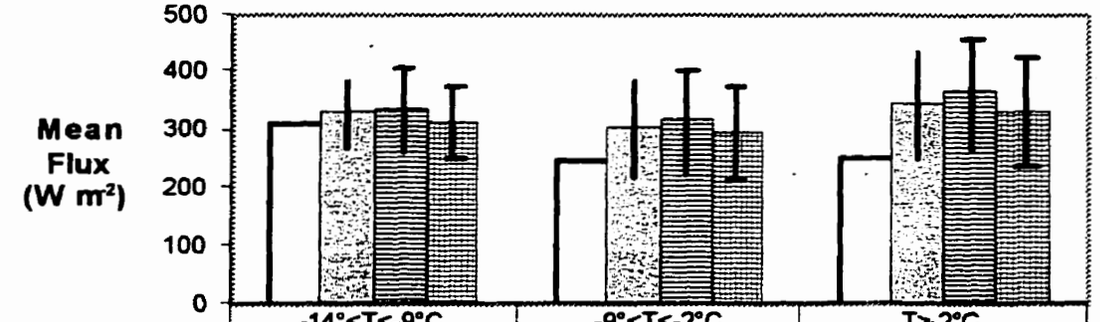



| | -14°<T<-9°C | -9°<T<-2°C | T>-2°C |
|---|---|---|---|
| obs | 378.9 | 387.3 | 434.7 |
| Shine | 388.9 | 406.3 | 461.6 |
| Bennett | 375.9 | 391.6 | 444 |

*Figure 4.11: Same as Figure 4.7 except for the ice-breaker. The coldest temperature category contained 7 data points, 26 in the next warmest, and 71 in the warmest category.*

The ice-breaker results for $K\downarrow$ all-sky show larger RMSE relative to the clear-sky case (Table 4.2). The mean errors and RMSE in the ice breaker results are also much

greater than the fast-ice. The cause is likely the result of marine clouds being optically thicker than the fast-ice environment. There is a positive seasonal bias in the ice breaker-derived fluxes where fluxes are overestimated more so as the ambient temperatures increase (with 99% confidence) (Figure 4.12). If cloud optical depth is increased from 7.0 to 9.0 for $-14°C < T_a \leq -9°C$, the all-sky Shine/Shine mean error dropped to the Shine clear-sky mean error for the same temperature category. Similarly, for $-9°C < T_a \leq -2°C$, the cloud optical depth increases to 14 and further increases to 20 for $T_a > -2°C$. These optical depths are consistent with typical arctic stratus clouds (Herman and Curry, 1984; Leontyeva and Stamnes, 1993). This is also different than the fast-ice site where an optical depth of 7.0 works well for warmer temperatures. Once again I assume there is no seasonal bias in the presence of mid and upper level cloud when they are obscured by low cloud during ground observations.

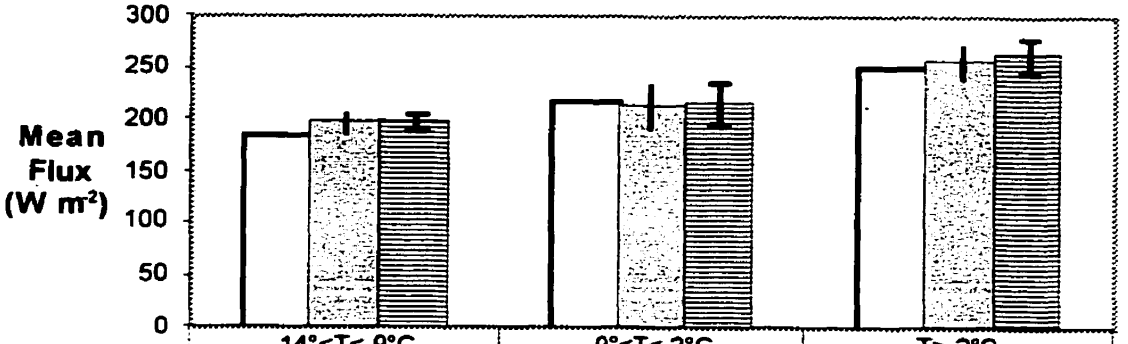



| | -14°<T<-9°C | -9°<T<-2°C | T>-2°C |
|---|---|---|---|
| obs | 310.4 | 245.6 | 251.1 |
| Shine/Shine | 326.9 | 302.1 | 342.8 |
| Bennett/Jacobs | 334.4 | 314.2 | 363.9 |
| Bennett/Shine | 313.5 | 294.8 | 332.1 |

*Figure 4.12: Same as Figure 4.8 except for the ice-breaker. The coldest temperature category contained 9 data points, 139 in the next warmest, and 407 in the warmest category.*

### *4.3.1.2.2 Incident Long-wave Fluxes*

The MC and Efimova $L\downarrow_{clr}$ overestimate smaller fluxes (i.e., $L\downarrow$<240 $W \cdot m^{-2}$), consistent with results from Table 4.3. The error terms are improved when applying the fast-ice adjustment to the MC and Efimova schemes. Both the MC and Efimova schemes do not contain the seasonal bias (with 99% confidence) over the polynya (Figure 4.13). Recall that a slight positive bias was found in the Efimova scheme for the fast-ice.

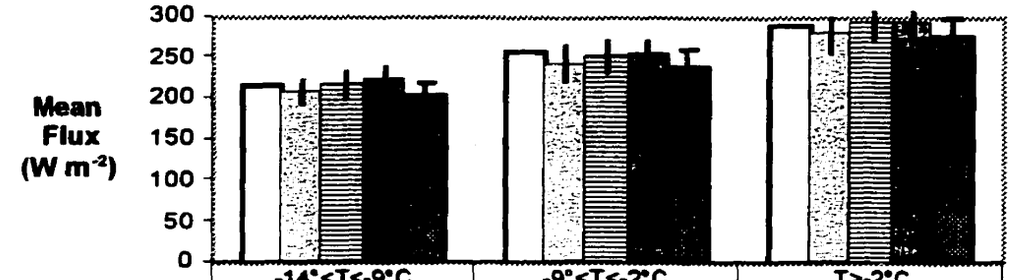



| | -14°<T<-9°C | -9°<T<-2°C | T>-2°C |
|---|---|---|---|
| obs | 186.5 | 217.5 | 252.2 |
| MC | 197.1 | 213.7 | 256.2 |
| Efimova | 197.7 | 215.5 | 263.2 |

*Figure 4.13: Same as Figure 4.9 except for the ice-breaker. The coldest temperature category contained 15 data points, 35 in the next warmest, and 81 in the warmest category.*

The ice breaker $L\downarrow$ all-sky errors improve when using the modified schemes from section 4.3.1.1 (comparing Tables 4.5 and 4.7; see also Figure 4.14). However, there is more variation in the mean errors than the fast-ice results with slightly larger RMSE and smaller $R^2$ (Table 4.7). Note the MBE is more negative than the fast-ice site (Table 4.7).

This may be due to 1) warmer mean ambient temperatures (and later data collection dates) accompanying the ice breaker data, 2) the cloud emissivity may need to be increased in a full marine environment (with optically thicker clouds) compared to the fast-ice regime, and 3) cloud base heights being lower in the marine environment. The first reason can be ruled out since the ice-breaker mean error is consistently more negative than the fast-ice mean errors for similar temperatures and data collection dates. This suggests the cloud emissivity needs to be increased for the marine schemes and/or the cloud base heights were lower in the marine environment. The K↓ all-sky results suggest a greater cloud optical depth in the polynya implying higher within-cloud precipitable water and in turn higher cloud emissivity. A higher cloud emissivity is required to match the parameterization schemes with measured flux density within the polynya. Cloud base heights would also contribute to the differences.




| | -14°<T<-9°C | -9°<T<-2°C | T>-2°C |
|---|---|---|---|
| obs | 214.5 | 258.1 | 289.9 |
| MC/MC | 207.7 | 241.8 | 278.4 |
| Efimova/Jacobs | 216.3 | 250 | 291.4 |
| MC/Jacobs | 221.2 | 253.8 | 292.8 |
| Efimova/MC | 203.1 | 238.3 | 277.1 |

*Figure 4.14: Same as Figure 4.10 except for the ice-breaker. The coldest temperature category contained 29 data points, 256 in the next warmest, and 572 in the warmest category.*

Combinations using the Jacobs cloudy-sky scheme are more consistent over the seasons and have the smallest mean errors. The Jacobs scheme was developed from data during the warm season to early winter (June-December) that may contribute to its better performance in a marine setting. The modified MC/MC and Efimova/MC scheme combinations were best in the fast-ice and terrestrial sites but would need to be further adjusted for the marine data. Hanafin and Minnett (2001) have built upon these results to improve the MC/MC $L\downarrow$ all-sky scheme in the polynya setting.

Given that the open water of the polynya is the major source of atmospheric water vapor in the area, it is to be expected that the water vapor burden and distribution will show differences between open water and land. The dominant wind direction in the area is from the NNW (Barber et al., 2001) and so the terrestrial and fast-ice data are predominantly under the influence of winds blowing over the land surface of Ellesmere Island; the ice-breaker measurements include a wide range of conditions, in terms of sea-ice cover and open water, that are unparameterized in the formulations used here. In particular the water vapor distribution is a strong function of fetch over open water for situations of off-ice winds. This has the potential for not only increasing the variability of the parameterization uncertainties, but also introducing bias errors.

While there are convincing physical reasons to explain the observed increases in uncertainty in the parameterized fluxes when compared to ship-based measurements, as opposed to measurements from fixed platforms, there is always the possibility that these are a result not of shortcomings in the parameterizations, but of flawed measurements. Of obvious concern is ship motion. In an attempt to reduce ice breaker tilting of the

radiometers, they were mounted on gimbals. This means that, on average, they are level; but at any moment may be tilted as a result of ship motion. Observations showed that in most cases the amplitude of the oscillation was only a few degrees, with some infrequent large amplitudes when the ship was breaking thick ice. The consequence of tilt is severe in the $K\downarrow$ measurements, as this has the effect of changing the apparent Z. However, while mean tilts can lead to significant errors, especially in clear skies, small-amplitude oscillating tilts do not significantly degrade the data (MacWhorter and Weller, 1991). The $L\downarrow$ measurement is inherently less sensitive to tilts of the pyrgeometer, but there is an error source caused by the temperature contrast between the sky and land, sea or ice. As the pyrgeometer tilts, it receives radiation from a warmer source, below the horizon, than the sky. Oscillating tilts do not cancel out but combine to produce a positive bias error. The effect on this would be to make a parameterization of $L\downarrow$ appear to predict fluxes low compared to the measurements, and this is indeed observed in some cases. Such an effect would be indistinguishable from the environmental factors discussed above. Clearly the issue of making accurate $K\downarrow$ and $L\downarrow$ measurements from ships requires further attention, but it is certainly not clear that these instrumental effects are of sufficient magnitude to dominate the error characteristics of the results discussed here.

## 4.4 Conclusions

Incident radiation fluxes are critical parameters within the SEB, and accurate representation of these fluxes is necessary in all types of model simulations of sea ice. Many of the current simple incident radiation schemes in sea ice models have not been validated in the environments investigated in this chapter and yet they are central areas for many biological species. The purpose of this chapter was to assess the performance of simple incident short-wave and long-wave radiation parameterizations used in several thermodynamic ice models and to investigate any inherent temporal and or spatial biases in these parameterizations (question 1). The second question was related to offering improved versions of these parameterizations where possible.

### *4.4.1 Question 1*

The fast-ice and terrestrial regimes showed very similar characteristics (within ± 2%) due to their close geographic proximity and snow covered surfaces. Differences arose when comparing the terrestrial/fast-ice results to the marine environment sampled by the ice-breaker.

The fast-ice preferred $K\downarrow_{clr}$ scheme was Shine (1984) that contained no solar zenith angle (Z) bias, unlike the scheme of Bennett (1982). However, the Shine scheme contained a positive seasonal bias where it underestimated fluxes in the cold season and

overestimated fluxes in the warm season. The positive bias was more dramatic in the marine data. Hanafin and Minnett (2001) built on these results to improve the Shine clear-sky scheme over the polynya. Likely causes and errors for the differences were offered.

The preferred fast-ice $K\downarrow$ all-sky combination scheme was the Shine (1984) clear-sky and Shine (1984) cloudy-sky. The Jacobs (1978) cloudy-sky scheme depleted too much radiation. The Shine cloudy-sky scheme depletes too much radiation in the colder season and not enough in the warm season, especially in the marine warm season. An improved implementation of the Shine scheme was made by varying the cloud optical depth. If cloud optical depth is allowed to vary seasonally from 1 to 7 in the fast-ice environment (late March to early June) and 9 to 20 in the marine setting (early May to mid July), the Shine cloudy-sky fluxes become closer to observed values. These optical depths are within the limits of arctic clouds (Herman and Curry, 1984; Shine et al. 1983; Shine, 1984). Sea ice models should allow for seasonal cloud optical depth variations with respect to incident short-wave radiation, even if they are used in a climatological sense.

The preferred fast-ice $L\downarrow_{clr}$ parameterization is Maykut and Church (1973) after adjusting (decreasing) the clear-sky emissivity to account for a less emissive atmosphere at colder temperatures in NOW. This correction was sufficient for the marine conditions. The Maykut and Church scheme did not contain a seasonal bias once this correction was made.

The preferred fast-ice $L\downarrow$ all-sky combination was Maykut and Church clear-sky and Maykut and Church cloudy-sky. I increased the cloudy-sky emissivity to account for under-estimations when clouds were present which also alleviated a slight seasonal bias.

The Maykut and Church exponential dependence of cloud fraction (equation (4.9)) was also found to be important.

Different $L\downarrow$ all-sky results were found for the marine environment. Overall, the ice-breaker consistently had more negative mean errors compared to the fast-ice site. This was not due to warmer ambient marine temperatures. The all-sky short-wave flux results suggested the marine environment had greater cloud optical depths compared to the fast-ice and terrestrial sites and is consistent with the long-wave results where a higher cloud emissivity may be required; although cloud base height differences could also be a factor. The Jacobs cloudy-sky scheme is more consistent over a seasonal basis and has the smallest mean errors. The Maykut-Church cloudy-sky scheme was further improved by Hanafin and Minnett (2001) for the marine setting.

### *4.4.2 Question 2*

I recommend using 1) the modified Maykut-Church $L\downarrow$ clear and cloudy sky schemes for a polynya fast-ice and terrestrial environment, 2) the modified Efimova/Jacobs or Maykut-Church/Jacobs schemes for the Arctic marine environment and 3) the modified Efimova/Jacobs scheme for application to all three environments simultaneously since it was the most consistent for this purpose. The modified Maykut-Church and Efimova $L\downarrow_{clr}$ formulations are, respectively:

$$L\downarrow_{clr} = 0.729\ \sigma\ T_a^4 \qquad (4.10)$$

$$L\downarrow_{clr} = \sigma\,(0.7 + 0.0066\ ea)\ T_a^4 \qquad (4.11)$$

The modified Maykut-Church and Jacobs L↓ all-sky formulations are, respectively:

$$L\downarrow_{all} = L\downarrow_{clr}\,(1 + 0.32\ c^{2.75}) \qquad (4.12)$$

$$L\downarrow_{all} = L\downarrow_{clr}\,(1 + 0.275\ c) \qquad (4.13)$$

## 4.5 Summary

In summary, the incident short-wave and long-wave parameterizations investigated here estimate these fluxes reasonably well in some cases and not so well in others for a polynya environment. Seasonal biases were found to be the major problem as well as environmental conditions (i.e. near-shore fast-ice conditions were much different than the full marine environment). An attempt was made to correct these biases and different forms for various parameterizations were presented. An important consideration for the incident short-wave fluxes under cloudy skies is the surface albedo due to multiple reflections between the surface and clouds. Chapter 5 explores the spatial variations in surface albedo during the spring melt season over FYI given the dearth of albedo information during this time of year. This will provide insight into the important surface cover types and their associated albedo for better representation in sea ice model albedo parameterization schemes.

# CHAPTER 5: Local and Regional Albedo Observations of Arctic First Year Sea Ice During Melt Ponding

## 5.1 Introduction

In this chapter, I expand on the examination of surface short-wave albedo during the melt season that was initially looked at in Chapter 3. The surface albedo is also linked to Chapter 4 through the incident short-wave radiation. The surface albedo is a critical climatological parameter controlling short-wave responses in the surface energy balance, and has been a major focus in several modeling and observational studies (see for example, Shine and Henderson-Sellers, 1985; Ebert and Curry, 1993; Robinson et al., 1986; Ross and Walsh, 1987; De Abreu et al., 1994; Grenfell and Maykut, 1977; Grenfell and Perovich, 1984). Albedo representation in sea ice models is also critical to modeled sea ice evolution (Shine and Henderson-Sellers, 1985) and still require improvement, especially during melt pond formation to break-up due to multiple surface types that affect albedo during this period. This chapter is designed to investigate the spatial aspects of FYI albedo and provide more insight to albedo characteristics from chapter 3. I also examine the sensitivity of sea ice ablation to percent pond fraction and the associated spatial variability in surface albedo using the 1-D thermodynamic sea ice model discussed in chapter 3.

The research questions are designed to examine the spatial aspects of fractional surface cover types with surface albedo and to determine the role melt ponds play in

these processes over land-fast FYI in the Canadian Arctic Archipelago. Because of the paucity of observational data on melt pond surface albedo, particularly in the Archipelago, the following research questions are addressed:

1) What are the local spatial variations (sub-km scale) of both broadband and spectral albedo using *in situ* observations?
2) How do the surface albedo measurements in (1) scale up to semi-regional (10's km) and regional (100's km) albedos over FYI as related to issues of larger scale albedo estimation during spring melt conditions?
3) How sensitive is sea ice ablation to percent pond fraction and the associated spatial variability in surface albedo?

The primary purpose of deriving larger scale albedo is to provide climate model-scale first-year sea ice albedo during the melt season to see which surface features become important at these scales. The core material in this chapter has previously been published in Hanesiak et al. (2001c).

## 5.2 Methods

### *5.2.1 Site Description*

Data consist of surface albedo measurements made on snow covered smooth first-year sea ice (FYI), aircraft video and satellite observations. Data were collected during the second intensive observation period of C-ICE97 (Collaborative - Interdisciplinary

Cryospheric Experiment) conducted between Year Day 175 (YD175; June 24) and YD193 (July 12), 1997 in Wellington Channel and Lancaster Sound, Nunavut. Standard meteorological observations from Resolute, Nunavut along with data from the first intensive observation period of C-ICE97 are used to initialize and force the sea ice model simulations between YD119 (April 29) to YD212 (July 31). The intensive field program was based on the east coast of Cornwallis Island near Read Bay (75° 3.073' N; 93° 33.193' W) about 65 km NE of Resolute, Nunavut (Figure 5.1). A secondary camp (ice camp) installation was located approximately 4 km ESE of the land camp on a triangular-shaped patch of a thin snow covered (10 cm) smooth FYI (Figure 5.1). The ice camp was the primary albedo sampling site with supplemental locations visited by helicopter.

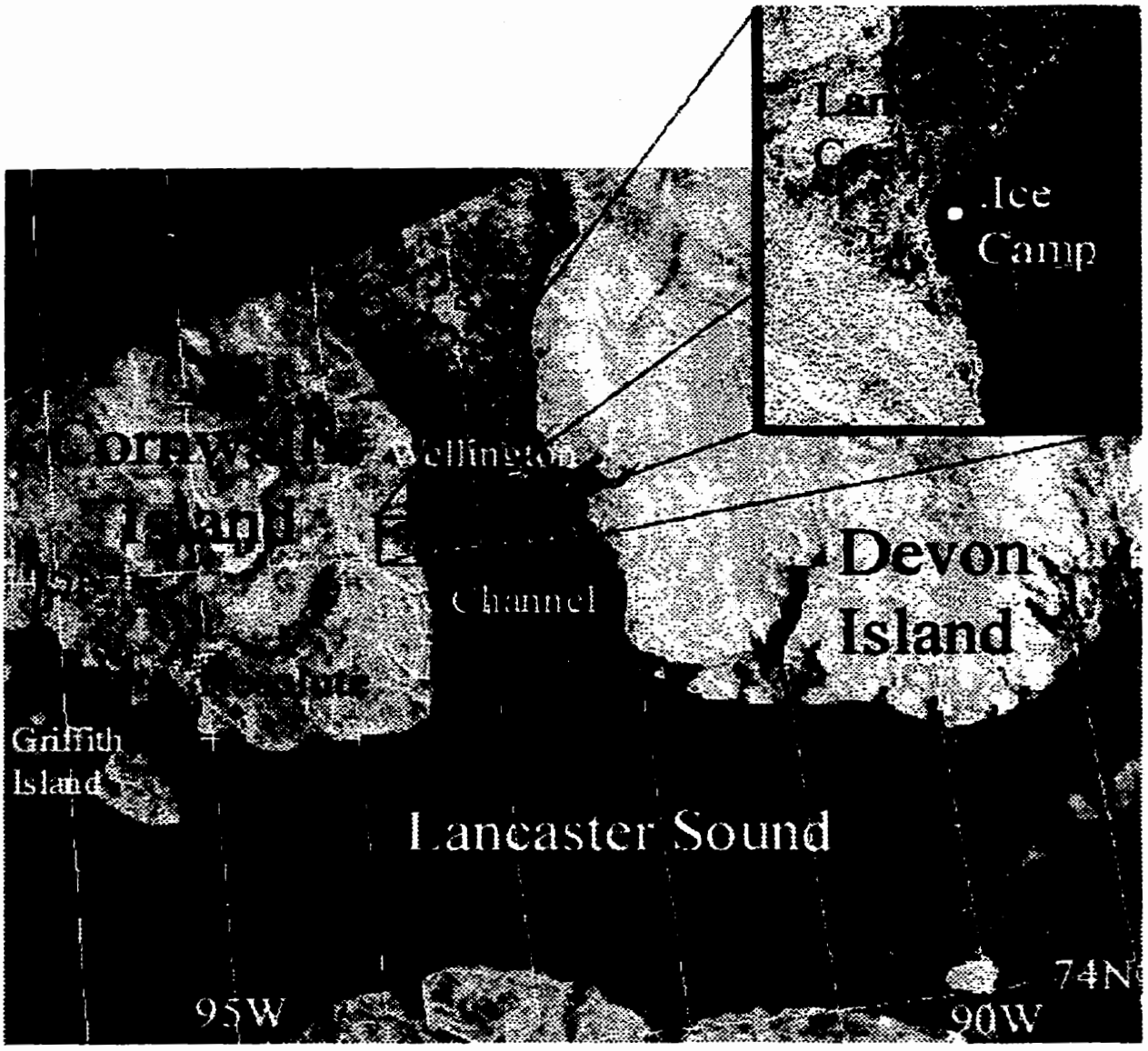



*Figure 5.1: Geographic locations of the land camp (black dot), ice camp (white dot) and ice conditions within Wellington Channel and Lancaster Sound depicted by ERS2 Synthetic Aperture Radar (SAR) on April 3, 1997. Note the arc of rafted and rubbled first-year ice (FYI) with multi-year ice (MYI) embedded north of the ice camp in Wellington Channel (north of 75N). Also note the slightly rafted and rubble FYI ice southwest of Devon Island at the southern end of Wellington Channel and marginally rougher FYI as one proceeds from Griffith Island toward the ice edge in Lancaster Sound. The ice edge is aligned SW-NE beginning at 74N, 90W with open water to the east.*

Upon our field site arrival (YD175; June 24), snow conditions were in the mid-stages of melt with snow depths ranging from 5 – 20 cm and ice thickness near 1.9 m on

the smooth FYI site; no melt ponds were present. By the end of the experiment (YD193; July 12), poor ice conditions and little terrestrial snow made access to the FYI site difficult with full melt pond advancement of about 75% coverage and initial stages of pond drainage. The top layer of the ice surface became crusty, thin and very porous adjacent to the melt ponds from advanced snowmelt stages.

Ice conditions within Wellington Channel and Lancaster Sound were generally FYI between 1.7 – 2.0 m thick. This ice was not as smooth as the FYI ice sampling site according to visual observations and active microwave satellite images (see Figure 5.1). Various pressure ridges and multi-year ice inclusions existed near the coast of Cornwallis and Devon Islands (see Figure 5.1). Highly variable snow depths ranged from 10 – 60 cm, with deeper drifted snow near rougher ice (ie: near pressure ridges). Melt ponds were not present at these locations upon arrival, but developed into similar pond conditions as the FYI site over the 2.5 week period of the experiment.

### *5.2.2 Surface Measurements of Albedo*

Prior to melt pond formation, spectral and broadband albedo measurements were made over various surfaces (see section 5.3) at a consistent location and other locations on the smooth FYI site. Snow physical property measurements (density, wetness, salinity, layer descriptions, and depth) were made but were discontinued beyond YD180 (June 29) due to either the disappearance of snow or development of layers impenetrable to the sampling equipment. Snow density and salinity were collected in 2 cm vertical layers with

a density sampler 2 x 3 cm and salinities measured by a refractometer once the samples melted, accurate to within 0.5 ppt. When conditions permitted, snow wetness between 1% – 10% water by volume (Wv) was measured in 2 cm layers with a capacitance plate precise to within 0.5%. Snow grain measurements were obtained episodically due to melting conditions. Once melt ponds formed, location-consistent and feature-driven albedo sampling was conducted. Broadband surface-based albedo measurements were made along a 1 km long linear transect aligned E-W at the FYI ice camp. These measurements were collected to assess the spatial variability in albedo over local scales. Surface-based albedo was also measured along the aircraft transects for surface validation (section 5.3.1) over a portion (150 m to 300 m) of the full transect distance.

Broadband albedo were measured using a portable Middleton EP16 pyrano-albedometer spanning 300-3000 nm. Measurement accuracy is factory calibrated to within 1.6%. The instrument was mounted on a metallic conduit 0.75 m in length and placed on a tripod approximately 1 m above the surface and leveled.

Spectral albedo was measured with an Analytical Spectral Device (ASD) VIR Spectrometer. High-resolution (1.4 nm bandwidth) spectral measurements were made between 330 - 1064 nm, however, precision of the unit permitted reliable data between 380 – 950 nm. Detector noise was removed (via dark current correction) from each measurement. The unit allows for real-time graphical display and manipulation of spectral albedo data. The foreoptics used included 26° field of view (FOV) and calibrated remote cosine receptor (RCR) with 180° FOV mounted on a tripod 1 m above the surface and leveled. Foreoptics are independent FOV instrumentation that are connected to the ASD

unit. The 26° foreoptic provides directional hemispheric reflectance to focus the sensor over a particular smaller surface as opposed to the RCR which covers a much broader area of measurement over all hemispheric angles. Down-welling irradiance was measured directly with the RCR, which was then inverted to measure up-welling irradiance to measure spectral albedo. A standard white reference barium-sulphate panel was used to convert the 26° foreoptic radiance measurements to albedo. The 26° foreoptic was directed at each surface type in order to sample the most homogeneous parts to ensure no other surface types were part of the measurements. All reflectance measurements are accurate to 1-2% and made under overcast sky conditions. Clear sky sampling led to specular reflection errors, especially using the 26° foreoptic, that could not be corrected since the angular distribution of up-welling radiance was not measured. The unit and white reflector were properly calibrated by the manufacturer (ASD) before the field season.

### *5.2.3 Airborne-Derived Albedo*

Direct broadband albedo measurements were made using a helicopter (Bell Jetranger 206B) on YD185 by mounting the EP-16 albedometer to the pontoon (Figure 5.2) and manually logging the data. Two horizontal flights were made at 152 m and 244 m (500 ft and 800 ft above ground level (AGL)) within 10 minutes of each other under complete overcast conditions (300 m ceiling) near the ice camp. The albedometer was very stable during the flights and the down-welling short-wave component was derived

from surface measurements since this component may have been biased by the helicopter. The albedometer was mounted low enough so the helicopter frame was not hindering the up-welling short-wave radiance. Errors associated with these measurements are unknown, however, instrument stability, variability in measurements (±5%) and a stable solar illumination suggest that measurements were made under the best possible conditions.

Indirect albedo measurements were made using aircraft aerial survey (Figure 5.3) video that were digitized and classified into albedo categories according to surface measurements. A Hi-8 video camera was mounted in the belly of a Twin Otter aircraft looking 'near-to-vertical' flying 610 m (2000 ft) AGL at a speed of 100 knots. The aerial surveys (linear transects) began at 74° 15' N, 95° 20' W (start of transect #1) and proceeded eastward to 74° 15' N, 95° 0' W (end of Transect #1) near the ice edge in Lancaster Sound (bottom right corner in Figure 5.3). Positional information was annotated directly from an onboard Global Positioning System (GPS) where GPS (±100 m) readings were taken at the beginning and end points and notable features along each line. Transects were flown at 5' latitude increments north through Wellington Channel with the last line over dense rubble and multi-year ice regions (transect #17) along 75° 35' N. The first 5 transects had the same length (from 95° 20' W to the ice edge) and the remainder began on Cornwallis Island, ending on Devon Island (Figure 5.3). The first 6 lines were flown on YD181 and the following 11 lines on YD184. Poor weather prevented further surveys.

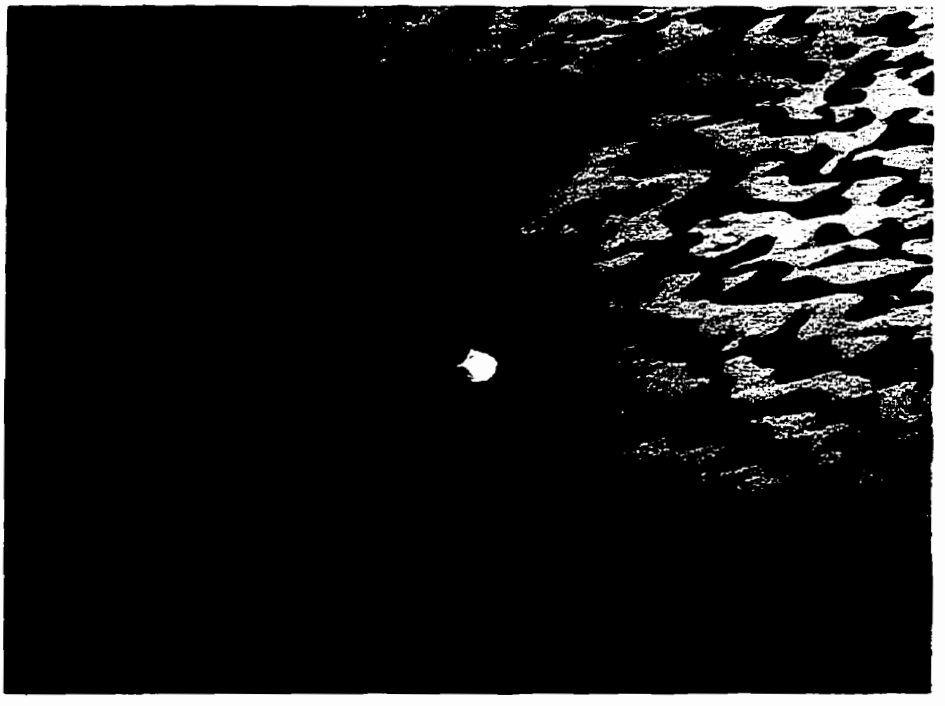


*Figure 5.2: In-cabin helicopter photo of the albedometer mounted to the pontoon. The height above ground was approximately 330 m and total pond fractions were near 35%.*

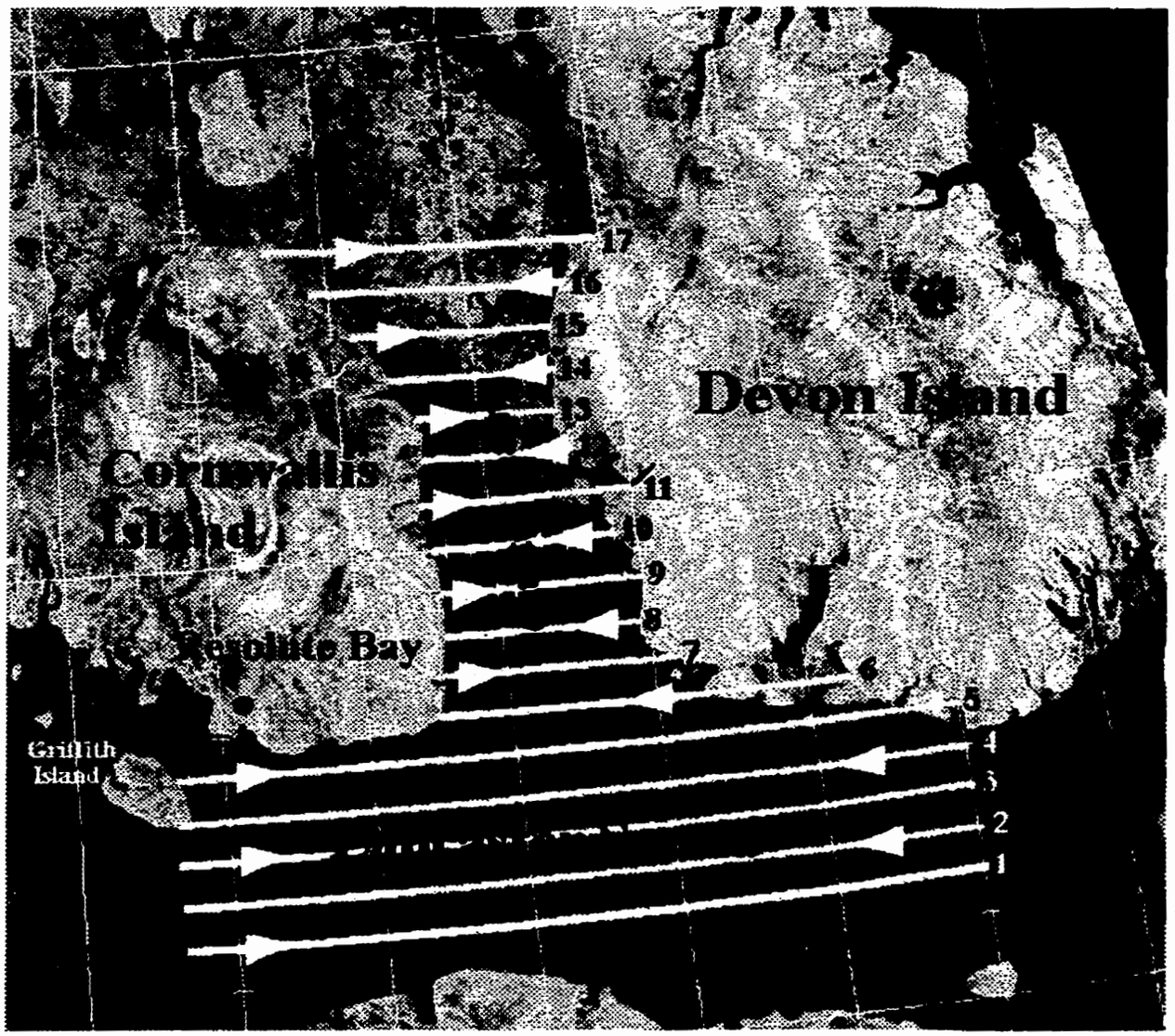



*Figure 5.3: Geographic locations of the aircraft video surveys (linear transects) within Lancaster Sound and Wellington Channel (SAR image date is April 5, 1997). The arrows indicate the direction flown by the aircraft for each transect line. The first 6 lines were flown on YD181 (June 30) and the following 11 lines on YD184 (July 3). Only the first 10 transect lines were used to derive surface albedo from the video.*

Video data were digitized using an automated routine that separated the continuous analog flight record into adjacent frames of 8-bit grey-scale images. Frame spacing was dependent on aircraft speed. Each frame represented an area average of 1.06 km wide x 1.2 km long with a spatial resolution of 16.56 $m^2$ on the ground and was standardized by the overall mean from the entire aerial survey. An unsupervised cluster analysis was performed on a full cross-section of cover types. This type of analysis is common for aerial video and a detailed description of the image analysis scheme can be found in Barber and Yackel (1999). Cluster analysis results consistently identified four classes, corresponding to the four cover types identified in the field; 1) wet snow, 2) a blend of water/snow mixture or highly granulated, wet surface and high/low density bubbled ice surface beneath a very shallow pond perimeter, 3) light colored ponds, and 4) dark ponds. All images were then reclassified into one of four cover types using standard digital image classification techniques.

