PREDICTION OF AGRICULTURAL DROUGHT FOR THE CANADIAN PRAIRIES USING CLIMATIC AND SATELLITE DATA

BY

VIJENDRA KUMAR

A Thesis

Submitted to the Faculty of Graduate Studies
in Partial Fulfillment of the Requirements for the Degree of

DOCTOR OF PHILOSOPHY

Department of Geography
University of Manitoba
Winnipeg, Manitoba

0-612-45005-8

THE UNIVERSITY OF MANITOBA

FACULTY OF GRADUATE STUDIES

*****

COPYRIGHT PERMISSION PAGE

Prediction of Agricultural Drought for the Canadian Prairies Using Climatic and Satellite Data

BY

Vijendra Kumar

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

Dedicated to
my revered parents

Late Shri Bhullar Singh (*father*)
and
Smt. Shanti Devi (*mother*)

TABLE OF CONTENTS

Page

ABSTRACT v

ACKNOWLEDGMENTS vii

LIST OF FIGURES ix

LIST OF TABLES x

LIST OF ABBREVIATIONS xi

CHAPTER ONE: INTRODUCTION 1

1.1 Drought Vulnerability of the Prairies 3

1.2 Agricultural Drought and Crop Yield 5

1.3 Wheat Yield Estimation: Current Approach in Practice 7

1.4 Statement of Research Problem 8

CHAPTER TWO: LITERATURE REVIEW 9

2.1 Geographic Perspectives of Drought 9

2.2 Drought Prediction 13

2.3 Drought Prediction Using Pattern Recognition 13

2.3.1 Formation of a drought vector 14

2.3.2 Error Correction procedure 16

2.4 Yield Forecasting Using Time Series Analysis 18

2.4.1 Series categorization: stationary or nonstationary 19

2.4.2 The case of the stationary series 21

2.4.2.1 Simple moving averaging 21

2.4.2.2 Simple exponential smoothing 22

2.4.2.3 Box-Jenkins technique 23

2.4.3 The case of the nonstationary series 26

2.4.3.1 Trend analysis 26

2.4.3.2 Double moving averaging 27

2.4.3.3 Double exponential smoothing 30

2.5 Yield Estimation Using Weather Data 31

2.5.1 Palmer Drought Severity Index 36

2.5.2 Accumulated Basin Storage 37

2.5.3 Moisture Adequacy Index 37

2.5.4 Yield estimation for the Canadian Prairies 38

2.6 Satellite Data in Yield Estimation 40

2.6.1 Biometeorological time scale model 45

2.7 Research Objectives and Hypotheses 47

2.7.1 Objectives 47

2.7.2 Hypotheses 47

2.8 Study Area and Data Procurement 49

2.8.1 Study area 49

2.8.2 Some ecological features of Saskatchewan 49

2.8.3 Ecological features of Prairie Ecozone of Saskatchewan 58

2.8.4 Selection of crop districts 62

2.8.5 Data procurement and derivation 71

CHAPTER THREE: PATTERN RECOGNITION IN DROUGHT PREDICTION **73**

3.1 Hypothesis I and Methodology 73

3.2 Results and Discussion 78

CHAPTER FOUR: YIELD FORECASTING USING TIME SERIES ANALYSIS **79**

4.1 Hypothesis II and Methodology 79

4.2 Evaluation of the Forecasting Techniques 81

CHAPTER FIVE: EFFECT OF TEMPORAL RESOLUTION OF CLIMATE DATA ON YIELD ESTIMATION **87**

5.1 Hypothesis III and Methodology 87

5.2 Monthly versus Daily Data 89

CHAPTER SIX: DEVELOPMENT OF A HYBRID MODEL **94**

6.1 Hypothesis IV and Methodology 94

6.2 Derivation of the NDVI-based variables 95

6.3 Development of the Hybrid Model 96

6.4 Evaluation of the Hybrid Model 100

CHAPTER SEVEN: CONCLUSIONS AND RECOMMENDATIONS **101**

7.1 Conclusions 102

7.2 Recommendations for Further Research 105

REFERENCES **107**

ABSTRACT

Wheat export is a significant component of the Canadian economy. In normal (nondrought) years, the export is as high as 30 million tonnes, but it is reduced to about 20 million tonnes in drought years. This significant reduction in exports not only reduces direct profits but may also upset export targets and prices that are set in advance, if droughts are not accurately predicted.

In this thesis, prediction of agricultural drought is attempted from both long-term and short-term perspectives. The long-term prediction refers to predicting wheat yield (production per unit area) prior to wheat planting; and, under the short-term prediction, wheat yield is estimated around harvesttime. Predictive analysis was performed on five crop districts of Saskatchewan (1b, 3bn, 4b, 6a , and 9a) using climate data (monthly and daily temperature and precipitation) from nine weather stations. In addition, Normalized Difference Vegetation Index values generated from NOAA (National Oceanic and Atmospheric Administration) / AVHRR (Advanced Very High Radiometric Resolution) satellite data were used.

The long-term prediction was made by fitting various time series techniques (trend, moving average, exponential smoothing, and autoregressive integrated moving average) to the yield series in a district. The technique providing minimum prediction-error was selected. The short-term prediction was made in both qualitative and quantitative forms. The qualitative prediction was attempted using the error correction procedure of pattern recognition. The quantitative prediction involved modification of the computer program currently being used by the Canadian Wheat

Board (CWB) to estimate wheat yield. The CWB program employs only monthly temperature and precipitation and determines a drought index for a weather station. A hybrid model that employs daily climate data and a NDVI-based variable was developed. Among various NDVI-based variables, the average NDVI during the entire growing period was found to be the best predictor of yield in the case of district 3bn. For the remaining four districts, the average NDVI during the heading stage was the most reliable predictor. The commencement and termination of the heading stage were determined using a biometeorological time scale model that required planting dates, daily maximum and minimum temperatures and the photoperiod. When evaluated, the hybrid model was found to have significantly higher predictive capability than the model currently in use; the values of r^2 were 0.79, 0.96, 0.83, 0.95, and 0.39 (in the case of hybrid model) as opposed to 0.20, 0.71, 0.57, 0.58, and 0.00 (in the case of the current model) for the districts 1b, 3bn, 4b, 6a, and 9a, respectively.

ACKNOWLEDGMENTS

Bringing a Ph.D. thesis to a close is a significant achievement in one's life. This thesis has made me feel so. When I look back at the past years of my Ph.D. program (full of both joy and frustration) to enlist those who helped me reach the present stage, many faces come to my mind distinctly. Firstly, my supervisor, Dr. C.E. Haque, Department of Geography, who guided me through the program deserves my sincere thanks. Dr. S. Halli, Department of Sociology, Dr. L.P. Stene, Department of Geography, and Dr. Carl Shaykewich, Department of Soil Sciences, University of Manitoba, served on my Ph.D. committee and made constructive comments, which helped improve the quality of the thesis. I am thankful to all of them. Dr. Danny Blair, Department of Geography, University of Winnipeg, acted as an external examiner and offered useful comments for which I thank him immensely.

Dr. Paul Bullock, the former Director, Crop and Weather Surveillance division of the Canadian Wheat Board (CWB), Winnipeg, helped me in more than one way. He provided the source code of the computer program being currently used by the Canadian Wheat Board to estimate wheat yield for the Prairies; I modified the CWB program in my study. Besides, he provided the climate data (monthly as well as daily) for the required weather stations, and spent time with me discussing the CWB model and the related issues. Many thanks to Paul.

Dr. Gordon Reichert, Manager, Mapping and Spatial Data Analysis division of Statistics Canada, Ottawa, sent me the satellite data and clarified related issues as and when required. His great assistance is thankfully acknowledged.

In order to determine a biometeorological time scale, I collected planting dates for wheat from Dr. Rick Raddatz, Winnipeg Climate Centre, Environment Canada. In addition, I had approached Dr. Raddatz a few times in regards to collecting data which were missing in my original datasets. His prompt response deserves my deep appreciation.

I thankfully acknowledge the financial support provided by the University of Manitoba in terms of its Graduate Students Fellowship for the period from May 1996 to December 1998, to carry out the present study.

At the end, I thank my wife, Sangeeta for her patience, understanding, and support throughout my program. A special hug to my son, Astitva, and to my daughter, Agrini, at the conclusion of the journey which they completed along with me, with my attention perhaps less than they deserved.

LIST OF FIGURES

Page

1.1 The Canadian Prairie Provinces.. 4

2.1 An example of yield vectors.. 15

2.2 Variation of autocorrelation and partial autocorrelation coefficients with the number of lags.. 25

2.3 Wheat yield variation in the selected crop districts of Saskatchewan during 1975-96.. 29

2.4 NDVI profile under nondrought conditions for district 3bn, Saskatchewan.......... 42

2.5 NDVI profile under drought conditions for district 3bn, Saskatchewan.............. 43

2.6 Distribution of Chernozems in Saskatchewan Prairies...................................... 50

2.7 Mean annual temperature (^{0}F) and permafrost boundary line in Saskatchewan..... 53

2.8 Mean June to August precipitation (inches) in Saskatchewan......................... 54

2.9 Vegetation of Saskatchewan.. 55

2.10 Ecozones of Saskatchewan.. 57

2.11 Seasonal water deficits at 50% probability, May-Sept. 1931-60.................... 60

2.12 Drought-based agro-climatic zones of Saskatchewan.................................... 63

2.13 Crop districts and weather stations in Saskatchewan.................................... 66

2.14 Soil types in selected districts in Saskatchewan.. 67

2.15 Seasonal water deficits at 50% probability (May-Sept. 1931-60) in selected districts in Saskatchewan.. 68

2.16 Vegetation types in selected districts in Saskatchewan.................................. 69

LIST OF TABLES

Page

2.1 Vectorizing the wheat yield for district 3bn, Saskatchewan.................................. 17

2.2 Coefficients of the biometeorological time scale model for Swift Current,
Saskatchewan.. 46

2.3 Spring wheat yield variation in Saskatchewan during 1975-96............................ 64

2.4 Some agroclimatic characteristics of selected districts in Saskatchewan............ 65

2.5 Collinearity in data recorded at different weather stations within
selected districts.. 72

3.1 Yield vectorization for district 3bn, Saskatchewan.. 75

3.2 Some basic variables used to characterize yield (or drought) vectors
for district 3bn, Saskatchewan.. 77

4.1 The τ statistics for yield series for the selected crop districts........................... 80

4.2 The coefficients of the time series models that fitted the yield series
for the selected districts.. 82

4.3 Estimated coefficients of the ARIMA models developed
for the selected districts.. 83

4.4 Evaluating performance of various forecasting techniques on the
selected crop districts... 84

5.1 Drought indices derived from monthly data for weather stations
within the selected districts.. 90

5.2 Drought indices derived from daily data for weather stations
within the selected districts.. 91

5.3 Average drought indices for the selected districts using monthly
and daily data... 92

5.4 Yield models based on monthly and daily data for selected districts................ 93

6.1 Coefficient of determination, r^2, pertaining to linear regression model
developed between the reported yield and various NDVI-based variables......... 97

6.2 Hybrid models for the selected districts.. 98

6.3 Dataset used to develop hybrid models for the selected districts...................... 99

6.4 Comparison of various models on the basis of coefficient of
determination, r^2.. 100

LIST OF ABBREVIATIONS

ABS	Accumulated Basin Storage
AE	Actual Evaporation
ARIMA	Auto Regressive Integrated Moving Average
ARMA	Auto Regressive Moving Average
AvgNDVI	Average Normalised Difference Vegetation Index
AvgNDVI_Jul	Average NDVI during July
AVHRR	Advanced Very High Resolution Radiometer
CoefCor_NDVI	Coefficient of correlation between a NDVI profile and the reference NDVI profile
CV	Coefficient of Variation
CV1-T	Coefficient of variation in temperature from May to July
CV-P	Coefficient of variation in precipitation during summer – from May to August
CV-T	Coefficient of variation in temperature during summer – from May to August
CWB	Canadian Wheat Board
CWSI	Crop Water Stress Index
DW	Durban Watson
EC	Error Correction
EJ	Emergence to Jointing
ENSO	El Nino Southern Oscillation
ET	Evapotranspiration
FAO	Food and Agriculture Organization
HS	Heading to Soft dough
JH	Jointing to Heading
LAI	Leaf Area Index
MAI	Moisture Adequacy Index
MAPE	Mean Absolute Percent Error
MaxNDVI	Maximum NDVI

May-P	Total precipitation in May
May-T	Average temperature in May
MRSC	Manitoba Remote Sensing Centre
NDVI	Normalised Difference Vegetation Index
NOAA	National Oceanic and Atmospheric Administration
PDSI	Palmer Drought Severity Index
PE	Planting to Emergence
PET	Potential Evapotranspiration
PNA	Pacific North American
PR	Pattern Recognition
SDD	Stress Degree Days
SO	Southern Oscillation
SR	Soft dough to Ripe
StdevNDVI_Jul	Standard deviation in NDVI data during July
StdNDVI	Standard Deviation in NDVI
TSD	Temperature Stress Day
WCWY	Western Canada Wheat Yield
Win-P	Total precipitation during winter – from November to March
Win-Tavg	Average temperature during winter – from November to March
WR	Water Requirement

CHAPTER ONE

INTRODUCTION

This thesis focuses on prediction of agricultural drought for the Canadian Prairies by improving estimation of spring wheat yields. This would help devise a better marketing strategy for wheat export to achieve optimum profits.