The fractional cover type mean and standard deviation of percent cover was calculated for each transect for spatial comparisons. Each cover type was assigned an albedo according to surface observations (see section 5.3.2) to assess the spatial variation of surface albedo and make inter comparisons along each transect. Albedos were only derived from transects 1-10 since there were no surface albedo measurements made over rougher ice and multi-year floes. Image data most likely contained shadow contamination due to low arctic sun angles, however, the surveys were conducted during peak daylight hours with an SZA near 52° and removing any shadow effects is difficult. The majority of the data were collected over smoother first-year sea ice where minimal shadow effects

took place. The most significant errors include: 1) misidentification of surfaces under marginal transition conditions, 2) unstable aircraft flight (roll, altitude and latitudinal drift) and GPS error resulting in area and linearity fluctuations of the transects, and 3) assignment of cover type albedo (using ground measurements) with a single albedo value, especially the mixed cover type. Quantifying these errors is difficult, however, albedo magnitudes are expected to be accurate within ±5% based on helicopter validation. Aircraft video-derived albedos were then compared to the Advanced Very High Resolution Radiometer (AVHRR) albedo to corroborate aircraft results and to further upscale albedo measurements to regional estimates (see section 5.3.2).

### *5.2.4 AVHRR-Derived Albedo*

Surface albedos were estimated using the Advanced Very High Resolution Radiometer (AVHRR) onboard the National Oceanic and Atmospheric Administration (NOAA) 11 and 12 satellites. AVHRR narrow band VIS and NIR radiances [channel 1 (0.58-0.68 μm) & channel 2 (0.73-1.1 μm)] are typically used to estimate top of the atmosphere (TOA) albedos. After cloud-screening the AVHRR data, only one NOAA 14 AVHRR scene, on YD180 (June 29 at 04:09:00 CDT), was considered clear enough to accommodate comparison with airborne albedo measurements. The AVHRR albedo are therefore valid under clear-sky conditions. To compare surface broadband albedo measurements with AVHRR values, the up-welling signal recorded by the satellite at the

TOA must be adjusted to account for: 1) sun-sensor view geometry, 2) calibration drift, 3) atmospheric attenuation, 4) anisotropic reflectance, and 5) the limited spectral response of the narrow band sensor. A multistep procedure was developed and is discussed in detail in De Abreu (1996). A brief description of the procedure is given here pertaining to the YD180 NOAA 14 scene.

The pre-launch calibration must be adjusted due to sensor degradation (Holben et al., 1990). Updated calibration coefficients for the YD180 data, provided by NOAA/NESDIS, were used to convert the raw data to up-welling TOA radiances (Kidwell (1991). NOAA 12 data were not considered here due to the lack of appropriate calibration update information. TOA radiances ($L_{sat}$) were then converted to surface narrowband AVHRR albedos ($\alpha_s$) by the following:

$$\alpha_s^i = \frac{L_{sat}^i \pi - a^i f E_{toa}^i \cos\theta_z}{b^i f E_{toa}^i \cos\theta_z} \tag{5.1}$$

where $L_{sat}$ is the satellite measured up-welling radiance, $f$ is the anisotropic reflectance factor (ARF) of Taylor and Stowe (1984) taken from the ARF sea ice dataset developed by Lindsay and Rothrock (1994), $E_{TOA}$ is the inband irradiance at the top of the atmosphere, $\theta_z$ is the solar zenith angle and $i$ denotes AVHRR channel 1 or 2. The coefficients $a$, $b$ describe the atmospheric attenuation of the radiances and are provided for AVHRR 1 and 2 in Koepke (1987). For AVHRR 1, these coefficients are dependent on the solar zenith angle, the atmospheric aerosol optical depth and columnar ozone amount. Coefficients for AVHRR 2 are dependent on the solar zenith angle, aerosol

optical depth and atmospheric water vapour amount expressed as a precipitable water amount. In the absence of appropriate *in situ* data, I used the following climatological estimates of these variables to select the appropriate coefficients: aerosol optical depth 0.05 (@ 0.55 μm), precipitable water content 2 cm, ozone amount 0.36 NTP (Normal Temperature and Pressure). The surface narrowband albedos were converted to broadband albedo (0.15-4.0 μm) via the following model taken from Li and Leighton (1992):

$$\alpha_A = 0.0453 + 0.389\alpha_s^1 + 0.452\alpha_s^2 \qquad (5.2)$$

This method was found to estimate broadband surface albedo in this geographic region with an accuracy of ± 5%. I suspect the accuracy to be degraded somewhat due to: 1) the lack of *in situ* measurements, especially for aerosol optical depth, and 2) the large average solar zenith angle of 78° of the study scene including specular reflection that can amount to errors within ±0.1 albedo units.

### *5.2.5 Sea Ice Model Simulations*

The one-dimensional thermodynamic sea ice model (Flato and Brown, 1996; Hanesiak et al., 1999), described in chapter 3, was used to show how changes in pond fractions and their associated albedo can alter sea ice thermodynamics and thereby its mechanical strength. The melting ice albedo parameterization in the model is based on Arctic lake ice observations of Heron and Woo (1994) that uses ice thickness as a proxy

for time which attempts to capture melt pond evolution. The model's albedo parameterization may produce a quicker temporal decline in surface albedo than observations indicate (shown in chapter 3). Differences in forcing data between the C-ICE97 location and Resolute can cause ice evolution differences, however, surface albedo is of primary importance during late spring ice decay according to field observations.

The ice model was forced with hourly air temperature ($T_a$), relative humidity (RH), wind speed (u), cloud fraction (c), and snowfall (sfall) from standard meteorological observations at Resolute, Nunavut between the first intensive observation period of C-ICE97 (YD119; April 29) to complete modeled ice melt (YD200; July 19). Unfortunately, no continuous set of forcing data were available from the C-ICE97 experiment so Resolute forcing was required. The model was initialized with snow and ice data from C-ICE97 on YD119 were the snow depth = 14 cm, ice thickness = 1.75 m, $T_a$ = -18°C, RH = 77%, u = 4.7 m $s^{-1}$, c = 0.9, and sfall = 0 mm/h. The objective of this modeling exercise is to illustrate the model's effectiveness in recreating observed C-ICE97 snow and ice thickness evolution as well as the seasonal changes in albedo for this particular field year.

The next set of modeling experiments intends to assess how the thermodynamic state and mechanical strength of sea ice is affected by variations in fractional pond cover since thermodynamics determine the ice's susceptibility to break-up (Barber et al., 1998). The sea ice volume reaches a thermodynamic state to a point where the brine volumes in the ice are sufficiently large that sea ice breakup can occur with either oceanic or atmospheric forcing (Wakatsuchi and Kawamura, 1987; Cox and Weeks, 1988). This

point is referenced as 'breakup potential' as it relates to the susceptibility of the sea ice to break-up. To show how discreet changes in pond fraction (PF) affects ice thermodynamics, a set of experiments were conducted using the FB sea ice model and consistent changes in PF. A week-long model simulation beginning on YD164 (start of continuous snow melt) and ending YD170 was performed using a constant PF (and albedo) over the week beginning with a 10% PF and ending with an 80% PF in steps of 10%. A week-long simulation was used instead of a longer time period since it is more realistic to have similar pond fractions over a one week period but still run the model long enough to produce noticeable changes in the ice volume. PF's of 10, 20, 30, 40, 50, 60, 70, and 80% corresponded to albedos of 0.66, 0.62, 0.58, 0.54, 0.50, 0.46, 0.42, and 0.38 respectively. The albedos were estimated using the formula,

$$\alpha_{total} = f_{snow} * \alpha_{snow} + PF * \alpha_{pond} \qquad (5.3)$$

where $\alpha_{total}$ is the resulting albedo, $f_{snow}$ is the snow fraction (wet snow and mixed cover type), $\alpha_{snow}$ is the snow albedo (combination of wet snow and mixed cover type) = 0.7, PF is the pond fraction (light and dark ponds), and $\alpha_{pond}$ is the pond albedo (combination of light and dark ponds) = 0.3. The snow (pond) albedo was computed as an average of the wet snow and mixed cover type (light and dark ponds) weighted more toward the wet snow (light pond) since these cover types are dominant over their counterparts in early ablation stages.

The effects of pond fraction on ice thermodynamics is gauged upon how much the vertical ice temperature gradient changes (ITGC) (warmed) between the bottom of the ice

and the 25 cm depth (from the snow/ice interface) from the initial condition to the end of the week-long simulation. This interval of depth was used since it is felt to represent the average ice temperature profile between the ice surface and ice bottom and also to exclude any near-ice-surface temperature waves that can diurnally bias the profile. The simulations were forced using the same data as the simulations above between YD164-YD170 and conducted with an initial snow depth = 0 since ponds are assumed to be present.

## 5.3 Results

### *5.3.1 Surface Measurements of Albedo*

Spectral albedo and snow physical property measurements were made at the ice camp over different surfaces from melt onset to full melt pond advancement described in section 5.2.1. Sampled surfaces (Table 5.1) ranged from a 22 cm deep snow cover to 20 cm deep melt ponds. The snow pack underwent several freeze-thaw cycles prior to and during the sampling period. Hence, all of the snow albedo sampling was done with some liquid water in the snow pack. Most or all of the brine within the snow pack drained to the bottom layer prior to YD175. All measurements were made under complete overcast skies (unless otherwise stated) to minimize specular reflection and since no bi-directional reflectance data were obtained.

*Table 5.1: Surface type samples of spectral and broadband albedo conducted during C-ICE97 between YD175 (June 24) and YD193 (July 12), 1997. The "S" and "B" refers to either spectral or broadband measurements taken, respectively. All surfaces that were spectrally measured were done so using the 26° and 180° foreoptics. Most measurements were done under cloudy skies unless otherwise stated in the text.*

| Surface Type | Depth Range | Spectral/Broadband |
|---|---|---|
| Snow | 3 – 22 cm | S and B |
| Dense Snow (after rainfall) | 0.5 – 2 cm | S and B |
| Saturated Snow | 0.5 – 2 cm | S and B |
| Snow/Water Mix | 0 – 0.5 cm | S and B |
| High Density Bubble Pond Perimeter | 1 – 3 cm | S |
| Low Density Bubble Pond Perimeter | 4 – 6 cm | S |
| Light Melt Pond | 2 – 8 cm | S and B |
| Dark Melt Pond | 10 – 22 cm | S and B |
| Shallow Pond w/ Ice Skin | 4 – 7 cm (Pond); 0.5 mm (Ice Skin) | S |
| Shallow Pond w/ No Ice Skin | 4 – 7 cm | S |

Rapid snow albedo changes were observed over 3 days (Figure 5.4) during extended warm periods (daily average $T_a$ near 2°C between YD177 and YD179) using the RCR foreoptic. The spectral albedo decreases not only in the NIR due to grain size changes and increased liquid water (by increasing the effective grain size and absorption) (Wiscombe and Warren, 1981) but also in the VIS. VIS changes were caused by the snow surface becoming visibly porous with greater air inclusions as the snow melted along with increased snow grain sizes and likely greater airborne dust contamination from strong winds off nearby land, although this was not directly measured. At the top of the snow layer (top 4 cm), snow densities increased from 380 kg $m^{-3}$ to 405 kg $m^{-3}$ and wetness values increased from 5%Wv to over 10%Wv over the 3 day period. Spectral curves in Figure 5.4 are similar to measurements made by Grenfell and Maykut (1977), Grenfell and Perovich (1984), De Abreu et al. (1995) and Radionov et al. (1997).

Figure 5.5 shows the variability in directional hemispheric spectral albedo using the 26° foreoptic typical of melting FYI surfaces (Table 5.1) in the spring. The spectral curves represent a typical progression if one walked from snow patches to the deeper part of a melt pond. The 26° foreoptic data are shown to highlight spectral albedo differences between various pond depths. The albedo difference between a deeper (22 cm) moist snow and shallow (3 cm) moist snow is roughly 0.1 across the optical spectrum (Figure 5.5). Differences are due to increased liquid water (5% Wv to >10%Wv at the surface), grain sizes (0.75-1 mm to 1-2.5 mm at the surface) and the underlying ice surface in the shallow snow case. Once the snow becomes very shallow and saturated, the ice surface that is blue in color becomes more visible (relative albedo maximum at blue wavelenghts 450-480 nm) and liquid water plays a greater role in the NIR, making an albedo difference of 0.25 across the spectrum compared to shallow moist snow (Figure 5.5). The albedo of melt ponds are dictated by scattering of the underlying ice below 500 nm, a sharp decrease between 500-725 nm caused by increasing liquid water absorption with increasing wavelength and very strong absorption by liquid water beyond 725 nm (Figure 5.5). Similar FYI full-hemispheric (180°) spectral measurements are found in De Abreu et al. (1995) and a few measurements made in other Arctic regions (Grenfell and Maykut, 1977; Grenfell and Perovich, 1984).

In early to mid melt pond stages, pond perimeters have high and low density bubbled ice surfaces. The higher density bubble regions had shallower water depths (0 – 4 cm) and lower density bubbles were beneath 4 – 7 cm of pond water. Inspection of many ponds (>35) revealed that the bubbled regions were non-existent where pond depths were

>7 cm. This is likely due to radiative processes and mechanical erosion of wave action on the pond "shore" releasing the bubbles. Typical spectral albedo differences are 0.1-0.13 albedo units over the optical spectrum between the high bubble density and low bubble density pond surfaces and have 0.05-0.2 higher albedo magnitudes compared to deeper parts of the ponds. The extent or fractional cover of the bubbled regions with respect to the overall pond area varied between 25% - 75% for young ponds 1-4 days old and 5% - 20% for more developed ponds 5-6 days old. By the late stages of ponding (≥1 week), the "shoreline" of the ponds no longer contained bubbled regions where radiative processes and waves made the pond perimeter slope fairly steep with a >7 cm drop from the snow-pond edge to the first few cm in the pond.

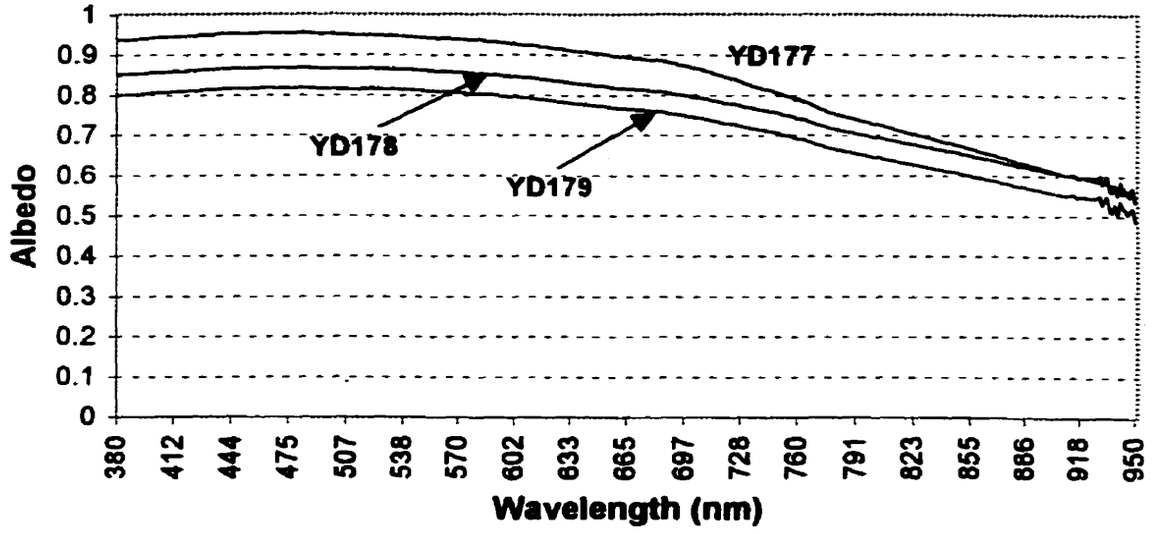



*Figure 5.4: Temporal spectral albedo (using the RCR foreoptic) evolution of a 22 cm moist snow pack undergoing rapid melt between YD177 (top curve), YD179 (middle line), and YD180 (bottom line). Measurements were made under overcast skies near solar noon.*

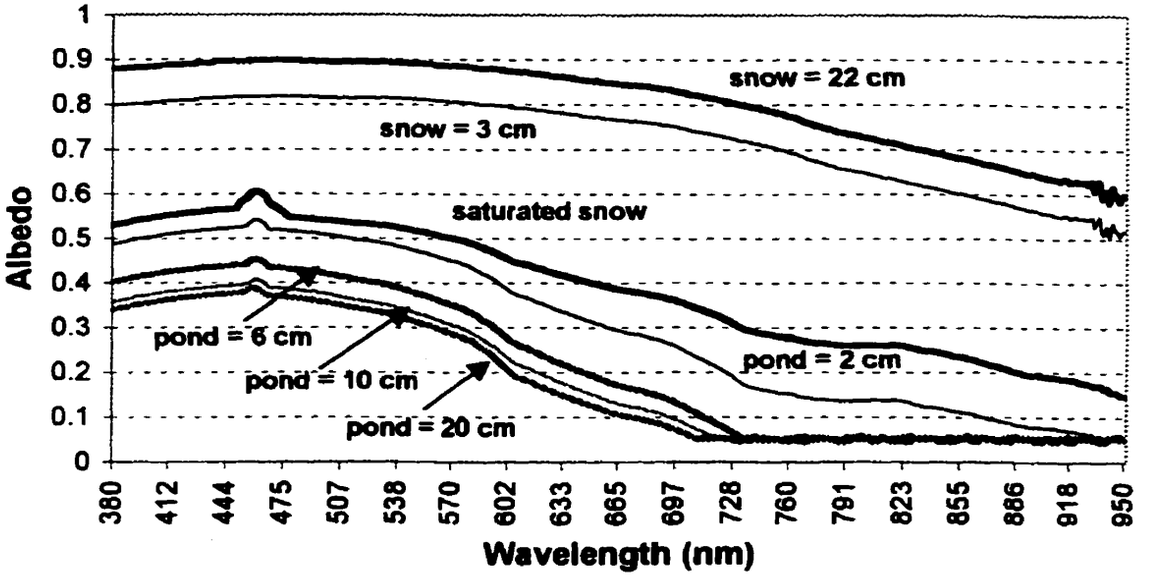



*Figure 5.5: Directional hemispheric spectral albedo of various surfaces over FYI (26° foreoptic) on YD185 under overcast skies at a solar zenith angle near 54°. Surfaces include: melting snow (22 and 3 cm deep), saturated snow (0.5 cm deep), high density bubble pond perimeter (2 cm water depth), low density bubble pond perimeter (6 cm water depth), shallow light pond (10 cm water depth), deeper dark pond (20 cm water depth).*

Overnight cooling can create a thin ice layer on the pond. The RCR spectral albedo of melt ponds 4 and 7 cm deep with a thin ice layer 0.5 mm thick can be very different from ponds with no ice layer (Figure 5.6). The increase in the VIS (0.06-0.07 albedo units for 4 cm pond depth; 0.03-0.04 albedo units for 7 cm pond depth) when ice is present is mainly due to backscatter from tiny air bubbles trapped within the ice layer. A greater density of bubbles in the shallower pond ice layer creates the greater change in the VIS compared to the deeper pond. There is less of a change in the NIR wavelengths (0.02-0.03 albedo units) when a thin layer of ice is present. These results are similar to partially refrozen ponds of Grenfell and Maykut (1977), however, they do not show relative changes in spectral albedo over precisely the same ponds. The ponds sampled here also contained a bubbled underlying ice surface which may account for the higher spectral albedos for ponds with no ice layer compared to that of Grenfell and Maykut (1977) and De Abreu et al. (1995).

Broadband albedo (300 – 3000 nm) sampled surfaces during clear skies with an SZA near 53° (Figure 5.7) range from deeper moist snow (22 cm depth) to dark melt ponds (23 cm water depth). Figure 5.7 shows the mean albedo for each surface from ten to fifteen samples of each with only slight deviations of ± 0.02 albedo units. A steady decrease in albedo occurs as a function of surface type between deeper moist snow (near 0.75) to dark ponds (0.21-0.25) between the range of surfaces that were sampled. The albedo of dark ponds 5 to 23 cm in water depth range slightly between 0.25 to 0.21, similar to other observations (Grenfell and Perovich, 1984). Deeper moist snow here (albedo = 0.73) is similar to the melting snow surface (0.75) of Perovich (1996). The

shallow and saturated snow surfaces (albedos 0.52 – 0.65) in Figure 5.7 are similar to the melting white ice surface (0.56 – 0.68) of Perovich (1996). The melting (blue) ice albedo (0.33) and FYI ponds (0.21) of Perovich (1996) are similar to the Figure 5.7 light pond albedo (0.35) and dark pond albedos (0.21 – 0.25), respectively.

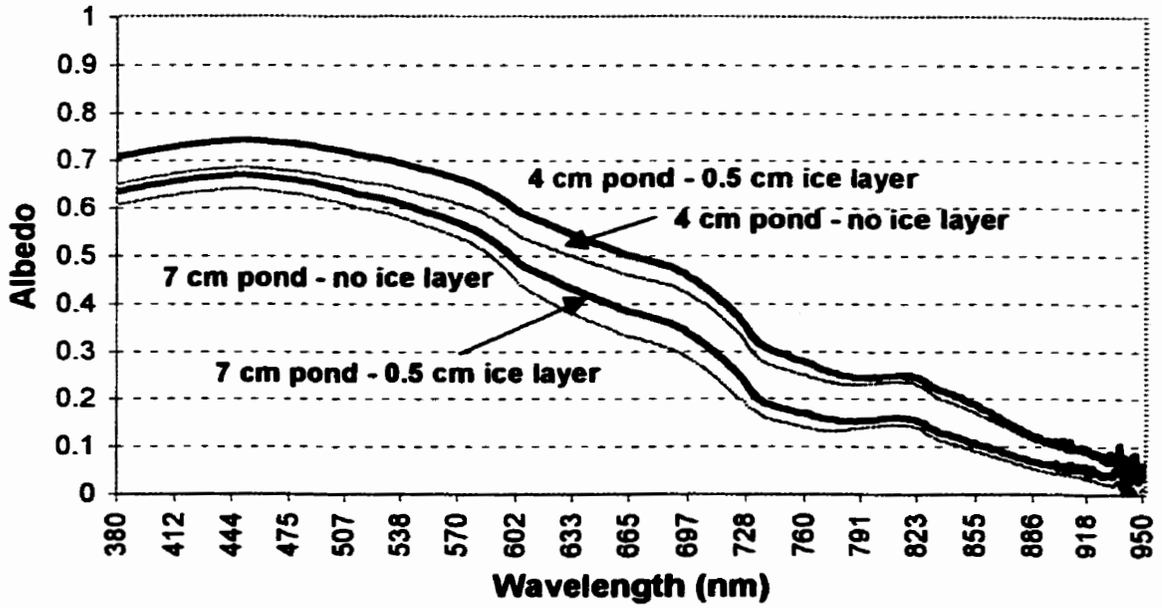



*Figure 5.6: Spectral albedo of a 4 cm deep early season pond and 7 cm deep early season pond with and without a thin (0.5 cm) ice layer. Measurements were made under overcast skies using the RCR foreoptic near 1030 CDT.*

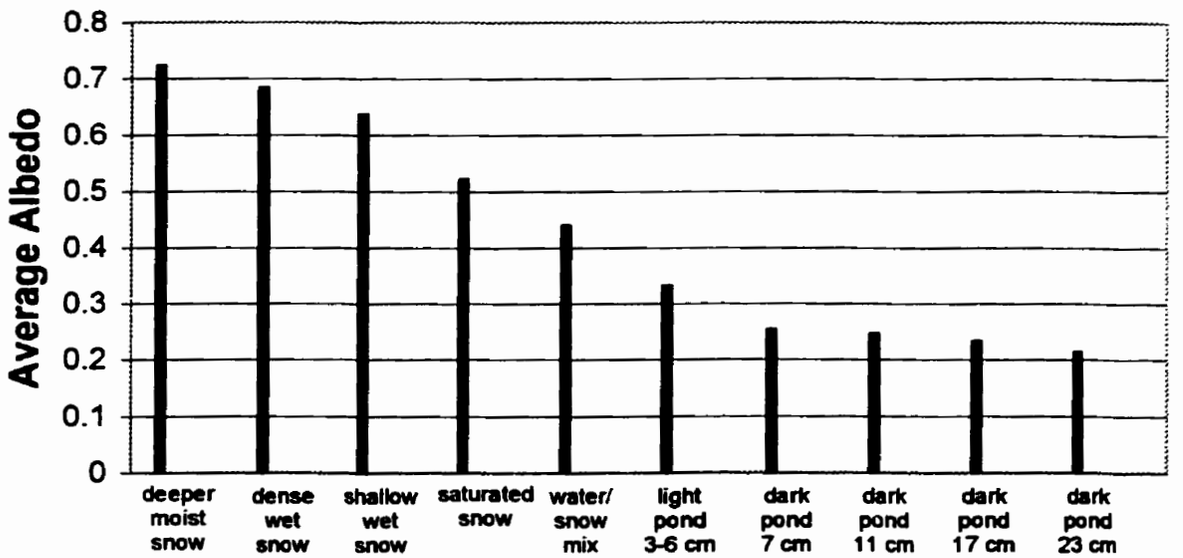



*Figure 5.7: Broadband albedo of various surfaces over FYI. Surfaces include: 1) deeper moist snow (4-6 cm), 2) dense snow (refrozen after a rainfall; 1-2 cm), 3) shallow wet snow (2-3 cm), 4) saturated snow (0.5 – 2 cm), 5) water/snow mixture (0.5 – 1 cm), 6) light colored pond (3-6 cm water depth), and 7) dark ponds (8, 11, 17, 23 cm water depth).*

The spatial variation in surface broadband albedo from ground linear transect measurements shows the changing nature in the distributions of snow patches and melt ponds (Figure 5.8) with albedos ranging between 0.28 over young ponds and 0.75 over moist snow patches with an overall mean of 0.53. Figure 5.8 depicts measurements taken along the 1 km transect at the ice camp on YD183.

Ground linear transect measurements of broadband albedo along aircraft transects 9 - 7 made on YD184 showed slightly different ice conditions compared to the ice camp region (not shown). Melt ponds over the aircraft transects were not as advanced as the ice camp. Albedos ranged between 0.7 for snow and 0.43 for very young bubbled ice surface ponds, with a mean of 0.55 over transect 9, 0.75 to 0.35 with a mean of 0.55 over transect 8, and 0.75 to 0.35 with a mean of 0.6 over transect 7. Melt pond advancement south of the ice camp was delayed by at least 2 days compared to the ice camp. The 2-day lag inference was based on a visual inspection of greater fractional pond cover over the ice camp compared to the other aircraft transect sites and is manifested in the aircraft video analysis (see next section). The greater pond cover at the ice camp was due to less snow accumulation where the very smooth ice (less surface roughness) inhibited snow accumulation.

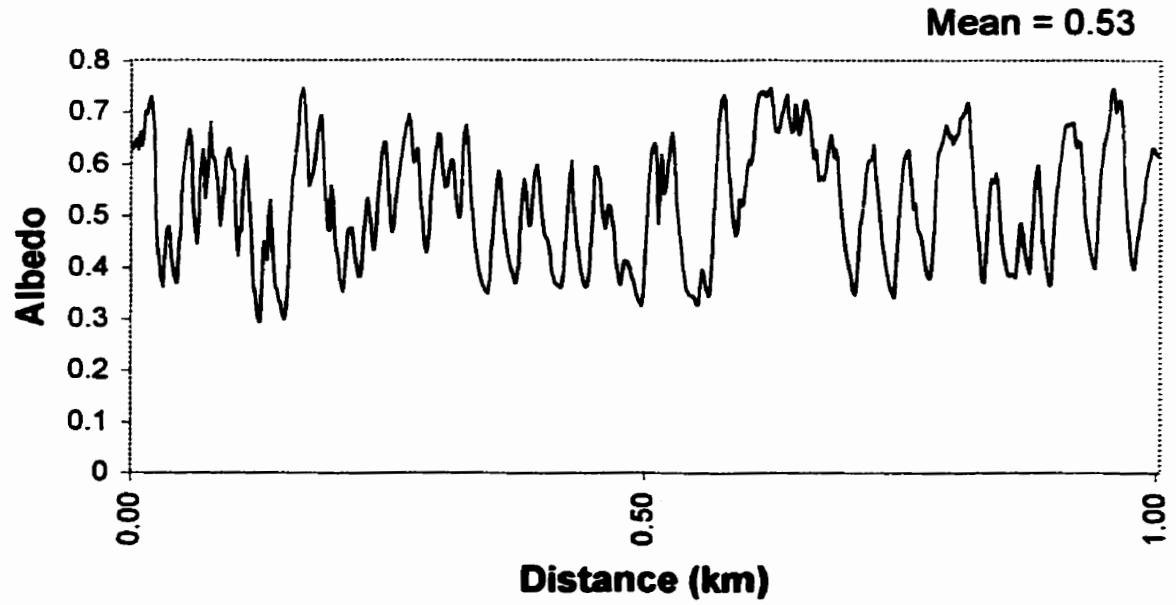



*Figure 5.8: Ground-based broadband albedo along the 1 km transect at the ice camp on YD183.*

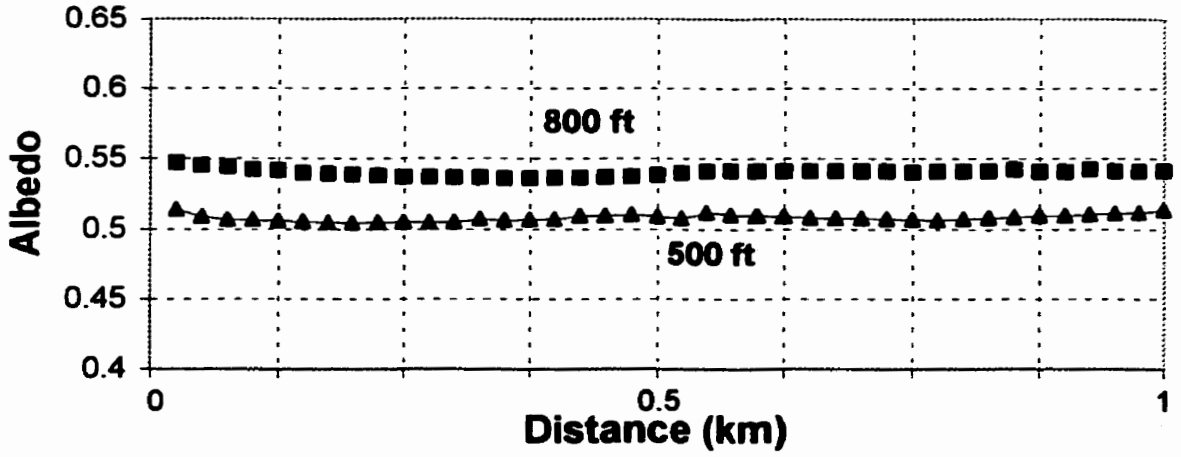



*Figure 5.9: Helicopter horizontal flight albedo at 152 m (bottom line) and 244 m (top line) AGL along the 1 km transect at ice camp on YD185.*

### *5.3.2 Upscaling Surface Albedo to Sub-Regional and Regional Scales*

Helicopter measurements of broadband albedo were made during a horizontal flight described in section 5.2.2 and were conducted to compare to and scale-up from surface-based observations (section 5.3.1). All measurements were made on YD185 under overcast conditions (ceiling near 300 m AGL). Broadband albedo along the 1 km ice camp transect showed fairly consistent magnitudes at 152 m AGL and 244 m AGL (Figure 5.9). The average albedo at 152 m was 0.51 ± 0.015, and 0.54 ± 0.016 at 244 m. The higher albedo at 244 m is the result of more distant, whiter surfaces being measured within the albedometer FOV, including more of the cloudy sky. For example, the albedo at and below 152 m were similar but increased by 0.01 points per 30.8 m between 152 m and 300 m. The helicopter albedos were similar to the ground-based average albedo that was measured two days earlier (0.53). The albedo did not change much over these days due to extended cloud cover and sub-zero air temperatures in contrast to previous days (YD177-179). The albedometer field of view (FOV) was much greater in the helicopter measurements compared to ground measurements averaging out all small scale surface variations, making Figure 5.9 much smoother than Figure 5.8.

Next, aircraft video-derived broadband albedo were estimated over the transects discussed in section 5.2.2. Given the four cover types that were identified in section 5.2.2 (wet snow, a blend of water/snow mixture and high/low density bubbled ice surface beneath a very shallow pond perimeter, light ponds, and dark ponds), each cover type was assigned an albedo based on surface measurements. For convenience, the blend of a

water/snow mixture and high/low density bubbled ice surface beneath a very shallow pond perimeter cover type will be referred to as mixed cover type. Wet snow cover types were assigned an albedo of 0.7, 0.4 for the mixed cover type (average of water/snow mix (0.45) and light ponds (0.35) from ground measurements), 0.35 for light colored ponds and 0.23 for dark ponds (average dark pond albedo from ground measurements).

The percent cover mean and standard deviation for each cover type and transect line are shown in Figure 5.10 with the mean and standard deviation of albedos along each aircraft transect line (Table 5.2 and Figure 5.11). Over FYI, snow patch fractions stayed relatively constant between transects (53-57% ± 5%) with light and dark pond fractional covers increasing from transect 7 - 10 at the expense of the mixed cover type surface (Figure 5.10). This increase in pond cover may be the result of transects 7 - 10 being flown 3 days later than transects 1 - 6 with further melt progressing over that time. The mixed cover type fractional cover varied between 30-38% ± 10% between transects 1 - 6 and decreased to near 20% ± 5% by transect 10. Light pond fractional cover stayed near 10% ± 6% between transects 1 - 6 and increased to near 20% ± 5% by transect 10. Dark pond fractional cover also remained fairly constant near 1% ± 2% between transects 1 - 6, increasing to about 5.5% ± 3% by transect 10.

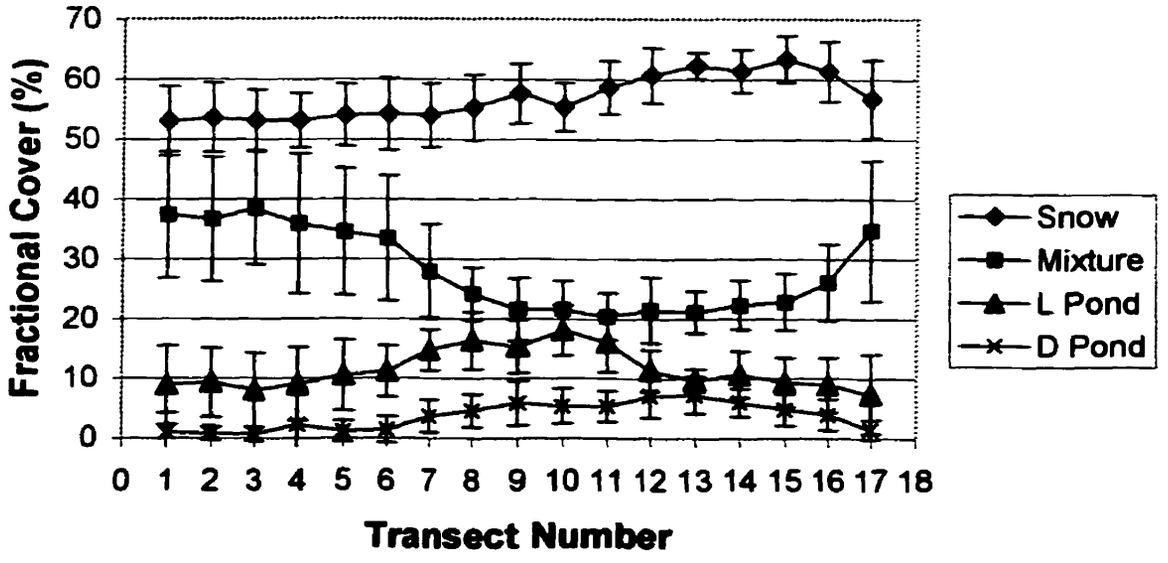



*Figure 5.10: The mean and standard deviation (error bars) fractional coverage of each cover type over each aircraft transect line. All land contamination has been omitted from the data.*

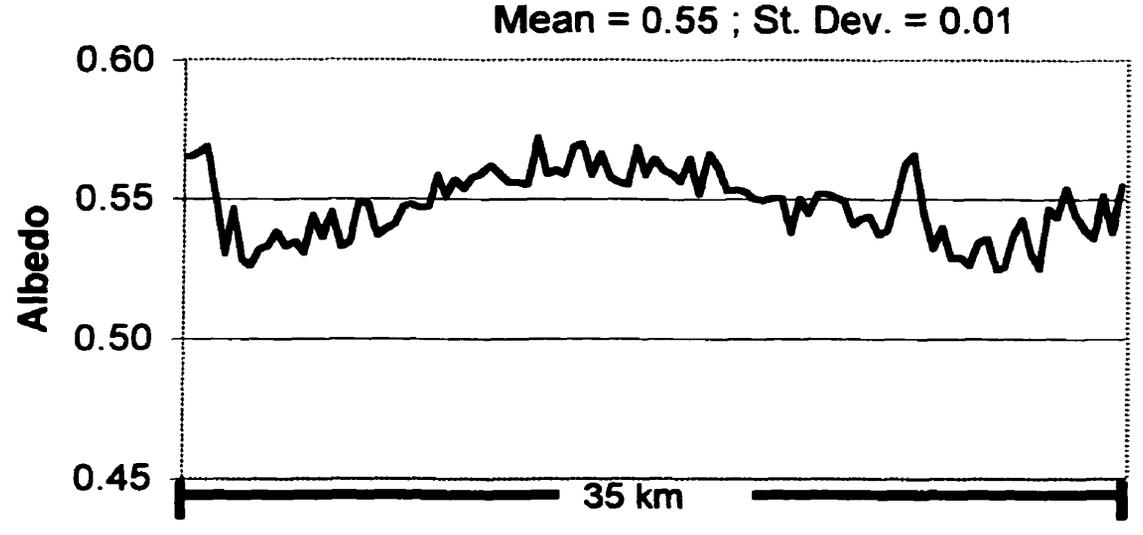



*Figure 5.11: Example of aircraft video-derived broadband albedo along transects #10. Note that the data is displayed in the direction of the flight (that is, data at the left side of the figure represents the beginning of the transect).*

*Table 5.2: Comparison between aircraft and AVHRR-derived mean broadband albedo and standard deviations along each of the aircraft transects 1-10.*

| | Aircraft | | AVHRR | |
|---|---|---|---|---|
| Transect # | Mean Albedo | St. Dev. | Mean Albedo | St. Dev. |
| 1 | 0.553 | 0.018 | 0.50 | 0.05 |
| 2 | 0.555 | 0.016 | 0.53 | 0.07 |
| 3 | 0.555 | 0.015 | 0.54 | 0.09 |
| 4 | 0.551 | 0.014 | 0.56 | 0.10 |
| 5 | 0.555 | 0.013 | 0.56 | 0.09 |
| 6 | 0.555 | 0.012 | 0.60 | 0.08 |
| 7 | 0.548 | 0.014 | 0.61 | 0.05 |
| 8 | 0.549 | 0.017 | 0.62 | 0.05 |
| 9 | 0.555 | 0.014 | 0.61 | 0.08 |
| 10 | 0.548 | 0.012 | 0.60 | 0.07 |
| **MEAN** | **0.552** | **0.015** | **0.573** | **0.074** |

Over rougher and multi-year ice (transects 11-17), as expected, the snow fractional cover is higher than over FYI about 62% compared to 52-53%. Melt pond fractions decrease from nearly 10% to 5% with an increase of multi-year and rougher ice as one proceeds north. The increase in fractional cover of the mixed cover type with increasing multi-year and rough ice concentration is most likely due to the brighter color of very shallow ponds and a greater occurrence of shallow water/snow mixtures in these ice regions according to visual observations.

The resulting albedos along each transect line indicate an overall mean value near 0.55, ranging between 0.548 and 0.555 with standard deviations between 0.012 and 0.018 (Table 5.2; Figure 5.11). This appears that the albedo did not change appreciably across individual transects. However, the maximum and minimum range of albedo derived from

video were 0.60 to 0.47 depending on the transect. As expected, relatively higher albedos correspond to rougher ice conditions since melt pond cover was less prominent there. The video-derived albedo within the ice camp region (between 0.52-0.55; 1st quarter of figure 5.11) is equivalent to the helicopter albedo (0.51-0.54) and the ground-based transect mean (0.53 on YD183). Once again, video-derived albedo shows greater spatial variability than helicopter measurements since the helicopter measurements are highly auto-correlated. Video-derived albedos were purposely less auto-correlated by not overlapping sequential images in the processing (see Section 5.2.2).

Video-derived mean albedo along transects 7 – 9 (Table 5.2) are also very similar to the *in situ* ground-based measurements along the same transects (recall they were 0.60 for #7; 0.55 for #8 and #9). Ground-based albedo averages along transect 8 and 9 are nearly identical to the video-derived mean albedo (0.55). The ground-based mean albedo (0.6) measured at transect 7 is slightly higher than the video-derived albedo mean (near 0.55), however, the ground measurements were taken near the vicinity of higher local video-derived albedos (not shown).

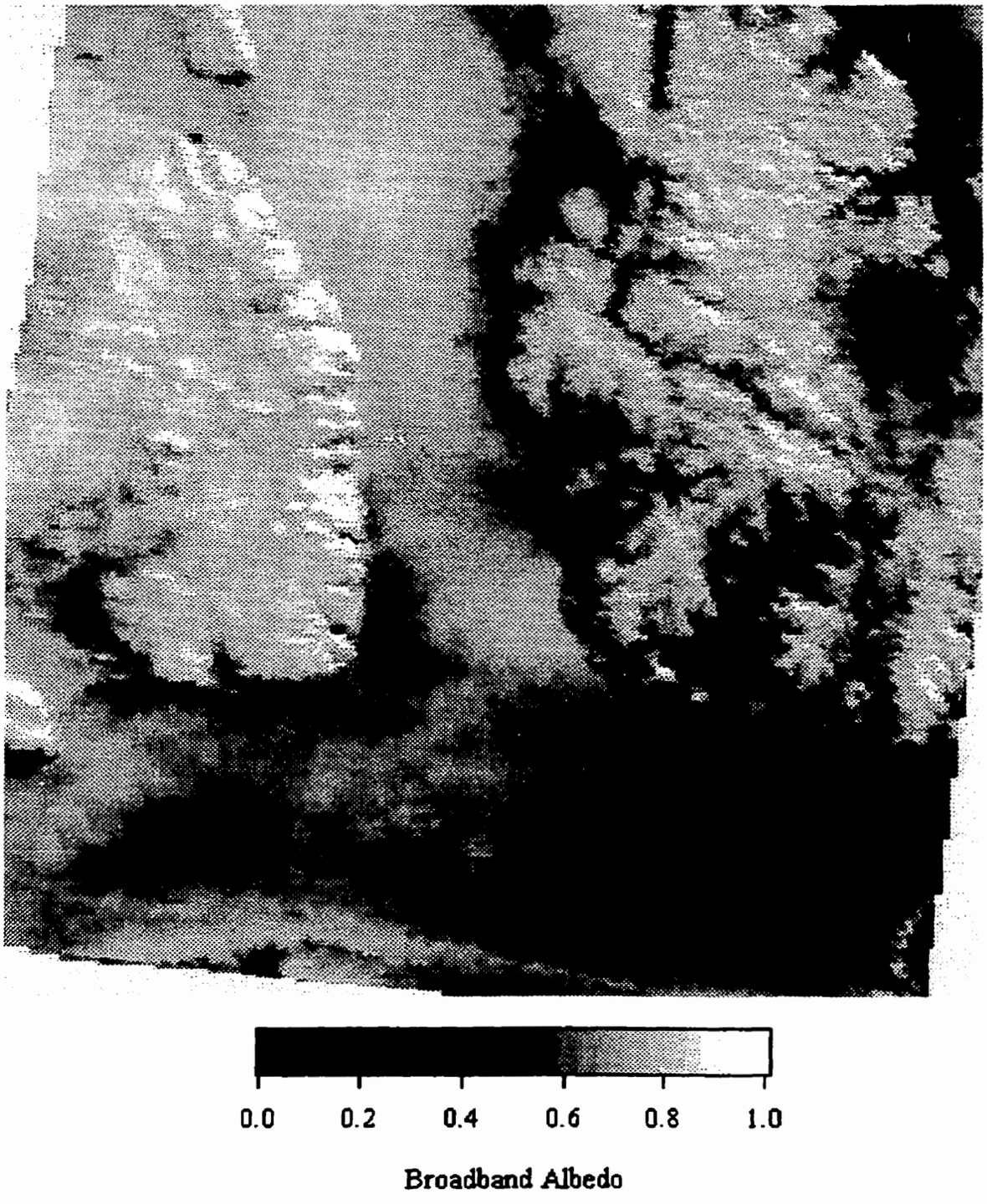



*Figure 5.12: NOAA-14 AVHRR-derived albedo on YD180 (June 29) at 0909 UTC (0409 CDT) in Lancaster Sound and Wellington Channel, one day prior to the first 6 aircraft video survey transects. Cloud contamination was present southwest of Griffith Island but did not contaminate the study region at this time. Note the ice edge in the bottom right corner of the figure. The legend indicates grey-shading corresponding to broadband surface albedo.*

Regional albedo derived from AVHRR imagery shows similar overall ice conditions as Figure 5.1 with generally higher albedo in rafted and rubble ice regions (Figure 5.12). In fact, the defined rubble zone (arc of rougher ice) in Wellington channel, north of transect #10, is typified by an arc of higher AVHRR-derived albedos within and north of the same region as the SAR imagery (Figure 5.1). The rougher FYI zone depicted by SAR imagery south-west of Devon Island at the southern end of Wellington Channel is also apparent with higher AVHRR-derived albedo (Figure 5.12). Slightly higher AVHRR-derived albedo occurs in Lancaster Sound as one proceeds from the ice edge toward Griffith Island, similar to video-derived albedo along transects #1-4 (not shown). Two features may contribute to this spatial albedo pattern; 1) slightly rougher FYI in east Lancaster Sound contained greater amounts of snow compared to the smoother FYI in west Lancaster Sound, leading to more extensive ponding and deeper ponds in east Lancaster Sound, 2) warm air advection from the open water near the ice edge would promote greater surface melt near the ice edge.

Surface albedos were extracted from the AVHRR image nearly geographically coincident to the first 10 aircraft video transects. The geographic location of the transects were subjectively derived after geo-correcting the AVHRR image to correspond as close as possible to the aircraft transect GPS data. The error associated with displacements between actual aircraft transects and AVHRR transects are estimated to be approximately 1-3 km. However, this is expected to be a minor problem due to the relatively homogeneous ice surface over most of the region. Biases may arise under rougher ice

conditions such as the south-west corner of Devon Island where it is difficult to estimate these errors.

The mean values of the AVHRR-derived albedo transects are similar to the aircraft transects in the southern half of the region (transects 1-5) but are slightly greater than the aircraft data by 5% in the northern part (transects 6-10) (Table 5.2). The variability in the AVHRR data is also greater than the video albedo by roughly 0.06 (Table 5.2) making the AVHRR albedo slightly higher in Wellington Channel compared to Lancaster Sound. The variations in cloud cover between albedo measurement platforms would cause some differences between 0.03 to 0.06 albedo units based on our measurements. However, it is possible that Wellington Channel's melt progression was slower than Lancaster Sound primarily due to open water warm air advection in Lancaster Sound. I suspect that if the aerial surveys were conducted all in one day (YD181; June 30), the video data may have shown the spatial albedo differences between Lancaster Sound and Wellington Channel similar to the AVHRR data. The most significant errors in the AVHRR albedos arise from a high SZA and lack of *in situ* aerosol and ozone data at the time of the image. The high SZA induces specular reflection and artificially increases albedos in higher pond fraction regions and increased shadows (lower albedo) on the west coast of the islands, particularly Devon Island. Localized low level undetected morning fog may have also increased albedos, especially in extreme southern Lancaster Sound. Incorrect aerosol and ozone amounts would also affect albedo estimates depending on whether there were more or less of each constituent compared to what was assumed.

### *5.3.3. Effect of Pond Fractions on Ice Thermodynamics*

In this section I examine how the thermodynamic state (and mechanical strength) of the sea ice is affected by variations in fractional pond cover (my third question in this chapter). In effect I am asking the question: "How important is it that the magnitude and spatial pattern of surface albedo (as examined via *in situ* and aerial survey approaches) is correctly parameterized within the annual cycle of sea ice ablation?" First, if we assume the evolution of surface albedo is unknown, the modeled albedo would have to be used to simulate it. A model simulation (standard run) was produced using the model's existing albedo parameterization with the exception of changing its dry snow albedo from 0.75 to 0.83 and the melting snow albedo from 0.65 to 0.73 based on measurements cited here and previous studies. In this case, the standard run produced complete ice melt by YD200 (July 19), 16 days earlier than the observed break-up date (YD217; August 5, indicated by the Canadian Ice Service charts) where *ice still existed* but was mechanically weak enough to break-up.

I define "mechanically weak ice" to be a near-isothermal vertical ice temperature profile near −1.8°C based on previous work (Barber et al., 1998). This also corresponds to a bulk brine volume of about 5% where the brine drainage channels interconnect (Wakatsuchi and Kawamura, 1987) thereby severely reducing the mechanical strength of the sea ice (Cox and Weeks, 1988). In addition, the model produced albdeos near 0.28 on the same dates (YD181-185) where field measurements showed values between 0.53-0.55

(Figure 5.13) (also see section 5.3.2). This suggests the model over estimates the temporal decline in albedo and ice thickness over the melt period in this case.

The time between initial continuous snow melt to the final stages of ice ablation and mechanically weak ice ranges between 3-12 weeks in the Wellington Channel/Lancaster Sound fast-ice region based on 9 years of fieldwork and recorded break-up dates. On average then, it takes about 7 weeks from continuous snowmelt to mechanically weak ice that easily breaks up. The 1997 field season was a "typical" sea ice year based on ice and atmospheric climatology. An albedo progression over this 7 week period is shown in Figure 5.13 ("typical" curve) based on typical fractional cover type changes over this period (Table 5.3) and assuming the last week (in Figure 5.13) includes a thinner ice albedo progression to open.

*Table 5.3: The "typical" or climatological fractional cover types and resulting total albedo over a 7 week period (used in Figure 5.13) produced from several years of field experiments over first-year sea ice.*

| Week | Wet Snow | Mixed | L. Pond | D. Pond | Total Pond | Total Albedo |
|---|---|---|---|---|---|---|
| 1 | 85 | 15 | 0 | 0 | 0 | 0.68 |
| 2 | 70 | 20 | 10 | 0 | 10 | 0.63 |
| 3 | 55 | 30 | 13 | 2 | 15 | 0.56 |
| 4 | 35 | 25 | 30 | 10 | 40 | 0.47 |
| 5 | 0 | 20 | 70 | 10 | 80 | 0.35 |
| 6 | 0 | 50 | 48 | 2 | 50 | 0.37 |
| 7 | Decay to | Open | Water | (0.15) | Or constant | at 0.37 |
| $\alpha$ = | 0.73 | 0.4 | 0.35 | 0.23 | | |

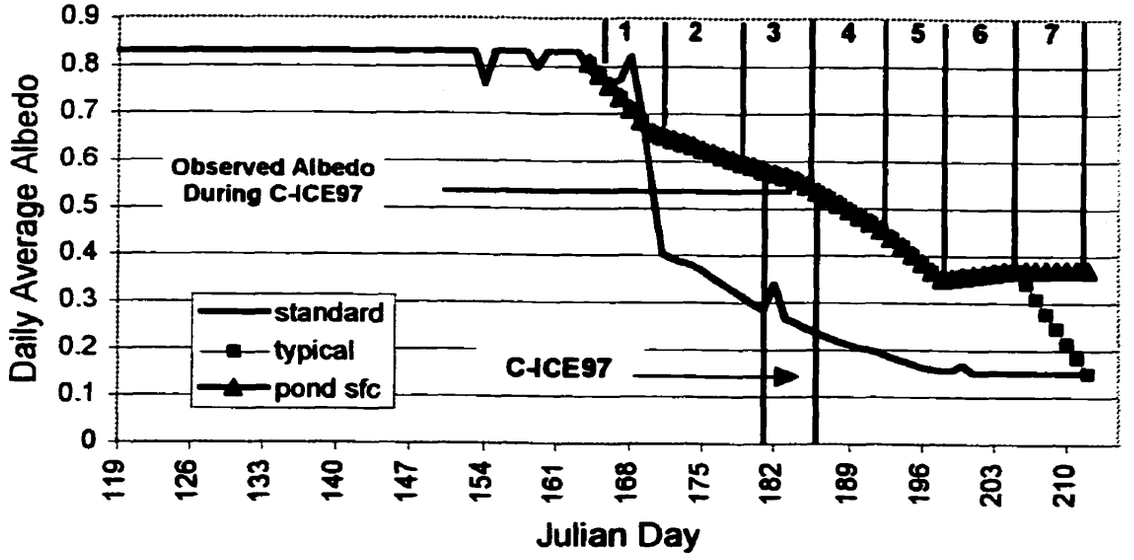



*Figure 5.13: Daily average albedo between YD119 and YD212, 1997 using the model albedo parameterization (standard), a "typical" albedo progression over a 7 week period, and a "pond surface" (pond sfc) albedo progression (same as "typical" curve except using a constant albedo (0.37) for the 7th week) - See text. The "typical" and pond sfc albedo curves were derived by linearly interpolating between values in Table 5.3 over 7 weeks (labeled 1-7). The observed albedo between YD181-185 during C-ICE'97 is also indicated.*

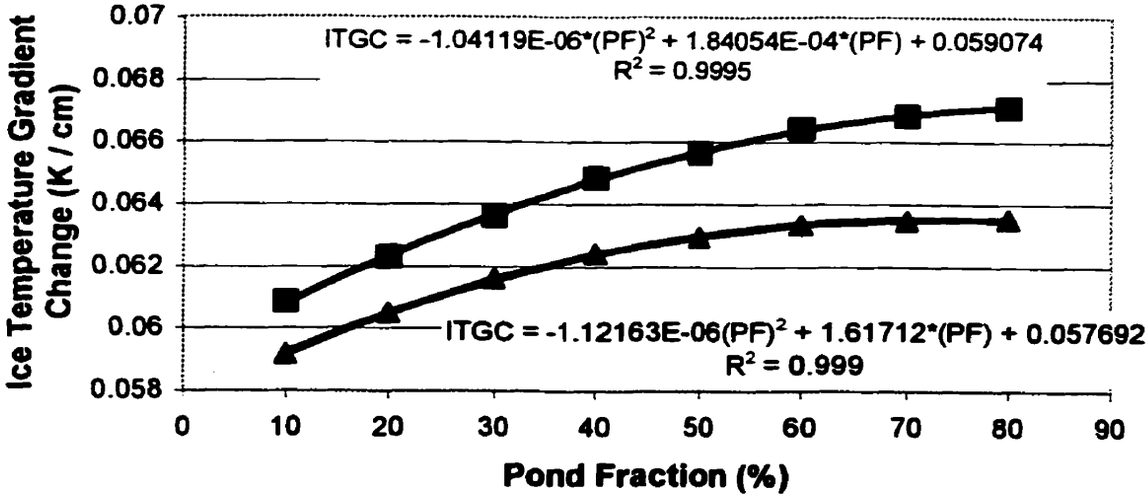



*Figure 5.14: Ice temperature gradient change (ITGC) versus pond fraction (PF) after 4 days (bottom curve) and 7 days (top curve) of model simulation beginning on YD164. The quadratic equations and their $R^2$ describing the ITGC and pond fraction relationship for each curve are also shown.*

Week 6 (in Figure 5.13) showed a slight increase in albedo to simulate typical pond drainage effects. If we use this "typical" albedo curve in the standard model simulation instead of using the model's albedo parameterization, the ice does not completely disappear by the end of the seventh week (YD212) but reaches a thickness near 0.7 m with a near-isothermal vertical ice temperature profile at –1.8°C. Since the ice is still thick by YD212, it is better to assume a stable albedo over week 7 instead of declining the albedo curve to open water values (0.15) (pond surface curve in Figure 5.13). Using this curve further increases the final ice thickness to 0.8 m with an ice vertical temperature profile close to becoming near-isothermal by YD212. These ice conditions are closer to the actual break-up date conditions compared to the standard simulation since ice still existed during actual break-up. Although this experiment is somewhat arbitrary (since we decouple the surface albedo from snow and ice surface changes), the point is to diagnostically illustrate how the knowledge of changes in pond fraction can affect the thermodynamic state of sea ice during its ablation stage.

The next set of modeling experiments assesses how the thermodynamic state and mechanical strength of sea ice is affected by variations in fractional pond cover (outlined in section 5.2.4) since thermodynamics determine the ice's susceptibility to break-up (Barber et al., 1998). The results of the sensitivity tests show that the ITGC (warming) can be predicted using a quadratic equation with respect to pond fraction (Figure 5.14). The quadratic relationship appears due to the physical quadratic dependence of sea ice heat capacity on temperature. The difference between the two curves in Figure 5.14 shows how the quadratic relationship changes (and the ITGC increases) between the

middle (4 days) and end (7 days) of the week-long simulation. As PF increases, the ITGC increases, moving the ice toward an isothermal vertical profile quicker (Figure 5.14).

The greatest acceleration in the ITGC toward an isothermal profile (greatest change in ITGC) is when ponds first form ($10\% \leq PF \leq 40\%$) and is explained by the following. As the ice approaches an isothermal profile, there is less opportunity for a temperature increase to occur in the ice (limited to near-melting temperatures), causing the ITGC to flatten out as PF increases. The quadratic relationship will change according to external forcing such as available short-wave energy (time of year and cloud cover), time the surface energy balance has to act on the ice, and to a lesser degree other atmospheric forcing variables. Pond fraction can have significant effects on the timing of near-isothermal ice profiles (Figure 5.15; using data from Figure 5.14). In this case, over a 4-day model simulation, an 80% PF can produce a near-isothermal ice profile whereas any $PF \leq 70\%$ will not. Over a 1 week simulation, PF's $\geq 50\%$ produced near-isothermal ice profiles and PF's $\leq 40\%$ did not (Figure 5.15). At these rates of warming, the difference in timing to reach an isothermal ice profile (and mechanically weak ice) is >3 weeks between a surface with 10% PF compared to 80% PF. This shows the sensitivity of sea ice temperatures and its susceptibility to break-up according to variations in pond fraction.

These model experiments strongly suggest that it is preferable to either better parameterize or remotely measure (by satellite) pond fractions and their associated albedo for use in annual operational or diagnostic model simulations. Due to the dynamics of how the seasonal evolution combines with snow distribution and sea ice topographic

snow catchment, I consider the remote estimation of pond fractions a preferred approach for model input under operational or diagnostic settings.

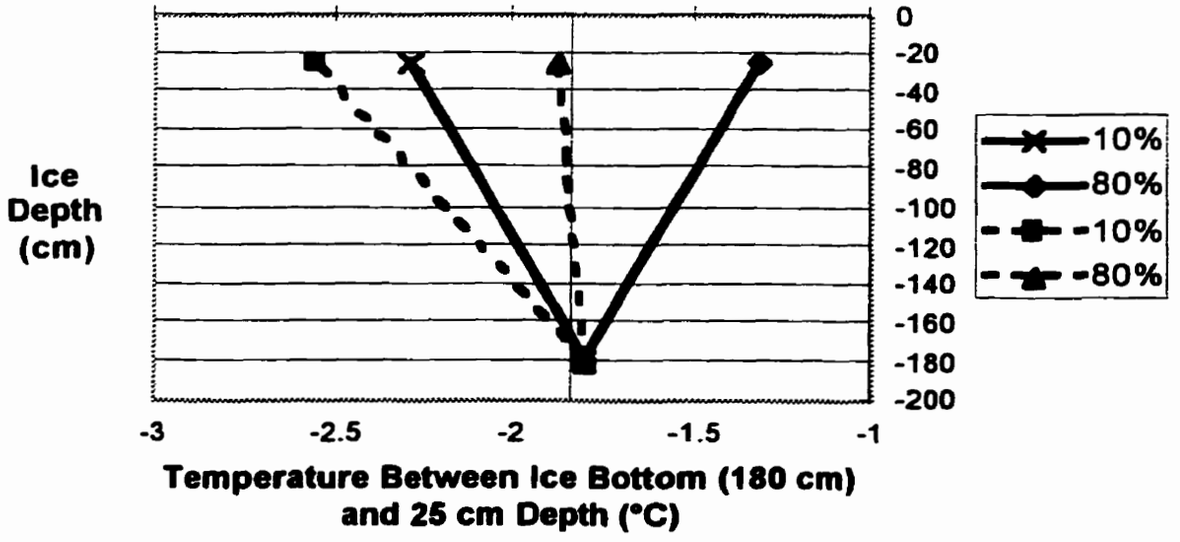



*Figure 5.15: Ice depth (cm) versus temperature (°C) between the ice-ocean interface and 25 cm from the top ice surface. Dashed (solid) lines are the final ice temperature profiles after 4 days (7 days) of model simulation using 10% and 80% pond fractions.*

## 5.4 Conclusions

The main theme of this chapter has been to examine the spatial aspects of fractional surface cover types and surface albedo and to determine the role melt ponds play in thermodynamic evolution of land-fast FYI in the Canadian Arctic Archipelago. I outlined three basic questions toward this theme: 1) what are the surface albedos over FYI and how do they vary over local scales?, 2) how do the surface measurements scale up to

regional scales for larger scale albedo mapping?, and 3) how sensitive is sea ice ablation to percent pond fraction and the associated spatial variability in surface albedo?

### *5.4.1 Question 1*

Results show large variations in broadband and spectral albedos associated with various surface features that one would encounter during the FYI melt season. Albedo varied between 0.75 and 0.21 for a deeper moist snow cover to dark older ponds, respectively. Intermediate surfaces include dense wet snow (0.67), shallow wet and saturated snow (0.52-0.65). Light colored melt ponds with 3-6 cm water depths had albedos which range between 0.32-0.36. Our observations were similar to Grenfell and Perovich (1984) and Perovich (1996), over comparable surfaces.

### *5.4.2 Question 2*

The second set of results used the surface albedo measurements, applying them to aircraft video-derived surface cover type identification over a large region to obtain regional scale (100's km) albedo to compare them to satellite-derived AVHRR albedo. Four basic cover types and their statistics (% cover) were identified. Snow patch fractions ranged between 53-58%, mixed cover type fractions 20-38%, light pond fractions 6-18%, and dark pond fractions 1-5%. Over rougher and multi-year sea ice

(MYI), snow fractions were higher than over the smooth FYI (62-65% compared to 53-58%) and MYI melt pond fractions were half those over FYI. These fractional cover types were then converted to overall albedo revealing similar regional albedos over smoother FYI ranging between 0.54 – 0.56 with standard deviations of 0.01 – 0.02. These were similar to helicopter-measured albedos (0.52-0.55), ground-based measurements and the overall AVHRR-derived mean albedo (0.57) over the same region. Relatively higher albedos corresponded to rafted/rubbled ice conditions since melt pond cover was less prominent there.

### *5.4.3 Question 3*

For the final question, I used a one-dimensional thermodynamic sea ice model to show how the various fractional cover types can affect the thermodynamic state and hence the ice's susceptibility to break-up. Mechanically weak ice can be defined by a near-isothermal vertical temperature profile near –1.8°C (Barber et al., 1998). This experiment also revealed the importance of being able to monitor melt pond fractions and surface albedo through the course of the melt season to indicate when the ice is most likely to break up. This is of particular importance to climate change research for improved understanding of sea ice processes, Arctic shipping, and indigenous people's safety who use the ice for traveling and hunting. First, if we do not know the albedo or pond fraction throughout the model simulation, the use of an albedo parameterization is required. Using the model's albedo parameterization resulted in a 16-day earlier break-up

than observed with complete ice ablation taking place. Complete ablation did not take place in actuality since the ice broke up before this could happen. The model produced a more rapid decrease in albedo and ice thickness over time than was observed. I then assumed a climatological temporal albedo progression produced by several years of fieldwork in the region under study. This resulted in more realistic ice conditions (physically and thermodynamically) when break-up actually occurred. The experiment is somewhat arbitrary since we decoupled the albedo from modeled snow and ice changes, however, the point was to diagnostically show how the albedo affects these kinds of simulations.

The simulations above set the stage for another set of experiments designed to explicitly show how changes in melt pond fraction can alter the ice's thermodynamic state and thus susceptibility to break-up. It was found that the vertical ice temperature gradient change (ITGC) increased as a quadratic with respect to linear increases in pond fraction (PF) in this case. Using the ITGC, an ice surface with 80% pond fraction can produce a near-isothermal ice profile more than 3 weeks earlier than an ice surface with 10% pond fraction. Factors that alter the ITGC relationship with PF are external forcing such as available short-wave energy (time of year and cloud cover), time the surface energy balance has to act on the ice, and to a lesser degree other atmospheric forcing variables. The modeling exercise shows the sensitivity of sea ice temperatures and its susceptibility to break-up according to variations in pond fractions. The model experiments also show that we require either better parameterizations or remotely measured (by satellite) pond fractions and their associated albedo for use in model simulations.

## 5.5 Summary

In summary, I have added to the number of albedo measurements made over FYI during the critical melt pond period and scaled up these surface measurements to regional scales. This was done by understanding which surface cover types become the most important features for modeling large scale albedo in FYI sea ice models. The results will ensure the SEB short-wave responses are appropriately handled in FYI models during the melt pond period (thereby addressing the second issue of my dissertation). I have also linked the changes in short-wave response by means of pond fraction to the thermodynamic state of FYI. That is, if the pond fraction is known, one may determine how long it will take for FYI to become mechanically weak given certain atmospheric forcing, assuming the atmospheric conditions and pond fraction do not change appreciably. This information is useful for climate change research for improved understanding of sea ice processes, arctic shipping (operational ice forecasting), and indigenous people's safety who use the ice for traveling and hunting.

This chapter (Chapter 5) primarily focused on the melt pond period when snow has virtually entirely melted. The following chapter looks at the period of time prior to melt pond formation by outlining the importance of the snow layer in annual FYI cycles from a modeling perspective. However, several albedo measurements of melting snow from Chapter 5 can be applied to results in Chapter 6 in the future. Chapter 6 focuses on improving the handling of the snow layer within thermodynamic models of sea ice. If the models do not realistically recreate the snow layer evolution, the simulated melt pond

period may occur too soon or too late in the year, ultimately leading to unrealistic FYI evolution. The next chapter compares snow and ice simulations from a FYI model that treats the snow layer as a single bulk homogeneous slab (as in Chapter 3) and a coupled snow sea-ice model that treats the snow layer in a much more rigorous and realistic fashion. Results will suggest whether using a more sophisticated model is warranted for annual cycles of FYI. The coupled model is also a necessary step for linking thermodynamic modeling to microwave remote sensing in Chapter 7.

# CHAPTER 6: Coupled 1-D Thermodynamic Snow Sea-Ice Model: Climate Processes

## 6.1 Introduction

This chapter is devoted to illustrating the importance of the snow layer and its processes on FYI over annual cycles using numerical modeling techniques. This is done by advancing the modified 1-D sea ice model of Flato and Brown (1996) used in Chapter 3 by coupling it to a 1-D mass and energy balance snow model called SNTHERM (see Jordan, 1991; Jordan et al., 1999). The purpose is to illustrate snow and ice evolutionary differences between the simpler 1-D sea ice model (in Chapter 3) and the more sophisticated coupled snow sea-ice model and compare them to *in situ* observations. This will suggest whether there is an advantage of using a more sophisticated model for annual snow-covered FYI. The other purpose for developing and testing a coupled snow sea-ice model is for microwave remote sensing applications (Chapter 7).

SNTHERM was adapted to simulate snow thermal and mass balance processes over sea ice by Jordan et al. (1999) but lacks other physical processes (such as an ocean mixed layer) to be used exclusively over FYI that ablates every summer and reforms in the fall. The coupling of SNTHERM to the Flato and Brown ice model allows the coupled model to simulate full FYI annual cycles with more sophisticated layered snow processes instead of a single bulk homogeneous snow slab.

The snow layer portion of the coupled model (SNTHERM) is designed to numerically simulate within-snow temperatures, bulk density, snow grain diameter, liquid water content, and other mass and energy exchanges for each snow layer (or node) over time. It does this by accounting for many of the physical processes outlined in Chapter 2

including the surface energy balance (SEB), snow metamorphism, compaction, melt effects, settling from age and overburden, liquid water filtration, blowing surface snow and wind packing, and newly fallen snow evolution to name a few. All of these processes are important for determining the optical, thermal and mass balance properties in the snow. Some of the model limitations have been described in section 2.3.3.

Salinity in snow is also an important feature that is typical of snow over FYI. Salinity plays a large role in snow thermodynamics (Papakyriakou, 1999) and microwave dielectric properties (see for example Barber et al., 1994; Barber and Nghiem, 1999). Thermal conductivity and specific heat parameterizations have been implemented in the coupled model to account for salinity affects on snow thermal properties (see section 6.2) since SNTHERM did not account for salinity in this manner. Note however the physical presence of salt in snow is not parameterized with regards to its affects on partial fractions of ice, air and liquid (brine) which can further affect the thermal snow state. The result of this simplification will be seen throughout Chapters 6 and 7. A description of how salinity was parameterized, the model coupling process, and model forcing is given in the following section.

In summary, the main objective of this chapter is to demonstrate the basic differences between a coupled 1-D snow sea-ice model and a similar model that uses only a single bulk property snow layer. In situ observations from SIMMS'92-93 (LeDrew and Barber, 1994) are used to quantify model performance. I also explore coupled model sensitivity to salinity in the snow layer through its effects on thermal conductivity and specific heat using parameterizations outlined in Papakyriakou (1999). Crocker (1984) demonstrated that salinity in snow and its effect on thermal conductivity improved a

simple fast-ice growth model, although his study did not include a full annual cycle for comparisons. Other studies suggest the importance of snow thermal conductivity on sea ice in a global climate context (see for example, Fichefet et al., 2000) but do not account for salinity effects directly. The specific research questions include:

1) Are there model differences in seasonal snow and ice thickness evolution and how do they compare to observations?
2) What are the model thermodynamic differences and how do they compare to observations?

## 6.2 Observational Data and Numerical Models

### *6.2.1 Observational Data*

The field observations used here are the same as those in Chapter 3 that were collected in the spring/summer periods during the SIMMS (Seasonal Ice Monitoring and Modeling Site) experiments between 1992-1993 near Resolute, Nunavut. The 1992-93 sea ice cycle was selected to perform the numerical model simulations since the non-coupled model simulations during this period were previously conducted in Chapter 3 appearing in Hanesiak et al. (1999). These two spring periods also offered distinctly different sea ice and meteorological conditions (see Chapter 3).

The field observations included incident and reflected short-wave radiation ($K\downarrow$, $K\uparrow$), net short-wave ($K^* = K\downarrow - K\uparrow$) downwelling and upwelling long-wave radiation ($L\downarrow$, $L\uparrow$), net long-wave ($L^* = L\downarrow - L\uparrow$) and net radiation ($Q^* = K^* + L^*$). The observations also include vertical temperature profiles within the snow/ice volumes and

snow and ice thickness. Instantaneous on-the-hour data were used since this corresponds to the temporal forcing used in the model simulations. Errors in snow and ice temperature measurements are roughly 0.2 - 0.5°C but may be higher in snow during the melt period. Snow and ice thickness were measured periodically throughout each field season with typical errors near 1 cm and 8 cm for snow and ice, respectively.

Other surface observations include hourly standard meteorological observations taken at Resolute, Nunavut between YD107, 1992 to YD170, 1993 by the Meteorological Service of Canada (MSC). The pertinent data include air temperature ($T_a$), relative humidity (RH), wind speed (u), precipitation (sfall), and 3 cloud layer fractions (c).

The snow layer salinity profiles used in the model simulations were derived from measurements in Barber et al. (1995b). A third-order polynomial was fit along the data to produce an average (AVG) profile and conservative upper limit (UPR) profile (Figure 6.1) given by

$$S\,(AVG) = 17.3882 - 172.212\,z + 571.259\,z^2 - 634.732\,z^3 \qquad (6.1)$$

$$S\,(UPR) = 25.5346 - 249.507\,z + 821.958\,z^2 - 913.287\,z^3 \qquad (6.2)$$

where z is the increasing snow depth above the ice surface (in meters) and S is salinity (ppt). Note that snow depths >30 cm will produce unrealistic negative salinities. A provision in the model was inserted to ensure negative salinities did not occur. The use of more accurate (but more computationally expensive) cubic spline fits did not offer advantages over the third-order polynomials and were not used. The salinity profiles were held constant in time for simplicity and is considered a valid assumption until significant liquid water filtration occurs within the snow (i.e. brine flushing; see Barber et al. 1994).