Droughts commonly refer to the scarcity of water that can cause significant reduction in production in various economic sectors (e.g., agriculture, hydroelectric generation, water supply, and industries) existing in a region. It means different things to different people. In Saudi Arabia and Libya, droughts are recognized after 2 to 3 years without significant rainfall (Dracup et al., 1980; Sen, 1990) while, in Bali, Indonesia, any period of 6 days or more without rain is considered drought. In Egypt, any year the Nile does not flood is a drought, regardless of rainfall (Chow, 1964). It would therefore be appropriate to say that drought is chiefly a matter of perception that varies from one region to another (Heathcote, 1969). This variability in perception makes it difficult to develop an analytical definition of drought essential for its analysis.

Depending upon the criteria and approaches to defining a drought, there are more than 150 published definitions of drought (Krishnan, 1979; Wilhite and Glantz, 1987). Nevertheless, droughts can be broadly clustered into four types:

meteorological, hydrological, agricultural, and socioeconomic (Wilhite and Glantz, 1987) whose generalized definitions are presented below.

Meteorological drought: is defined as "a period of more than some specified number of days with precipitation less than some specified amount" (Great Britain Meteorological Office, 1951).

Hydrological drought is concerned with the effects of dry spells (periods without precipitation) on surface or subsurface water resources. As a result, water levels in rivers, lakes, reservoirs, and ground water resources decline.

Agricultural drought occurs when a plant's demand for water, which is dependent on prevailing meteorological conditions, biological characteristics of the plant, its stage of growth, and the physical and biological properties of soil, is not met. As a result, the crop yield (production per unit of area) is significantly reduced.

Socioeconomic drought relates to the features of the socioeconomic effects of meteorological, hydrological, and agricultural droughts. These effects may include price inflation, famine, population migration, and political upheaval.

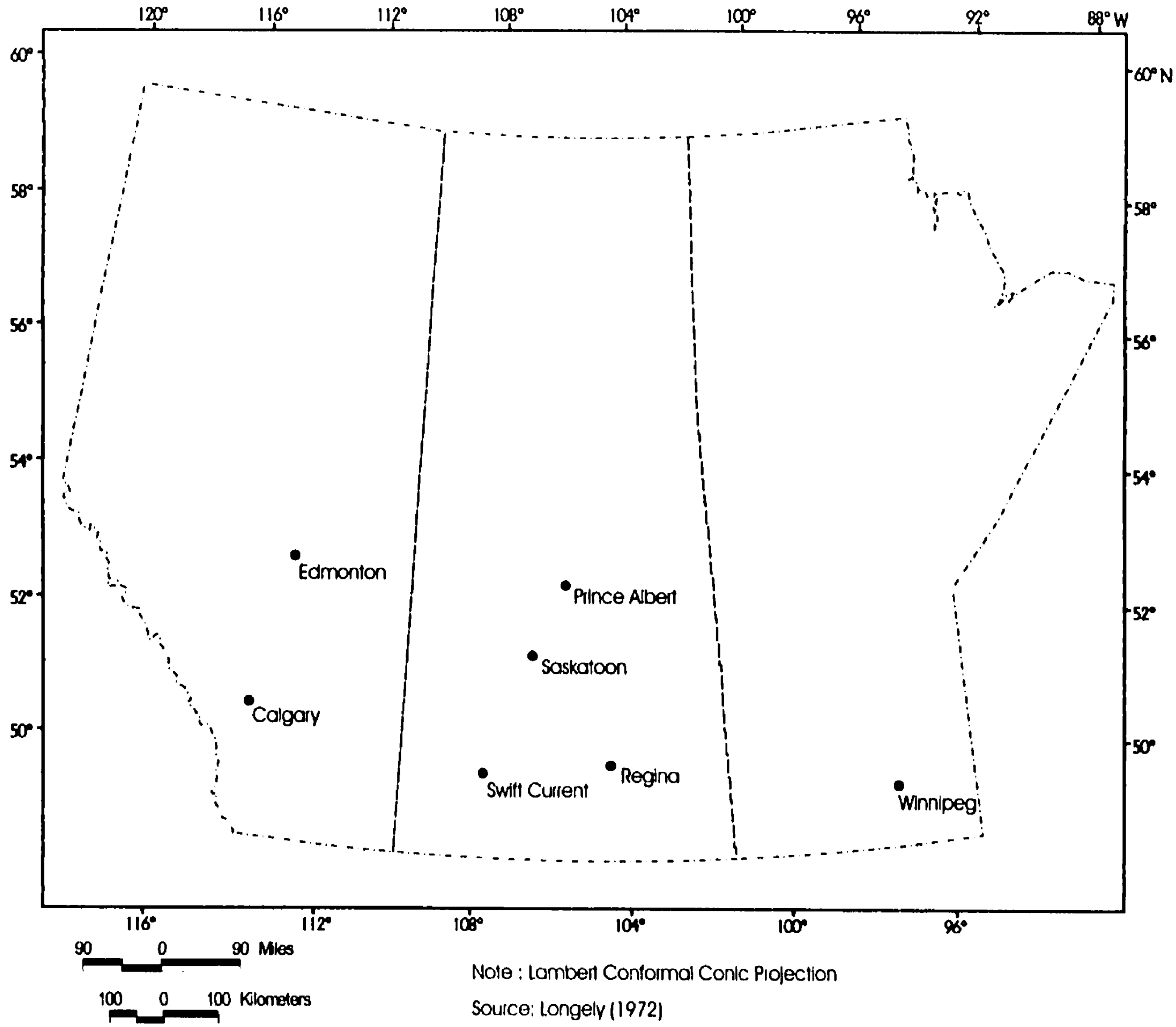
The significance of each type of drought to a region mainly depends on its agroclimatic and socioeconomic characteristics. In the case of Canadian Prairies, the agricultural drought (hereafter referred to as drought) plays an important role.

1.1 Drought Vulnerability of the Prairies

The Prairies which occupy a vast area extending northwards from 49^0 N (Canada-US border) to 54^0 N, and east-west from eastern Manitoba, across Saskatchewan to western Alberta, approximately between 96^0 W and 114^0 W longitudes (Fig. 1.1) are vulnerable to droughts. The Prairie region is a part of the Great Plains[1] and constitutes a physiographic subdivision which combines the Canadian Shield and a thin Cordilleran slice in western Alberta. In the Prairies, the main effect of drought is on agriculture (Rosenberg, 1980). In the past one hundred years, severe droughts occurred in the Prairies in the 1890s, 1910, 1914, between 1917 and 1920, 1924, 1929, and the 1930s. More recently, droughts occurred in Manitoba in 1961, 1967, 1976, 1979, 1980, 1984, and 1988 (Environment Canada, 1989). Alberta and Saskatchewan also faced severe droughts in 1984 causing millions of dollars of loss in agricultural production (Grace and Johnson, 1985). As a result of such droughts, export of agricultural products is reduced, directly impacting the Canadian economy.

The total prevention of drought is not possible as droughts form an essential component of a climatic system. Nevertheless, if droughts can be predicted the losses attributed to them can be minimized by better planning. But the drought is a creeping phenomenon and the estimation of its onset and termination is a difficult task (Gillette, 1950). Identification of a variable (i.e., determinant variable) which characterizes the drought is the first step in drought prediction.

[1] The Great Plains extend from Texas in the southern United States to the Canadian Prairies and are bounded by the Rocky Mountains in the west and by the western portions of Ohio and Missouri in the east.



Note : Lambert Conformal Conic Projection

Source: Longely (1972)

Figure 1.1 The Canadian Prairie Provinces

As the drought is linked with loss of agricultural production, crop yield can be chosen as a drought indicator. By predicting reduced crop yields, therefore, one can predict droughts (Morgan, 1985; Sinha et al., 1992).

1.2 Agricultural Drought and Crop Yield

In the event of agricultural drought, crop yields are reduced due to scarcity of rainfall. The relationship between rainfall and yield is a complex phenomenon (Smart, 1983). For an optimum yield, a crop requires a specific amount of water during its various phenological stages: emergence, jointing, heading, soft dough, and ripening (Robertson, 1968). During droughts, it is likely that yields of all crops are reduced depending on the severity of drought, though the degree of reduction in yields may vary from crop to crop due to the spatial variability in rainfall and crop-to-crop water requirements. For a region such as the Prairies, a single dominant crop can be chosen to attempt drought prediction.

Major crops grown in the Prairies are spring wheat, canola, and barley. In 1998, the harvested area under wheat, canola, and barley was 10.4, 5.4, and 4.0 million hectares, respectively (Statistics Canada, 1998). Wheat deserves special attention since its acreage is highest when compared with other crops in the Prairies. Approximately 75 percent of the total wheat produced in the Prairies is exported, which contributes significantly to the Canadian economy (McKay, 1983; Walker, 1989). The task of the wheat export is performed by the Canadian Wheat Board (CWB) -- an international grain marketing agency.

In years of favourable weather conditions (e.g., nondrought conditions), wheat export has been as high as 30 million tonnes. But, in drought years when the average temperature is significantly above normal and/or precipitation is significantly below normal, export is reduced to about 20 million tonnes. The CWB requires estimates of the wheat amount that will be available for export in the year ahead. Such an estimation is an important component to the overall export marketing strategy. If these estimates are not accurate, export decisions may prove to be less efficient. Therefore, the CWB desires the pre-harvest estimates of wheat to have a great degree of accuracy.

To estimate the amount of production, the data pertaining to both planting area and yield are required. While the area estimates are guided by the predictable economic factors influencing farmers' decisions in regards to land allocation for wheat planting, the yield estimates primarily depend on the weather conditions during the wheat growing period (the end of April or early May to the end of August or early September) and are difficult to predict. Hence, the yield estimates are likely to be more error-prone than the area estimates. In addition to being less error-prone, the area estimates can even be improved significantly by the field-visit reports that are available soon after the wheat is planted. The emphasis should therefore be given to improving the yield estimates in order to achieve greater accuracy in the production estimates.

Both short-term and long-term estimates are required by institutions like the CWB, for export planning. The long-term estimates are required during October-November preceding the wheat planting when the CWB sets broad targets for the volume of wheat to be exported in the year ahead. The short-term estimates are made around harvesttime and these estimates serve to refine the broad targets set earlier. The current practice adopted by the CWB for developing the above two types of estimates is discussed in the following section.

1.3 Wheat Yield Estimation: Current Approach in Practice

The CWB employs different approaches to obtain the long-term and short-term estimates. Long-term estimates are generated using a time series technique (i.e., trend analysis) while the short-term estimates are obtained using a model, namely, the Western Canada Wheat Yield (WCWY) model developed by Walker (1989). Both the long-term and short-term estimates are obtained for the Prairies as a single unit. The Prairies, however, consists of 8, 12, and 20 crop districts in Alberta, Manitoba, and Saskatchewan, respectively. A significant degree of variation in agroclimatic conditions and drought-tolerance level exists among these districts. It is possible that the yield variation patterns differ from one district to another, warranting the employment of techniques other than just the trend-analysis to better model the district-level yield variation. This would improve the long-term yield estimates. In order to improve the short-term estimates, additional variables need to be included in the model; at present, only two primary variables -- namely, monthly temperature and precipitation -- are used in the WCWY model.

1.4 Statement of the Research Problem

As discussed in Section 1.3, an attempt can be made to improve both long-term and short-term wheat yield estimates for the Canadian Prairies. More accurate yield estimates will lead to improved drought prediction by knowing the extent, if any, to which the yields are reduced. This will help design a more efficient export plan. In addition, the development of an improved model would enhance the understanding of additional factors impacting wheat yield. In order to construct appropriate hypotheses to address the above research problem, the pertinent literature is reviewed in the following chapter. The hypotheses are tested in subsequent chapters.

CHAPTER TWO

LITERATURE REVIEW

2.1 Geographic Perspectives of Drought

The study of drought in the contemporary disciplines, particularly in agro-climatology, geography, sociology, and anthropology, has been honed from a wider interest in climate impact studies. Drought issues are therefore addressed from various perspectives, differing widely in their content and emphasis. Recent drought studies have primarily focussed on four major areas: (i) meteorological and hydrological causes of drought, (ii) perception of drought, (iii) physical, agricultural, and socioeconomic impacts of drought, and (iv) prediction of drought. The work of geographers has in general concentrated on the first two problem areas, demonstrating that there is room for further advancement in methods and techniques to predict drought from a geographic perspective.

The problem of defining and spelling out various types of droughts is quite common in the literature. In defining agricultural drought, Heathcote (1974) suggests that the phenomenon should be viewed in terms of the shortage of water harmful to agricultural activities. Drought takes place as an interaction between agricultural activity (i.e., the demand) and natural events (i.e., the supply) which results in a water volume or quality inadequate for plant and/or animal needs. Hewitt (1983), in his seminal work, questions the benefits of a technocratic approach of classifying or dividing natural hazards; drought is one of the natural hazards as it is primarily caused by deficient precipitation, a natural phenomenon. The problem here is tackled by an extreme narrowing of the range of interpretation and acceptable empirical evidence. Such reductionist approaches are both conceptually and methodologically constrained due to three features: (i) the signficant extent to which a natural hazard is not explained by, nor uniquely dependent upon, the geophysical processes that may trigger damage; (ii) the significant extent to which human awareness of and responses to the hazard are not associated with the geophysical conditions; and (iii) the important degree to which a natural disaster, its causes, internal features, and impacts are not explained by conditions and behaviour peculiar to calamitous events (Hewitt, 1983: 24-25).

As noted above, many geographical investigations of drought occurrences attempted to identify their causes. In studying drought in northeast Brazil, Hall (1978) found out that it was attributed to rainfall shortage as well as the delay in the arrival of winter rains which were essential for growing subsistence crops. Liverman (1990) also reported that rainfall variation was the major cause of drought. Both the amount of annual rainfall, and its temporal and spatial distributions are important for drought analysis (Kelly and Wright, 1978; Garcia, 1981).