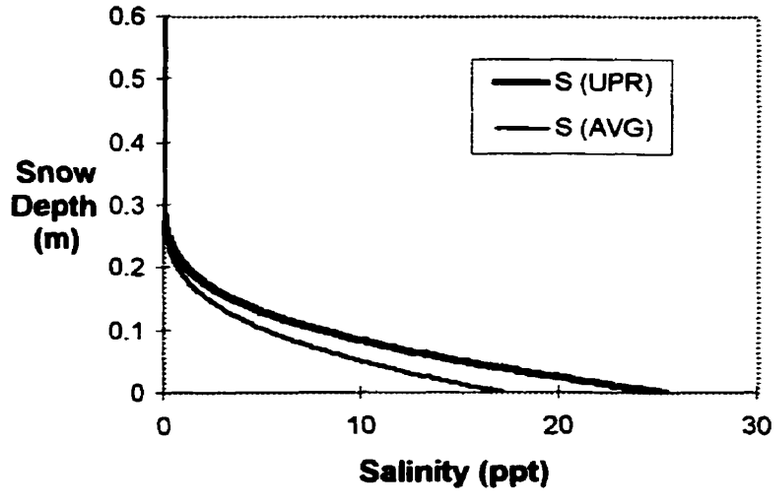



*Figure 6.1: The snow salinity (ppt) versus depth above the ice surface (m) third-order polynomials fit along average (AVG) and upper limit (UPR) observations made by Barber et al. (1995b) for a 60 cm snow depth.*

### *6.2.2 Numerical Models*

The modified FYI model of Flato and Brown (1996) (in Chapter 3 and Hanesiak et al., 1999) is coupled to SNTHERM (Jordan et al., 1999); a type of coupled model that has previously been unavailable. The non-coupled model is the modified FYI model of Flato and Brown (1996) in Chapter 3. The ice and snow models are coupled at the top ice layer with similar thermal properties (temperature, density, thermal conductivity and heat capacity). This was done to preserve the snow model handling of mass balance at the snow-ice interface, i.e. excess liquid water that is not re-frozen at the base of the snow pack is artificially drained to simplify mass exchanges (Jordan et al., 1999). Heat fluxes from the snow model are passed into the ice model that computes the new ice temperatures that are in turn passed back into the snow model at each time step. The time step is variable to a minimum of 5 s during water flow within the snow. The SEB is

controlled by the snow portion of the model when snow is present and by the ice model if snow is absent. This was done since the turbulent exchange parameterization in the snow component of the model is more sophisticated than the ice model. When snow is not present during the melt pond season, the surface properties such as albedo are better (although very simply) handled in the sea ice model.

The coupled and non-coupled models are forced with basic hourly meteorological data; air temperature, relative humidity, wind speed, incident short-wave and long-wave radiation along with measured surface albedo. If incident radiation and albedo are not measured, they can be parameterized with additional cloud cover input data. The coupled model is initialized with snow and ice physical profiles (densities, grain sizes, layer thickness) and temperature profiles. The non-coupled model is similarly initialized but without the internal snow physical properties. Fifty ice layers are used here in both models that remain constant in number over time. A variable number of snow layers exist in the coupled model depending on snow layer characteristics (see Jordan et al. 1999). Specific information about the snow and ice components of both models can be found elsewhere (Flato and Brown, 1996; Jordan et al., 1999; Hanesiak et al., 1999).

The snow thermal conductivity ($k_s$) and specific heat ($C_s$) parameterizations are discussed next since they are used for the inclusion of salinity within the snow. Note that the physical presence of salt in snow is not parameterized with regards to its affects on partial fractions of ice, air and liquid (brine) which can further affect the thermal snow state. The result of this simplification will be discussed later. Salinity inclusions significantly reduces the snow thermal conductivity and increases the snow specific heat (when excluding salinity from the mass balance) depending on salt content and

temperature (Papakyriakou, 1999). The snow effective thermal conductivity ($k_s$) without inclusion of salinity is estimated from density ($\rho_s$) by Jordan et al. (1999) as

$$k_s = k_a + (7.75 \times 10^{-5}\ \rho_s + 1.105 \times 10^{-6}\ \rho_s^2)\ (k_i - k_a) \qquad (6.3)$$

where $k_a$ and $k_i$ are the thermal conductivities of air (0.023 W $m^{-1}$ $K^{-1}$) and ice where $k_i$ varies with temperature as

$$k_i = (780 / T) - 0.615 \qquad (6.4)$$

where $k_i$ is in W $m^{-1}$ $K^{-1}$, $\rho_s$ in kg $m^{-3}$ and T is in Kelvin.

The effective thermal conductivity for saline snow ($k_{ss}$) is estimated from the modified Pitman and Zuckerman (1967) model by Crocker (1984) and further by Papakyriakou (1999). Conservative limits of this model are between snow salinities (S) of 0 to 40 ppt, $\rho_s$ between 100 to 450 kg $m^{-3}$, and liquid water contents up to about 8% Wv (water by volume). The $k_{ss}$ scheme was used in place of the original $k_s$ parameterization even for non-saline snow layers. A lower limit to $k_{ss}$ was arbitrarily set to 0.001 W $m^{-1}$ $K^{-1}$ in snow to prevent unrealistically low values at or near melting temperatures and preserve model stability during these events.

The apparent specific heat ($C_s$) (sensible and latent heating) for non-saline snow is parameterized as in Jordan et al. (1999) by

$$C_s = C_t + L_{li}\ (\gamma_w / \rho_t)\ (\partial f_l / \partial t) \qquad (6.5)$$

where $C_t$ is the combined specific heat in terms of mass fractions (ice and water) = $(1/\rho_t)($ $\gamma_i\ C_i + \gamma_l\ C_l)$, $\rho_t$ is the combined density, $\gamma_i$ is the bulk ice density, $\gamma_l$ is the bulk liquid density, $\gamma_w$ is the combined bulk density, $C_i$ is the specific heat of ice (= -13.3 + 7.8 T),

$C_l$ is the specific heat of water (4217.7 J $kg^{-1}$ $K^{-1}$), $L_{li}$ is the latent heat of fusion for ice (3.335 x $10^5$ J $kg^{-1}$), and ($\partial f_l$ / $\partial t$) is the time rate of change for mass liquid-water fraction ($f_l = \gamma_l$ / $\gamma_w$). Mass-liquid water fraction ($f_l$) takes an empirical form from Guryanov (1985) and is outlined in detail in Jordan (1991). All units are in standard SI.

The apparent specific heat for saline snow ($C_{ss}$) was empirically derived from experimental measurement via Ono (1966) given by

$$C_{ss} = 2.114 + 0.0075\ T + 18052.0\ S / T^2 - 3.35\ S + 0.84\ S\ T \qquad (6.6)$$

where $C_{ss}$ is in units of kJ $kg^{-1}$ $°C^{-1}$, T is in °C and S in kg $kg^{-1}$ (ppt/1000). The $C_{ss}$ scheme replaced the original $C_s$ parameterization, similar to the thermal conductivity. An upper limit to $C_{ss}$ was arbitrarily set to $5x10^4$ J $kg^{-1}$ $C^{-1}$ in snow to prevent extreme high values at or near melting temperatures and preserve model stability. A comparison between $k_s$, $k_{ss}$, $C_s$, and $C_{ss}$ for the non-saline and saline cases is discussed in chapter 7 where a more detailed investigation of the coupled model is performed.

## 6.3 Methods

To examine the first question, the non-coupled (NC), coupled model without salinity (C), and coupled model with salinity (UPR; using the UPR salinity curve in equation 6.2) were forced with Resolute meteorological data between YD107, 1992 to YD170 1993. The models were initialized and forced with exactly the same snow/ice, radiative and meteorological data. Further to question 1, sensitivity runs were conducted to examine detailed differences in snow and ice evolution and ice growth between the

models in a fall freeze-up scenario. This involved similar model runs as above except they were initialized with an ice thickness equal to 1 cm and 8 cm snow layer on YD195. Because of significant snow thickness differences between the coupled and non-coupled models (see section 6.4), the NC was forced to use precisely the same snow thickness evolution as C for the freeze-up case. This allowed a very detailed look at snow and ice evolution differences between C, NC, and UPR.

To examine the second question, NC, C and UPR were forced with SIMMS'92 *in situ* radiative and meteorological data between YD107 to YD177, 1992. The models were initialized with YD107 SIMMS'92 snow and ice data. Snow and ice thickness evolution variations were examined as above. In addition, the coefficient of determination ($R^2$), mean bias error (MBE), and root-mean-squared error (RMSE)) were computed for thermodynamic variables including the snow surface temperature ($T_{sfc}$), ice surface temperature ($T_{isfc}$), and net surface heat flux ($Q_{net}$). $Q_{net}$ is defined here to include the net radiation and turbulent heat fluxes and was included since it directly impacts the thermodynamics. This allowed a detailed look at model differences given similar forcing and high temporal resolution *in situ* data to gauge model performance.

## 6.4. Results

### *6.4.1 Seasonal Snow and Ice Evolution*

The C snow and ice thickness evolution is closer to observations compared to NC for both years that were simulated (Figure 6.2). This is partially due to the blowing snow parameterization in C; NC does not parameterize blowing snow. The fractional stress

gradient in the blowing snow scheme of C was altered to provide a more realistic snow evolution according to the wind conditions at the SIMMS site (see Chapter 2 and Jordan et al., 1999). If C produced a deeper snow cover, its simulated snow decay would be closer to observations (see section 6.4.2). In general, differences between actual and parameterized incident radiation as well as meteorological conditions between Resolute and the SIMMS site can create discrepancies between modeled and observed snow/ice evolution; this can be seen in section 6.4.2 and has been noted previously (Hanesiak et al., 1999). It is interesting to note the 1992 break-up (BU) and freeze-up (FU) dates of C (BU = YD198; FU = YD293) lie between the hourly simulation and the daily simulation from Chapter 3, within the BU and FU periods indicated by the Canadian Ice Service (CIS) charts (see Figure 3.2). Also note the NC simulation is not the same as the hourly simulation from Chapter 3 due to differences in forcing (i.e. NC was forced with incident radiation derived from C).

The differences between C and UPR were fairly small so only UPR is shown in the figures. For instance, salinity delayed complete snow ablation by 1 day and only 2 days for the ice cover in 1992. This resulted in a 1 day sooner freeze-up date for UPR compared to C, and the final ice thickness for UPR was only 4 cm thicker than C by the end of SIMMS'93. However, if snow would have fallen during the final 2 days of ice melt in UPR, this would have, 1) prolonged the ice cover existence, 2) induced earlier freeze-up and, 3) increased the final ice thickness by the end of SIMMS'93. It should be noted the results may have been slightly different if salinity were treated in the UPR mass balance since liquid water (brine) would be generated sooner in the spring period. Future work will investigate this.

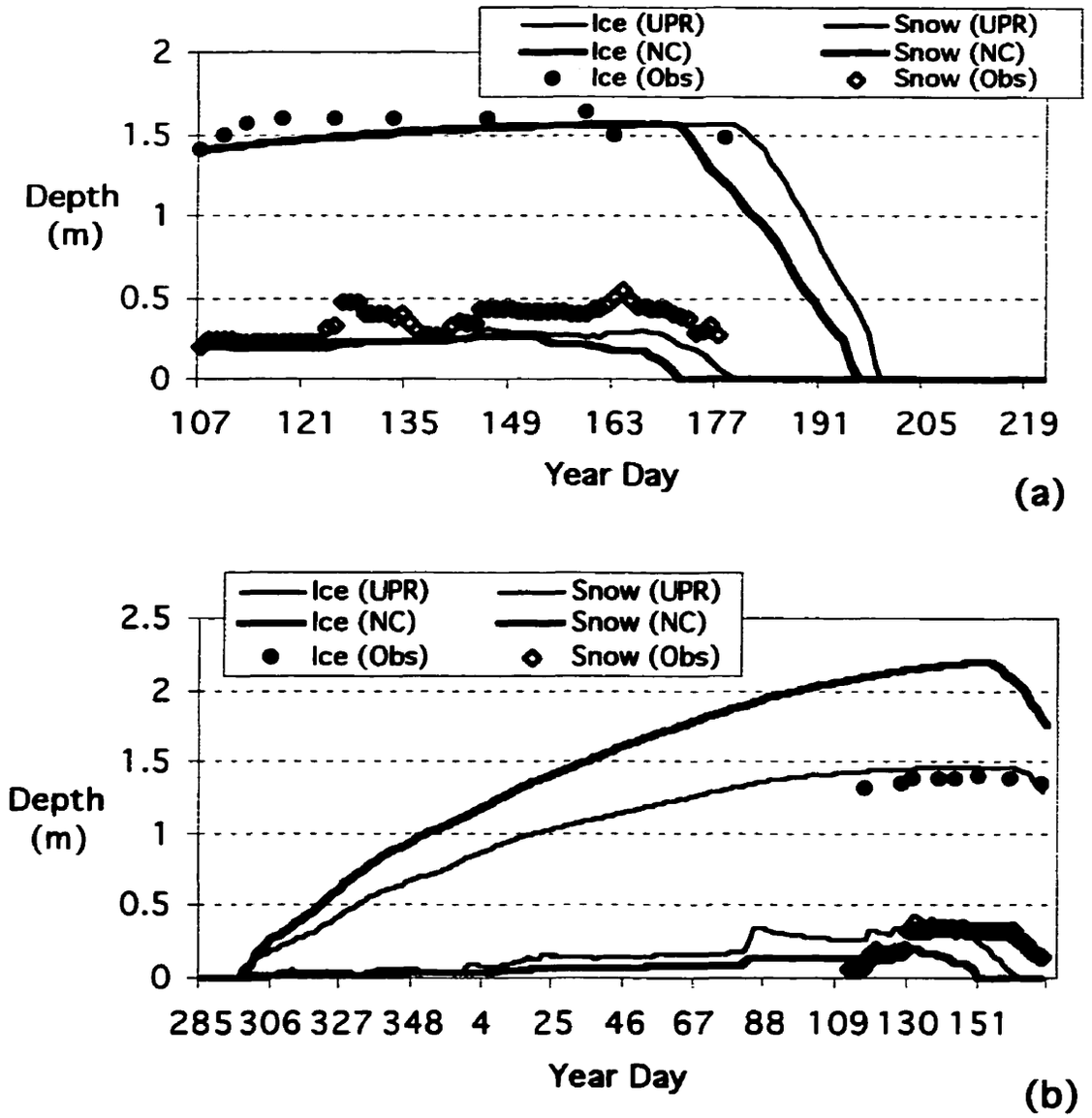



*Figure 6.2a, b: Observed (obs), UPR and NC snow and ice thickness evolution from a) the beginning of SIMMS'92 to break-up, and b) freeze-up to the end of SIMMS'93 using Resolute meteorological forcing from YD107, 1992 to YD170, 1993. Note that C and UPR were very similar, hence only UPR is shown.*

To illustrate a more detailed look at model differences on ice growth and seasonal snow and ice evolution, a freeze-up sensitivity scenario was conducted (discussed in section 6.3). In this case, maximum ice thicknesses were 1.4 m and 2.05 m for C and NC, respectively; a difference of 65 cm (Figure 6.3). Once again, snow thickness evolution

between C and NC was also significant due to blowing snow (Figure 6.3). If NC is forced to have the same snow thickness evolution as C, the maximum ice thickness of NC is reduced to 1.95 m (FS curve in Figure 6.3), making the difference between C and FS 55 cm. Adding salinity to the snow did not alter the freeze-up simulations appreciably, hence only the UPR curve is shown (Figure 6.3). Ice growth with salinity was slightly reduced (by a few cm) in the cold period but ice ablation in the spring period was also reduced effectively nullifying $k_s$ and $C_s$ salinity effects on the ice seasonal evolution differences. The maximum ice thickness difference between C and UPR was only 3 cm.

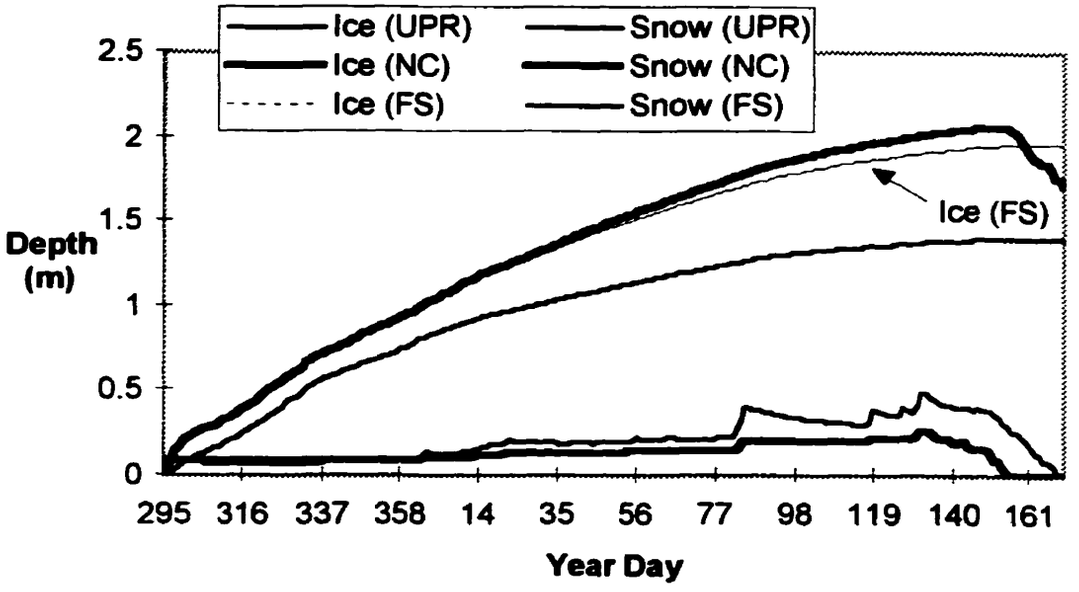



*Figure 6.3: UPR and NC snow and ice thickness evolution for the freeze-up sensitivity study between YD195, 1992 to the end of SIMMS'93. The "ice (FS)" curve is the ice thickness evolution for NC when it is forced to have the "snow (UPR)" curve. The "snow (FS)" and "snow (UPR)" are overlapping in the figure. Note that C and UPR are virtually the same for snow and ice.*

The above results clearly show the coupled snow sea-ice model produces thinner ice and more realistic snow and ice conditions using this data set. As expected, salinity

effects on $k_s$ and $C_s$ act to slightly reduce both ice growth rates in the cold season and ablation rates in the spring melt period. Once again, the results may have been slightly different if salinity were accounted for in the UPR mass balance. Predicting these effects over an entire annual ice cycle is difficult due to the liquid water (brine) influence on snow density and thermal properties. Further work is required in this regard.

### *6.4.2 Thermodynamic Evolution*

To examine the second question, NC, C and UPR were forced with *in situ* data over the SIMMS'92 duration (discussed in section 6.3). The snow thickness evolution of C (and UPR) follows the observed progression fairly well (Figure 6.4) for reasons discussed in section 6.4.1. The result of using *in situ* radiative and meteorological forcing instead of parameterized radiation and Resolute data is revealed when comparing Figures 6.2a and 6.4 and the ice conditions; i.e. snow and ice ablation are delayed and slower than those in section 6.4.1. Ice thickness changed from 1.4 m (initial value) to 1.53 m for C and 1.4 m to 1.58 m for NC over the 70 day simulation; observed values at the end of the experiment were 1.5 ± 0.08 m. There was no difference between C and UPR so only UPR is shown (Figure 6.4).

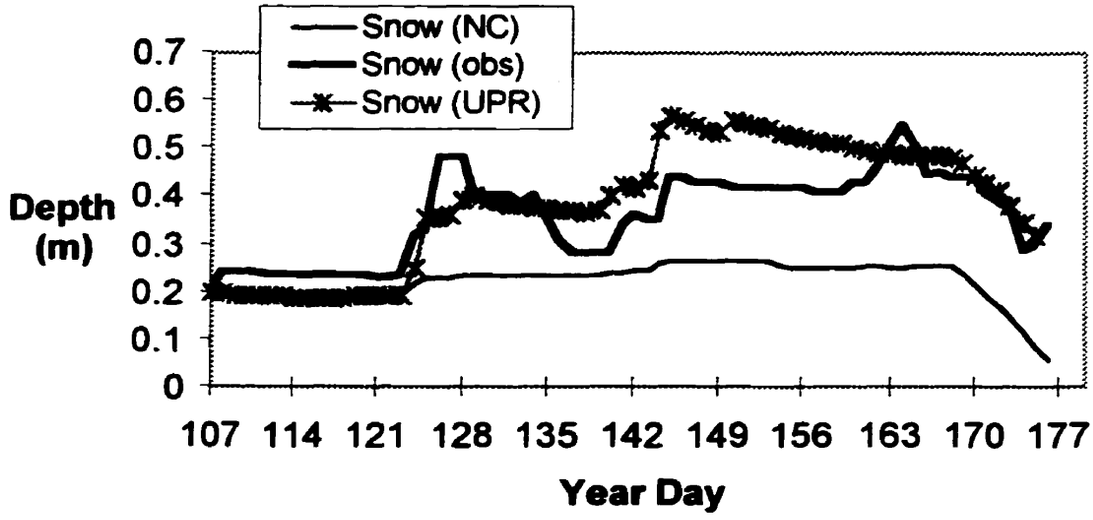



*Figure 6.4: The observed (obs), NC and UPR snow thickness evolution between YD107 and YD177, 1992 using SIMMS'92 in situ data as model forcing. The UPR snow evolution is different than "snow UPR" in Figure 6.2a due to greater wind speeds on the ice compared to Resolute.*

*Table 6.1 Non-coupled (NC) model, coupled model without salinity (C), and coupled model with salinity (UPR) $R^2$, mean bias error (MBE) and root-mean-squared error (RMSE) compared to observations of $Q_{net}$ (W m-2), $T_{sfc}$ (°C), and $T_{isfc}$ (°C). NC model data are adapted from Hanesiak et al. (1999). Note there was no difference between C and UPR for $Q_{net}$ and $T_{sfc}$. Observed mean values are shown for reference.*

| | $Q_{net}$(NC) | $Q_{net}$(C) | $T_{sfc}$(NC) | $T_{sfc}$(C) | $T_{isfc}$(NC) | $T_{isfc}$(C) | $T_{isfc}$(UPR) |
|---|---|---|---|---|---|---|---|
| $R^2$ | 0.56 | 0.74 | 0.98 | 0.99 | 0.98 | 0.99 | 0.99 |
| MBE | -2.00 | +3.00 | -0.10 | -0.06 | -0.58 | +0.95 | +0.98 |
| RMSE | 25.0 | 21.0 | 1.7 | 1.3 | 0.7 | 0.5 | 0.5 |
| Mean Obs | -1.4 | -1.4 | -9.8 | -9.8 | -7.2 | -7.2 | -7.2 |

In addition, the statistical comparison between $T_{sfc}$ and $Q_{net}$ for C and NC indicates that C reproduces $Q_{net}$ more accurately but little difference is apparent in $T_{sfc}$ (Table 6.1). It can be expected that differences in $T_{sfc}$ would be small between C and NC

due to thermal balancing between the snow surface and atmosphere. However, since C (and UPR) use very different turbulent exchange parameterizations than NC, one may expect to see larger differences in their respective $Q_{net}$. The C $T_{sfc}$ is very close to the observed $T_{sfc}$ evolution, even during the critical melt period (Figure 6.5). There was no difference in $T_{sfc}$ evolution between C and UPR.

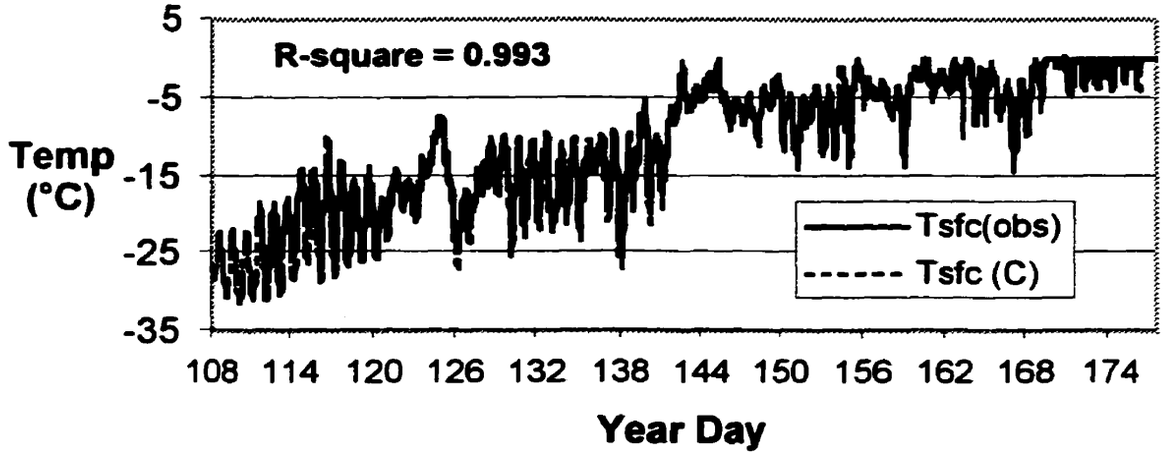
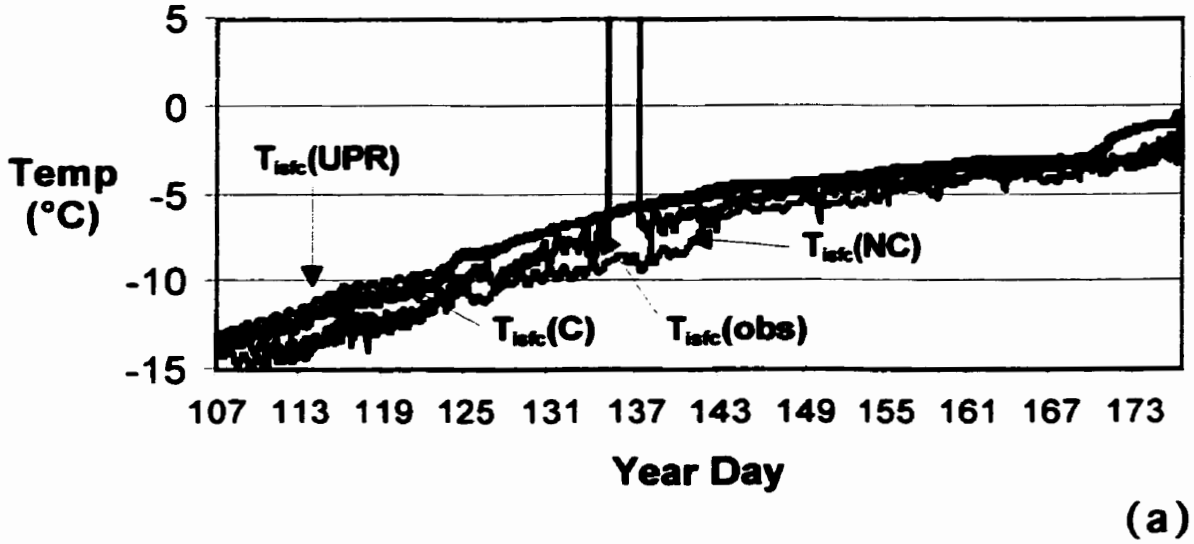



*Figure 6.5: The C snow surface temperature ($T_{sfc}$) versus observations (obs) between YD107 and YD177, 1992 using SIMMS'92 in situ data as model forcing. The $R^2$ is also shown. Note there was no difference between C and UPR.*

The simulated ice surface temperatures ($T_{isfc}$) in C are slightly too warm and NC are slightly too cold for the middle period of SIMMS'92 (Figure 6.6a and Table 6.1). The addition of salinity to $k_s$ and $C_s$ acts to increase $T_{isfc}$ in the colder season and decrease $T_{isfc}$ in the melt season, with more dramatic effects in the melt season (i.e. the decrease in snow thermal conductivity and increase the snow specific heat are more pronounced during melt). The warming effect of salinity in the colder season may or may not be counteracted if salt were parameterized in the UPR mass balance since the complicated

effects of liquid water (brine) and latent heat due to phase changes would be generated sooner in the spring period. The NC $T_{isfc}$ is closer to observations in the cold season compared to C. It will be shown (Chapter 7) that snow density during the cold season may be under estimated in C and UPR at several layers. This would artificially decrease snow thermal conductivity and effectively cause C and UPR $T_{isfc}$ to be too warm as shown here. Effects of salinity in the mass balance would also complicate the process, as explained above.

An interesting observation is the diurnal cycling (waves) of temperatures within UPR are better reproduced compared to C and NC (Figure 6.6b). Cross-correlation analysis suggests NC $T_{isfc}$ diurnal waves are about 5-6 hours out of phase with observations, whereas C was 2-3 hours out of phase during the colder season; the inclusion of salinity in the snow $k_s$ and $C_s$ further decreased this phase difference between 0-1 hour. This suggests NC does not damp out air temperature changes sufficiently; this is evident beyond YD124 where air temperatures significantly declined along with NC $T_{isfc}$, whereas the observations, C and UPR do not (see Figure 6.6b). The complex problem of snow thermal conductivity, specific heat, thermal diffusivity and salinity effects on mass balance are still largely under debate and very detailed. I will only state that more work is required in these areas.



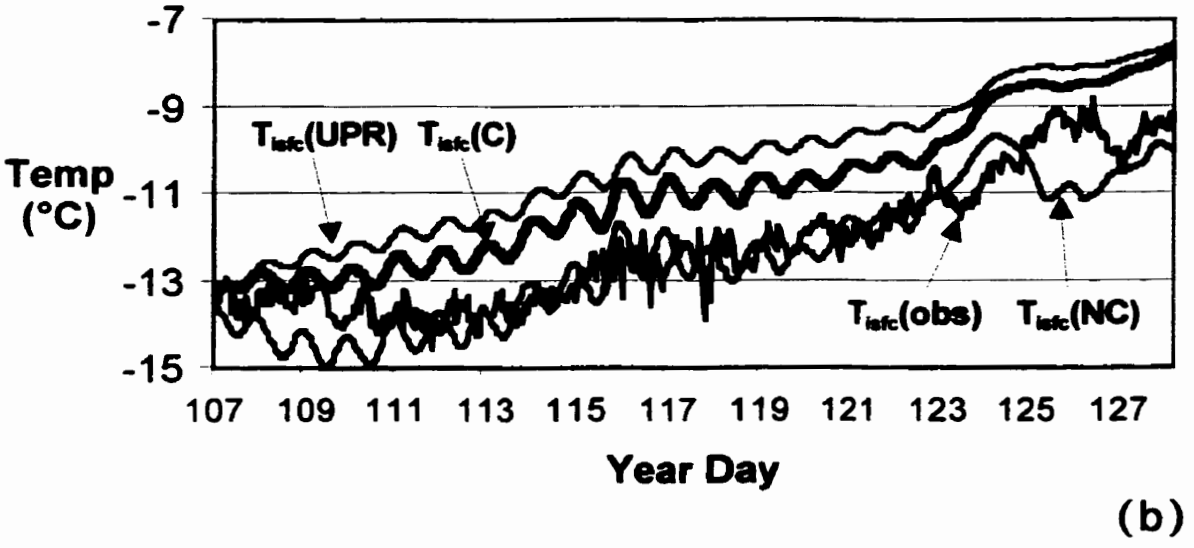




*Figure 6.6a,b: The NC, C, and UPR simulated ice surface temperature ($T_{isfc}$) versus observations (obs) between, a) YD107 to YD177 and b) YD107 to YD128. Observed data between YD135 - YD138 are missing.*

## 6.5. Conclusions

The goal of this chapter is to examine whether there are advantages of using a more sophisticated treatment of the snow layer for modeling the annual cycles of snow-covered FYI. This was accomplished by comparing model simulations of a 1-D coupled snow sea-ice model (C) and another 1-D sea ice model that uses a single bulk property snow layer (NC). The C model was developed in light of previous chapters outlining the importance of the snow layer and the surface energy balance on FYI annual cycles. The coupled model not only treats the snow layer in a much more rigorous fashion in terms of mass and thermal balance, but also includes a better handling of blowing snow processes, turbulent exchanges, and albedo when snow is present. I also illustrated the effects of salinity on the snow and ice physical and thermal evolution by including snow thermal conductivity ($k_s$) and specific heat ($C_s$) parameterizations that account for salinity in snow (given in Papakyriakou, 1999). Saline snow, which is common over FYI, can influence snow and sea ice thermal evolution and microwave remote sensing properties (seen in Chapter 7). Two specific research questions were designed to address the goal of this chapter.

### *6.5.1 Question 1*

In the first question, I investigated differences between NC and C in terms of the seasonal evolution of snow and ice thickness. The C snow and ice thickness evolution is better simulated compared to observations partially due to the coupled model's blowing

snow parameterization. Maximum ice thickness differences between C and NC can be between 55 - 65 cm from freeze-up to the early summer melt period, which is significant for annual FYI simulations. It is unknown whether these differences can be accounted for in NC since C and NC are very different in terms of snow layer treatment, blowing snow, turbulent exchange, albedo and incident radiation. It can be concluded that using the coupled model is advantageous for seasonal snow-covered sea ice, however, it is more computationally expensive and has a longer run time than NC. For example, it takes nearly one minute for NC to produce a simulation between YD107, 1992 and YD170, 1993, whereas C takes an order of magnitude longer (~10 minutes). It takes between 13-16 minutes for UPR. The addition of salinity in the snow $k_s$ and $C_s$ reduces ice growth rates in the cold season and ablation rates in the spring melt season. These thermal variations do not appear to be significant in terms of overall annual snow cover and FYI evolution. Results may be different if salinity were included in the snow mass balance.

### *6.5.1 Question 2*

In the second question, the thermodynamic comparisons between NC and C showed there is very little difference in simulated surface temperatures, however, C reproduces the net surface flux better due to its more sophisticated turbulent exchange parameterization. It will been shown the coupled model also does a reasonable job simulating other snow layer temperatures (see Chapter 7). Differences also arise at the ice surface where C temperatures are consistently slightly too warm and NC slightly too cold in the transitional season compared to observations presented here. This may be due to

several factors, some of which include errors in snow density, snow thermal properties and the neglect of salinity in the mass balance. The addition of salinity in the snow $k_s$ and $C_s$ for C further makes the ice surface temperatures too warm in the cold season but improves them during the melt season. This is due to the overall salinity effect that slightly warms the ice in winter and cools it during the melt period; the cooling effect is more significant during melt due to larger departures between snow thermal conductivity and specific heat when comparing snow with and without salinity. The net warming effect of salt in snow $k_s$ and $C_s$ may or may not be counteracted if salts were accounted for in the modeled mass balance.

## 6.6 Summary

In summary, the snow layer is very important for FYI evolution and thus modeling its processes effectively are also critical. The coupled model developed in this chapter simulates many of the critical aspects of a seasonally evolving snow-covered FYI volume. This type of model with its sophistication and application to FYI has previously been unavailable and one that will provide a better means to address future arctic sea ice climate process studies. In addition, the sophistication of the coupled model is required for application to microwave remote sensing in Chapter 7. The focus of Chapter 7 is to use the coupled snow sea-ice model to predict the snow and ice physical and thermal characteristics that give rise to their respective microwave dielectric properties.

# CHAPTER 7: Electro-Thermophysical Model of the Snow Sea-Ice System (ETSSIS) for Microwave Remote Sensing

## 7.1 Introduction

One-dimensional (1-D) thermodynamic models of snow and sea ice are valuable tools for examining polar climate processes over a variety of ice types and temporal scales (see for example, Ebert and Curry, 1993; Flato and Brown, 1996; Hanesiak et al., 1999; Jordan et al., 1999). In addition, satellite microwave remote sensing is also useful for inferring snow and sea ice physical and thermal properties (besides ice type and concentration) over a wide range of seasonal spatial and temporal scales (see for example, Livingstone and Drinkwater, 1991; Shuchman et al., 1996; Jefferies et al., 1997; Barber and Nghiem, 1999; Barber and Yackel, 1999). Microwave forward and inverse scattering models have also advanced to a state as supplementary tools for understanding the remote observations (see reviews by Golden et al., 1998a,b).