Saarinen (1966) pioneered perception studies on droughts in North America, and examined the different factors which influence farmers' perception in the Great Plains. He refuted the notion that people are strongly attached to the image of a place even though that place may be known to be hazardous to scientists and objective observers (Jackson and Mukerjee, 1974; Oliver-Smith, 1982). Saarinen asserts that such perception varies in accordance with the potential hazardousness of a place; for example, people in regions of high drought hazard potential will show a remarkable sense of drought perceptivity, while those living far from drought prone regions will have lower levels of perception (Burton et al., 1978; Brooks, 1972; 1973). The populations of the intermediate area will tend to fluctuate between two extremes in their perception. Kates (1967) has demonstrated that people might be aware of a hazard but have different interpretations of the nature of that hazard. Likewise some persons are well aware of the possibility of hazards but regard the potential danger as less significant than their attachment to the place under threat. In addition, the recall of past experiences may be less than perfect; Kirkby (1974), for example, has shown that the memory of farmers for specific rainfall events tends to be limited in duration and is mainly restricted to the largest rainfalls.

Drought impact has many dimensions and it has received attention in the geographical literature. Heathcote (1974: 135) noted that the knowledge of drought in South Australia suggested that drought has many different effects, ranging from environmental (i.e., heat waves, water salinity, and dust stroms) to economic (short falls in sales and business activities), social (community concerns and appeals for help), and demographic (population migrations). Crop loss due to drought is a common observation throughout the world although effects of such loss are felt differently in the developed and the developing world. In a study of historic droughts in the United States, Borchert (1971) identified four years of extensive drought during the 19^{th} and 20^{th} centuries: 1892, 1912, 1934, and 1953. In these years annual rainfall dropped to 15-25 percent of normal rainfall. More specifically, he discovered that the July-August rainfall had been reduced to 25-50 percent of the normal. As a result, soil moisture and ground water recharge depleted significantly, which in turn adversely affected agriculture and urban water supply in the grassland regions. Further, farm income was reduced considerably through loss of crops and livestock. In a study of Mallee, South Australia, Heathcote (1974) reported a remarkable drop in crop yields due to drought: wheat yield declined by 80 percent. Wisner and Mbithi (1974) analysed the 1970-71 drought in Kenya and observed that, in many parts of the country, rainfalls were only 50-75 percent of the long-term average. Serious production loss in the drought affected regions created pervasive starvation and social tensions. Among the social and demographic consequences of drought, mass migration has been cited by many analysts (e.g., Brooks, 1972; 1973; Warrick, 1983; Heathcote, 1974). Such sudden, unplanned, mass movements of populations often cost both the sending and host communities.

Prediction of drought has always been a peripheral issue in geographic approaches to the phenomenon, and since the aspect is a necessary component to spatial and socioeconomic planning, it deserves special attention. In particular, exploration into prediction of drought and crop yield estimation from a geographic perspective would possibly enrich both the disciplinary contribution to this area of study and the techniques and tools to address the problem.

2.2 Drought Prediction

For predicting drought, there are two approaches: (i) qualitative, and (ii) quantitative. With the qualitative approach, drought is predicted only in terms of whether or not it would occur. However, the quantitative approach focuses on predicting drought on the basis of quantitative yield-estimation of the main crop in a region (Kumar, 1998). If the estimated yield of the main crop is found to be significantly lower than its long-term average, it is predicted that drought will occur (Sinha et al., 1992). In the following section, both qualitative and quantitative methods of drought prediction are evaluated.

2.3 Drought Prediction Using Pattern Recognition

Pattern Recognition (PR) is a computerized process used to classify objects that can be characterized using numerical data. Applications of pattern recognition are found in numerous fields such as image processing, medical engineering, criminology, speech recognition, and signature identification (Duda and Hart, 1973). In geography, pattern recognition can be applied to predict various spatial phenomena including drought (Kumar and Panu, 1994; Kumar et al., 1997). Defining a pattern vector is a first step in the process of pattern recognition; a pattern vector is a linear equation based on the variables characterizing the pattern (or phenomenon).

In the case of drought, a pattern vector can refer to a drought vector, as explained in the following.

2.3.1 Formation of a drought vector

In order to form a drought vector, a quantitative definition of drought is adopted first and then the nominal variables (related to a drought definition) are chosen as elements of the drought vector. There are number of ways in which a drought can be defined. One of the common ways is based on the comparison of yield of the main crop in the area with its long-term average. If the yield is significantly lower than the average yield, the yield can be termed as a case of a drought.

With crop yield being a basis to define drought, it would be appropriate to define a yield vector prior to defining a drought vector. In the case of a yield vector, the elements can be yield-influencing variables. Various variables (e.g., farming techniques, fertilizers, pesticides, irrigation, seed variety, and weather) influence yield, but, in the Prairies, weather (e.g., temperature and precipitation) is most significant. A yield vector may consist of two or more elements (or variables). Just for the sake of an example, two elements of a yield vector may be: (i) the total precipitation during the growing period, and (ii) the average temperature during the growing period. Such a yield vector can then be shown as a point in a two-dimensional display (Fig. 2.1).

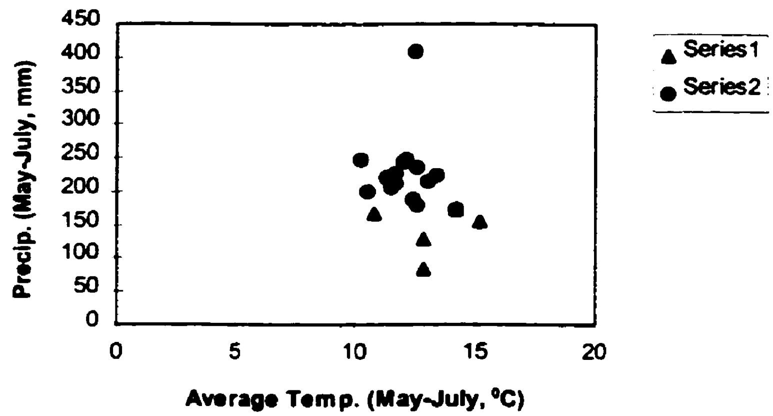


Figure 2.1 An example of yield vectors;
Series_1: Drought, Series_2: Nondrought

In the above figure some of the points refer to drought while the remaining points indicate the cases of nondrought, depending on whether or not the corresponding yield is significantly lower than the average yield. But how to specify the term 'significantly lower'? To address the problem of defining a drought, a percent value may be selected. In Fig. 2.1, for example, a value of 25 percent is selected, just to demonstrate the process of pattern recognition. Based on this criterion, if the yield is lower than 75 percent of the average yield, the related yield vector is labeled as a drought vector. The error-correction procedure of pattern recognition is then applied to test if a linear dividing line can be determined to separate the two categories. In the event of the test concluding positively, the dividing line (a vector) can be used to classify an unknown vector as belonging to one of the two categories. In other words, if the elements of a yield vector are known, prediction could be possible prior to the crop harvest as to whether or not a drought will occur. Existence of the vector linearly separating drought and nondrought categories can be tested using the Error Correction (EC) procedure as described in the following section.

2.3.2 Error Correction Procedure

The EC procedure of pattern recognition is applied to distinguish two different patterns -- in the present case, drought and nondrought patterns. Before applying this procedure, derivation of data is required, as shown in Table 2.1.

Table 2.1 contains annual yield data for 1975-91 period, leaving the 1992-96 data for testing. The yield in each year is categorized as drought or nondrought on the basis of a chosen definition of drought. Elements of yield vectors are also decided. Then the EC procedure is applied in three steps: (i) an additional element of 1 is included into the elements of all of the yield vectors, (ii) all of the elements in every vector of the second category (i.e., drought vector) are multiplied by -1 (Table 2.1), and (iii) a solution vector, W, is determined such that the product of W with any yield vector, Y_i, exceeds zero. That is,

$$Y_i W > 0 \qquad \text{for all } i \qquad [2.1]$$

where i, which is used to identify a vector, ranges from 1 to the total number of vectors (Duda and Hart, 1973).

Table 2.1 Vectorizing the wheat yield for district 3bn, Saskatchewan

Yield	Yield (t/ha)	% Deviation from mean yield	Average Temp. during May-Aug. (^{o}C), Sum_Tavg	Total Precip. During May-Aug. (mm), Total-P	Augmenting Element
1975	1.67	-2.77	15.1	198.2	1
1976	2.21	28.67	16.6	245.2	1
1977	2.04	18.77	15.7	245.1	1
1978	1.71	-0.44	16.3	159.7	1
1979	1.53	-10.92	-15.9	-167.1	-1
1980	1.54	-10.34	-15.9	-216.0	-1
1981	1.8	4.80	16.1	239.2	1
1982	2.21	28.67	14.8	307.3	1
1983	1.88	9.46	16.5	223.5	1
1984	1.28	-25.48	-16.9	-159.0	-1
1985	0.91	-47.02	-16.1	-97.5	-1
1986	2.08	21.10	16.2	243.7	1
1987	1.86	8.29	16.5	184.2	1
1988	0.5	-70.89	-18.8	-167.3	-1
1989	1.67	-2.77	16.6	274.5	1
1990	2.08	21.10	16.0	233.2	1
1991	2.23	29.83	16.7	354.9	1

[Source: Statistics Canada for yield data; weather variables are derived from the monthly data collected from Environment Canada]

To start the process of determining W, W is assumed to be a unit vector (i.e., every element is 1). Then the product of W with the individual yield vector is computed. If the condition (Equation 2.1) is not satisfied, the W is corrected as follows:

$$W_{k+1} = W_k + c/k * Y_i \qquad [2.2]$$

where c is greater than zero (usually chosen as 1); k, whose initial value is zero for unit vector, W, is increased by 1 every time a correction in the W is required. Y_i is the yield vector whose product with the W does not exceed zero, and as a result, a correction in the W is sought. This process of correcting the W continues until Equation 2.1 is satisfied. If a W is determined, it would be possible to distinguish drought vectors from nondrought vectors. In other words, it would be possible to predict drought occurrence.

With the quantitative approach to drought prediction, yield of the main crop in the region is estimated first. Then the drought severity is gauged based on the deviation of the estimated yield from the average yield (Kumar, 1993; Kumar and Panu, 1997). In the present context, techniques of quantitative yield estimation can be broadly classified into two groups: (i) techniques that can be employed for long-term estimation of yields (i.e., those techniques that do not employ weather data during the growing period because such data are not available at the time of yield estimation), and (ii) the techniques that employ weather data during the growing period to obtain short-term estimates of yield. Time series analysis holds potential for obtaining the long-term yield estimates. This is discussed in the following section.

2. 4 Yield Forecasting Using Time Series Analysis

Time series analysis is performed in various academic fields to forecast various events (Box and Jenkins, 1976; Prankatz, 1983); nevertheless, such methods have not been extensively explored for predicting drought or forecasting yields.

To develop a time series model for forecasting, the variable to be forecasted (yield, in the present case) is modeled as a function of time (Equation 2.3; Abraham and Ledolter, 1983), using past yield data. Thus,

$$Y_t = f(t) + \varepsilon_t \qquad [2.3]$$

where Y_t is the yield at time *t* (i.e., a specific year), f (t) is a function of time *t* and ε_t refers to errors.

Once a functional relationship between yield and time is established, the yield can be forecasted for the next year. The process of developing a time series model begins by testing whether the time series under study is stationary or nonstationary. By determining this, one can choose specific groups of techniques meant to model the respective series.

2.4.1 Series categorization: stationary or nonstationary

A time series can be broadly categorized as: (i) stationary, or (ii) nonstationary. If the statistical properties (mean, variance and autocorrelation) of the series are independent of time, the series is categorized as stationary. Otherwise, it is referred to as nonstationary.

Various methods for testing whether a series is stationary or not are available in the literature, but the Unit Root Test has been widely applied in recent years (Gujarati, 1995). To execute this test, the following regression equation is developed using the series data:

$$Y_t = \rho Y_{t-1} + u_t \qquad [2.4]$$

where Y_t is yield at time t, Y_{t-1} is yield at time t-1, ρ is a coefficient, and u_t is the stochastic error that follows the classical assumptions, namely, it has zero mean, a constant variance, and is nonautocorrelated. A null hypothesis ($\rho = 1$) is then tested using Augmented Dickey-Fuller test[1]. If the hypothesis is accepted, it is concluded that the series can be treated as stationary.

Such categorization of series is a prerequisite to the selection of techniques to model the series. Nonetheless, a series tested as nonstationary may be transformed to a stationary series. Such a transformation will pave the way for application of techniques otherwise extraneous to the nonstationary series. In addition, the transformation will enable employment of an increased number of techniques to model the series, which is desirable since the more models developed to describe a series, the better the selection of the most appropriate model for the series.

An easy and commonly used method for transforming a nonstationary series to a stationary one is the method of differencing. By applying this method to the present case, a nonstationary yield series is transformed to another series by subtracting Y_{t-1} from Y_t (when the level of differencing is 1), Y_{t-2} from Y_t (when the level of differencing is 2), and Y_{t-3} from Y_t (when the level of differencing is 3), and Y_{t-n} from Y_t (when the level of differencing is n). To begin the process of

---

[1] In order to test the null hypothesis, the absolute Dickey-Fuller τ statistic (i.e., estimated ρ divided by its standard error) is compared with absolute critical τ statistic. If the computed τ statistic exceeds the critical τ statistic (-3.13, at the 10 percent confidence level which is a default value in SHAZAM software), the null hypothesis is not rejected and, in turn, the series is inferred as stationary. Conversely, if the τ statistic falls short of the critical τ statistic, the series is concluded as being nonstationary.

transformation, the level of differencing is chosen as 1 and the resulting series is tested for stationarity following the procedure described earlier. If the transformed series is still tested as nonstationary, the next higher level of differencing is chosen and the process of transformation is repeated to see if the series is transformed to a stationary series. If not, the next higher level is chosen and the rest of the procedure is repeated. Usually, up to a third level of differencing is sufficient to transform the series.