In this chapter I argue for the integration of a thermodynamic model and time series microwave remote sensing data in the development of an 'electro-thermophysical' relationship of snow covered sea ice. In this framework the intent is to develop a model which couples knowledge of the scattering physics of the surface with a 1-D thermodynamic model of the sea ice/snow system. The advantage of this approach is that the geophysical and electrical properties of the surface give rise to scattering and emission in the microwave portion of the spectrum; changes in the geophysical and or electrical state of the volume are driven by the thermodynamics across the interface. A model of these relationships would allow researchers to measure both the geophysical

state of sea ice at any particular time ($t_1$) and also evaluate the temporal evolution of this scattering over a period coinciding with the seasonal evolution of the surface ($t_n$). Since the surface energy balance (SEB) also evolves directly from the thermodynamic and physical state of the system, it is expected to be able to determine proxy measures of certain SEB state variables such as surface temperature, albedo, long-wave flux, etc. This information can also be useful in determining the timing of accretion and ablation and producing estimates of the mechanical strength of sea ice and/or timing of fast ice break-up.

Chapter 1 and 2 outlined issues that need to be addressed before fully exploiting thermodynamic modeling toward microwave scattering. First, typical 1-D sea ice models treat the snow layer as a bulk homogeneous slab with very few internal processes which limits their usefulness for microwave remote sensing applications. This is because the variables that contribute to microwave scattering are not simulated in the models such as ice surface roughness, snow liquid water content, basal snow grain sizes, and brine volumes in the bottom snow and top ice layers (Barber and Nghiem, 1999; Nghiem et al., 1998; Barber et al., 1995b)). Second, most detailed snow models do not consider salinity in snow which is a common occurrence in snow over FYI (Barber et al. 1995a). This limits their applicability to microwave remote sensing since salinity plays a large role in snow thermodynamics (Papakyriakou, 1999) and dielectric properties (see for example Barber et al., 1995b; Barber and Nghiem, 1999). The coupled snow sea-ice model in Chapter 6 was specifically developed for application to microwave remote sensing, although it has its use for arctic sea ice climate processes shown in Chapter 6.

The intent of this chapter is to begin the initial steps of linking thermodynamic modeling with microwave remote sensing in a forward sense, that is, determine microwave dielectrics and scattering given modeled estimates of the thermophysical variables that control dielectrics through the electro-thermophysical relationship. I develop and test a working numerical model that enables one to interpret satellite microwave remote sensing (active or passive) in a more physical sense. This will lead to better ways of monitoring the thermophysical properties of the snow/sea ice system within the context of climate variability and sea ice related processes.

Output from the coupled snow sea-ice model in Chapter 6 provides the necessary input to a microwave dielectric model of the Debye form (Barber et al., 1995b); the amalgamation of the models represents an Electro-Thermophysical model of the Snow Sea-Ice System that I will refer to as ETSSIS. The research questions to be addressed in this chapter are designed to utilize *in situ* observations to validate the model's ability to simulate the critical variables that control microwave scattering and emission.

1) Is ETSSIS capable of reproducing snow and ice thermodynamic characteristics of a seasonally variable FYI snow/sea ice system (both with and without salinity in the snow)?
2) Is ETSSIS capable of reproducing the physical characteristics of a seasonally variable FY snow/sea ice system?
3) Can ETSSIS be used toward the interpretation of the microwave time series scattering over a FYI snow/sea ice system?

It is not the intent to fully validate the snow sea-ice model since this has been done by others in the non-coupled forms (see Jordan et al., 1999; Flato and Brown, 1996; Hanesiak et al., 1999). However, it is necessary to illustrate the model's ability to simulate the primary variables that control microwave scattering and emission using *in situ* data utilized here.

## 7.2 Field Data and Numerical Model

### *7.2.1 Observational Data*

Field observations illustrated in previous chapters from SIMMS'92 are utilized here as well. SIMMS'92 had a variable snow depth up to a maximum near 50 cm and operated between Year Day (YD) 107 (April 18) to YD177 (June 27).

Measurements include SEB parameters for incident and reflected short-wave radiation ($K\downarrow$, $K\uparrow$), incident and emitted long-wave radiation ($L\downarrow$, $L\uparrow$), and net surface heat flux ($Q_{net}$), vertical temperature profiles within the snow/ice volumes, snow and ice thickness, and physical sampling of the snow and ice (snow density, salinity, wetness, grain diameter). $Q_{net}$ is defined here to include the net radiation and turbulent fluxes at the surface. Instantaneous hourly data were used for model forcing and snow physical sampling were measured periodically.

Some noteworthy errors associated with the measurements include the following. First, snow density measurement accuracy is between 30-60 kg $m^{-3}$ (Barber et al., 1995b). Second, snow wetness (i.e. water volume fractions) may not represent real conditions due to instrument limitations when snow salinity is roughly >3 ppt (parts per thousand);

measurement accuracy is roughly 0.1% water by volume. Third, snow salinity is considered to be both precise and accurate to within ±0.5 ppt (Barber et al., 1995b).

### *7.2.2 Coupled 1-D Snow Sea-Ice Model*

ETSSIS was forced with SIMMS'92 air temperature, relative humidity, wind speed, incident short-wave and long-wave radiation along with measured surface albedo between YD107 (April 18) to YD177 (June 27). Initialization was done using measured snow and ice physical profiles (densities, grain sizes, layer thickness) and temperature profiles. Model runs were conducted using the non-saline version of ETSSIS (C), the average salinity profile (AVG) from chapter 6, and the upper salinity profile (UPR) in chapter 6.

### *7.2.3 Snow and Sea Ice Dielectric Modeling*

Methods of modeling the microwave dielectric constants and penetration depths of snow and sea ice have been discussed in various studies (see for example, Barber, 1993; Drinkwater, 1989; Drinkwater and Crocker, 1988; Barber et al., 1995b), hence only a brief discussion is given here.

To model the dielectric constant of the snow and or sea ice one must separate the material into its host and inclusion components. In the lower portion of the snow pack there is a considerable amount of salinity in the form of brine. This brine is treated as an 'inclusion dielectric' within a dry snow 'host dielectric', following the approach of

Matzler et al. 1984b and Drinkwater and Crocker, 1988. The mixture model is expressed as (7.1).

$$\Delta\varepsilon^{*}_{mix} = \chi V_b \left\{ \frac{\varepsilon^{*}_b - \varepsilon^{*}_{ds}}{1 + [\frac{\varepsilon^{*}_b}{\varepsilon^{*}_{ds}} - 1]A_0} \right\} \tag{7.1}$$

Where $\varepsilon^{*}_{ds}$ and $\varepsilon^{*}_b$ are expressed in complex terms and represent the dielectric constants of dry snow and brine, $\chi$, $A_0$, and $V_b$ are: a coupling factor representing the fraction of brine accounted for by $A_0$; the dominant depolarization factor; and the volume of brine contained within the saline snow layer; $\chi$ is set at 0.66667, a value appropriate for modeling the brine inclusions as isotropically oriented oblate spheroids (Drinkwater and Crocker, 1988). The depolarization factor is set at 0.053, following Denoth, 1980, and the brine volume is computed as in Barber (1993).

In (7.1) the host dielectric is considered as a dry snow matrix, consisting of ice grains and air. The parameter $\varepsilon^{*}_{ds}$ is computed using an empirical model attributable to Hallikainen et al. (1986). The permittivity of dry snow was computed as (7.2) and the loss as (7.3).

$$\varepsilon'_{ds} = 1.0 + 1.9\rho_{ds} \tag{7.2}$$

$$\varepsilon''_{ds} = \frac{0.34 V_i \varepsilon''_i}{(1 - 0.417 V_i)^2} \tag{7.3}$$

Where $\rho_{ds}$, $\varepsilon_i''$, and $V_i$, are the density of dry snow, the dielectric loss, and the volume fraction of freshwater ice in the snow matrix.

In (7.1) the inclusion dielectric is considered to consist of brine pockets held within the interstices of the snow grains. The parameter $\varepsilon_b^*$ is defined as the complex dielectric constant of brine computed from (7.4 and 7.5) based on models of the Debye form, following Ulaby et al. (1986).

$$\varepsilon_b' = \varepsilon_{w\infty} + \left[\frac{\varepsilon_{b0} - \varepsilon_{w\infty}}{1 + (2\pi f \tau_b)^2}\right] \tag{7.4}$$

$$\varepsilon_b'' = (2\pi f \tau_b) \cdot \left[\frac{(\varepsilon_{b0} - \varepsilon_{w\infty})}{1 + (2\pi f \tau_b)^2}\right] + \left[\frac{\sigma_b}{2\pi f \varepsilon_0}\right] \tag{7.5}$$

Where $\varepsilon_{w\infty}$; $\varepsilon_{b0}$; $f$; $\tau_b$; $\sigma_b$; $\varepsilon_0$ are: the high frequency limit of the dielectric constant of pure water; the static dielectric constant of pure water; the frequency of electromagnetic energy (Hz); the relaxation time of the brine; the ionic conductivity of the brine solution ($S \cdot m^{-1}$); and the permittivity of free space ($F \cdot m^{-1}$). These parameters are estimated using a series of polynomial models which will not be reproduced here (see Stogryn, 1971; Ulaby et al. 1986; and Hallikainen and Winebrenner, 1992).

When free water becomes available within the snow pack the dielectric properties change considerably. The relationships between $\varepsilon'$ and $\varepsilon''$ for wet and dry snow have been determined empirically. Models for computation of $\varepsilon'$ and $\varepsilon''$ of wet snow relative to the values for dry snow, have been developed by Tiuri et al., (1984) and are presented in (7.6 to 7.8).

$$\Delta\varepsilon'_{ws} = \varepsilon'_{ws} - \varepsilon'_{dry} \tag{7.6}$$

$$\Delta\varepsilon'_{ws} = \varepsilon'_{w}(0.1W_{v} + 0.8W_{v}^{2}) \tag{7.7}$$

$$\varepsilon''_{ws} = \varepsilon''_{w}(0.1W_{v} + 0.8W_{v}^{2}) \tag{7.8}$$

Where: $\varepsilon'_{dry}$ is the permittivity of dry snow from (7.2); Wv is the volumetric liquid water content measured *in situ*; and $\varepsilon'_{ws}$ and $\varepsilon''_{ws}$ are the permittivity and loss of wet snow. The complex dielectric constant of water ($\varepsilon'_{w}$ and $\varepsilon''_{w}$) are computed using the Debye model for liquid water (7.9) and (7.10).

$$\varepsilon'_{w} = \varepsilon_{w\infty} + \left[\frac{\varepsilon_{w0} - \varepsilon_{w\infty}}{1 + (2\pi f \tau_{w})^{2}}\right] \tag{7.9}$$

$$\varepsilon''_{w} = (2\pi f \tau_{w}) \cdot \left[\frac{(\varepsilon_{w0} - \varepsilon_{w\infty})}{1 + (2\pi f \tau_{w})^{2}}\right] \tag{7.10}$$

Where the parameters are the same as in (7.4) and (7.5) but for liquid water rather than brine.

Through use of these dielectric mixture models it is possible to compute the penetration depth (δp) of various frequencies of electromagnetic energy into a seasonally dynamic snow covered sea ice volume (7.11), following Drinkwater (1989).

$$\delta_p = \frac{\lambda}{4\pi}\left\{\left[\left(1+\left(\frac{\varepsilon''}{\varepsilon'}\right)^2\right)^{\frac{1}{2}} - 1\right]\frac{\varepsilon'}{2}\right\}^{\frac{-1}{2}} \quad (7.11)$$

Where $\lambda$ is the SAR wavelength in meters, $\varepsilon'$ and $\varepsilon''$ are the dielectric permittivity and loss given a particular water volume within the snow pack. I have computed the maximum penetration depths for selected snow and ice layers for a 5.3 GHz frequency (typical of active microwave sensors) for illustrative purposes. A similar analysis can be performed for other microwave frequencies such as passive sensors (for example 19 GHz).

## 7.3 Analysis Methods

To quantify ETSSIS's ability to simulate the thermodynamic, physical and electrical properties of the snow and ice volumes, the coefficient of determination ($R^2$), mean bias error (MBE), and root-mean-squared error (RMSE)) were computed from modeled and observed values. The statistics were calculated for the non-saline model case (C), average saline case (AVG) and upper limit saline case (UPR). Thermodynamic variables included temperatures at various levels in the snow and ice surface, including the snow surface ($T_{sfc}$) and ice surface ($T_{isfc}$), net surface heat flux ($Q_{net}$), and snow and ice thickness. $Q_{net}$ and snow and ice thickness are included since they impact, or are the direct result of the thermodynamics. Physical variables included snow density ($\rho_s$), grain diameter (d), liquid water fraction (Wv), salinity (S), and brine volume (Vb). Although salinity is held constant in the model, it is useful to illustrate differences between

modeled and observed values for error analysis. The microwave dielectric variables included permitivity ($\varepsilon'$), loss ($\varepsilon''$) and penetration depth ($\delta p$) at various levels in the snow and ice. The $\varepsilon'$ and $\varepsilon''$ were estimated by the dielectric mixture model in section 7.2.3 using snow and ice T, $\rho$, Wv, and S as input. "Observed" $\varepsilon'$ and $\varepsilon''$ are defined as values computed by the dielectric model using observed thermodynamic and physical variables as input. Microwave maximum penetration depths were computed from $\varepsilon'$ and $\varepsilon''$ at 5.3 GHz to define which snow layers become important for microwave remote sensing interpretation over the various seasonal stages of snow-covered FYI. In particular, we are interested in the dry snow (cold season) scenario, when water in liquid phase appears in snow (the pendular regime with Wv between 1-7% water by volume), and full snow melt (funicular regime with Wv between 8-15% water by volume). These seasons were defined according to *observed* snow Wv measurements (see section 7.5).

The $R^2$ was only computed for model and observed data in continuous hourly time increments since they were more consistent in time (i.e. $Q_{net}$, $T_{sfc}$, and $T_{isfc}$). The MBE and RMSE were computed for other variables with periodic daily samples. The thermodynamic and physical variables used in the analysis were selected based on their importance to microwave dielectrics.

## 7.4Results

### *7.4.1 Thermodynamic Evolution*

This section gives an overview of how well ETSSIS models the important thermodynamic variables for microwave remote sensing (question 1). A more detailed look at which model errors may or may not be important will be discussed in section 7.5.

The modeled temporal evolution of the snow thermal conductivity and specific heat over SIMMS'92 for various snow layers shows a progressive spread between the non-saline model run (C) and saline model runs (AVG and UPR) (Figure 7.1). This is mainly due to the increase in snow temperature over time in conjunction with salinity effects. The maximum difference in thermal conductivity is between 0.1 - 0.15 W $m^{-1}$ $K^{-1}$ (15-25%) and an order of magnitude (> 1.2 x $10^4$ J $kg^{-1}$ $C^{-1}$) for specific heat before the onset of significant snow melt (prior to YD170). Beyond YD170, significant separation between saline and non-saline values occurs because of the extreme nature of $k_{ss}$ and $C_{ss}$ during melt.

The model's snow and ice thickness evolution are fairly well simulated compared to observations (see Figure 6.4); the snow thickness evolution is mainly due to ETSSIS's blowing snow parameterization. The improved snow thickness distribution should allow for more realistic snow and ice thermodynamic evolution. Ice thickness changed from 1.4 m (initial value) to 1.53 m for the C simulation and 1.4 m to 1.51 m for the UPR simulation over the 70 day period; observed values at the end of the experiment were 1.5 ± 0.07 m. Qualitatively, the model simulates the snow and top ice layer temperatures

fairly well (Figure 7.2). Statistically, the comparisons between observed and modeled $T_{sfc}$ and $Q_{net}$ indicates the model reproduces these parameters very well (Table 7.1).

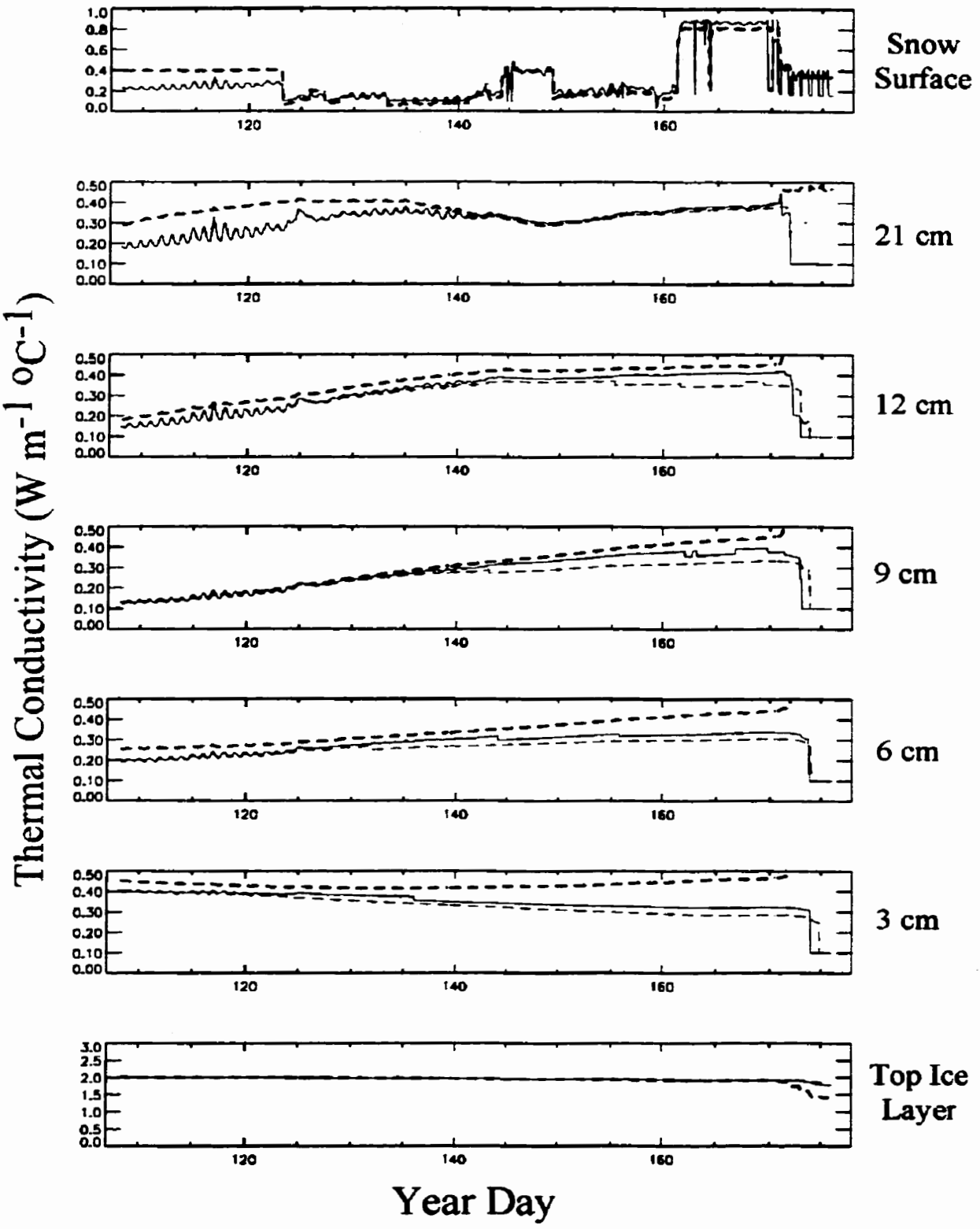



*Figure 7.1a: Non-saline (C – heavy dashed line) and saline (AVG – solid; UPR – dashed line) ETSSIS thermal conductivity (W $m^{-1}$ $K^{-1}$) for various snow and top ice layers.*

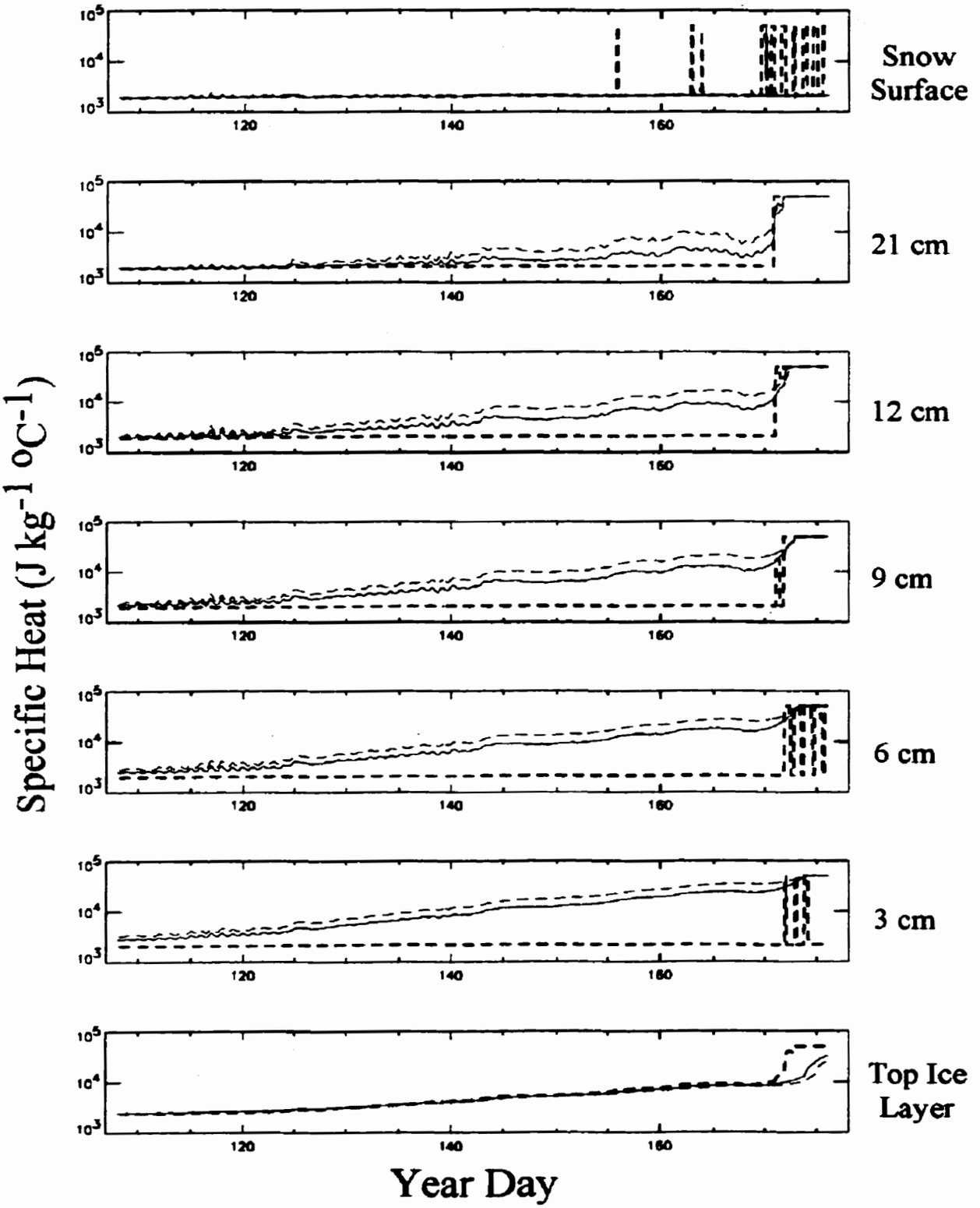



*Figure 7.1b: Non-saline (C – heavy dashed line) and saline (AVG – solid; UPR – dashed line) ETSSIS specific heat (J kg$^{-1}$ C$^{-1}$) for various snow and top ice layers.*

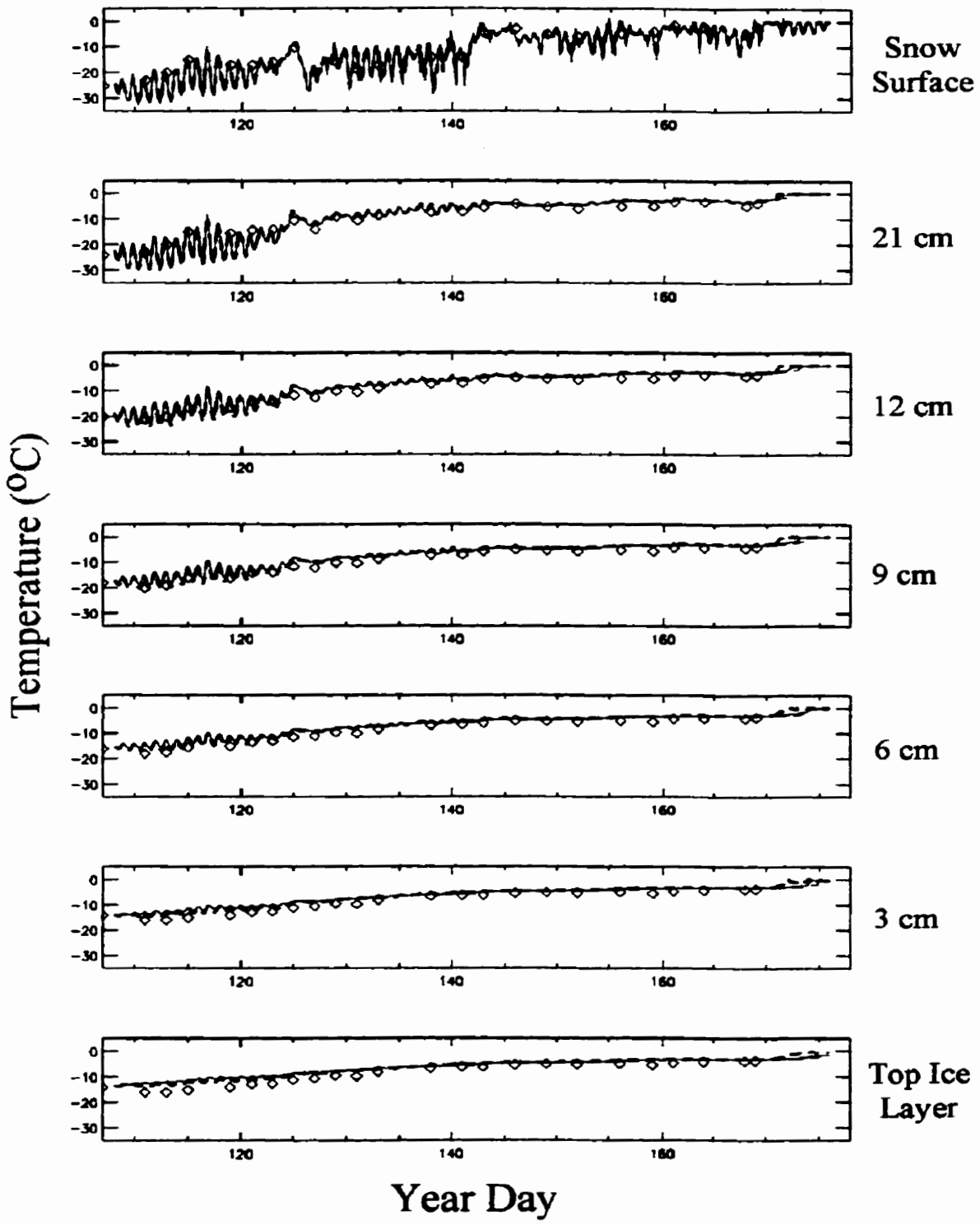



*Figure 7.2: Modeled and observed temperature (°C) evolution between YD107 and YD177 of various snow layers for C (heavy dashed line), AVG (solid line), UPR (not noticeable beneath solid line), and observed data (triangles). Observed values represent instantaneous morning temperatures corresponding to days where snow sampling took place.*

A detailed look at the modeled $T_{sfc}$ shows a close resemblance to the observed $T_{sfc}$ evolution, even during the critical melt period (see Figure 6.5). The melt period is important to simulate well not only for microwave dielectrics and penetration depths but for climate process studies. There was no change in the $T_{sfc}$ and $Q_{net}$ results when comparing the C, AVG and UPR simulations due to very little or no salinity in the top of the snow pack.

*Table 7.1: ETSSIS without salinity (C), average salinity (AVG) and upper limit salinity (UPR) $R^2$, mean bias error (MBE) and root-mean-squared error (RMSE) compared to observations of $Q_{net}$ (W $m^{-2}$), $T_{sfc}$ (°C) and $T_{isfc}$ (°C). Positive MBE implies the model overestimates observations. Observed mean values are also included for refence.*

| | $Q_{net}$ (C) | $T_{sfc}$ (C) | $T_{isfc}$ (C) | $T_{isfc}$ (AVG) | $T_{isfc}$ (UPR) |
|---|---|---|---|---|---|
| $R^2$ | 0.74 | 0.99 | 0.99 | 0.99 | 0.99 |
| MBE | +3.00 | -0.06 | +0.96 | +0.98 | +1.04 |
| RMSE | 21.0 | 1.3 | 0.49 | 0.74 | 0.79 |
| Mean Obs | -1.4 | -9.8 | -7.2 | -7.2 | -7.2 |

The simulated top ice layer temperatures ($T_{isfc}$) are slightly too warm (by a maximum of 2°C) in all model runs but improve during the warm season when salinity is included (see Figure 6.6a). The over prediction in the cold and mid-spring seasons may be the result of middle snow layer densities being under estimated and/or the neglect of brine drainage (see section 7.4.2). That is, higher snow densities and liquid water increase the thermal conductivity to promote cooling. Overall thermal effects of including salinity in snow $k_s$ and $C_s$ are to raise temperatures in the colder season (up to YD130) and reduce them during the warmer season (beyond YD130) (Figure 7.3). This effect would be complicated if the physical presence of salt was parameterized with regard to snow

mass balance since liquid water (brine) would be generated much sooner in the spring season (see section 7.4.2).

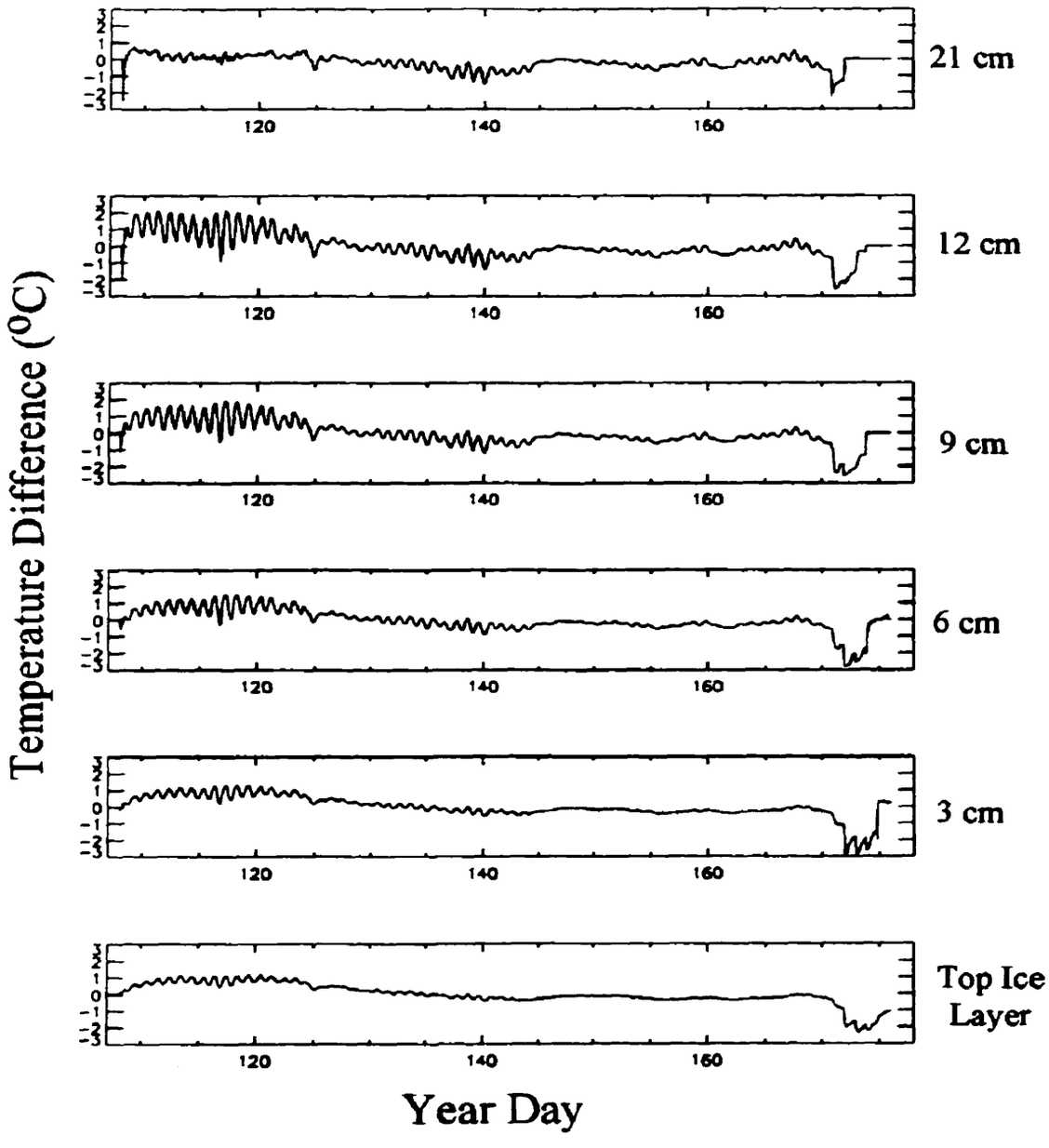



*Figure 7.3: Temperature difference (°C) between non-saline (C) and saline (UPR) ETSSIS simulations for various snow layers and top ice layer between YD107 to YD177. Positive values indicate warmer UPR simulated temperatures.*

An interesting thermal feature is that diurnal cycling (waves) of $T_{isfc}$ within the model are better reproduced when the effect of salt is considered in the snow $k_s$ and $C_s$ (see Figure 6.6b). Diurnal temperature cycling can be important for cold season microwave dielectrics that affect the appearance of time-series diurnal SAR image pairs and their physical interpretation (Barber and Nghiem, 1999; Nghiem et al. 1998). Cross-correlation analysis suggests the C simulation $T_{isfc}$ diurnal waves are at least 3 hours out of phase with observed values during the colder season but are reduced to as little as 0 hours in UPR. This suggests the C simulation snow thermal diffusivity may be too large by not damping out air temperature changes sufficiently or other factors related to $k_s$ and $C_s$. The complex problem of snow $k_s$, $C_s$, and thermal diffusivity are still largely under debate and too detailed to be included here, hence I will only state that more work is required in this regard.

Given these results and various small thermodynamic errors in the model and observations, results in section 7.5 will show whether these errors significantly impact microwave dielectrics.

### *7.4.2 Physical Evolution*

This section gives an overview of how well ETSSIS models the important physical variables for microwave remote sensing (question 2). A more detailed look at which model errors may or may not be important is presented in section 7.5.

The difference between C, AVG and UPR are very small (<5%) in terms of bulk density and grain sizes until water in liquid phase appears (Figures 7.4 to 7.6). The differences in density and grain sizes amongst the model simulations becomes 30-75% due to the different timing of the first appearance of water in liquid phase. The initial appearance of water in liquid phase is delayed with higher snow salinities which cools overall snow temperatures in the warm season; C snow liquid water appears 3 days earlier than AVG (YD170 compared to YD173) and 4 days earlier than UPR (YD170 compared to YD174). These differences would be greater if salt were explicitly parameterized in the mass balance.

Time series modeled versus observed $\rho_s$, d and Wv indicates that $\rho_s$ is simulated fairly well but slightly under estimated at 21 cm and the snow surface where larger discrepancies arise from fresh snow falls (see Figure 7.4 and Table 7.2). Modeled $\rho_s$ prior to YD120 at 9 and 12 cm is also under estimated. This may contribute to the warmer ice surface temperatures in the cold season by artificially reducing the snow thermal conductivity. Also note that field sampling of $\rho_s$ may introduce errors between modeled and observed values. Grain sizes are generally simulated reasonably well over most of the period except near the bottom part of the snow pack during the melt season where observed liquid water contents increase (see Figures 7.5 and 7.6 and Table 7.2). Grain growth can accelerate rapidly with available liquid water. The non-existence of simulated liquid water prior to YD170 inhibits modeled grain growth and increases discrepancies between modeled and observed values. Note that snow grains do not affect microwave dielectrics but become physically important for the total volume scattering contributions.