After confirming whether a yield series (original or transformed) is stationary or nonstationary, techniques can be selected to model the series for the purpose of yield forecasting. Cases for the stationary and nonstationary series are discussed below.

2.4.2 The case of the stationary series

Stationary series can be modelled using simple moving averaging, simple exponential smoothing, and the Box-Jenkins techniques (Box and Jenkins, 1976). If the series is or can be transformed to a stationary series, the following techniques can be applied to the transformed series.

2.4.2.1 Simple moving averaging

Averaging methods are of two types: simple, and moving averaging methods. In simple averaging methods the mean of all the data is used to forecast the next period.

In contrast, in moving averaging methods the analyst is more concerned with recent observations. A constant number of data points is specified at the outset and a mean is computed for the specified number of the most recent observations. As the new observation becomes available, a new mean is computed by dropping the oldest variable and including the newest one. The moving average is used to forecast the next period as expressed by the following equation:

$$M_t = Y_{t+1} = (Y_t + Y_{t-1} + \ldots . Y_{t-n+1}) / n \qquad [2.5]$$

where M_t = moving average at time t, Y_{t+1} = forecast value for next period i.e., $t+1$ Y_t = actual value at period t, and n = number of terms in the moving average.

In averaging methods, equal weights are given to all observations. If more weights are to be given to the most recent observations, exponential smoothing is selected to model the series.

2.4.2.2 Simple exponential smoothing

This method is based on averaging past values of a series in a decreasing (exponential) manner. The weights used are α (for the most recent observation), $\alpha(1-\alpha)$ for the next observation, $\alpha(1-\alpha)^2$ for the next, and so on. The constant α is termed as the smoothing constant and is assumed to range from 0 to 1. The actual value of α determines the extent to which the most current observation is to influence the forecast. The resulting exponential smoothing leads to the following equation:

$$\hat{Y}_{t-1} = \propto Yt + (1-\propto)\hat{Y}_t \qquad [2.6]$$

where, $\hat{Y}_{t-1}$ = new forecast (smoothed value) for the next period, Y_t = new observation or actual value of series in period *t*, and $\hat{Y}_t$ = old forecasted value.

In order to estimate the smoothing constant, an iterative procedure that minimizes the mean squared error in the forecasts is used. Forecasts are computed for α equal to 0.1., 0.2,...., 0.9 and the sum of squared forecast error is computed for each. The value of α producing the smallest error is chosen for use in generating a future forecast.

2.4.2.3 Box-Jenkins technique

In the Box-Jenkins models, Y_t (yield at time *t*) is modelled as a linear combination of the past yields. If the series at time *t* is found to have regressed on itself at lagged time (Equation 2.7), it is said that an autoregressive (AR) component exists in the series.

$$Y_t = \sum_{j\geq 1} \pi_j Y_{t-j} + \varepsilon_t \qquad [2.7]$$

where π is a coefficient, *j* is an index, and Y_t and ε_t are the same as explained in Equation 2.3.

Such an autoregressive representation can lead to models with many parameters that may be difficult to interpret. However, the autoregressive models can be approximated by Auto Regressive Moving Average (ARMA) models of the form:

$$Y_t = \phi_1 Y_{t-1} + \ldots + \phi_p Y_{t-p} + \varepsilon_t - \theta_1 \varepsilon_{t-1} - \ldots - \theta_q \varepsilon_{t-q} \qquad [2.8]$$

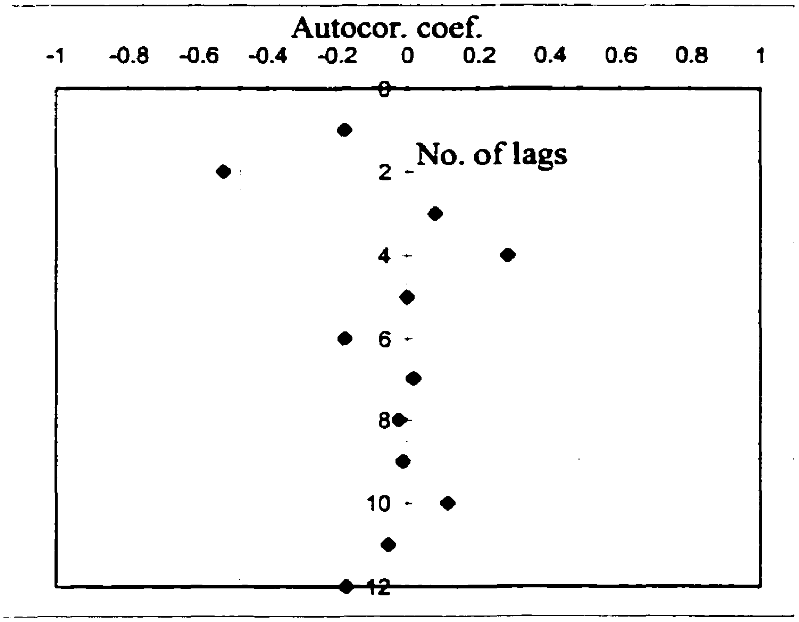
where p and q refer to the orders of the autoregression and of the moving averages, respectively; ϕ and θ are coefficients.

If the original series is nonstationary and is transformed to a stationary series after *d* levels of differencing, the ARMA model is termed as the ARIMA (Auto Regressive Integrated Moving Average) model and is specified as ARIMA (p, d, q). Examination of the possibility of an ARIMA model describing the series is accomplished in two steps: (i) identification of model specification, and (ii) diagnostic test. In the first step, the orders (*p* and *q*) of the ARIMA are estimated, and the level of differencing *d* is also determined following the procedure described earlier. In order to estimate *p* and *q*, the correlogram[2] is utilized.

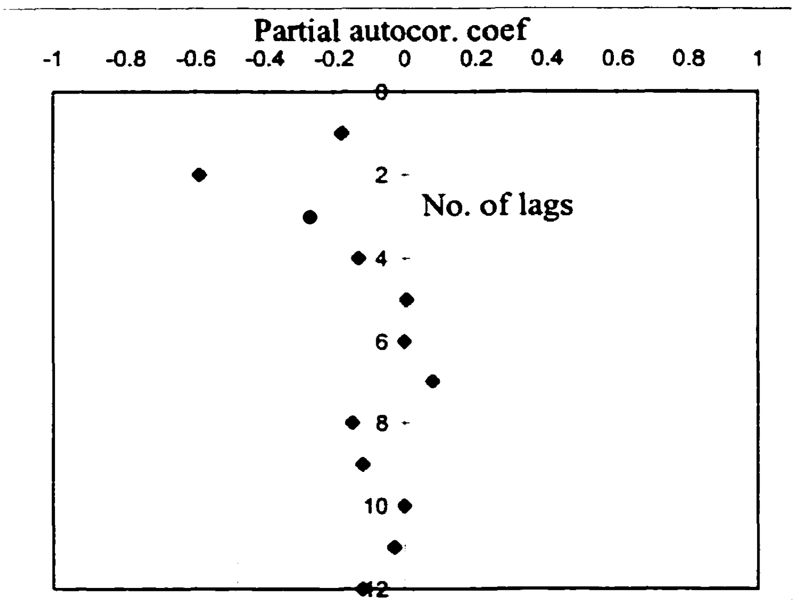
After identifying an ARIMA model fitting the series, a diagnostic test is conducted to examine if the model meets the assumptions regarding the residuals. Autocorrelation coefficients and partial autocorrelation coefficients of residuals are determined and plotted against lag-k, similar to Fig. 2.2.

If these coefficients are not statistically significant at any lag, it is concluded that the residuals are purely random and, in turn, the identified ARIMA model is concluded as fitting the series. Otherwise, the model is modified and the diagnostic test is repeated.

[2] Correlogram is a graph showing variation in the autocorrelation coefficient or partial autocorrelation coefficient with number of lags as illustrated in Fig. 2.2 (Gujarati, 1995).



(i)



(ii)

Fig. 2.2 Variation of autocorrelation and partial autocorrelation coefficients with the number of lags

2.4.3 The case of the nonstationary series

The following are some of the techniques that are relevant to the present case and can be applied to model a nonstationary series.

2.4.3.1 Trend analysis

Because of numerous factors (e.g., technological, fertilizer/pest management practices, and use of improved seed varieties), the yield series may show a linear or nonlinear trend. The linear and nonlinear (i.e., quadratic) trends are defined by Equation 2.9 and Equation 2.10, respectively.

$$Y_t = \beta_0 + \beta_1 t + \varepsilon_t \qquad [2.9]$$

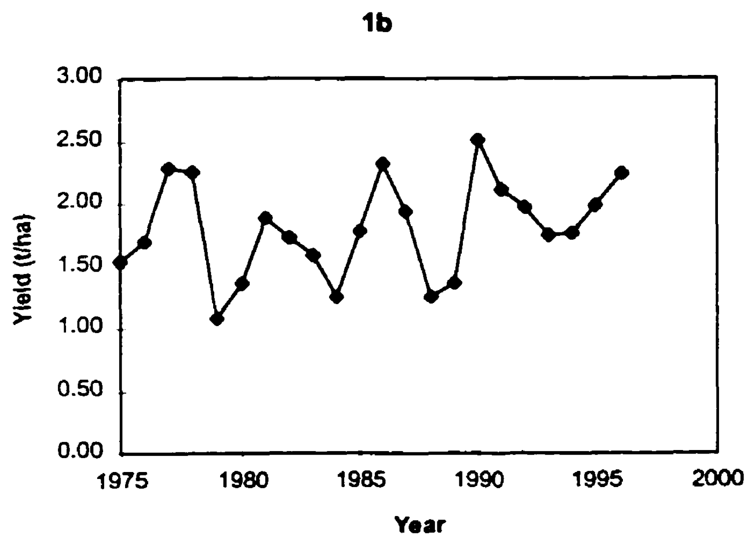
$$Y_t = \beta_0 + \beta_1 t + \beta_2 t^2 + \varepsilon_t \qquad [2.10]$$

where β_0, β_1, and β_2 are coefficients; Y_t , t, and ε_t are the same as defined in Equation 2.3.

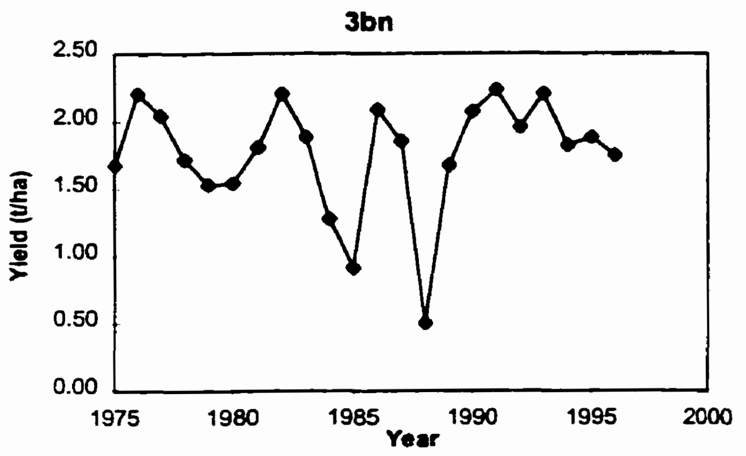
In addition to the trend, a series may have a cyclic variation (wave like fluctuation within the trend). In the present case, however, a cycle of constant amplitude and time period is not apparent, but a cycle of irregular amplitude and time period does appear (Fig. 2.3). To model such a nonstationary cycle, a double moving averaging method is required (Hanke and Reitsch, 1995).

2.4.3.2. Double moving averaging

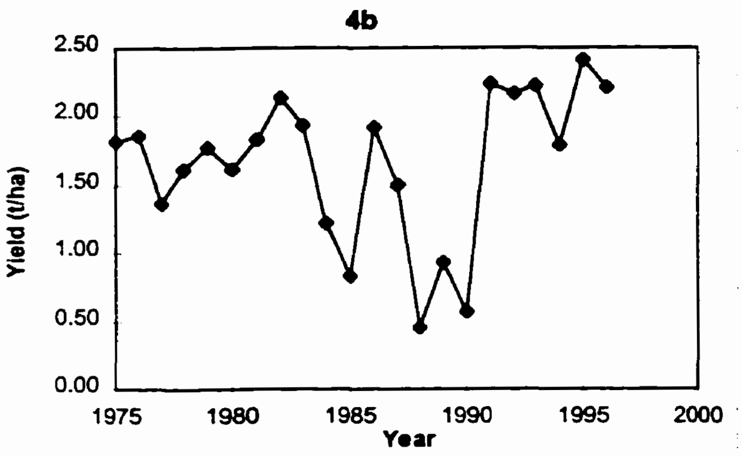
A double moving averaging method is used when a nonstationary series has a linear trend. In this method, one set of the moving averages is computed first and then a second set of the moving averages is computed from the first one. In summary, Equations 2.11, 2.12, 2.13, 2.14, and 2.15 are utilized to compute the first moving average, the second moving average, the difference between the two, the slope, and the forecasts for *m* periods, respectively.