Grain sizes are particularly important in development of algorithms for estimating snow water equivalent (SWE) on sea ice (Markus and Cavalieri, 1998).

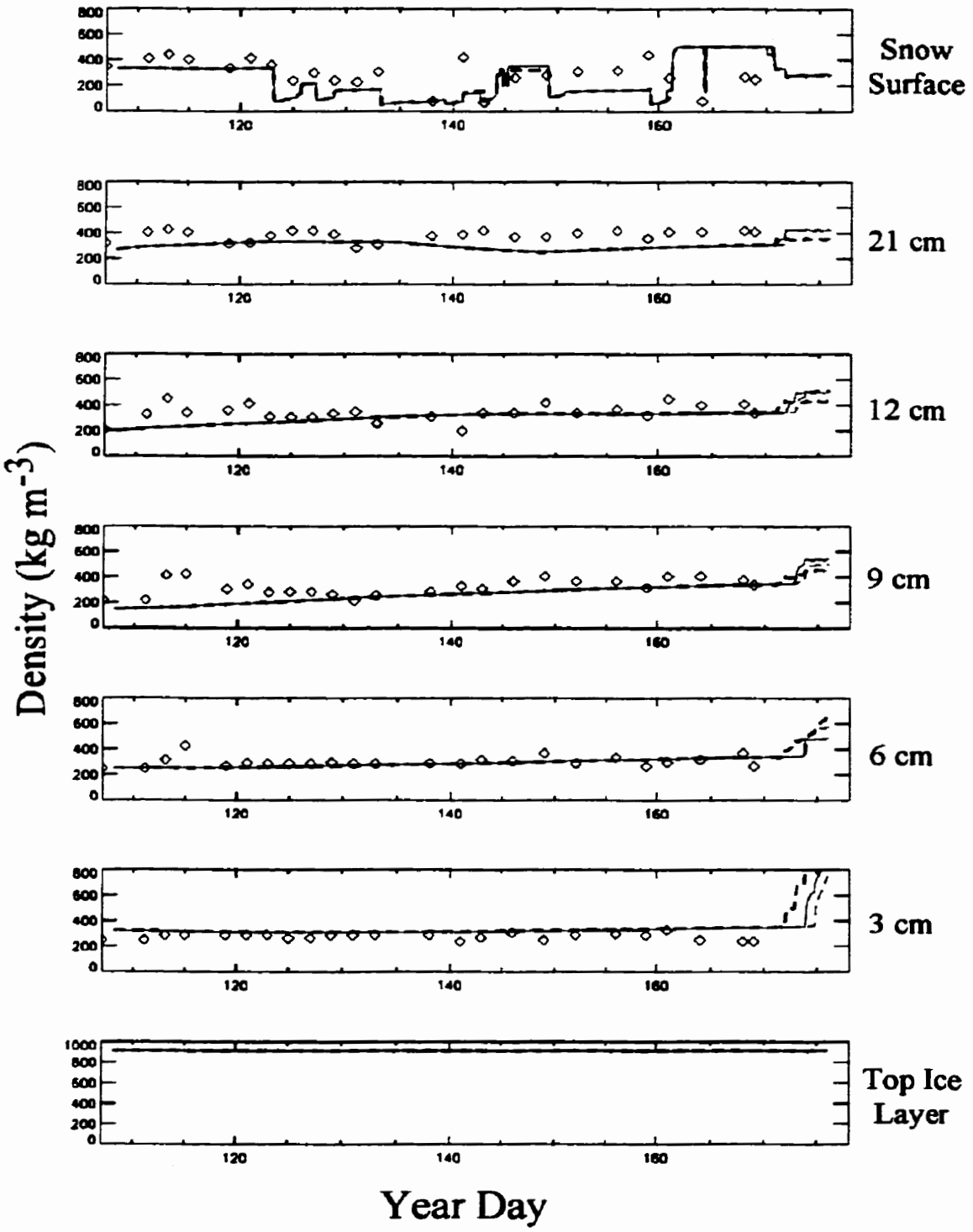



*Figure 7.4: Modeled and observed density (kg $m^{-3}$) evolution between YD107 and YD177 of various snow layers for C (heavy dashed line), AVG (solid line), UPR (dashed line), and observed data (triangles). Observed values represent averages as in Barber et al. (1995b) corresponding to days where snow sampling took place.*

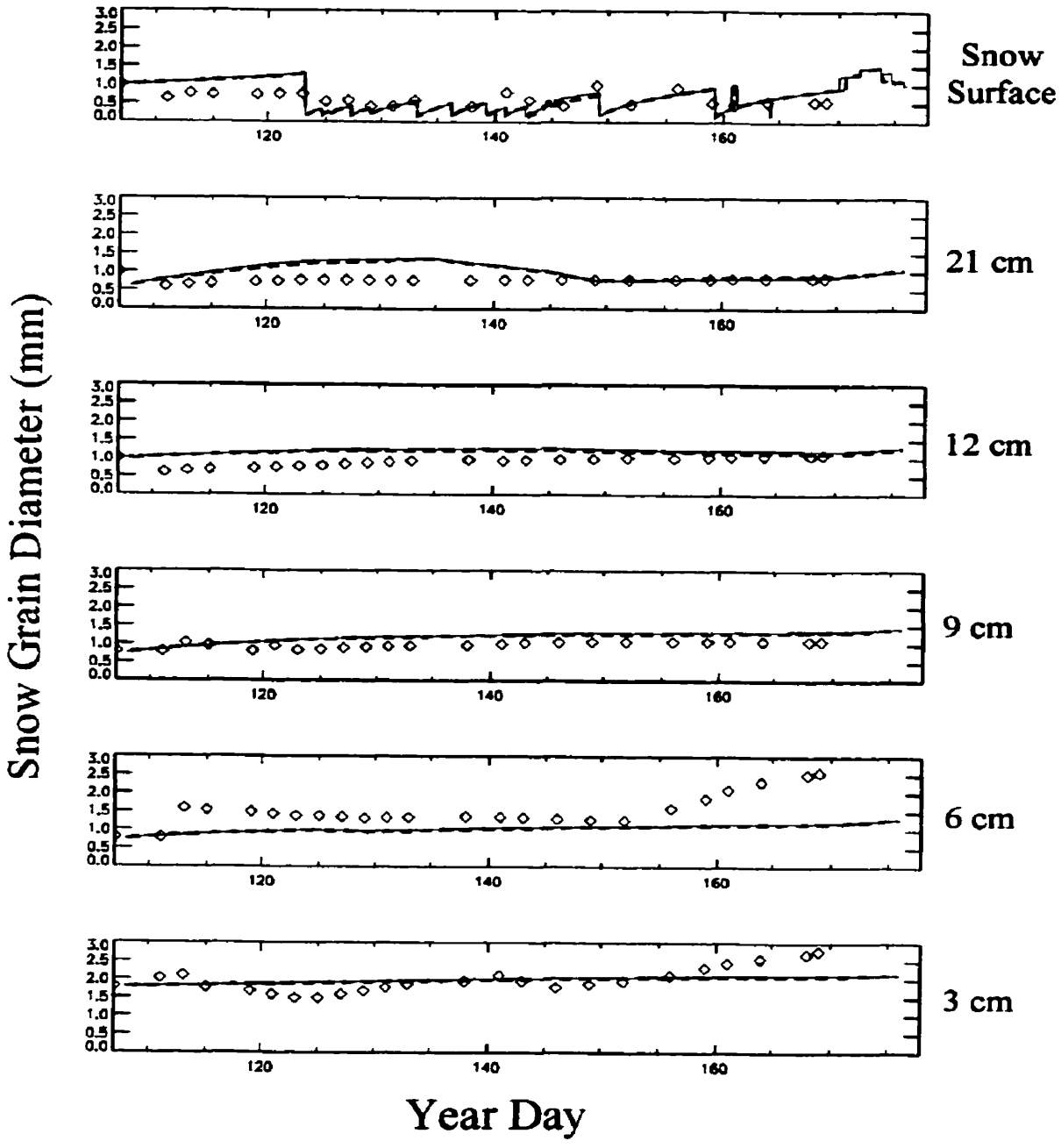



*Figure 7.5: Modeled and observed grain size (mm in diameter) evolution between YD107 and YD177 of various snow layers for C (heavy dashed line), AVG (solid line), UPR (not noticeable beneath solid line), and observed data (triangles). Observed values represent averages as in Barber et al. (1995b) corresponding to days where snow sampling took place.*

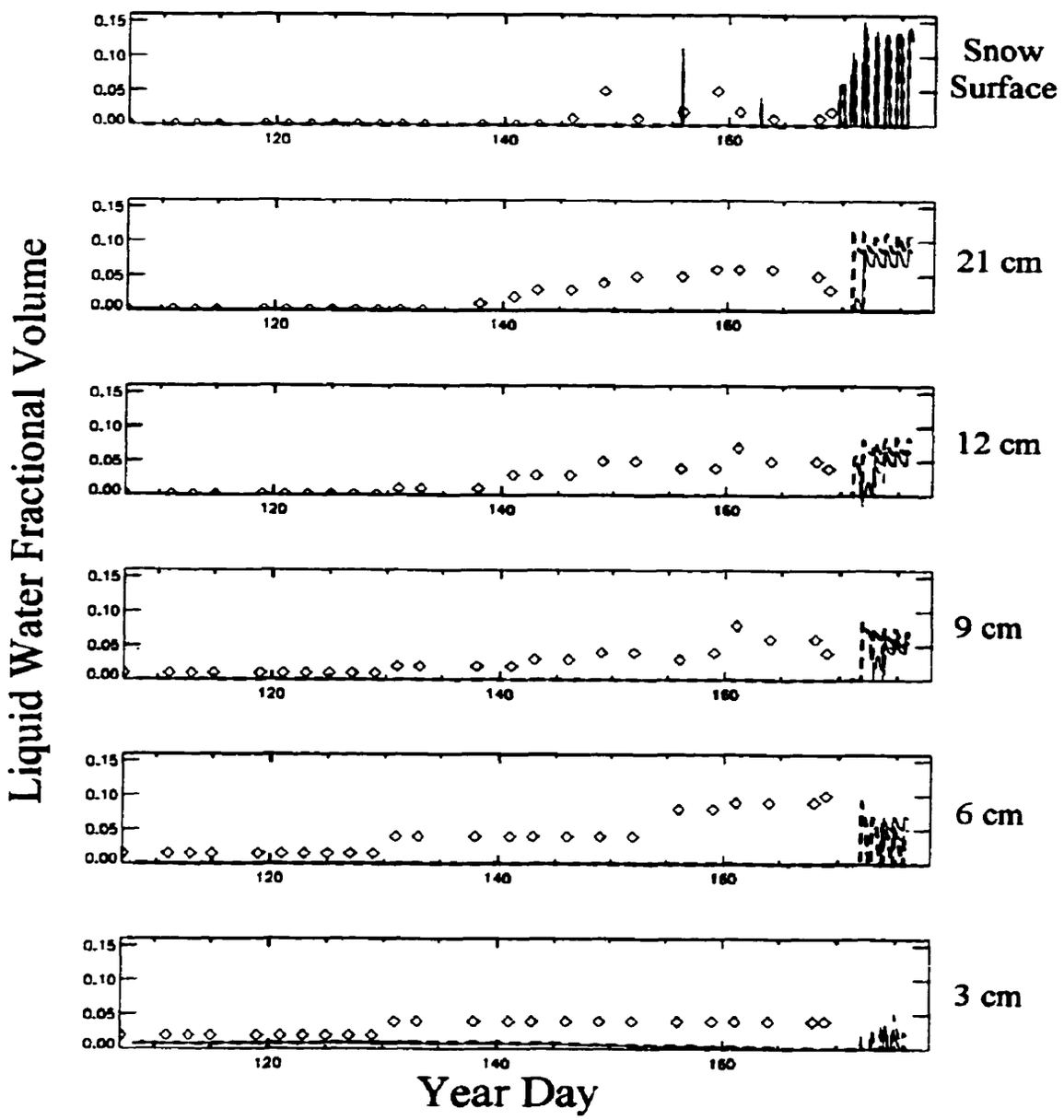



*Figure 7.6: Modeled and observed fractional water volume evolution between YD107 and YD177 of various snow layers for C (heavy dashed line), AVG (solid line), UPR (dashed line), and observed data (triangles). Observed values represent averages as in Barber et al. (1995b) corresponding to days where snow sampling took place. Observed water volumes with significant salinity (>3 ppt) may not represent actual conditions due to instrument limitations.*

The modeling of Wv clearly needs improvement in most snow layers (Figure 7.6 and Table 7.2) which impacts the dielectric properties (shown later). This is due to not explicitly including the presence of salt in the snow mass balance of the model; future work will address this issue. Salinities are also not simulated very well since the model does not account for snow pack desalination over time as liquid water drains freely toward the ice surface in the late season (Figure 7.7). The model holds salinity constant over time since a physical parameterization for snow desalinization is currently not available. Note here that errors in sampling, especially $\rho_s$ and Wv can have an impact on the analysis presented.

*Table 7.2: ETSSIS without salinity (C) mean bias error (MBE) and root-mean-squared error (RMSE) snow density, grain sizes, water volume fraction for various snow depths compared to observations. The first (second) value in the salinity column represents the AVG (UPR) simulation. Positive MBE implies the model overestimates observations. Observed values (in brackets) represent averages as in Barber et al. (1995b) corresponding to days where geophysical snow sampling took place. Instantaneous morning model data were used for comparison to observations. Water volume fraction observations may not represent real conditions when salinities are >3 ppt.*

| | Density (kg m$^{-3}$) | | Grain Size (mm) | | Water Vol. Frac. | | Salinity (ppt) | |
|---|---|---|---|---|---|---|---|---|
| | MBE | RMSE | MBE | RMSE | MBE | RMSE | MBE | RMSE |
| Surface | -14 (310) | 107 | +0.1 (0.65) | 0.3 | +0.01 (0.02) | 0.02 | -0.02/-0.01 (0.1) | 0.02/0.01 |
| 21 cm | -97 (395) | 41 | -0.3 (0.73) | 0.2 | +0.03 (0.035) | 0.04 | +0.5/+0.8 (1.1) | 1.0/1.2 |
| 12 cm | -15 (340) | 38 | +0.2 (0.75) | 0.3 | +0.02 (0.037) | 0.04 | +1.2/+3.1 (1.9) | 2.5/4.9 |
| 9 cm | -25 (355) | 45 | +0.2 (0.8) | 0.3 | +0.03 (0.042) | 0.04 | +1.1/+4.2 (4.7) | 4.2/6.5 |
| 6 cm | -8 (300) | 17 | -1.1 (1.6) | 1.8 | +0.05 (0.051) | 0.07 | +1.5/+5.2 (8.1) | 4.4/5.8 |
| 3 cm | +13 (280) | 20 | -0.1 (1.9) | 1.2 | +0.03 (0.038) | 0.04 | -7.8/+0.6 (17.1) | 8.9/3.5 |

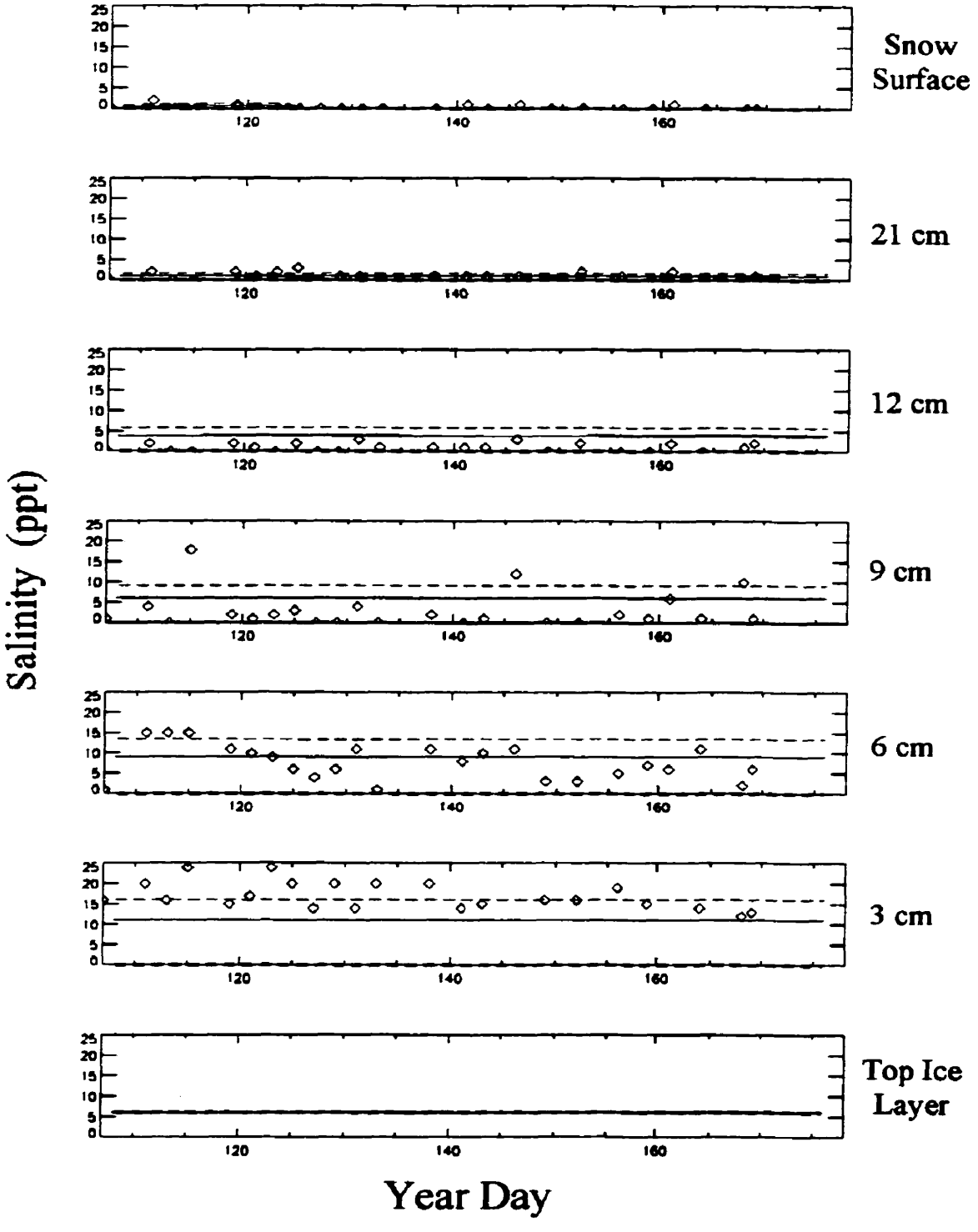



*Figure 7.7: Modeled and observed salinity (ppt) evolution between YD107 and YD177 of various snow layers for C (always zero), AVG (solid line), UPR (dashed line), and observed data (triangles). Observed values represent averages as in Barber et al. (1995b) corresponding to days where snow sampling took place.*

Brine volumes were estimated as in Barber et al. (1995b) using observed and modeled time series snow temperatures, densities, salinity and wetness. The overall modeled brine volume evolution for C, AVG and UPR (Figure 7.8) for various snow layers is somewhat different than those estimated using the observed input data. Desalination in the mid and lower snow layers over time is the main cause of the larger discrepancies between modeled and observed estimates of Vb, especially in the warmer season. The major sources of error include inadequately modeled liquid water evolution and related feedbacks along with measurement error in snow temperature (up to ±1°C in the warmest part of the season), $\rho_s$, S and Wv. A sensitivity analysis of snow Vb given the relative errors in input parameters discussed in section 7.2 suggests that Vb is virtually insensitive to a 1% error in Wv measurement and 0.5°C error in temperature. Vb changes by only 1% for a 60 kg $m^{-3}$ error in density and 0.5 ppt error in salinity. With all of these errors combined, Vb changes by about 2% - 4% with the largest effects seen near the base of the snow pack.

Given these results and various errors in the model and observations, section 7.5 will show whether these errors significantly impact microwave dielectrics.

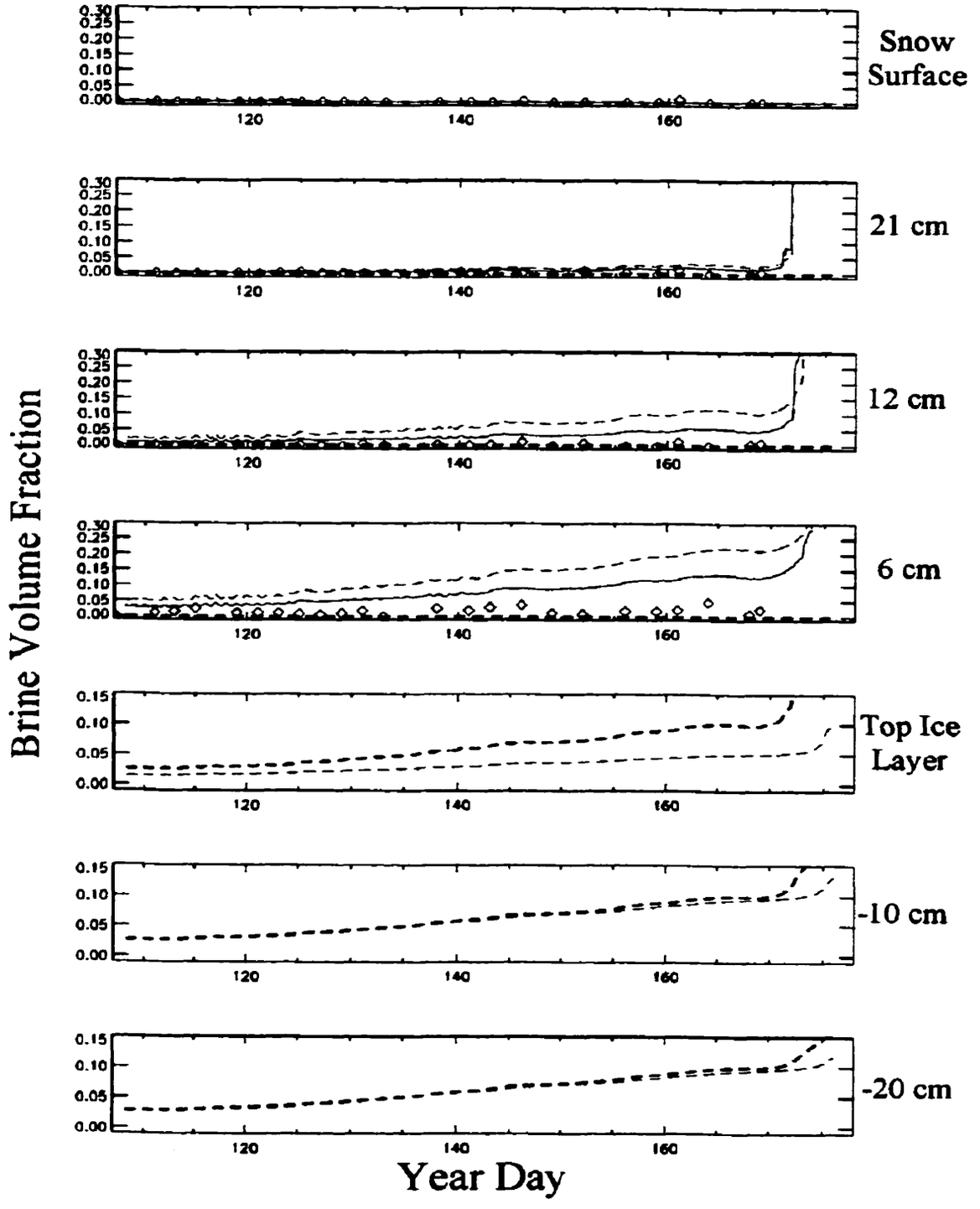



*Figure 7.8: Brine volume evolution for various layers of snow (positive depths) and ice (negative depths) between YD107 to YD177 for C (heavy dashed line), AVG (solid line), UPR (dashed line) and "observed" estimates (triangles). Only snow layers showing significant differences between them are depicted and C and UPR are shown for ice. Observed values represent estimates derived from input variable averages as in Barber et al. (1995b) corresponding to days where snow sampling took place.*

## 7.5 Microwave Dielectrics

In this final section I investigate differences in the snow and ice layers microwave dielectric properties and penetration depth given ETSSIS simulations without salinity (C) and with salinity (AVG and UPR) compared to estimations from available observations. The purpose is to show ETSSIS's direct application to microwave remote sensing (question 3). Due to the lack of physical observational data in the ice layers, only modeled results are presented. I present results according to the various observed snow evolutionary seasons (i.e. cold dry snow; pendular regime; funicular regime). A sensitivity analysis of snow dielectrics to observational measurement errors is also conducted to illustrate their effects on results.

### *7.5.1 Cold Dry Snow*

The temporal evolution of dielectric permitivity (ε') and loss (ε") estimated from C, AVG, UPR and those from observations for various snow depths reveal that both are estimated well by ETSSIS at all snow depths in the colder, dry snow season (YD107 to about YD135). This is due to ETSSIS's ability to simulate snow temperature and density (which are most important for ε') during the cold season. Inclusion of salinity in snow improves ε' and ε" (as well as δp) at the snow pack base (Figures 7.9 to 7.11 and Table 7.3). The results suggest ETSSIS is capable of reproducing low snow dielectric properties in the cold season which in turn makes the ice volume the dominant microwave scatterer and emitter according to penetration depths (Figure 7.11). The snow volume is virtually

transparent (at 5.3 GHz) in this season since the penetration depths at each snow layer are much larger than the thickness of the snow at each layer (Figure 7.11). Penetration depths within the ice are roughly 15 to 20 cm which means active microwave sensors would primarily "see" ice characteristics originating at these depths within the ice; this is why ice data down to 20 cm is presented in Figures 7.9 - 7.11. Modeled penetration depth errors would be caused by inaccurate ice temperatures, salinities and density. ETSSIS ice temperatures below 4 cm were within ± 1°C at all levels (not shown), whereas salinity and density in the model are typical of FYI (see Figures 7.4 and 7.7; top ice layer). Therefore, errors in ice penetration depths are expected to be minimal.

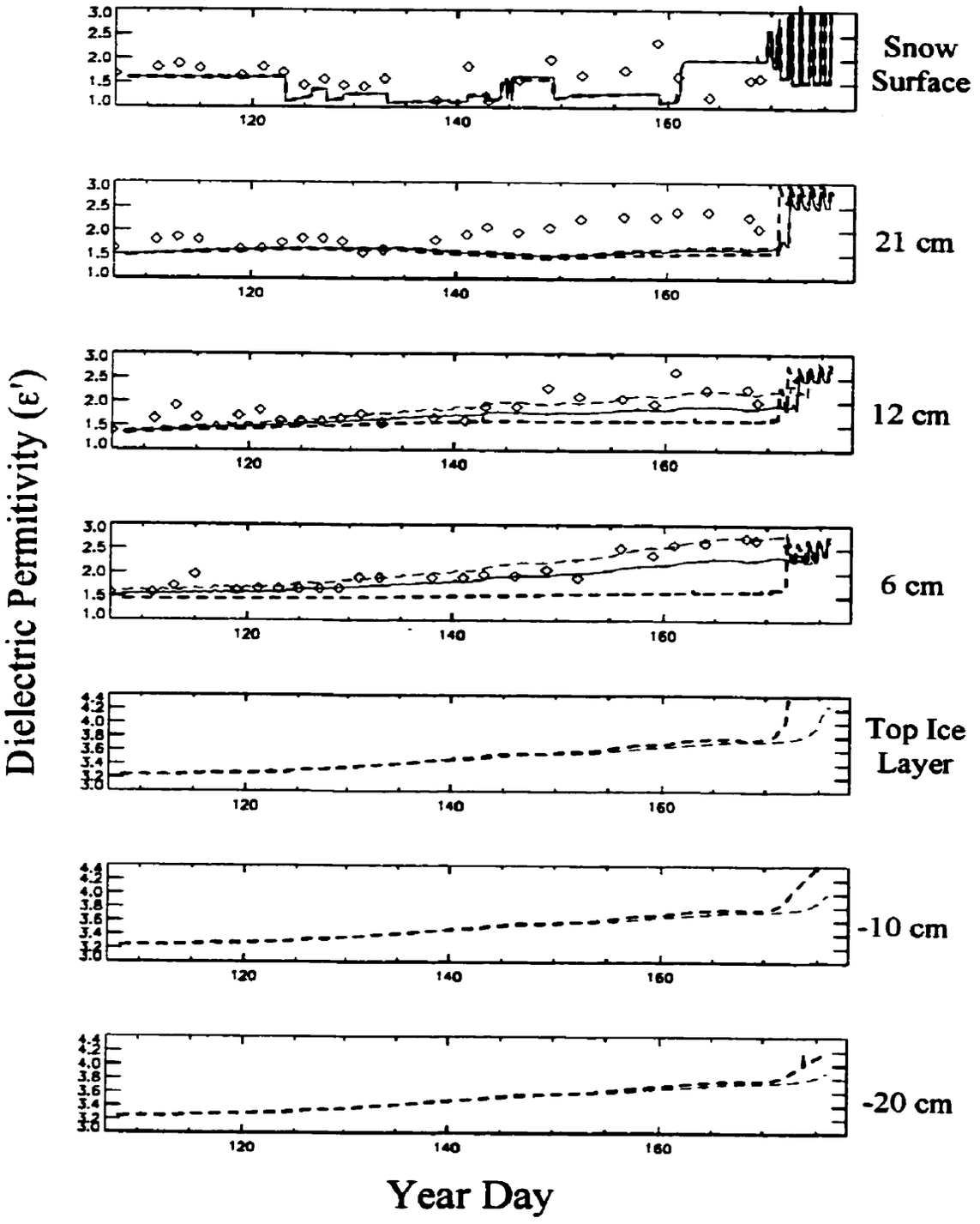



*Figure 7.9: Dielectric permitivity (ε') evolution for various layers of snow (positive depths) and ice (negative depths) between YD107 to YD177 for C (heavy dashed line), AVG (solid line), UPR (dashed line) and "observed" estimates (triangles). Only snow layers showing significant differences between them are depicted and C and UPR are shown for ice. Observed values represent estimates derived from input variable averages as in Barber et al. (1995b) corresponding to days where snow sampling took place.*

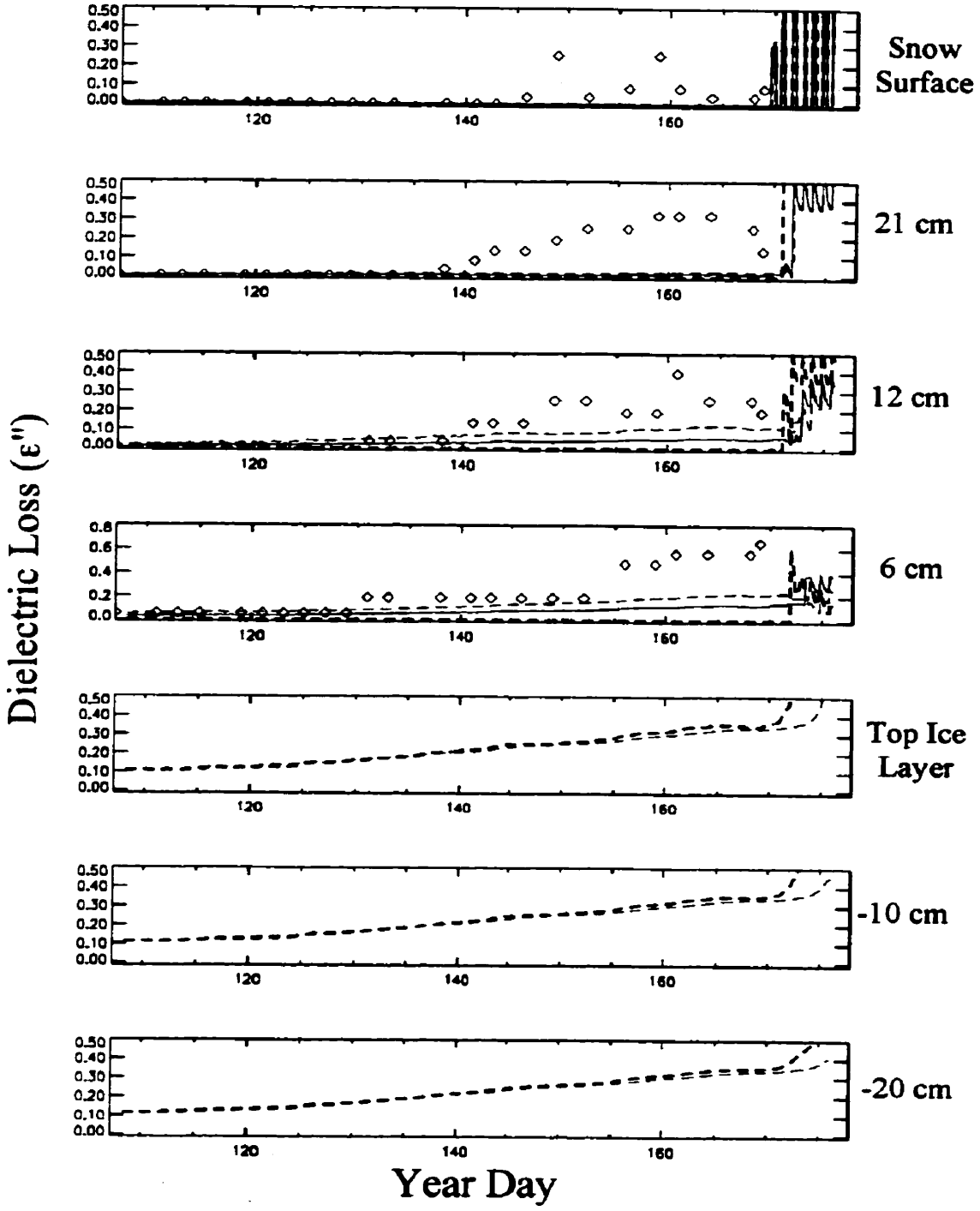



*Figure 7.10: Dielectric loss (ε") evolution for various layers of snow (positive depths) and ice (negative depths) between YD107 to YD177 for C (heavy dashed line), AVG (solid line), UPR (dashed line) and "observed" estimates (triangles). Only snow layers showing significant differences between them are depicted and C and UPR are shown for ice. Observed values represent estimates derived from input variable averages as in Barber et al. (1995b) corresponding to days where snow sampling took place.*

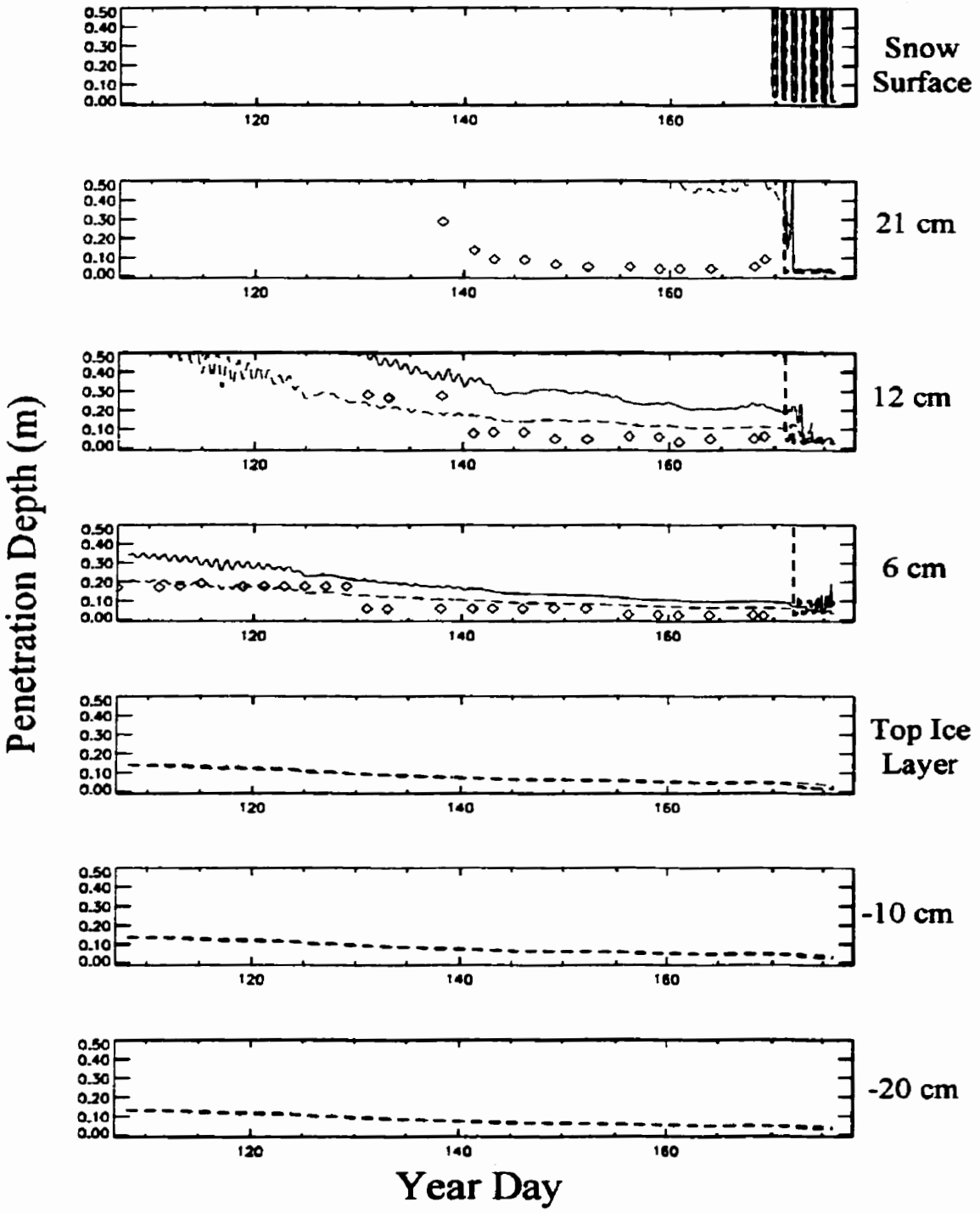



*Figure 7.11: Penetration depth (δp – in meters) evolution for various layers of snow and ice between YD107 to YD177 for C (heavy dashed line), AVG (solid line), UPR (dashed line) and "observed" estimates (triangles). Only snow layers showing significant differences between them are depicted and C and UPR are shown for ice. Observed values represent estimates derived from input variable averages as in Barber et al. (1995b) corresponding to days where geophysical snow sampling took place.*

*Table 5.3a-c: ETSSIS C, AVG and UPR mean bias error (MBE) and root-mean-squared error (RMSE) permitivity (ε'), loss (ε") and penetration depth (δp) compared to observational estimates for various snow depths. Positive MBE implies the model overestimates observations. Observed values (Mean Obs) represent estimates derived from input variable averages as in Barber et al. (1995b) corresponding to days where geophysical snow sampling took place.*

**Permitivity (ε')**

| | Mean Obs | MBE (C) | MBE (AVG) | MBE (UPR) | RMSE (C) | RMSE (AVG) | RMSE (UPR) |
|---|---|---|---|---|---|---|---|
| Surface | 1.5 | -0.2 | -0.2 | -0.2 | 0.6 | 0.6 | 0.6 |
| 21 cm | 2.0 | -0.8 | -0.8 | -0.7 | 0.5 | 0.5 | 0.4 |
| 12 cm | 1.9 | -0.8 | -0.4 | -0.1 | 0.6 | 0.4 | 0.3 |
| 6 cm | 2.0 | -0.9 | -0.3 | +0.1 | 0.6 | 0.3 | 0.2 |

**(a)**

**Loss (ε")**

| | Mean Obs | MBE (C) | MBE (AVG) | MBE (UPR) | RMSE (C) | RMSE (AVG) | RMSE (UPR) |
|---|---|---|---|---|---|---|---|
| Surface | 0.08 | -0.12 | -0.12 | -0.12 | 0.07 | 0.07 | 0.07 |
| 21 cm | 0.17 | -0.20 | -0.19 | -0.18 | 0.13 | 0.12 | 0.12 |
| 12 cm | 0.15 | -0.25 | -0.20 | -0.11 | 0.16 | 0.11 | 0.09 |
| 6 cm | 0.5 | -0.38 | -0.27 | -0.20 | 0.52 | 0.41 | 0.34 |

**(b)**

**Penetration Depth (m)**

| | Mean Obs | MBE (C) | MBE (AVG) | MBE (UPR) | RMSE (C) | RMSE (AVG) | RMSE (UPR) |
|---|---|---|---|---|---|---|---|
| Surface | 2.9 | +0.15 | +0.12 | +0.10 | 0.18 | 0.15 | 0.13 |
| 21 cm | 0.35 | +0.90 | +0.60 | +0.45 | 0.42 | 0.35 | 0.30 |
| 12 cm | 0.2 | +0.95 | +0.27 | +0.11 | 0.24 | 0.13 | 0.09 |
| 6 cm | 0.09 | +0.97 | +0.10 | +0.04 | 0.30 | 0.10 | 0.05 |

**(c)**

### *7.5.2 Pendular Snow Regime*

The temporal evolution of dielectric permitivity (ε') estimated from C, AVG, UPR and those from observations for various snow depths reveal that ε' is estimated well by ETSSIS at most snow depths in the pendular regime (YD135 to YD155). This is mainly due to ETSSIS's ability to simulate snow temperature and density (which are most important for ε') during this season. Inclusion of salinity in snow dramatically improves ε' and δp at the snow pack base (Figures 7.9 and 7.11 and Table 7.3). Because ETSSIS has difficulty reproducing liquid water in the observed pendular regime, it does not reproduce ε" very well (Figure 7.9) and will continue to make the ice volume the dominant microwave scatterer and emitter since penetration depths within the snow remain higher than observed values (Figure 7.11). Observations suggest the snow cover would contribute between 75-90% of the total microwave scattering and emission during the pendular regime in which case snow grain size would become an important variable to simulate for volume scattering considerations (for active sensors).

### *7.5.3 Funicular Snow Regime*

The temporal evolution of dielectric permitivity (ε') estimated from C, AVG, UPR and those from observations for various snow depths reveal that ε' is estimated well by ETSSIS at most snow depths in the funicular regime (YD155 to YD177). Once again this is due to ETSSIS's ability to simulate snow temperature and density during this season. Inclusion of salinity in snow improves ε' and δp at the base and middle portions

of the snowpack (Figures 7.9 and 7.11 and Table 7.3). Once ETSSIS generates water in liquid phase (near YD170), the simulated ε" becomes close to expected values (Figure 7.9) and will begin to make the snow volume more of a dominant microwave scatterer (active) or emitter (passive), illustrated by the penetration depths at 5.3 GHz (Figure 7.11). Observations up to YD169 suggest the top 5 - 10 cm of the snow cover will contribute 100% of the total microwave scattering during the funicular regime and ETSSIS suggests similar magnitudes once it generates liquid water beyond YD170.

### *7.5.4 Dielectric Sensitivity to Measurement Errors*

Finally, errors in input parameters made from observations that control ε' and ε" (similar to Vb) were investigated to examine their effects on the results. The measurement errors for snow temperature, density, liquid water content, and salinity are applied to the observed time series data shown in Figures 7.2, 7.4, 7.6 and 7.7. It should be noted that conducting an error analysis using the differentials of the various parameters may lead to larger errors than those cited here.

There is virtually no change in ε' with a 0.5°C error in temperature or a 0.5 ppt error in salinity; ε' changes by 0.1 units for a 1% Wv error and 60 kg $m^{-3}$ error in density. With all errors included simultaneously, ε' changes by roughly 0.25 units. There is virtually no change in ε" with similar errors above in temperature, salinity and density, however, ε" changes by 0.05 units with a 1% error in Wv. With all errors induced simultaneously, ε" of course also changes by 0.05 units. Thus, the microwave model sensitivity analyses of ε' and ε"suggests the *in situ* measurement errors in $\rho_s$ and Wv can

contribute to small errors in the dielectric snow properties and the discrepancies between modeled and "observed" $\varepsilon$' and $\varepsilon$". The combined errors in $\varepsilon$' and $\varepsilon$" result in snow layer penetration depth uncertainties up to 30 cm in the cold season and only a few cm in the pendular and funicular regimes. With very high snow penetration depths in the cold season, the 30 cm error does not significantly affect the results nor does a few cm in the pendular and funicular regimes. Thus, errors in $\varepsilon$' and $\varepsilon$" resulting from observational sampling errors in Wv, $\rho_s$, salinity and temperature do not significantly affect microwave penetration depth estimates here.

## 7.6Conclusions

The intent of this chapter is to begin the initial steps of linking thermodynamic modeling with microwave remote sensing in a forward sense, that is, determine microwave scattering given the variables that control scattering through an electro-thermophysical relationship. The overall goal was to develop a working numerical model (ETSSIS) that enables us to interpret satellite microwave remote sensing signatures (active or passive) in a more physically meaningful way leading to better ways of monitoring the arctic climate and interpreting arctic climate change via microwave satellite remote sensing. Three specific questions were designed to utilize *in situ* field observations from SIMMS'92 to validate ETSSIS's ability to simulate the critical variables that control microwave scattering.

### *7.6.1 Question 1*

The first question was to examine ETSSIS's ability to simulate the snow sea-ice thermodynamic characteristics that are related to microwave remote sensing. Results suggest the snow and upper ice layer thermal evolution are simulated fairly well (but slightly too warm in winter) even with the new salinity parameterization in $k_{ss}$ and $C_{ss}$. Bottom snow and top ice layer temperatures are dramatically improved, relative to observations, during the melt season when salinity is included in $k_{ss}$ and $C_{ss}$. The overall effect of the addition of salinity in the snow is to increase temperatures in the cold season and decrease temperatures in the warm season by its relative effects on snow $k_{ss}$ and $C_{ss}$. If salt were explicitly treated in the snow mass balance of ETSSIS, salinity effects on the thermal evolution would be further complicated due to liquid water (brine) and latent heat release due to phase changes in this period. This would result in different thermodynamic effects within the snow pack in terms of vertical temperature structure.

### *7.6.2 Question 2*

The second question was to examine the model's ability to simulate the snow's physical characteristics which are important for the interpretation of both active and passive microwave remote sensing. ETSSIS handles the relevant physical characteristics fairly well in most cases. However, since salinity effects on ETSSIS's mass balance are not parameterized, ETSSIS does not generate free liquid water (brine) within the snow soon enough in the spring season. This in turn affects modeled snow brine drainage

(brine volumes) and grain growth in the late season. Despite this, the modeled snow density ($\rho_s$) is simulated fairly well at all levels but may under estimate density in the colder season at mid levels. Snow grain size (d) trends are also realistic, however, d is underestimated near the base of the snow pack in the late season since liquid water is not generated soon enough.

### *7.6.3 Question 3*

The final question was designed to illustrate the use of ETSSIS for direct application to microwave remote sensing through the first two questions and the electro-thermophysical relationship over the course of various snow stages (cold dry snow; pendular regime; funicular regime). This was done by examining "observed" and modeled microwave dielectric permitivity (ε'), loss (ε") and microwave penetration depth (δp) for various snow and ice layers. In the cold, dry snow season, ETSSIS reproduces the low snow ε' and ε" well, making the ice volume the dominant microwave scatterer and emitter with large penetration depths into the ice (i.e. the snow is virtually transparent in this season). This makes ETSSIS a valuable tool for interpreting microwave satellite signatures during the cold season that have potential for deriving snow depth estimations over thick smooth FYI (e.g., Barber and Nghiem, 1999; Markus and Cavalieri, 1998).

During the pendular snow regime, ε' is again simulated well by ETSSIS, however, since ETSSIS does not generate free liquid water (brine) soon enough in this regime, it continues to maintain a low snow ε" during this period and too high δp. Liquid water in snow makes the snow a more lossy material to microwaves and hence a larger ε"

and smaller δp would occur during the pendular regime. This implies ETSSIS would not produce enough of a snow layer contribution toward the total microwave scattering and emission. Once ETSSIS's mass balance considers the physical presence of salt in snow during the warmer spring season (snow temperatures ≥ -7°C), it will improve liquid water (brine) generation in this period and modeled ε" and δp.

Once liquid water is generated within the snow layer in ETSSIS, it produces ε' and ε" magnitudes that would be expected in the funicular snow regime. This also improves δp magnitudes within the snow and would allow a much greater proportion of the total scattering and emission to originate from snow.

## 7.7 Summary

In summary, this chapter has developed a 1-D electro-thermophysical model of the snow sea-ice system (ETSSIS) that attempts to predict snow-covered sea ice dielectric properties from late winter to advanced snow melt in late spring. This is a valuable tool for interpretation of microwave remote sensing signatures over this type of surface. It will allow researchers to better understand the climate processes occurring within the surface that cause the signatures viewed in microwave satellite remote sensing imagery given a certain environmental forcing. In a perfect scenario, the need for *in situ* observation (besides environmental forcing for the model) is alleviated and many aspects of arctic sea ice climate monitoring can be conducted through microwave satellite remote sensing alone. ETSSIS is a step toward this goal. ETSSIS reproduces many of the critical physical and thermodynamic variables that alter the dielectric properties of the snow and

sea ice volumes. However, like every model, it has limitations and needs further improvements, particularly in part of the spring season. Previous chapters have not only provided insight into some of the processes necessary for ETSSIS to be used for microwave remote sensing, but have also eluded to processes that require further attention. The following chapter summarizes the goals and results of my dissertation and outlines future work geared toward improving some of the processes within ETSSIS.

# CHAPTER 8: Conclusions and Future Research

## 8.1 Conclusions

The Canadian Arctic is experiencing changes in its atmosphere and cryosphere which can ultimately lead to widespread biosphere changes. Evidence of arctic climate change has not only been scientifically indicated, but has also been voiced by the Inuit who rely on their environment as a way of life. Arctic climate change is expected to be significantly enhanced over the next hundred years due to various feedbacks in its climate system. World leaders have become aware of these climate change issues that affect the polar regions and the entire global climate system. However, in light of the most drastic changes projected in the polar regions, world authorities are now beginning to focus more attention on the polar regions.

Polar researchers have commonly used satellite remote sensing, numerical modeling and *in situ* observation to better understand polar climate processes and monitor any changes occurring in the polar climate. Recent research has shown there is an electro-thermophysical relationship between the electrical properties of snow-covered FYI that control microwave remote sensing signatures and the thermophysical evolution of snow-covered sea ice. The dielectric (or electrical) properties of the snow and sea ice are dictated by their thermophysical evolution which is driven by the surface energy balance (SEB). Numerical models of snow-covered sea ice can simulate the SEB and the associated

thermophysical evolution of the snow and ice. Dielectric models can simulate snow and sea ice electrical properties that in turn suggest the scattering signals we see in microwave satellite data. A model that couples the thermophysical evolution of snow-covered FYI with its electrical properties can simultaneously predict all of these characteristics over the various seasons.

The overall goal of my dissertation is to begin the process of linking thermodynamic snow sea-ice modeling to microwave remote sensing in a forward sense. This entails the development of a coupled one-dimensional thermodynamic snow sea-ice model where its output provides the necessary input to a snow and sea ice microwave dielectric model; the amalgamation of the models produces a new approach to modeling which I call the Electro-Thermophysical model of the Snow Sea Ice System (ETSSIS). The utility of this model allows one to predict the microwave scattering and emission within snow-covered first-year sea ice. The scientific application of such a model is to essentially by-pass *in situ* observation by utilizing ETSSIS over various FYI locations throughout the arctic to predict the thermophysical evolution and mechanical state of the ice. This in conjunction with actual microwave satellite data at similar locations will guide validation of the accuracy of the model within real conditions. If not, sensitivity analyses of various model runs may suggest other thermophysical reasons for the scattering that is taking place. In addition, with the advent of new microwave satellites (passive and active) with ever increasing spatial resolutions, ETSSIS can be a valuable tool for interpretation of these data sets over FYI.

The practical application of ETSSIS is to use its thermophysical information to predict locations where the ice is thermally and mechanically weak (or high break-up potential). This is not only useful for marine navigation but for the Inuit people and others who use the ice for travel, hunting and scientific investigations.

In a global climate context, the applications above spanning several years can lead to a better understanding of the FYI thermophysical changes that may take place in a global climate warming scenario. That is, relationships between arctic climate change, the surface energy balance and thermophysical changes within FYI can be made over several years of investigation. Furthermore, the detailed thermophysical processes of ETSSIS (i.e. climatic processes) can be used to conduct sensitivity studies that can suggest ways of scaling up the important processes to GCMs. This would improve current GCMs by ensuring they properly handle the critical climate processes that pertain to the polar regions.

As the development of this model progressed, several issues needed to be addressed and along with them, newer findings and improved understanding of snow-covered sea ice climate processes followed. These issues included: 1) ensuring the existing one-dimensional sea ice models treat the SEB and snow/ice thermodynamics in the appropriate time scales we see occurring in field experiments, 2) ensuring the snow and ice thermodynamics are not compromised by differences in environmental and spatial representation within components of the SEB, 3) ensuring that the snow layer is properly handled in the modeling environment, and 4) how satellite microwave remote sensing data can be used within the model environment.

In chapter 3, the first two issues were addressed by three research questions concerning, 1) the differences between diurnal time scales as opposed to daily averages in modeling seasonal snow covered fast-ice, 2) the representation of shorter time scale processes in the model's parameterizations, and 3) the differences between using actual "on ice" meteorological data as forcing as opposed to using "land-based" data as forcing within the model. I showed there were significant variations in snow and ice evolution using diurnal time scales when compared to daily averages, a fact that previously has never been addressed. Reasons for the differences are due to non-linear responses of the SEB within the snow cover with 1) the diurnal distribution of short-wave energy exchange within the snow cover accounting for 33% of the difference, 2) turbulent flux exchanges between the snow surface and atmosphere accounting for 17% of the difference, and 3) the positive albedo feedback mechanism of enhanced short-wave absorption during melt events accounting for 50% of the difference. This result highlighted the importance of accounting for diurnal processes within sea ice models which have also been cited as being predominant in time-series microwave satellite signatures (see for example, Barber and Nghiem, 1999; Barber et al., 1999).

The simple incident radiation parameterizations are capable of representing these energy fluxes over diurnal time scales in most cases. The radiation parameterizations require somewhat accurate knowledge of cloud cover and cloud optical depth with tropospheric aerosols and solar zenith angle dependencies being secondary factors. The current albedo parameterization of the FB sea ice model requires further attention, especially during the late snow melt and melt pond periods. The albedo scheme decreases

magnitudes too quickly during these periods causing the snow and fast-ice to decay too rapidly. I found that the biases between meteorological conditions on land and sea ice are significant enough to cause differences in fast-ice model simulations (i.e. spatial variations become important). Typically, land-based meteorological measurements are used to drive model simulations, however, I found that this can cause an abnormally quick decay of seasonal snow-covered sea ice. This is primarily due to variations in incident radiation, larger wind speeds over the ice surface (increased turbulent exchange), and differences between modeled and observed surface albedo. The results suggest that improvements in surface albedo parameterizations during the melt season and that atmospheric information over the sea ice surface are required to improve sea ice model simulations. This would also provide better feedback for deciphering microwave satellite signals that are dictated by SEB and surface changes.

In chapter 4, I expanded on addressing the first two issues and Chapter 3 results by addressing two research questions that further concentrated on incident radiation parameterizations, in light of their critical importance to the SEB. In particular, I critically assessed selected $K\downarrow$ and $L\downarrow$ parameterizations against *in situ* field measurements to identify any seasonal or environmental/spatial biases and offered improved representations of these parameterizations. Many of the parameterizations were developed from arctic coastal stations in the northwest arctic and Canadian Archipelago and were never validated over other important arctic locations such as polynyas or near large open water regions. Unique data collected within and around an arctic polynya suggests the $K\downarrow$ clear-sky schemes produce a positive seasonal bias where they

underestimated fluxes in the cold season and overestimated fluxes in the warm season. The positive bias is more dramatic in the full marine environment. The K↓ cloudy-sky schemes deplete too much radiation in the colder season and not enough in the warm season, especially in the marine warm season. An improved implementation of the schemes was made by varying the cloud optical depth. The L↓ clear-sky fluxes were improved after adjusting (decreasing) the clear-sky emissivity to account for a less emissive atmosphere at colder temperatures. This correction was sufficient for the marine conditions and corrected any seasonal bias. The L↓ cloudy-sky emissivity also needed to be increased for the near-shore fast-ice regime to account for under-estimations when clouds were present. This also alleviated a slight seasonal bias. An exponential dependence of cloud fraction was found to be important for the L↓ all-sky fluxes. Different L↓ all-sky results were found for the marine environment. Overall, the marine region consistently had more negative mean errors compared to the fast-ice site. This was not due to warmer ambient marine temperatures. The all-sky short-wave flux results suggested the marine environment had greater cloud optical depths compared to the fast-ice and terrestrial sites and is consistent with the long-wave results where a higher cloud emissivity may be required; although cloud base height differences could also be a factor. I concluded by offering newer forms of the L↓ clear-sky and all-sky parameterizations.

In Chapter 4 I highlighted the effects of various arctic environments on the simulation of incident short-wave and long-wave radiation. Improvements to the parameterization schemes will help reduce errors in these fluxes and improve sea ice

simulations for these environments. This will enhance our ability to better correlate microwave satellite signatures with sea ice model simulations.

In Chapter 5 I continued with a focus on the first two issues and built upon results from Chapters 3 and 4 by examining another important SEB variable; the surface albedo. The albedo is linked to Chapter 4 since it implicitly affects incident short-wave fluxes when clouds are present. Very few albedo observations and its spatial characteristics exist over FYI, especially during the melt season. Three research questions were designed to examine the spatial aspects of fractional surface cover types and surface albedo and to determine the role melt ponds play in these aspects over land-fast FYI. The results offered unique information on surface albedo during the melt pond season over FYI. In response to the first question, I included surface-based broadband and spectral albedo measurements over various cover types (deeper snow to deep melt ponds) during the melt pond season. The second question considered scaling up the surface measurements to regional scales by applying them to aircraft video data. These regional albedo estimates were corroborated by independent helicopter measurements. The last question involved a modeling experiment using the one-dimensional sea ice model (non-coupled) that showed that its albedo parameterization during the pond season progressed too quickly according to observations made in the first two questions. If a climatological albedo progression over a seven week period is used, simulated ice evolution improved. Further model experiments explicitly showed how changes in melt pond fraction can alter the ice thermodynamic state and thus its susceptibility to break-up. It was found that the vertical ice temperature gradient change (ITGC) increased as a quadratic with respect to

linear increases in pond fraction. Using the ITGC, an ice surface with 80% pond fraction can produce a near-isothermal ice profile more than 3 weeks earlier than an ice surface with 10% pond fraction.

Results from Chapter 5 contribute to the growing knowledge of albedo measurements over FYI for implementation into process models and GCMs. The four primary cover types identified during the melt season highlight their importance when attempting to model albedo over FYI and should be included in process models and GCMs. All of the modeling exercises showed the importance of improved melt pond representation in model albedo parameterizations and included a more realistic albedo progression over FYI in a climatological sense. These results will ultimately improve our ability to simulate the polar climate and interpreting microwave satellite observations.

Chapter 6 specifically focussed on the third issue of improving the sea ice model's representation of the snow layer. This chapter is implicitly linked to Chapters 3 to 5 by ensuring the snow layer is appropriately treated in FYI models. That is, errors in modeled snow layer characteristics will lead to errors in SEB parameters (e.g. albedo and hence incident short-wave radiation), ice growth and ablation. The model improvement involved coupling a detailed one-dimensional energy and mass balance model of snow to a one-dimensional thermodynamic first-year sea ice model. This type of coupled model has previously not been available. In addition, salinity effects in snow were parameterized with respect to snow thermal conductivity and heat capacity but were not parameterized within the snow mass balance. Salinity in snow is a common and important feature over

FYI due to its thermodynamic and microwave remote sensing implications. Until now, snow salinity has been ignored in sea ice models.

Two research questions investigated differences in modeled snow/ice evolution (question 1) and thermodynamic behavior (question 2) between the coupled model with and without salinity (coupled model without salinity - C; coupled model with salinity - UPR) and the non-coupled model with a single bulk property snow layer (non-coupled model - NC). Results suggest there can be significant differences in snow and ice evolution between C and NC. This is due to C treating the snow in a more rigorous and realistic fashion and its more sophisticated turbulent exchange, incident radiation, and snow albedo parameterizations. It was concluded that C was superior in modeling the snow and ice evolution and is advantageous to use over simpler models. The drawback of C is, of course, computational expense (run time).

Accounting for salinity in snow thermal conductivity and specific heat acts to reduce ice growth rates in the cold season and ablation rates in the spring melt season but does not appear to be a significant factor for annual sea ice cycles in most cases. However, results may be slightly different if salt were treated in the modeled mass balance since liquid water (brine) would be generated sooner in the spring period. Thermodynamically, there is no difference in simulated snow surface temperatures between NC and C but ice surface temperatures are slightly too warm in C; salinity amplifies this error in the cold season but significantly reduces the error during the melt season. However, it was not apparent that these inaccuracies affected FYI annual cycle simulations. It is unknown what effect the addition of salinity in the snow mass balance

would have on the thermodynamic results over the course of annual simulations due to complex feedbacks that would take place.

In Chapter 6 I highlight the importance of the snow layer when simulating FYI evolution and offer a new type of model for conducting polar climate processes. This will ultimately lead to future improvements in accounting for critical processes within GCMs as well as directly linking microwave remote sensing to the snow and ice models.

In the final chapter I examined the last issue by investigating the coupled snow sea-ice model and its use for passive and active microwave remote sensing applications. The coupled model was developed because current sea ice models do not provide enough physical information within the snow layer to be useful in the interpretation of microwave scattering. The coupled model provides the necessary output to be used as input to a microwave dielectric model of the Debye form; the amalgamation of the models produces an Electro-Thermophysical model of the Snow Sea Ice System (ETSSIS). This type of sophisticated forward modeling approach has previously never been conducted for microwave remote sensing of sea ice. The first two research questions looked at how well ETSSIS simulated the critical thermodynamic (question 1) and physical (question 2) variables required for input to the dielectric model by comparing the simulated variables to time-series observations. The last question illustrates how well ETSSIS predicts microwave dielectric properties and penetration depths.

Results suggest ETSSIS simulates the thermodynamic and physical properties reasonably well in most cases (with exception to salinity) but does not generate free liquid water within the snow pack early enough in the spring period. This is due to salinity not

being parameterized in the snow mass balance of the model. Salt embedded within and around the snow crystals causes internal melt at temperatures well below zero (near -8°C). Liquid water fractions during this melt typically range between 1 - 7% water by volume (i.e. pendular snow regime). The lack of modeled liquid water in this regime causes ETSSIS to under predict dielectric loss and over predict penetration depths (at 5.3 GHz), the result being that the snow layer would not contribute enough of the total microwave scattering and emission. Despite this drawback, ETSSIS is still very useful for applications in deciphering microwave remote sensing imagery of snow-covered sea ice in the winter and full melt seasons. In winter dry snow, ETSSIS produces snow and ice dielectric properties and penetration depths at expected magnitudes, making the ice volume the dominant microwave scatterer and emitter; snow is virtually transparent in this regime. In the full snow melt (funicular) regime, ETSSIS produces enough free liquid water in the snow pack (7 - 11% water by volume) to increase (decrease) the dielectric loss (penetration depths) to expected values making the snow pack the dominant scatterer and emitter.

Results from Chapter 7 provides an exciting new tool for linking the thermophysical characteristics and microwave remote sensing of snow covered FYI. The model can be applied to various locations throughout the polar regions to infer thermophysical changes from microwave satellite observations given certain environmental forcing. This will ultimately lead to better understanding the relationships between polar climate change and the associated electro-thermophysical changes that will accompany those changes.

## 8.2 Summary

Typically, microwave remote sensing and thermodynamic modeling of snow-covered FYI are used independently for studying and monitoring the arctic sea ice climate. The overall theme of my dissertation is to link thermodynamic modeling and microwave sensing of snow-covered FYI. Issues that needed to be addressed before this was possible have been discussed in various chapters. Results from these issues have added to the knowledge base of current sea ice climate processes as well as providing a new modeling tool for understanding these processes and interpretation of microwave remote sensing. Continuation of model improvements and exploiting other remote sensing/modeling techniques to further understand and monitor the arctic climate should be conducted in the future. This is based on the fact that many processes and feedbacks in the arctic climate system are still largely unexplored and unknown (e.g. cloud-precipitation-albedo feedbacks, and interdisciplinary feedbacks such as physical sea ice regimes and biological coupling). With the advancement of new microwave remote sensing satellites and improved modeling, further implementations of linking the two will provide us with better tools to understand these processes and feedbacks. In light of this, the last section provides some insight for further improving thermodynamic modeling efforts toward microwave remote sensing and sea ice climate processes.

## 8.3 Future Research

In light of the results contained in my dissertation, several areas for model improvement arise. Future arctic field work will address some of these modeling and remote sensing issues.

***1) Incorporate newer incident radiation parameterization schemes (i.e. radiative transfer models) that can be more accurate for shorter time scales.***

Total column atmospheric radiative transfer (RT) models have become computationally efficient enough to use within various scales of sea ice models for estimating incident short-wave and long-wave radiation. Implementation of these radiation schemes in the coupled snow sea-ice model would be relatively simple provided the appropriate input is available. Current RT schemes can reproduce incident radiative fluxes to within 3-9% for short-wave and 1-5% for long-wave (see for example, Li and Leighton, 1993; Allan, 2000) which are better than the simpler schemes in Chapter 4. Vertical profiles of pressure and humidity, ozone and aerosol optical depth, cloud amounts and optical depth, cloud top heights, and broadband surface albedo are typically required as input to RT schemes. The main problem is obtaining these data in arctic locations, unless one is fortunate enough to have upper air observations available. If the input parameters are not accurate, the RT estimates will also be inaccurate and potentially can be worse than those of the simpler schemes. The use of numerical weather prediction models or

satellite-derived parameters may provide most of the input required for arctic applications of RT models if upper air soundings are not available.

***2) Improve the temporal representation of melt ponds in current albedo parameterizations or ingest microwave satellite-derived albedo into the model for operational ice forecasting.***

A current limitation of all sea ice models is the temporal representation of the melt season (including melt ponds) in their albedo parameterizations. The fractional area of the critical cover types, and hence total surface albedo, (in Chapter 5) change dramatically in time and space over the course of the melt season. Currently, there are no physical models to temporally represent these surface cover and albedo changes. Future field work will be tailored toward temporal measurements of fractional cover type changes and the associated parameters that control these changes (the SEB and snow/ice melt input, precipitation, melt pond drainage mechanisms, mechanical wave forcing (by wind), and warm/cold air advection) in an attempt to develop new melt pond albedo parameterizations over FYI. Another approach is to use satellite microwave remote sensing-derived melt pond fraction that can be ingested by the sea ice model. Yackel and Barber (1999) have shown that melt pond fraction (and albedo) may be inverted from active microwave satellite imagery in windy conditions. This information can be input directly into the model without the use of (or supplement) albedo parameterizations.

***3) Include a more physically-based snow grain growth parameterization that includes wet snow growth processes (vapor growth and liquid water scavenging).***

Snow grain growth processes in many snow models include a combination of physical and empirical relationships. A major deficiency particularly exists during the pendular regime (liquid water fractions between 1-8%) when grain growth can rapidly accelerate by liquid water "scavenging" (illustrated in Chapter 7). Free liquid water in sub-freezing temperatures along with a vertical temperature gradient causes existing large grains to increase in size. Although this growth is very complex and difficult to measure, future field work and/or further theoretical studies will attempt to improve our physical understanding of these processes for implementation into the model.

***4) Include salinity in the mass balance of the model to account for salts (brine), ice and air volume fractions to improve the snow layer's thermal behavior between 0°C and -8°C. A physical parameterization of how salinity becomes introduced into the various snow layers over time is also required.***

A limitation of ETSSIS is its simplification of saline snow where salts (brine) can physically occupy space within the snow layer besides air, ice and liquid (if melting). In terms of mass balance, salt (brine) inclusions affect the overall snow mass, density, and physical make up of snow structure (grains and bonding). This ultimately affects the snow's thermal characteristics. The local freezing temperature of the snow pack is reduced depending on salt (brine) content and fractional volumes of air, ice and liquid. These factors are accounted for in the thermal conductivity and specific heat of the

current version of ETSSIS and the routine that does this can be applied to the mass balance of the model. That is, a routine to include salt (brine) in the mass balance of ETSSIS already exists, however, the parameters within the routine must replace existing mass balance parameters; this procedure is fairly involved and one that could not be implemented before the completion of this work. Future work will address this.

Currently, little is known about how salt gets introduced into the various snow layers. Initially, when young ice grows, salts are forced upward toward the atmosphere (Ono and Kasai, 1985) and can "contaminate" the snow layer near its surface. The physical mechanism of how salinity appears in the middle snow layers is uncertain. Future field work will explore this phenomena in an attempt to develop a physical relationship for implementation into ETSSIS. This would alleviate the need for specifying a snow salinity profile as is done in the current version of ETSSIS.

***5) Include the desalination process. That is, liquid water filtrates toward the ice surface during melt and flushes salt (brine) to the base of the snow pack where dilution also takes place.***

Once liquid water is generated in the snow pack, gravitational forces cause liquid water to drain to the bottom of the snow pack; the rate of drainage is dependent on the various amounts of liquid available from each snow layer. This process is accounted for within ETSSIS. If the snow is saline, brine is also flushed out of the middle snow layers and gets diluted by liquid water in lower layers. In effect, the middle snow layers entirely desalinate and the lower layers become less concentrated with salt (brine) over time until

the entire snow pack becomes desalinated. This is not accounted for in ETSSIS and is shown in Chapter 7. This is not only important for the snow's thermal evolution but for microwave dielectrics (see Chapter 7). Desalination processes could be tied to current liquid water filtration processes within ETSSIS. Future field work will attempt to monitor snow desalination processes by directly measuring snow salinity over short time scales (minutes to hours). These measurements would act as a proxy for monitoring liquid water filtration since direct measurement of liquid water movement is difficult, especially when snow is saline (due to instrument limitations). However, consistent *in situ* measurement of salinity both spatially and temporally is also difficult. Hopefully, new techniques for measuring snow salinity or liquid water will become available (either directly or remotely).

***6) Include a physical mechanism for brine/liquid water filtration across the snow-ice interface (this is also pointed out by Jordan et al., 1999).***

Once liquid water (or brine) reaches the ice surface from snow melt, it either refreezes or filtrates into the ice surface where it can refreeze within the ice or filtrate through the ice itself. ETSSIS refreezes and/or drains this liquid water once it reaches the ice surface. This simplification results in neglected latent and sensible heating associated with phase changes (and brine effects) at the ice surface and within the ice (if liquid (brine) flows through the ice volume). The process of liquid water allowing to pass through the snow-ice interface will be implemented in future versions of SNTHERM (R. Jordan, pers. Com., 2000) as well as ETSSIS. Attempting to validate and measure this

process in the field is difficult due to instrumentation limitations and knowing whether the liquid water at the ice surface is due to snow melt or ice surface melt.

***7) Include a physical representation of blowing snow to account for accumulation/erosion due to wind.***

If snow models were capable of recreating the physical characteristics of surface snow accurately (i.e. roughness characteristics and impact of airborne snow), natural accumulation or erosion according to wind speed would take place. The need for "tuning" the fractional stress gradient in ETSSIS would no longer be required to reproduce snow depths. All of the required snow characteristics are difficult to recreate in a 1-D model, such as downstream roughness and topography. Researchers involved in blowing snow processes are actively engaging in improving these types of processes in 2-D and 1-D modeling of snow (J. Pomeroy, pers. Com., 2000). Future field work is necessary toward this goal.

***8) Focus on incorporating remote sensing into snow sea-ice models such as ingesting real satellite data to guide model parameterizations. Another approach would be to expand the snow sea-ice model coupling to a forward or inverse full microwave scattering/emission model.***

Inversion of microwave satellite remote sensing can provide several snow and sea ice parameters that can be potentially ingested by ETSSIS (such as albedo). This data can either supplement or replace model parameterizations to ensure the model appropriately

represents these processes or parameters. This would ultimately allow for more accurate simulations of snow-covered sea ice. One difficulty arises with this approach since the remote sensing parameter that is ingested by the model is inherently decoupled from the model's physics (e.g. the satellite-derived albedo may represent a different surface than the model is creating). Another option is to couple a numerical microwave inverse or forward scattering model to ETSSIS. This approach is more physically attractive since we can explicitly model total microwave scattering and emission given the snow-sea ice thermodynamic and physical evolution and compare this to actual remote observations. This would elude to a) why we are seeing certain signatures in satellite imagery over a full annual sea ice cycle, and b) possibly suggest areas for future model improvement if we can not explain what we are seeing in the satellite imagery. Regardless of the approach, using microwave remote sensing and modeling of snow-covered sea ice in tandem can lead to feedbacks between the two. That is, limitations in one approach can be counteracted or built upon by the strengths of the other.