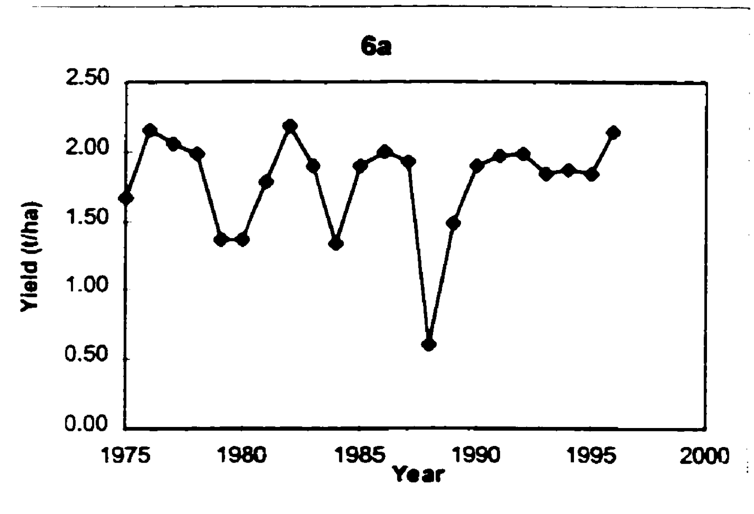
(i) Yield series for district 1b



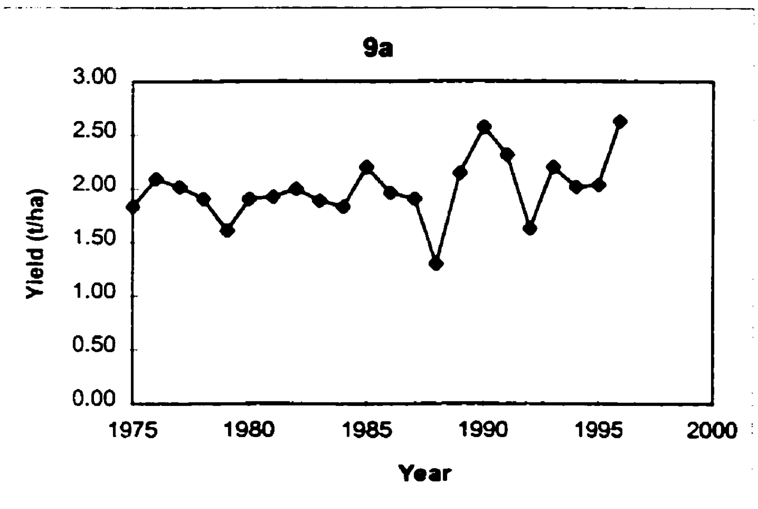
(ii) Yield series for district 3bn



(iii) Yield series for district 4b



(iv) Yield series for district 6a



(v) Yield series for district 9a

Fig. 2.3 Wheat yield variation in the selected crop districts of Saskatchewan during 1975-96
[Source for yield data: Statistics Canada]

$$M_t = (Y_t + Y_{t-1} + Y_{t-2} + \ldots + Y_{t-n+1}) / n \qquad [2.11]$$

$$M`_{t+1} = (M_t + M_{t-1} + M_{t-2} + \ldots + M_{t-n+1}) / n \qquad [2.12]$$

$$a_t = 2\,M_t - M`_t \qquad [2.13]$$

$$b_t = 2\,(M_t - M`_t) / (n-1) \qquad [2.14]$$

$$\hat{Y}_{t+p} = a_t + b_t p \qquad [2.15]$$

where M_t is the first moving average, average yield in a period of *n* years, n = number of periods in the moving average, $M`_t$ = second moving average, average of *n* terms of the first moving averages, a_t = difference between the first and second moving averages, b_t = slope i.e., the difference (between first and second moving averages) per year, and $\hat{Y}_{t+p}$ = forcasted yield for the year $t + p$.

2.4.3.3 Double exponential smoothing

If the series has a linear trend, Holt's two parameter method (Hanke and Reitch, 1995) can be used to model the series. Holt's technique makes the trend and slope smooth, directly by using a different smoothing constant for each. Holt's technique uses Equations 2.16, 2.17, and 2.18 for exponentially smoothing, trend estimating, and forecasting the series, respectively:

$$A_t = \alpha Y_t + (1-\alpha)(A_{t-1} + T_{t-1}) \qquad [2.16]$$

$$T_t = \beta(A_t - A_{t-1}) + ((1-\beta) T_{t-1} \qquad [2.17]$$

$$\hat{Y}_{t+p} = A_t + pT_t \qquad [2.18]$$

where A_t = new smoothed value, α= smoothing constant, Y_t = actual value of series in period t, β = smoothing constant for trend estimate (ranging between 0 and 1), T_t = trend estimate, p = period to be forecasted into the future (1 in the present case), and $\hat{Y}_{t-p}$ = the forecast for p periods.

The procedure outlined above (Section 2.2) to model yield series can be utilized to select an appropriate technique to obtain long-term yield estimates; the procedure requires only yield data in the past years. However, short-term yield estimates are obtained using techniques that require weather data.

2. 5 Yield Estimation Using Weather Data

Weather during the crop-growing season is the most important factor determining crop yield. There are two distinct approaches which employ weather data to estimate crop yield: deterministic and stochastic. The deterministic approach generally treats the dynamics of plants through a set of mathematical expressions tying together the interrelationship of plant, soil, and climatic processes. The stochastic approach, however, employs a sample of annual yield data together with the corresponding weather data, and relates them through statistical techniques such as multiple regression (Parry et al., 1988). Furthermore, the deterministic approach requires specific information on soil, plant parameters, planting dates, and certain agro-meteorological factors to develop even a simple model. Because of their extensive data requirements, the application of the deterministic approach is restricted to experiment-based studies. Over large regions, conducting the experiments needed to develop a deterministic model is complex and practically infeasible. In contrast, development of stochastic models is far simpler. Therefore, stochastic models have been commonly used in yield estimation (Bair, 1977, 1982; Sakamoto, 1981; Robertson, 1983; World Meteorological Organization, 1985).

The development of a stochastic model requires, as a first step, identification of possible factors influencing yield. According to van Diepen and Wall (1996), yield influencing factors can be categorized as: (i) abiotic factors such as soil water, soil fertility, soil texture, soil taxonomy class, and weather, (ii) farm management factors such as soil tillage, soil depth, planting density, sowing date, weeding intensity, manuring rate, crop protection against pests and diseases, harvesting techniques, post-harvest loss, and degree of mechanization, (iii) land development factors such as field size, terracing, drainage, and irrigation, (iv) socioeconomic factors such as the distance to markets, population pressure, investments, costs of inputs, prices of output, education levels, skills, and infrastructure, and (v) catastrophic factors that include warfare, flooding, earthquakes, hailstorms, and frost. Measurement or estimation of some of these factors is often not feasible and the influence of some other factors may be considered insignificant or constant in a local environment. As a result, a limited number of primary variables derived from simple and measurable factors is used for yield modeling. While few of these variables can be directly measured, others can be estimated.

Idso et al. (1978) used albedo measurements (i.e., ratio of reflected light to incident light on the crop) during the crop growing season for wheat yield estimation. The minimum value of albedo prior to grain ripening (when the albedo value dramatically increases) was found to be linearly related to the yield; that is to say, the lower the albedo prior to ripening, the greater the yield.

Sakamoto (1978) employed a moisture anomaly index (Z index) based on the difference between monthly observed precipitation and climatically appropriate precipitation to estimate wheat yield in south Australia. The climatically appropriate precipitation, P, was estimated from the following equation:

$$P = ET + R + RO - L \qquad [2.19]$$

where ET, R, RO, and L are evapotranspiration, recharge, runoff, and losses, respectively.

The Z index, together with temperature departure from the long-term average, was used as a variable in multiple regressions. The Z index was found to be a better indicator of yield than the ratio of ET to PET (potential evapotranspiration). Estimation of yield by this method requires derivation of P from parameters that are estimated on the basis of certain assumptions. Because of these reasons, such methods are found unsuitable for application to large areas.

For northwest Iran, Slabbers and Dunin (1981) developed the following equation to estimate wheat yield:

$$Y = 6.3\ E \qquad [2.20]$$

where E is the total crop water-use (total evapotranspiration during the critical stage of the growing period, from mid May to the end of June). Though this technique is fairly simple and requires only one parameter to estimate yield, a linear relationship between P and E may not be guaranteed under drought conditions. A complete distribution of crop water-use over the entire growing period may be more helpful.

Diaz et al. (1983) used canopy temperature indices to estimate evapotranspiration and yield of spring wheat in Utah. In particular, they utilized Stress Degree Days (SDD, the difference between air temperature and wheat foliage temperature), Temperature Stress Days (TSD, the difference between wheat foliage temperature and wheat foliage temperature of fully irrigated fields), and Crop Water Stress Index (CWSI, 1 - ratio of evapotranspiration to potential evapotranspiration). Their linear regression analysis showed that the SDD was the most suitable index for wheat yield estimation for the case study considered. The wheat foliage temperature as used in the SDD is an indicator of moisture present in the foliage. In moisture-stressed or drought conditions, the foliage temperature tends to approach the air temperature. Though the air temperature data are commonly available, the foliage temperature data are not. Further, it is rather too cumbersome to collect the foliage temperature data over large regions.

Rainfall (or precipitation) - a simple indicator of drought - has been widely used to predict droughts (Prouit et al., 1986; Sinha et al., 1992; Kumar, 1998). If the amount of rainfall during a cropping season is significantly less than its long-term average, droughts are likely to occur (Kelly and Wright, 1978; Grace and Johnson, 1985; Kumar, 1988). But such drought predictions require prior predictions of the seasonal rainfall, which is presently not possible because of complexities involved in land-ocean-atmosphere interactions (Shukla, 1985; Kates et al., 1985).

An oceanic event, the Southern Oscillation (SO), has been correlated with drought occurrence in some parts of the world (Lockwood, 1986; Bonsal and Chakravarti, 1993). The SO, first defined by Walker and Bliss (1932), is a slowly varying atmospheric pressure differential over the eastern and western regions of the tropical Pacific. When SO is combined with sea surface temperature anomalies over the equatorial central and eastern Pacific, the event is termed an El Niño-Southern Oscillation (ENSO) which has been found to cause global weather anomalies in the world. Garnett and Khandekar (1992) analyzed ENSO data and found it associated with droughts in Indian monsoon followed by low grain yields over south Asia and Australia and high grain yields over the North American prairies.

For north America, Wallace and Gutzler (1981) utilized the Pacific North American (PNA) index to make long-term weather forecasts. The PNA index is defined as a linear combination of normalized geopotential height anomalies at the 700 mb level at four selected locations across North America and the North Pacific. Knox and Lawford (1990) found a relationship between mean Canadian Prairie precipitation anomalies and contemporaneous anomalies of the Northern Hemisphere circulation at the 50 kPa level; dry months on the Prairies are associated with characteristic 50 kPa anomalous circulation in the Northern hemisphere. Garnett et al. (1998) used monthly values of the ENSO and PNA indices to predict summer season weather over crop growing region of the Canadian Prairie provinces. Their multiple regression analyses revealed distinctly different profiles of accumulated values of the indices for the hottest and coldest or driest and wettest summers over the Canadian Prairies.

The ENSO event may provide some qualitative indication of drought by predicting whether a seasonal rainfall would be less than the average rainfall. But for quantitative analysis of drought, the ENSO data may not suffice because the drought analysis is linked more to the distribution of rainfall over the cropping season than the total rainfall (Garcia, 1981; Smart, 1983).

Considering the complex relationship that exists between yield and rainfall, numerous indices have been formulated to analyze droughts. Some commonly used indices are presented in the following.

2.5.1 Palmer Drought Severity Index

Based on the soil-water budgeting, Palmer (1965) used the difference between actual precipitation and precipitation requirement, and determined an index, namely, the Palmer Drought Severity Index (PDSI) to evaluate drought severity in time and space.

This index is extensively and routinely used in the United States for agricultural planning and estimating crop yields (Rao and Padmanabhan, 1984). But the application of the PDSI to a large area is hindered by the difficulty in estimating factors such as runoff and evapotranspiration (Yevjevich et al., 1978). These factors must be estimated in order to determine the PDSI.

2.5.2 Accumulated Basin Storage

Booy and Lye (1986) developed a parameter, Accumulated Basin Storage (ABS), which was defined as equivalent to the total amount of water stored in the drainage basin at any time. Awumah et al. (1990) considered the ABS as a possible indicator of drought, and correlated it with crop yield. When compared with other drought indicators, such as summer precipitation, total runoff, and evaporation, the ABS was found to be the best indicator of drought; while water levels in streams, lakes and rivers are indicators of runoff, evaporation can be measured by various methods such as those described in Laycock (1964). Nevertheless, the ABS is applicable more to a drainage basin which is a natural topographical unit, and not to a vast geographic region such as the Prairies which comprise various differing topographical units.

2.5.3 Moisture Adequacy Index

According to the procedure adopted by the Food and Agriculture Organization (FAO) of the United Nations, the Moisture Adequacy Index (MAI) is defined as the ratio of the actual evaporation (AE) from the crop to the crop water requirement (WR) (Doorenbos and Pruitt, 1977). The procedure requires rainfall, evapotranspiration, and information on soil data to determine the periods of deficit or surplus soil moisture over the cropping season. Following this procedure, Kumar (1993) calculated the MAI values at weekly intervals during the growing period of pearl millet, one of the main crops in arid and semi-arid regions of India. The values of the MAI varied between 0 and 1. As long as the water requirement for the crop was fully met, the MAI remained at 1, otherwise it declined, reflecting soil moisture deficits during the growing season, perhaps leading to a drought. On the basis of regression analyses, the MAI was found to be a significant indicator of drought. The MAI can be a potential indicator of drought for large, but agro-climatically homogeneous, areas. This procedure needs only rainfall as a primary input; its other parameters can be conveniently derived following standard FAO procedures.

2.5.4 Yield estimation for the Canadian Prairies

In the Canadian Prairies, both stochastic and deterministic models have been utilized for crop yield estimation. Using total crop water-use (evapotranspiration from the planting date to the ripening date) as a variable, Campbell et al. (1988) used field experiments to develop the following equation for wheat yield prediction in Saskatchewan:

$$Y = 9.2\ \{(\textstyle\sum_{P}^{R} ET) - 71.8\} \qquad [2.21]$$

where ET is daily evapotranspiration, $\sum_P^R ET$ is total crop water-use, R is the Julian ripening date, and P is the Julian planting date. This technique is conceptually similar to one developed by Slabbers and Dunin (1981) using ET during the critical crop stage.

Using the concept of ET and days to maturity (R-P), Raddatz et al. (1994) developed the following model to estimate average yield for spring wheat over the Canadian Prairies:

$$Y = 10.8((\sum_P^R ET / \sum_P^R (cu.ET_p))(R-P))^{1.2} \qquad [2.22]$$

where *cu* stands for "consumptive use factor", the ratio of actual water demand of a crop growing in a field to potential evapotranspiration. For wheat, It ranges from about 0.3 to 1. The application of a power function of days-to-maturity, rather than a linear function, led to the marked improvement in yield prediction. Up to 69 percent of the variation in the observed yield of spring wheat could be explained by the model.

The yield estimation model presently being used by the Canadian Wheat Board was developed by Walker (1989) and is referred to as the Western Canada Wheat Yield (WCWY) model. The model determines a drought index (DI) for all of the weather stations across the Prairies and then computes an average for the entire Prairies. The average DI is regressed against average yield for the Prairies to develop the WCWY model. The DI basically reflects the extent to which soil moisture supply and transpiration demand limit crop growth, and is calculated as a function of the cumulative water supply ($\sum S$, all the water that is cumulatively available to the crop , starting in the preceding summer) and transpiration demand ($\sum T$). While the $\sum S$ is estimated using monthly precipitation, the $\sum T$ is determined using monthly temperature. Although the model uses monthly data, the drought index is computed on a daily basis and is accumulated over the growing period. Daily precipitation, maximum temperature, and minimum temperature are required to compute the daily DI. Daily precipitation is estimated by dividing the average monthly precipitation by the number of days in a month. Daily maximum and minimum temperatures are estimated by adding six degrees and subtracting six degrees from the average monthly temperature, respectively. It should be noted here that the daily data, thus estimated, remain constant for a month. This is at variance with the reality of the situation. Use of directly measured daily data is therefore likely to improve the model performance. Further improvement in the WCWY model may be possible by including additional parameters estimated from satellite data which are capable of reliably monitoring crop conditions at frequent time intervals.

2.6 Satellite Data in Yield Estimation

Monitoring physical conditions of crops helps in predicting droughts. Satellite data have been applied in crop-condition assessments because of their better spatial, spectral, and temporal characteristics, and their sensitivity to changes in crop-vigour/health (Curtis, 1978). For large areas, National Oceanic and Atmospheric Administration (NOAA) satellite data have been preferred because of their wide swath (about 2700 km). The NOAA satellite is equipped with the Advanced Very High Radiometric Resolution (AVHRR) sensor that collects data in five spectral channels. Only two channels, red (0.58-0.68 μ m) and infrared (0.725-1.10 μ m), have commonly been utilized to develop indices for monitoring vegetation conditions (Tucker, 1979; Tarpley et al., 1984; Gutman, 1987, 1991; Goward et al., 1991). As the agrometeorological conditions vary, a change in the numerical value of the vegetation index is expected. Some of the agrometeorological factors influencing the vegetation index are soil moisture, weather, fertility, irrigation (Thompson and Wehmanen, 1979; Barnett and Thompson, 1982; Brown et al., 1982; Asrar et al., 1985) and other crop-growth indicators (Ahlirchs and Bauer, 1983; Hatfield, 1983; Brown, 1986; Clevers, 1989).

Among various vegetation indices, the NDVI (Normalized Difference Vegetation Index), as defined below, has been preferred to monitor vegetation conditions over large areas.

$$NDVI = (IR-R)/(IR + R) \qquad [2.23]$$

where R and IR are the reflectance in the red and infrared bands of the NOAA/AVHRR sensor, respectively. The NDVI values tend to be higher for more healthy (i.e., green) and dense vegetation, and lower for less healthy and sparse vegetation. A typical profile of the NDVI (i.e., variation of NDVI from the beginning to the end of cropping season) under nondrought conditions is shown in Figure 2.4 for the Swift Current district, Saskatchewan.

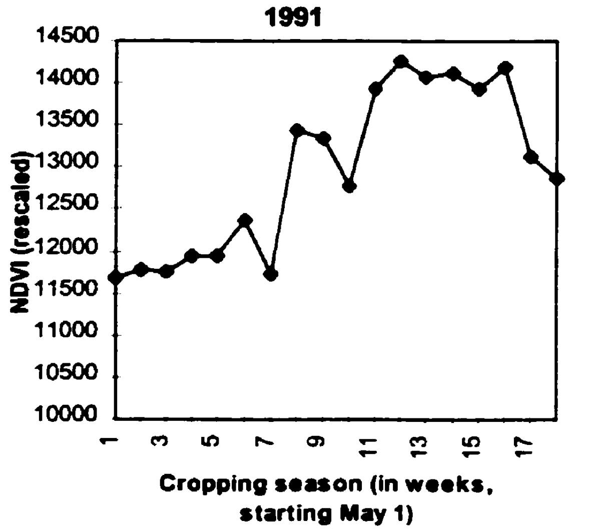


Figure 2.4 NDVI profile under nondrought conditions for district 3bn, Saskatchewan [Source for NDVI data: Spatial Analysis and Geomatics Applications division, Statistics Canada, Ottawa]

The nondrought conditions refer to crop conditions when the crop water requirement is fully met throughout the process of crop development. The complete crop development comprises various crop phenological stages, namely, emergence, jointing, heading, soft dough, and ripening (Robertson, 1968). The period from emergence until the onset of heading stage is usually referred to as the vegetative phase. During the vegetative phase, the green leaf area continuously expands. As a result, the NDVI values tend to rise during this stage. After the critical stage is passed, plant leaves start to senesce and the NDVI values tend to decrease.

When the crop experiences drought conditions (i.e., scarcity of soil moisture), the evapotranspiration from the cropped field is reduced, and consequently, the colour of the leaves turns from green to yellow. This process leads to a reduction in the NDVI values. Figure 2.5 shows a case of an NDVI profile under drought conditions for the Swift Current district, Saskatchewan. The NDVI data in Fig. 2.5 pertain to year 1988 when the yield was at the minimum for the 1975-96 period.

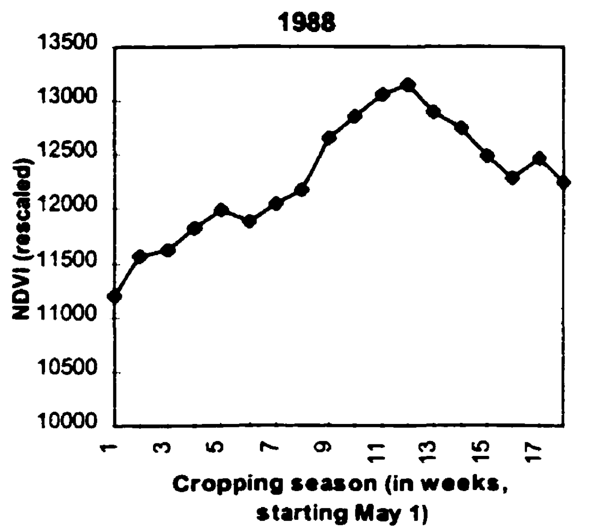


Figure 2.5 NDVI profile under drought conditions for district 3bn, Saskatchewan [Source for NDVI data: Spatial Analysis and Geomatics Applications division, Statistics Canada, Ottawa]

In order to relate the NDVI data to crop conditions, derivation of parameters characterizing physical conditions of the crop is required. Such parameters include the leaf area index (LAI, leaf area per unit of ground area), total biomass, and a crop canopy index (Tucker, 1979; Wiegand et al., 1979; Wiegand and Richardson, 1983, 1990; Aasae and Siddoway, 1981; Mass, 1988). However, it is the LAI that has most commonly been estimated from satellite data. While a linear relationship between LAI and crop yield has been developed to forecast crop yields (Tucker et al., 1980; Wiegand and Richardson 1990; Wiegand et al., 1991; Bullock, 1992), the relationship between LAI and yield can also be nonlinear. For example, Wiegand and Richardson (1983) reported this relationship to be logarithmic. The development of the NDVI-LAI-yield relationship makes it possible that NOAA/AVHRR data can be employed to estimate yield.

By comparing the estimated yield with the long-term average yield, drought can be predicted. However, the success of a drought prediction relies on the quality of NOAA/AVHRR data. As the NOAA/AVHRR data are not available under cloudy conditions, significant presence of clouds over the study area prevents the NDVI data from accurately reflecting crop conditions. Therefore, a weekly composite image is generated from the daily NOAA images using cloud-masking procedures in order to minimize the influence of clouds on the NDVI pertaining to the study area (Goward et al., 1985; Holben, 1986). The NDVI compositing has greatly enhanced the potential of the NOAA data in vegetation monitoring (Tucker et al., 1984) and hence in yield estimation.

Robertson et al. (1992) developed a system to generate geocoded, radiometrically correct and near cloud-free NDVI composites from NOAA/AVHRR data. This system is in use at the Manitoba Remote Sensing Centre (MRSC), Winnipeg, Manitoba, Canada. The MRSC sends the NDVI composites pertaining to the crop growing season to Statistics Canada, Ottawa. In order to facilitate the inter-annual comparison among the composited NDVI data, Statistics Canada (1996) transforms the composited NDVI data to a new scale.

The impact of the NDVI data on yield estimation can be studied to a greater depth if the commencement and termination of various phenological stages can be estimated. In this regard, Robertson (1968) developed various biometeorological time scale models as discussed below.

2.6.1. Biometeorological time scale model

The biometeorological time scale is related to estimation of commencement and termination of the phenological stages of a crop. Robertson (1968) compared four different models to determine the biometeorological time scales for wheat for various regions including the Prairie region (Swift Current, Saskatchewan). The model required data on the planting date, daily maximum and minimum temperatures, and photoperiod. Photoperiod, the duration between sunrise and sunset, can be estimated for a given location (Robertson and Russelo, 1968). The following model provided the best results:

$$1 = \sum_{S1}^{S2} [\{a_1(L-a_0)+a_2(L-a_0)^2\}\{b_1(T_1-b_0)+b_2(T_1-b_0)^2+c_1(T_2-b_0)+c_2(T_2-b_0)^2\}] \qquad [2.24]$$

where a_0, a_1, a_2, b_0, b_1, b_2, c_0, c_1, and c_2 are coefficients, L is the daily photoperiod (hours), T_1 = daily maximum (daytime) temperature (^{0}F), T_2 = daily minimum (nighttime) temperature (^oF), and S_1 and S_2 refer to the commencement and termination of a phenological stage.

Equation 2.24 can also be simplified as follows:

$$1 = \sum_{S1}^{S2} V_1 (V_1 + V_3) \qquad [2.25]$$

where, $V_1 = a_1(L-a_0) + a_2(L-a_0)^2$, [2.26]

$$V_2 = b_1(T_1 - b_0) + b_2(T_1 - b_0)^2, \text{ and} \qquad [2.27]$$

$$V_3 = c_1(T_2 - b_0) + c_2(T_2 - b_0)^2 \qquad [2.28]$$

Using an iterative regression, the coefficients used in Equation 2.24 were obtained as presented in Table 2.2.

Table 2.2 Coefficients of the biometeorological time scale model for Swift Current, Saskatchewan

Coefficient	Duration*				
	PE	EJ	JH	HS	SR
a_0	[V_1 =1]**	8.413	10.93	10.94	24.38
a_1		1.005	0.9256	1.389	-1.140
a_2		0	-0.06025	-0.08191	0
b_0	44.37	23.64	42.65	42.18	37.67
b_1	0.01086	-0.003512	0.0002958	0.0002458	0.00006733
b_2	-0.0002230	0.00005026	0	0	0
b_3	0.009732	0.0003666	0.0003943	0.00003109	0.0003442
b_4	-0.0002267	-0.000004282	0	0	0

* PE = planting to emergence, EJ = emergence to jointing, JH = jointing to heading, HS = heading to soft dough, and SR = soft dough to ripening.

**For determination of the PE, V_1 = 1 in Equation 2.25, thus not requiring use of a_0, a_1, and a_2 coefficients, separately. [Source: Robertson,1968]

2.7 Research Objectives and Hypotheses

2.7.1 Objectives

Based on the statement of the problem as introduced in Section 1.4 and the literature review presented above, the specific objectives of the present study are set forth as follows:

I. To develop an improved method for obtaining short-term and long-term wheat yield estimates for the Canadian Prairies, and

II. To improve understanding of factors contributing to drought prediction based on yield prediction for the Canadian Prairies.

2.7.2 Hypotheses

The following hypotheses have been formulated to address the research objectives.

I. Pattern recognition techniques can be applied to predict drought effectively on a short-term basis.

II. Long-term estimates of wheat yield can be obtained using time series analysis. Different types of time series models may best fit the yield series in different crop districts, disputing the current practice of only using a linear trend model for the Prairie region as a whole.

III. Direct use of daily climatic data, instead of monthly data, can lead to an improvement in short-term yield estimates.

IV. Incorporation of the NDVI data into the current wheat yield model (i.e., the Western Canada Wheat Yield model) can improve the performance of the model in obtaining short-term yield estimates.

These hypotheses are tested on several crop districts of the Prairies in the following section.

2.8 Study Area And Data Procurement

2.8.1 Study area

The selection of the study area was influenced by the availability of the data requirement to test the hypotheses and by few other aspects. The selection process is explained in the following.

The hypotheses formulated in the previous section need to be tested over the study area, i.e., the Prairies. However, considering that the focus of the study is on the spring wheat crop, Saskatchewan seemed to be a logical choice for the study area as it contributes the largest quantity of wheat (about 60% of the total) among the prairie provinces (Walker, 1989). Description on some ecological features (geology, climate, soils, water resources, and vegetation) of Saskatchewan follows.