## References


Aagaard, K., L.K. Coachman and E. Carmack, 1981: On the Halocline of the Arctic Ocean. *Deep-Sea Res.*, 28A, 529-545.

Alford, D., 1974: Snow. In: *Arctic and Alpine Environments,* Ed: J. Ives and R. Barry, 85-110.

Allan, R.P., 2000: Evaluation of simulated clear-sky longwave radiation using ground-based observations. *J. Climate*, 13, 1951-1964.

Anderson, E.A., 1976: A point energy and mass balance model of a snow cover. *NOAA Tech. Report NWS 19*, pp. 150.

Arking, A., O.M. Izakova and Y.E.M. Feygel'son, 1992: Use of satellite, ground, and aerological data for calculating IR fluxes. *Atmos. Oceanic Phys.*, 28, 283-287.

Bader, H. P. and P. Weilenmann, 1992: Modeling temperature distribution, energy and mass flow in a (phase-changing) snowpack. I, Model and case studies. *Cold Reg. Sci. Technol.*,, 20, 157-181.

Barber, D.G., D.D. Johnson and E.F. LeDrew, 1991: Measuring climatic state variables from SAR images of sea ice: The SIMS SAR validation site in Lancaster Sound. *Arctic*, 44, 108-121.

Barber, D.G., 1993: Assessment of the interaction of solar radiation with a seasonally dynamic snow covered sea ice volume, from microwave scattering. Ph.D. Thesis. University of Waterloo, Waterloo, Ontario, pp. 266.

Barber, D.G., T.N. Papakyriakou and E.F. LeDrew, 1994: On the relationship between energy fluxes, dielectric properties and microwave scattering over snow covered first year sea ice during the spring transition period. *J. Geophys. Res.*, 99(C11), 22,401-22,411.

Barber, D.G. and E.F. LeDrew, 1994: On the links between microwave and solar wavelength interactions with snow-covered first-year ice. *Arctic*, 47, 298-309.

Barber, D.G., T.N. Papakyriakou, E.F. LeDrew and M.E. Shokr, 1995a: An examination of the relation between the spring period evolution of the scattering coefficient and radiative fluxes over landfast sea ice. *Intl. J. Remote Sensing*, 16(17), 3343-3363.

Barber, D.G, S.P. Reddan, and E.F. LeDrew. 1995b. Statistical Characterization of the Geophysical and Electrical Properties of Snow on Landfast First-Year Sea Ice. *J. Geophys. Res. (Oceans).* 100(C2): 2673-2686.

Barber, D.G., 1997: Sea Ice Decay for Marine Navigation – Phase 1 Report: prepared for the Arctic Ice Regime Shipping System (AIRSS), Marine Safety Division, Transport Canada, pp. 112.

Barber, D.G., J.J. Yackel, R. Wolfe and W. Lumsden, 1998: Estimating the thermodynamic state of snow covered sea ice using time series synthetic aperture radar (SAR) data. *Proc. 8th Intl. Offshore and Polar Eng. Conf.*, Montreal, Canada, May 1998. Intl. Soc. of Offshore and Polar Engineering. Vol. III, 50-54.

Barber, D.G. and A. Thomas, 1998: The influence of cloud cover on the radiation budget, physical properties and microwave scattering coefficient of first-year and multi-year sea ice, *IEEE Trans. Geosci. Rem. Sens.*, 36(1), 38-50.

Barber, D.G., A. Thomas, and T.N. Papakyriakou, 1998: Role of SAR in surface energy flux measurements over sea ice. In: Analysis of SAR Data over the Polar Oceans. Ed: C. Tsatsoulis and R. Kwok. Published by Springer-Verlag Berlin Heidelberg, pp. 35-67.

Barber, D.G. and J. Yackel, 1999: The physical, radiative and microwave scattering characteristics of melt ponds on Arctic landfast sea ice. *Intl. J. Remote Sensing*, 20, 2069-2090.

Barber, D.G. and S.V. Nghiem, 1999: The role of snow on the thermal dependence of microwave backscatter over sea ice. *J. Geophys. Res.*, 104(C11), 25789-25803.

Barber, D.G., J.J. Yackel and J.M. Hanesiak, 1999: Perspectives on sea ice, Radarsat-1 and Arctic climate change. *Can. J. Remote Sensing*, in press.

Barber, D.G., J.M. Hanesiak, W. Chan and J. Piwowar, 2001: Sea ice and meteorological conditions within Baffin Bay and the North Water Polynya between 1979 and 1996, *Atmos.-Ocean* (in review).

Barber, D.G. et al., 2001: The North Water Polynya Project: Physical Processes, *Atmos.-Ocean* (in review).

Barry, R.G., A. Henderson-Sellers, and K.P. Shine, 1984: Climate sensitivity and the marginal cryosphere. In: Climate Processes and Climate Sensitivity, Ed: J.E. Hansen, T. Takahashi. Geophysical Monograph 29, Maurice Ewing Vol. 5, Am. Geophysical Union, pp. 221-237.

Barry et al., 1989: Characteristics of arctic sea ice from remote sensing data and their relationship to atmospheric processes. *Ann. Glaciol.*, 12, 9-15.

Barry, R.G. and J. Maslanik, 1989: Arctic sea ice characteristics and associated atmosphere-ice interactions in summer inferred from SMMR data and drifting buoys: 1979-1984. *Geo. Journal*, 18, 35-44.

Barry, R.G., 1996: The parameterization of surface albedo for sea ice and its snow cover. *Progress in Phys. Geog.*, 20(1), 63-79.

Bennett, T.J., 1982: A coupled atmosphere-sea ice model study of the role of sea ice in climatic predictability. *J. Atmos. Sci., 39, 1456-1465.*

Bennington, K.O., 1963: Some crystal growth features of sea ice. *J. Glaciol.*, 4(36), 669-688.

Bradley, C.C., R.L. Brown and T.R. Williams, 1977: Gradient metamorphism, zonal weakening of the snow-pack and avalanche initiation. *J. Glaciol.*, 19, 335-342.

Bergin, M.H., et al., 2000: Comparison of aerosol optical depth inferred from surface measurements with that determined by sun photometry for cloud-free conditions at a continental U.S. site. *J. Geophys. Res.*, 105(D5), 6807-6816.

Blanchet, J-P and R. List, 1983: Estimation of optical properties of Arctic haze using a numerical model. *Atmos.-Ocean*, 21, 444-465.

Boer, G.J., G.M. Flato and D. Ramsden, 2000: A transient climate change simulation with greenhouse gas and aerosol forcing: projected climate to the twenty-first century. *Climate Dynamics*, 16, 427-450.

Brown, R.J., 1995: Global change in agriculture. Canadian Remote Sensing Contribution to Understanding Global Change, Dept. of Geography Publication Series, No. 38, University of Waterloo, 125-150.

Brown, R.D. and P. Cote, 1992: Interannual variability of landfast ice thickness in the Canadian high Arctic, 1950-1989. *Arctic*, 45, 273-284.

Brown, R.D. and B.E. Goodison, 1994: The sensitivity of the Arctic climate system to snowfall: evidence from the Canadian high Arctic. *Proc. 51st Annual Meeting of the Eastern Snow Conf.*, Dearborn, Mich., June 15-16.

Brown, R.D., 2000: Northern hemisphere snow cover variability and change, 1915-97. *J. Climate*, 13, 2339-2355

Brun, E., E. Martin, V. Simon, C. Gendre and C. Colleou, 1989: An energy and mass model of snow cover suitable for operational avalanche forecasting. *J. Glaciol.*, 35, 333-342.

Brun, E., P. David, M. Sudul and G. Brunot, 1992: A numerical model to simulate snow-cover stratigraphy for operational avalanche forecasting. *J. Glaciol.*, 38, 13–22.

Budyko, M.I., 1966: Polar ice and climate. In: Proceedings of the symposium on the Arctic Heat Budget and Atmospheric Circulation. Ed: J.O. Fletcher, The Rand Corporation, Santa Monica, CA, RM-5233-NSF, pp. 3-22.

Burns, B.A., 1990: SAR image statistics related to atmospheric drag over sea ice. *IEEE Trans. Geosci. Rem. Sens.*, 28(2), 158-165.

Camiso, J.C., 1986: Characteristics of Arctic winter sea ice from satellite multispectral microwave observations. *J. Geophys. Res.*, 91(C1), 975-994.

Carmack, E.C., R.W. Macdonald, R.G. Perkin, F.A. LcLaughlin, and R.J. Pearson. 1995. Evidence of warming of Atlantic water in the Southern Canadian Basin of the Arctic Ocean. *Geophys. Res. Letters*. 22. 1061-1064.

Carsey F. Microwave Remote Sensing of Sea Ice. American Geophysical Union. 1993. Geophysical Monograph #68. 462 pp.

Cavalieri, D.J., 1994: A microwave technique for mapping thin sea ice. *J. Geophys. Res.*, 99, 12,561-12,572.

Colbeck, S.C., 1973: Theory of metamorphism of wet snow. *CRREL Res. Rep. 313.*

Colbeck, S.C., 1980: Thermodynamics of snow metamorphism due to variations in curvature. *J. Glaciol.*, 26, 291-301.

Colbeck, S.C., 1982: An overview of seasonal snow metamorphism. *Rev. Geophys. Space Phys.*, 20, 45-61.

Colbeck, S.C., 1983: Theory of metamorphism of dry snow. *J. Geophys. Res.*, 88(C9), 5475-5482.

Colbeck, S.C., 1986: Statistics of coursening in water-saturated snow. *Acta Metalurgica*, 34, 347-352.

Colbeck, S.C., 1987: Snow metamorphism and classification. In: *Seasonal Snowcovers: Physics, Chemistry, Hydrology.* New York: D. Reidel, 1-35.

Colbeck, S.C., 1989: Air movement in snow due to wind pumping. *J. Glaciol.*, 35, 209-213.

Colbeck, et al., 1990: The International Classification for Seasonal Snow on the Ground. Prepared by: The Intl. Commission on Snow and Ice of the Intl. Association of Scientific Hydrology and the Intl. Glaciology Society. pp. 24.

Colbeck, S.C., 1991: The layered character of snow covers. *Rev. Geophys.*, 21(1), 81-96.

Cox, G.F. and W.F. Weeks, 1988: Numerical simulations of the profile properties of undeformed first-year sea ice during the growth season. *J. Geophys. Res.*, 93(C10), 12449-12460.

Crocker, G.B, 1984: A physical model for predicting the thermal conductivity of brine-wetted snow. *Cold Reg. Sci. Technol.*, 10, 69-74.

Curry, J.A. and E.E. Ebert, 1992: Annual cycle of radiation fluxes over the Arctic Ocean: sensitivity to cloud optical properties. *J. Climate.*, 5, 1267-1280.

Curry, J.A., J. L. Schramm and E.E. Ebert, 1995: Sea ice-albedo climate feedback mechanism. *J. Climate.*, 8, 240-247.

DeAbreu, R.A., J. Key, J. Maslanik, M.C. Serreze and E.F. LeDrew, 1994: Comparison of in-situ and AVRHH-derived broadband albedo over Arctic sea ice. *Arctic*, 47(3), 287-297.

De Abreu, R.A., D.G. Barber, K. Misurak and E.F. LeDrew, 1995: Spectral albedo of snow covered first-year and multi-year sea ice during spring melt. *Annals of Glaciol.*, 21, 337-342.

DeAbreu, R.A. and E.F. LeDrew, 1996: Multispectral analysis of fast sea ice albedo using AVHRR data. *Proc. IGARSS'96*, 651-653.

De Abreu, R.A., 1996: In Situ and Satellite Observations of the Visible and Infrared Albedo of Sea Ice During Spring Melt. Ph.D. Thesis, Department of Geography, University of Waterloo, Waterloo, Ontario, 326 p.

Drinkwater, M.R. and G.B. Crocker, 1988: Modeling changes in the dielectric and scattering properties of young snow-covered sea ice and GHz frequencies. *J. Glaciol.*, 34(118), 274-282.

Drinkwater, M.R., 1989: LIMEX'87 Ice surface characteristics: implications for C-Band SAR backscatter signatures. *IEEE Trans. Geosci. Rem. Sens.*, 27(5), 501-513.

Ebert, E.E. and J.A. Curry, 1993: An intermediate one-dimensional thermodynamic sea ice model for investigating ice-atmosphere interactions. *J. Geophys. Res.*, 98, 10,085-10,109.

Efimova, N.A., 1961: On methods of calculating monthly values of net long-wave radiation. *Meteorol. Gidrol., 10, 28-33.*

Evans, K.F., 1998: The spherical harmonics discrete ordinate method for three-dimensional atmospheric radiative transfer. *J. Atmos. Sci.*, 55, 429-446.

Fichefet, T., B. Tartinville and H. Goosse, 2000: Sensitivity of the antarctic sea ice to the thermal conductivity of snow. *Geophys. Res. Letters*, 27, 401-404.

Fily, M. and D.A. Rothrock, 1986: Extracting sea ice data from satellite SAR imagery. *IEEE Trans. Geosci. Rem. Sens.*, GE-24, 849-854.

Flato, G.M. and W.D. Hibler, 1992: On modeling pack ice as a cavitating fluid. *J. Phys. Oceanogr.*, 22, 626-651.

Flato, G.M. and R.D. Brown, 1996: Variability and climate sensitivity of landfast Arctic sea ice. *J. Geophys. Res.*, 101(C10), 25767-25777.

Fowler, C., J.A. Maslanik and W. Emery, 1994: Observed and simulated ice motion for an annual cycle in the Beaufort Sea. *Proc. IGARSS'94.*

Fukami, H., K. Kojima and K. Aburakawa, 1985: The extinction and absorption of solar radiation within a snow cover. *Ann. Glaciol.*, 6, 118-122.

Fukusako, S., 1990: Thermophysical properties of ice, snow and sea ice. *Int. J. Thermophys.*, 11, 353-372.

Gabison, R., 1987: A thermodynamic model of the formation, growth, and decay of first-year sea ice. *J. Glaciol.*, 33, 105-119.

Golden, K.M. et al., 1998a: Forward electromagnetic scattering models for sea ice. *IEEE Trans. Geosci. Rem. Sens.*, 36, 1655-1674.

Golden, K.M., et al., 1998b: Inverse electromagnetic scattering models for sea ice. *IEEE Trans. Geosci. Rem. Sens.*, 36, 1675-1703.

Gower, J.F.R, 1995: Global change: Impact on oceans and fisheries, contributions of remote sensing to future studies. Canadian Remote Sensing Contribution to Understanding Global Change, Dept. of Geography Publication Series, No. 38, University of Waterloo, 111-124.

Grenfell, T.A. and G.A. Maykut, 1977: The optical properties of ice and snow in the Arctic basin. *J. Glaciol.*, 18(30), 445-463.

Grenfell, T.C. and D.K. Perovich, 1984: Spectral albedos of sea ice and incident solar irradiance in the Beaufort Sea. *J. Geophys. Res.*, 89, 3573-3580.

Grenfell, T.A., 1996: Microwave and thermal infrared emission from young sea ice and pancake ice. *Proc. IGARSS'96*, 959-961.

Gruelll, J.W. and T. Konzelmann, 1994: Numerical modeling of the energy balance and the englacial temperature of the Greenland ice sheet: Calculations for the ETH-Camp location. *Global Planet. Change*, 9, 91-114.

Guryanov, I.E., 1985: Thermal-physical characteristics of frozen, thawing and unfrozen grounds, *4th Intl. Symp. Ground Freezing*, Sapporo, Japan, 225-230.

Hanafin J. A. and P. J. Minnett, 2001. Cloud forcing of surface radiation in the North Water Polynya. *Atmos.-Ocean*, (in press).

Hänel, G, 1976: The properties of atmospheric aerosol particles as functions of the relative humidity at thermodynamic equilibrium with the surrounding moist air. *Advances in Geophysics*, 19, 73-188.

Hanesiak, J.M., 1998: Historical Perspective. In: *NOW'98 Sea Ice/Climate Dynamics Subgroup Field Summary*, Ed. Papakyriakou, T.N., C.J. Mundy, and D.G. Barber, Centre for Earth Observations Science, Department of Geography, University of Manitoba, CEOS tech 98-8-2, pp. 15-20

Hanesiak, J.M., D.G. Barber and G.M. Flato, 1999: The role of diurnal processes in the seasonal evolution of sea ice and its snow cover. *J. Geophys. Res.*, 104(C6), 13593-13604.

Hanesiak, J.M., D.G. Barber, T.N. Papakyriakou and P.J. Minnett, 2001a: Parameterization schemes of incident radiation in the North Water. *Atmos.-Ocean*, (in press).

Hanesiak, J.M., D.G. Barber, T.N. Papakyriakou and R.E. Jordan, 2001b: Utility of a coupled 1-D thermodynamic snow-sea ice model toward microwave remote sensing. *J. Geophys. Res.*, (in review).

Hanesiak, J.M., D.G. Barber, R.A. De Abreu and J.J. Yackel, 2001c: Local and regional observations of arctic first-year sea ice melt pond albedo. *J. Geophys. Res.*, 106(C1), 1005-1016.

Herman, G. F., 1986: Arctic stratus clouds. In, *The geophysics of sea ice*. Ed. Untersteiner, N. (Plenum press, New York) p. 9-164.

Herman, G.F. and J.A. Curry, 1984: Observational and theoretical studies of solar radiation in Arctic stratus clouds. *J. Climate and Appl. Meteorol.*, 23, 5-24.

Heron, R. and M.K. Woo, 1994: Decay of high Arctic lake-ice cover: Observations and modeling. *J. Glaciol.*, 40, 283-292.

Holben, B.N., Y.J. Kaufman, and J.D. Kendall,1990: NOAA-11 AVHRR visible and near-IR inflight calibration. *Intnl. J. Remote Sensing*, 11(8), 1511-1519.

Holland, D.M., L.A. Mysak and D.K. Manak, 1993: Sensitivity study of a dynamic thermodynamic sea ice model. *J. Geophys. Res.*, **98**(C2), 2561-2586.

IPCC (Intergovernmental Panel on Climate Change), 1996. Climate Change: The IPCC Assessment. J.T. Houghton, G.J. Jenkens, and J.J. Ephraums (eds.) Cambridge University Press, Cambridge, U.K.

Jacobs, J.D., 1978: Radiation climate of Broughton Island, energy budget studies in relation to fast-ice breakup processes in Davis Strait, edited by R.G. Barry and J.D. Jacobs, *Occas. Pap.* 26 pp. 105-120, Inst. Of Arctic and Alp. Res., Univ. of Colorado, Boulder.

Jeffries, M.O., K. Schwartz and S. Li, 1997: Arctic summer sea-ice SAR signatures, melt-season characteristics, and melt-pond fractions. *Polar Record*, 33, 101-112.

Jezek, K.C., D. Perovich, K.M. Golden, C. Luther, D.G. Barber, P. Gogineni, T. Grenfell, A. Jordan, C. Mobley, S. Nghiem, and R. Onstott. 1998 A broad Spectral, Interdisciplinary Investigation of the Electromagnetic Properties of Sea Ice. *IEEE Trans. Geosci. Rem. Sens.*. ONR ARI special issue. 36(5):1633-1641.

Jing, X. and R.D. Cess, 1998: Comparison of atmospheric clear-sky shortwave radiation models to collocated satellite and surface measurements in Canada. *J. Geophys. Res.*, 103(D22), 28817-28824.

Johannessen, O.M., M.W. Miles and E Bjorgo, 1996: Global sea ice monitoring from microwave satellites. *Proc. IGARSS'96*, 932-934.

Jordan, R.E., 1991: A one-dimensional temperature model for a snow cover: Technical documentation for SNTHERM.89, *Rep. 91-16*, 49 pp., Cold Reg. Res. and Eng. Lab, Hanover, NH.

Jordan, R.E., E.L. Andreas and A.P. Makshtas, 1999: Heat budget of snow-covered sea ice at North Pole 4. *J. Geophys. Res.*, 104(C4), 7785-7806.

Kattenburg, A. et al., 1996: Climate Models – Projections of Future Climate. In: *Climate Change 1995: The Science of Climate Change*. Cambridge University Press, pp. 285-357.

Key, J. R., R.A. Silcox and R.S. Stone, 1996: Evaluation of surface radiative flux parameterizations for use in sea ice models. *J.Geophys. Res.*, 101(C2), 3839-3849.

Kidwell, K.B., 1991: NOAA Polar Orbiter Data Users Guide. US Dept. of Commerce, NOAA NESDIS, Washington, D.C., pp. 250.

Koepke, P., 1989: Removal of atmospheric effects from AVHRR albedos. *J. Appl. Meteorol.* 28(12), 1341-1348.

Lake, R.A. and E.L. Lewis, 1970: Salt rejection by sea ice during growth. *J. Geophys. Res.*, 75(30), 583-597.

Langleben, M.P., 1971: Albedo of melting sea ice in the southern Beaufort Sea. *J. Glaciol.*, 10, 101-104.

LeDrew, E.F. and D.G. Barber, 1994: The SIMMS Program: a study of change and variability within the marine cryosphere. *Arctic*, 47, 256-264.

Leontyeva, E. and K. Stamnes, 1993: Estimations of cloud optical thickness from ground-based measurements of incoming solar radiation in the Arctic. *J. Climate*, 7, 566-578.

Li, Z. and H.G. Leighton, 1992: Narrow to broadband conversion of spatially autocorrelated reflectance measurements. *J. Appl. Meteorol.*, 31, 421-432.

Li, Z., H.G. Leighton, K. Masuda and T. Takashima, 1993: Estimation of SW flux absorbed at the surface from TOA reflected flux. *J. Climate*, 6, 317-330.

Lindsay, R.W. and D.A. Rothrock, 1993: The calculation of surface temperature and albedo of Arctic sea ice from AVHRR. *Annals of Glaciol.*, 17, 391-397.

Lindsay, R.W. and D.A. Rothrock, 1994: Arctic sea ice albedo from AVHRR. *J. Climate*, 7(11), 1737-1749.

Livingstone, C.E., R.G. Onstott, L.D. Arsenault, A.L. Gray and K.P. Singh, 1987: Microwave sea-ice signatures near the onset of melt. *IEEE Trans. Geosci. Rem. Sens.*, GE-25(2), 174-187.

Livingstone, C.E. and M.R. Drinkwater, 1991: Springtime C-band SAR backscatter signatures of Labrador Sea marginal ice: measurements versus modeling predictions. *IEEE Trans. Geosci. Rem. Sens.*, 29, 29-41.

Lock, G.S.H., 1990: The Growth and Decay of Ice. Cambridge University Press, New York, pp. 444.

Lofgren, G. and W.F. Weeks, 1969: Effect of growth parameters on the substructure spacing in NaCl ice crystals. *J. Glaciol.*, 8(52), 153-164.

Loth, B., H.F. Graf and J.M. Oberhuber, 1993: Snow cover model for global climate simulations. *J. Geophys Res.*, 98, 10451-10464.

Lynch-Steiglitz, M., 1994: The development and validation of a simple snow model for the GISS GCM. *J. Climate*, 7, 1842-1855.

Makshtas, A., 1991: The heat budget of arctic ice in the winter. Intl. Glaciological Society, Cambridge, UK, pp. 77.

Malinas, N.P. and R.A Shuchman, 1994: SAR derived sea ice thickness during ICEX'92. *IEEE trans. Geosci. Rem. Sens.*, 1756-1758.

Markus, T. and B.A. Burns, 1995: A method to estimate subpixel-scale coastal polynyas with satellite passive microwave data. *J. Geophys Res.*, 100, 4473-4487.

Markus, T. and D.J. Cavalieri, 1996: Comparison of open water and thin ice areas derived from satellite passive microwave data with aircraft measurements and satellite infrared data in the Bering Sea. *Proc. IGARSS'96*, 1523-1525.

Markus, T. and Cavalieri, D.J., 1998. Snow depth distribution over sea ice in the southern ocean from satellite passive microwave data. *Antarctic Sea Ice Physical Processes, Interactions and Variability*, Antarct. Res. Ser. 74, AGU, Washington, D.C., pp.19-39.

Maslanik, J.A. and H. Maybee, 1994: Assimilating remotely-sensed data into a dynamic-thermodynamic sea ice model. *IGARSS'94*, 1306-1308.

Maykut, G.A. and N. Untersteiner, 1971: Some results from a time dependent thermodynamic model of sea ice. *J. Geophys. Res.*, 76, 1550-1575.

Maykut, G.A. and P.E. Church, 1973: Radiation climate of Barrow, Alaska, 1962-66, *J. Appl. Met.*, 12, 620-628.

Maykut, G.A. and T.C. Grenfell, 1975: The spectral distribution of light beneath first-year sea ice in the Arctic Ocean. *Limnol. And Oceanog.*, 20(4), 554-563.

Maykut, G.A., 1978: Energy exchange over young sea ice in the central arctic. *J Geophys. Res.*, 83(C7), 3646-3658.

Maykut, G. A., 1982: Large scale heat exchange and ice production in the central Arctic. *J. Geophys. Res.*, 87, 7971-7984.

Maykut, G.A., 1985: *The Ice Environment*. Sea Ice Biota, Ed: R.A. Horner, CRC Press, Boca Raton, FL, pp. 21-82.

Maykut, G.A., 1986: The surface Heat and Mass Balance. In: *The Geophysics of Sea Ice*. Ed: N. Untersteiner, Plenium Press, New York, 395-464.

Minnett, P.J., 1999: The influence of solar zenith angle and cloud type on cloud radiative forcing at the surface in the Arctic. *J. Climate*, 12, 147-158.

Moritz, R. E., and D. K. Perovich, Ed, SHEBA, A Research Program on the Surface Heat Budget of the Arctic Ocean, Science Plan, Rep. No. 5, U. of Washington, 64 pp., 1996.

Nghiem, S.V. et al., 1998: Diurnal thermal cycling effects on backscatter of thin sea ice, *IEEE Trans. Geosci. Rem. Sens.*, 36(1), 111-124.

Ohmura, A., 1981: Climate and energy balance of the Arctic tundra. *Zurcher Geogr. Schr 3*, 448 pp. Geogr. Inst., Zurich, Switzerland.

Ono, N., 1966: Thermal properties of sea ice: III; On the specific heat of sea ice, *Low Temp. Sci.*, A24, 249-258.

Ono, N. and T. Kasai, 1985: Surface layer salinity of young ice. *Ann. Glaciol.*, 6, 298-299.

Ono, N. and M.S. Krass, 1993: Theoretical approach describing the thermal regime of snow-covered sea ice. *Ann. Glaciol.*, 18, 79-84.

Onstott, R.G., 1992: SAR and scatterometer Signature of Sea Ice. Chapter 5. In; Carsey F. (Editor) Microwave Remote Sensing of Sea Ice. (American Geophysical Union. 1992. Geophysical Monograph 68)

Owens, W.B. and P. Lemke, 1990: Sensitivity studies with a sea ice-mixed layer-pycnocline model in the Weddell Sea. *J. Geophys. Res.*, 95, 9527-9538.

Papakyriakou, T.N., 1999: An examination of relationships among the energy balance, surface properties and climate over snow covered sea ice during the spring season. Ph.D. Thesis, University of Waterloo, Waterloo, Ontario, pp. 364.

Parkinson, C.L. and W.M. Kellogg, 1979: Arctic sea ice decay simulated for a CO2-induced temperature rise. *Clim. Change*, 2, 149-162.

Parkinson, C.L., D. Cavalieri, P. Gloersen, H.Zwally and J. Comiso, 1999. Arctic sea ice extents, areas and trends, 1978-1996. *J. Geophys. Res.* 104(C9), 20837-20856.

Perovich, D.K., 1996: The Optical Properties of Sea Ice. USA Cold Regions Research and Engineering Laboratory, Hanover, NH, Monograph 96-1, 25 pp.

Petzold, D. E., 1977: An estimation technique for snow surface albedo. *Climat. Bull.*, 21, 1-11.

Pitman, D. and B. Zuckerman, 1967: Effective thermal conductivity of snow at -88, -27, and -5°C. *J. Appl. Phys.*, 38, 2698-2699.

Pomeroy, J.W., P. Marsh and D.M. Cray, 1997: Application of a distributed blowing snow model to the arctic. *Hydrol. Processes.*, 11, 1451-1464.

Preller, R.H., J.E. Walsh and J.A. Maslanik, 1992: The use of satellite observations in ice cover simulations. Microwave Remote Sensing of Sea Ice, American Geophysical Union, 385-404.

Radionov, V.F., N.N. Bryazgin and E.I. Alexandrov, 1997: The snow cover of the Arctic Basin. *Tech. Report APL-UW TR 9701*, Applied Physics Laboratory, University of Washington, Seattle, Washington, pp. 63.

Robinson, D.A., G. Scharfen, M.C. Serreze, G. Kulka and R.G. Barry, 1986: Snow melt surface albedo in the Arctic basin. *Geophys. Res. Letters*, 13, 945-948.

Robinson, D.A., G. Scharfen, M.C. Serreze, R.G. Barry and G. Kukla, 1992: Large-scale patterns and variability of snow melt and parameterized surface albedo in the Arctic basin. *J. Climate*, 5(10), 1109-1119.

Ross, B., and J.E. Walsh, 1987: A comparison of simulated and observed fluctuations in summertime Arctic surface albedo. *J. Geophys. Res.*, 92, 13,115-13,125.

Rothrock, D.A., Y. Yu, and G.A. Maykut, 1999: Thinning of the Arctic sea-ice cover. *Geophys. Res. Letters.*, 26, 1-5.

Semtner, A.J. Jr., 1976: A model for the thermodynamic growth of sea ice in numerical investigations of climate. *J. Phys. Oceanogr.*, 6, 379-389.

Semtner, A.J. Jr., 1987: A numerical study of sea ice and ocean circulation in the Arctic. *J. Phys. Oceanogr.*, 17, 1077-1099.

Serreze, M.C. et al. 2000: Observational evidence of recent change in the northern high-latitude environment. *Climate Change*, 46, 159-207

Shine, K.P, A. Henderson-Sellers and R.G. Barry, 1983: Albedo-climate feedback: The importance of cloud cryosphere variability, In: *New Perspectives in Climate Modeling.* Eds: A. Berger and C. Nicolis, published by Elsevier.

Shine, K.P., 1984: Parameterization of shortwave flux over high albedo surfaces as a function of cloud thickness and surface albedo. *Q. J. R. Meteorol. Soc.*, 110, 747-764.

Shine, K.P and A. Henderson-Sellers, 1985: The sensitivity of a thermodynamic sea ice model to changes in surface albedo parameterization. *J. Geophys. Res.*, 90(D1), 2243-2250.

Shirasawa, K. and R.G. Ingram, 1997: Currents and turbulent fluxes under the first-year sea ice in Resolute Passage, Northwest Territories, Canada. *J. Marine Systems.*, 11, 21-32.

Shuchman, R.A., R.G. Onstott, R.W. Flett and C.C. Wackerman, 1996: Satellite remote sensing of the Beaufort Sea during LEADEX'92. *IEEE trans. Geosci. Rem. Sens.*, 1015-1017.

St. Germain, K.M and D.J. Cavalieri, 1996: A microwave technique for mapping ice temperatures in the Arctic seasonal ice zone. *Proc. IGARSS'96*, 1208-1210.

Steffen, K. and T. DeMaria, 1996: Surface energy fluxes of Arctic winter ice in Barrow Strait. *J. Appl. Meteorol.*, 35, 2067-2079.

Sturm, M., 1991: The role of thermal convection in heat and mass transport in the subarctic snow cover. U.S. Army Core of Engineers Cold Regions Research and Engineering Lab, Hanover, New Hampshire, CRREL Report 91-19, pp. 52.

Sturm, M., J. Holmgren, M. Konig, and K. Morris, 1997: The thermal conductivity of seasonal snow. *J. Glaciol.*, 43, 26-41.

Svendsen, E., et al., 1983: Norwegian Remote Sensing Experiment: Evaluation of the Nimbus 7 Scanning Multichannel Microwave Radiometer for sea ice research. *J. Geophys. Res.*, 88, 2781-2791.

Taylor, V.R.and L.L. Stowe, 1984: Atlas of reflectance patterns for uniform Earth and cloud surfaces (NIMBUS-7 ERB -- 61 days). NOAA Technical Report NESDIS 10, 66 pp.

Tiuri, M.E. et al., 1984: The complex dielectric constant of snow at microwave frequencies. *IEEE J. Oceanic Engineering OE-9*, 5, 377-382.

Tucker, W.B., A.J. Gow and W.F. Weeks, 1987: Physical properties of summer sea ice in Fram Strait. *J. Geophys Res.*, 92(C7), 6787-6803.

Ulaby, F.T., R.K. Moore and A.K. Fung, 1986: Microwave Remote Sensing: Active and Passive. Vol 3, Norwood, MA: Addison-Wesley Publishing Company.

Vant, M.R., R.O. Ramseier and V. Makios, 1974: Dielectric properties of fresh snow and sea ice at 10 and 36 GHz. *J. Appl. Physics.*, 45, 4712-4717.

Vant, M.R., R.O. Ramseier and V. Makios, 1978: The complex dielectric constant of sea ice at frequencies in the range 0.1 to 40 GHz. *J. Appl. Phys.*, 49(3), 1264-1280.

Wakatsuchi, M. and T. Kawamura, 1987: Formation processes of brine drainage channels in sea ice. *J. Geophys. Res.*, 92(C7), 7195-7197.

Walsh, J.E., W.L. Chapman, and T. Shy, 1996. Recent decrease of sea level pressure in the central Arctic. *J. Clmate*. 9. 480-486.

Warren, S.G. and W.J. Wiscombe, 1981: A model for the spectral albedo of snow. I: Snow containing atmospheric aerosols. *J. Atmos. Sci.*, 37, 2735-2745.

Weaver, R.L., C. Morris, R.G. Barry, 1987: Passive microwave data for snow and ice research: planned products from the DMSP SSM/I system. *EOS (Trans. Amer. Geophys. Union)*, 68, 769, 776-777.

Weaver, R.L. and V.J. Troisi, 1996: Remote sensing data availability from the Earth Observation System (EOS) via the Distributed Active Archive Center (DAAC) at NSIDC. *Proc. IGARSS'96*, 73-77.

Weeks, W.F. and A.J. Gow, 1980: Crystal alignments in the fast-ice of arctic Alaska. *J. Geophys. Res.*, 84(C10), 1137-1146.

Wenshanan, M.R., T.C. Grenfell, D.P. Winebrenner and G.A. Maykut, 1993a: Observations and theoretical studies of microwave emission from thin saline ice. *J. Geophys. Res.*, 98(C5), 8531-8545.

Wenshanan, M.R., G.A. Maykut, T.C. Grenfell, D.P. Winebrenner, 1993b: Passive microwave remote sensing of thin sea ice using principle component analysis. *J. Geophys. Res.*, 98(C7), 12,453-12,468.

Wilson, K. et al, 2001: Validation of the TRACKER algorithm for sea ice motion. *Atmos.-Ocean* (in press).

Winebrenner, D.P., E.D. Nelson, R. Colony and R.D. West, 1994: Observations of melt onset on multi-year Arctic sea ice using the ERS-1 SAR. *J. Geophys. Res.*, 99, 22,425-22,442.

Winebrenner, D.P, 1996: Polarmetric backscatter at 23 cm wavelength from Antarctic lead ice and estimation of ice thickness. *Proc. IGARSS'96*, 941-943.

Wiscombe, W.J. and S.G. Warren, 1981: A model for the spectral albedo of snow. I: Pure snow. *J. Atmos. Sci.*, 37, 2712-2733.

Yackel, J.J., D.G. Barber and J.M. Hanesiak, 2000: Melt ponds on sea ice in the Canadian Archipelago: 1, variability in morphological and radiative properties, *J. Geophys. Res.*, 105(C9), 22049-22060.

Yackel, J.J. and D.G. Barber, 2000: Melt ponds on sea ice in the Canadian Archipelago: 2, on the use of RADARSAT-1 synthetic aperture radar for geophysical inversion. *J. Geophys. Res.*, 105(C9), 22061-22070.

Zhang, T., K. Stamnes and S.A. Bowling, 1996: Impact of clouds on surface radiative fluxes and snowmelt in the Arctic and Subarctic. *J. Climate*, 9, 2110-2123.