2.8.2 Some ecological features of Saskatchewan

Saskatchewan has two main geologic regions: (i) the Precambrian Shield, exposed in northern Saskatchewan, which comprises crystalline basement rocks and sedimentary rocks, and (ii) the Phanerozoic Basin, which occupies the southern part of the province, and comprises younger sedimentary rocks. Soils in the Precambrian Shield are generally sandy and stony, nutrient poor, acidic, and low in salts. In contrast, soils in the Phanerozoic Basin to the south generally have more clay and plant nutrients, and are more alkaline and sometimes saline.

There are four types of soils that are dominant in the province (Fig. 2.6). The entire Prairie region is dominated by Chernozemic soils, which is synonymous with grassland vegetation. Dark Gray Chernozemic soils prevail farther north, in the areas

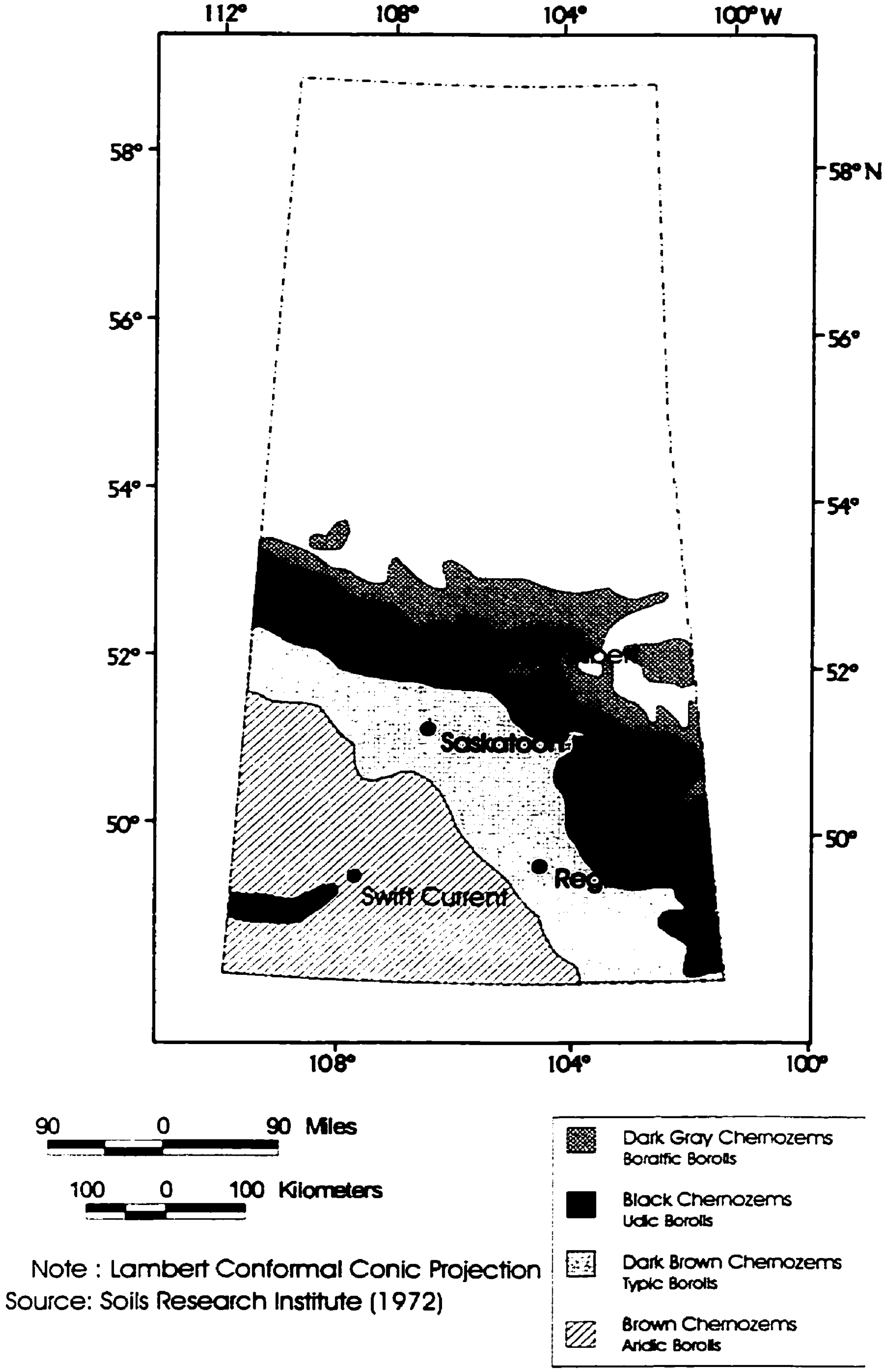


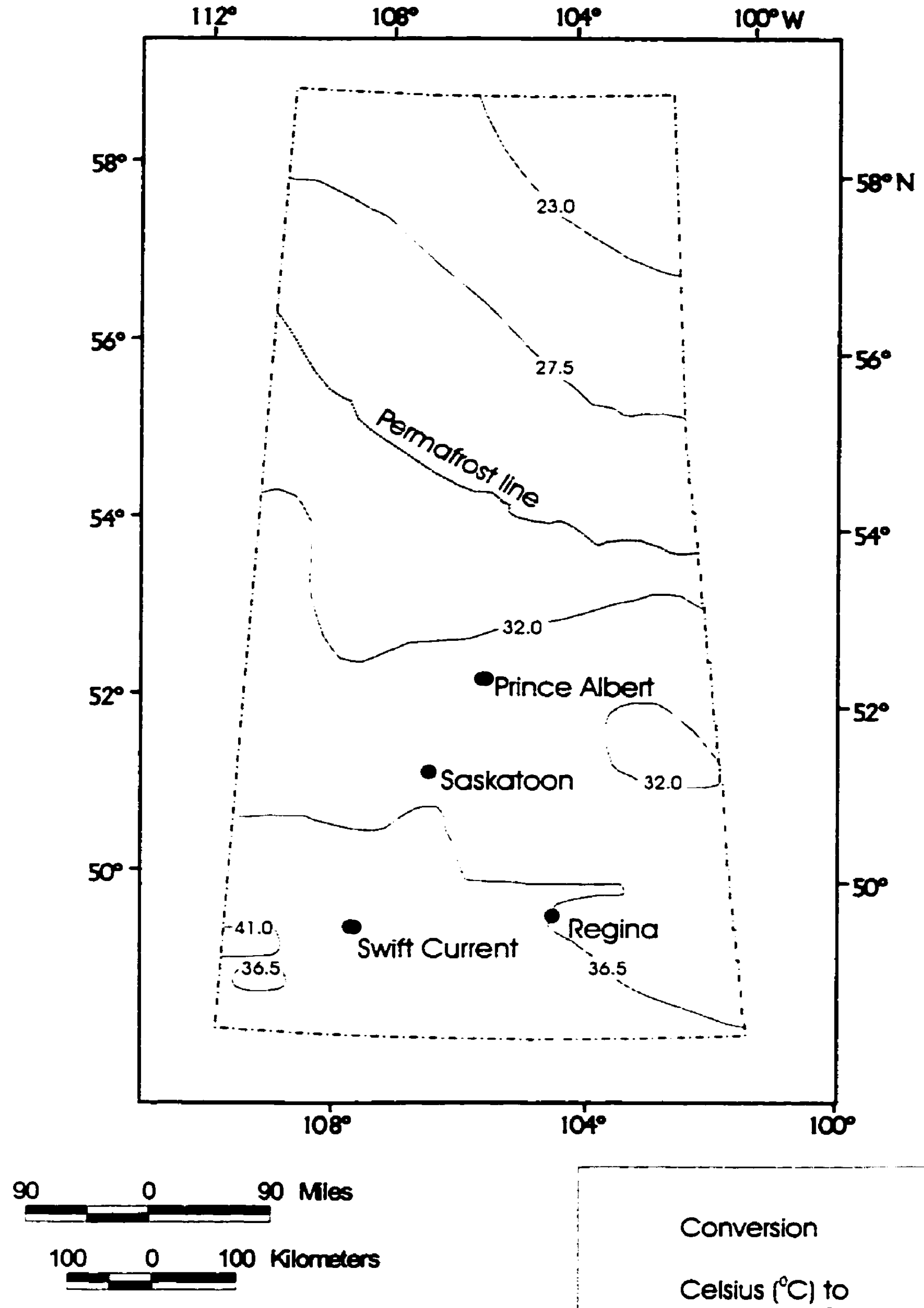
Figure 2.6 Distribution of Chernozems in Saskatchewan Prairies

transitional to the forest. Black Chernozemic soils have developed in the Aspen Parkland. Dark Brown Chernozemic soils occur in the more moist grassland regions of the north, and Brown Chernozemic soils prevail in the semiarid Mixed-Grass Prairie, where the amount of plant biomass produced is relatively low. Most of the soil organic matter in these grasslands originates from the decomposition of roots, and the products of decomposition are held in place by the mineral components of the soil. In general, soils are usually thinner on upper slopes than they are further down slopes. Bottoms of slopes, especially where glacial kettles prevail, are dominated by Gleysolic soils. These soils are dark coloured and saline where groundwater is discharging at surfaces. In arid areas of the southwest where salt content is high and precipitation is relatively low, solonetzic soils have developed. Regosolic soils occur on recent deposits such as alluvial flood plains and sand dunes, while Gleysolic soils exist in wetland areas.

The climate of Saskatchewan is semiarid to subhumid with long and cold winters, short and very warm summers, and cyclonic storms. The climate is influenced by air pressure and wind fields. The western Cordillera forms a significant obstacle to the mild, moist air currents from the Pacific. The region is readily accessible to cold Arctic high-pressure systems in the winter and to frequent incursions of warm, moist tropical air from the southeast, particularly in the summer. This in turn results in the higher precipitation in the summer than in the winter months; additional precipitation in the summer originate from cyclonic and convective thunder storms during the spring or early summer. At a given altitude, climate is warmest in the south and cools gradually northward. Temperatures are highest at lower elevation in the south, progressively decreasing with increasing latitude. Fig. 2.7 shows contours of mean annual temperature across the province. The temperatures vary from below 23°F (-5°C) in the north to above 41°F (5°C) in the south. It is only in the areas below the permafrost line that agriculture can be practiced. Fig. 2.8 presents contours of mean precipitation during the June-August period. The precipitation is generally low, but it increases slightly from south to north and more markedly from west to east (about 6" to 9").

Both surface and ground water resources are available in Saskatchewan. It is estimated that 81, 632 km^2 of surface water (approximately 12% of the total area of the province) exists in lakes and rivers, streams, ponds, man made reservoirs, and other bodies (Acton et al., 1998). Nearly all surface water originates from precipitation that falls within the confines of the provincial boundaries, with much smaller amount entering the province in rivers and streams from the west. Local runoff of snowmelt and rainfall is a major contributor to water in natural ponds and lakes. More than 90% of precipitation returns to the atmosphere through evpotranspiration.

Vegetation types prevalent in Saskatchewan include are Taiga, Boreal Forest, Transition Forest, Aspen Parkland, Mixed-Grass Prairie, and Short-Grass Prairie (Scott, 1995; Fig 2.9). While the northern portion of the province is occupied by Taiga, Boreal and Transition forests, Aspen Parkland, Mixed- and Short-Grass Prairie exist in the southern portion. The Aspen Parkland belt, representing a transition from grasslands in the south to the Boreal Forest to the north, extends diagonally from southeast to northwest across the southern part of the province.



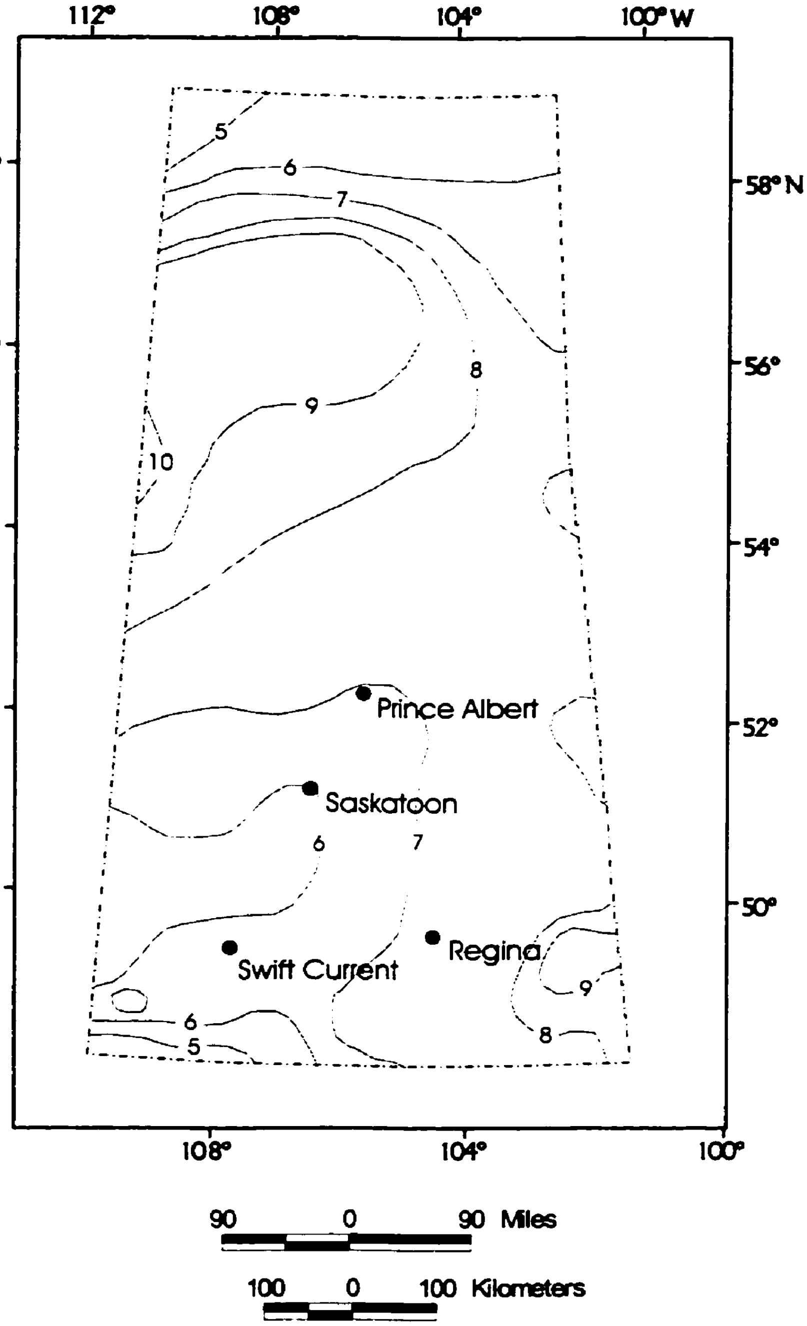
Note : Lambert Conformal Conic Projection
Source: Brown (1967)

Conversion

Celsius (°C) to
Fahrenheit (°F)

C/5 = (F-32)/9

Figure 2.7 Mean annual temperature (°F) and permafrost boundary line in Saskatchewan



Note : Lambert Conformal Conic Projection

Source: Longely (1972)

Figure 2.8 Mean June to August precipitation (inches) in Saskatchewan

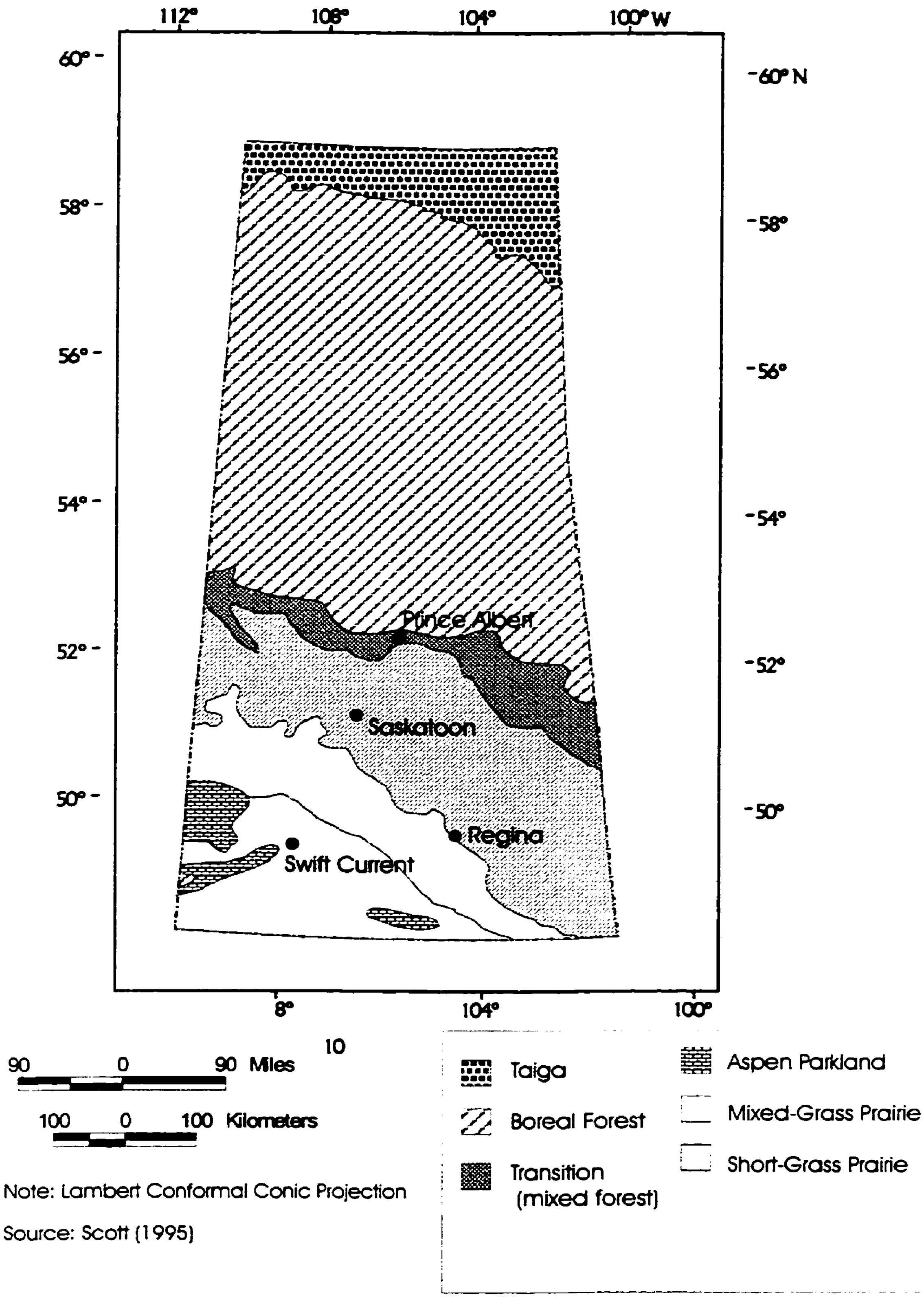


Figure 2.9 Vegetation of Saskatchewan

The Aspen Parkland belt, representing a transition from the grasslands to the south and the boreal forest to the north, extends diagonally from southwest to northwest across the southern part of the province. Summers are cooler in this region, winters are longer and colder, and snow cover is more continuos than in the regions to the south and west. Summer evaporation and precipitation rate in this area are almost equal which minimizes potential severity of the late summer moisture deficits. Here a mosaic of trembling aspen surrounds numerous wetlands. These form groves in the sea of plains rough fescue grasslands.

On the basis of various ecological features, the province of Saskatchewan can be divided into several ecozones (Fig. 2.10), although there are no fixed number of ecological features based on which to make the division (Bird and Rapport, 1986). It is suggested that ecozone is a discrete system that has resulted from the interplay of the landform, water, soil, vegetation, climate, wildlife and various human uses. The ecozones are highly generalized and boundaries between them are not always sharply defined. Nevertheless, interrelationships among ecological features (e.g., physiography, soil, climate, and vegetation) exist within an ecozone. For example, total plant growth, expressed as primary production, is strongly related to temperature, precipitation, and evapotranspiration (Emanuel et al., 1985; Aber and Melilo, 1991). Similarly, effects of the relative intensity of different soil-forming processes, influenced significantly by the climate characteristics, can be seen in a soil profile.

Based on the ecological features described in previous sections, Saskatchewan can be divided into four ecozones (Fig. 2.10): i) Taiga Shield, ii) Boreal Shield,

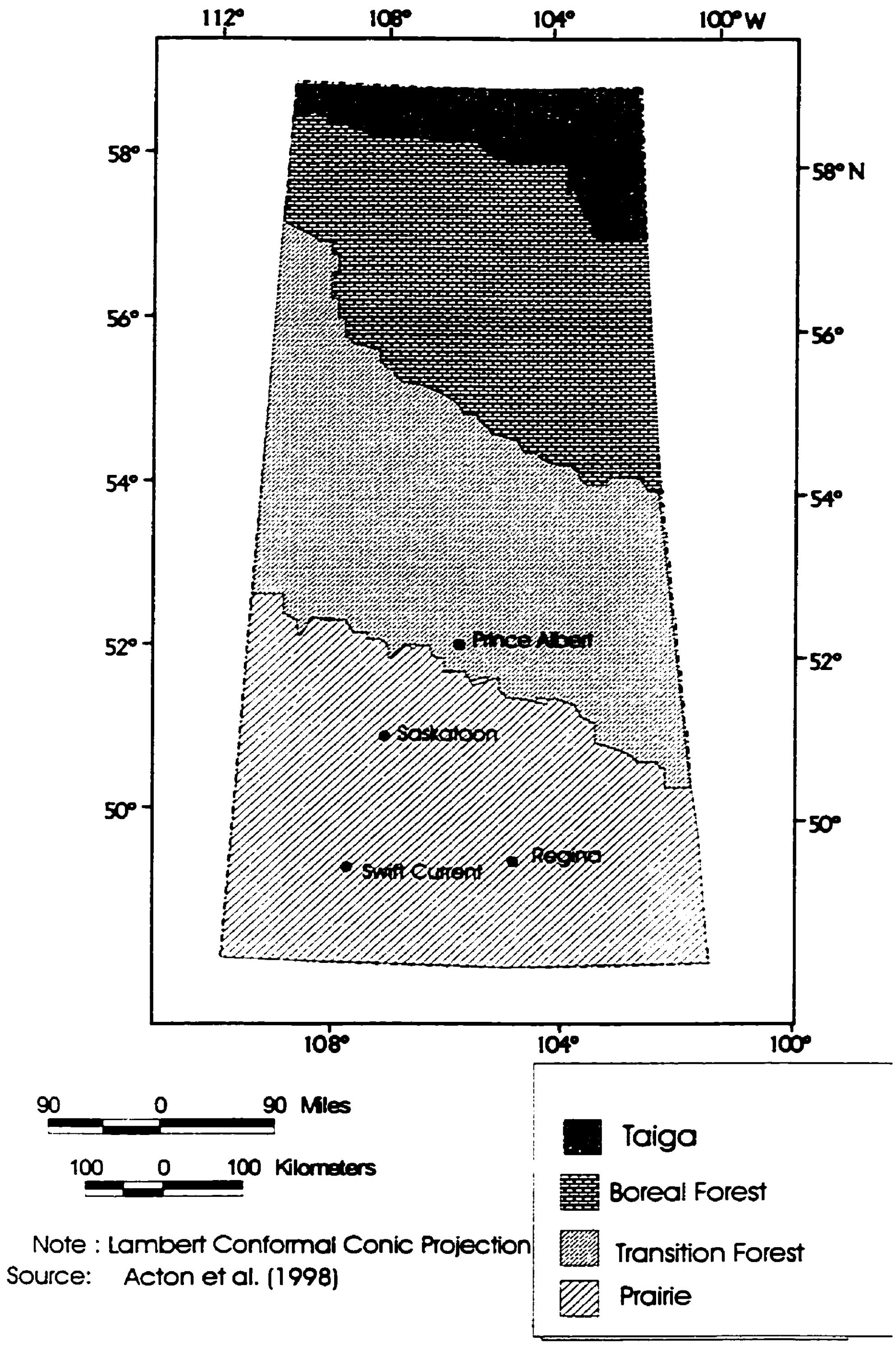


Figure 2.10 Ecozones of Saskatchewan

iii) Boreal Plain, and iv) Prairie (Acton, et al., 1998). The Russian term "taiga" refers to the northern edge of the boreal coniferous forest and stretches across part of Canada's subarctic north. Boreal Shield ecozone separates warmer boreal plains to the south from the colder taiga shield to the north occupying nearly 18.7 million hectares, one third of the total area of the province. Boreal Plain ecozone covers about 27 per cent of the province. Most of this region is characterized by northern boreal forest, although about 25% of this ecozone along its southern boundary is used for agricultural production. Prairie ecozone covers approximately 24 million hectares, southern one-third of Saskatchewan, extending from boundary with the United States to the Boreal Plain ecozone. Since most of Saskatchewan's agriculture is practiced in the Prairies, which is the focus of this study, the following section describes various ecological features of the Prairie ecozone.

2.8.3 Ecological features of Prairie Ecozone of Saskatchewan

The Prairie ecozone is essentially a level to gently rolling plain with numerous subdued uplands dispersed throughout most of its extent. Elevations are lowest in northeastern part of the area. Glacial deposits represent the surficial sediment throughout the ecozone. These deposits are thinner in the Prairie ecozone than in the Boreal Plain ecozone. Glacial and postglacial deposits have profound influence on the nature of the local landscape. Underfit streams with their characteristic floodplains, are active in channels that originated as melt water channels from the ice or spillways from glacial lakes. Sand dunes have been active since glaciation.

The climate in the Prairie ecozone ranges from semiarid to humid continental, with long and cold winters, short and very warm summers, and cyclonic storms. Temperature are highest at lower elevation in the south, progressively decreasing with increasing altitude and latitude. Precipitation is generally low but it increases slightly from south to north and more markedly from west to east. Moderately cold semiarid to subhumid conditions prevail on uplands in what is otherwise the driest part of the ecozone. The climate is further typified by periods of extensive droughts. The droughts are caused by summer moisture deficits which occur due to low precipitation and high evapotranspiration. Northward and eastward from the mixed grassland, moisture deficits are less severe and droughts are relatively short (Fig. 2.11).

Most landforms in the Prairie ecozone are of the glacial origin. Nearly level ground moraine (till plains), glaciolacustrine and glaciofluvial plains are major contributors to the "flat prairie" landscape. Valleys and coulees with enclosed lakes are most striking landscape features. Soil here strongly reflects climate and natural vegetation and associated landforms. Soils formed in glacial till, the sediment that constitutes ground moraine and hummocky moraine are usually loam textured, while those formed in glaciolacustrine deposits have higher proportion of silt and clay, and those formed in glaciofluvial deposits have more sand and gravel.

The Prairie ecozone is a grassland region. A mixed-grass community dominates the southwestern, warmer, and a more arid part of the ecozone, represented by the Mixed Grassland ecoregion (Acton et al., 1998). A late summer moisture deficit, caused by low precipitation and high evaporation, and periods of

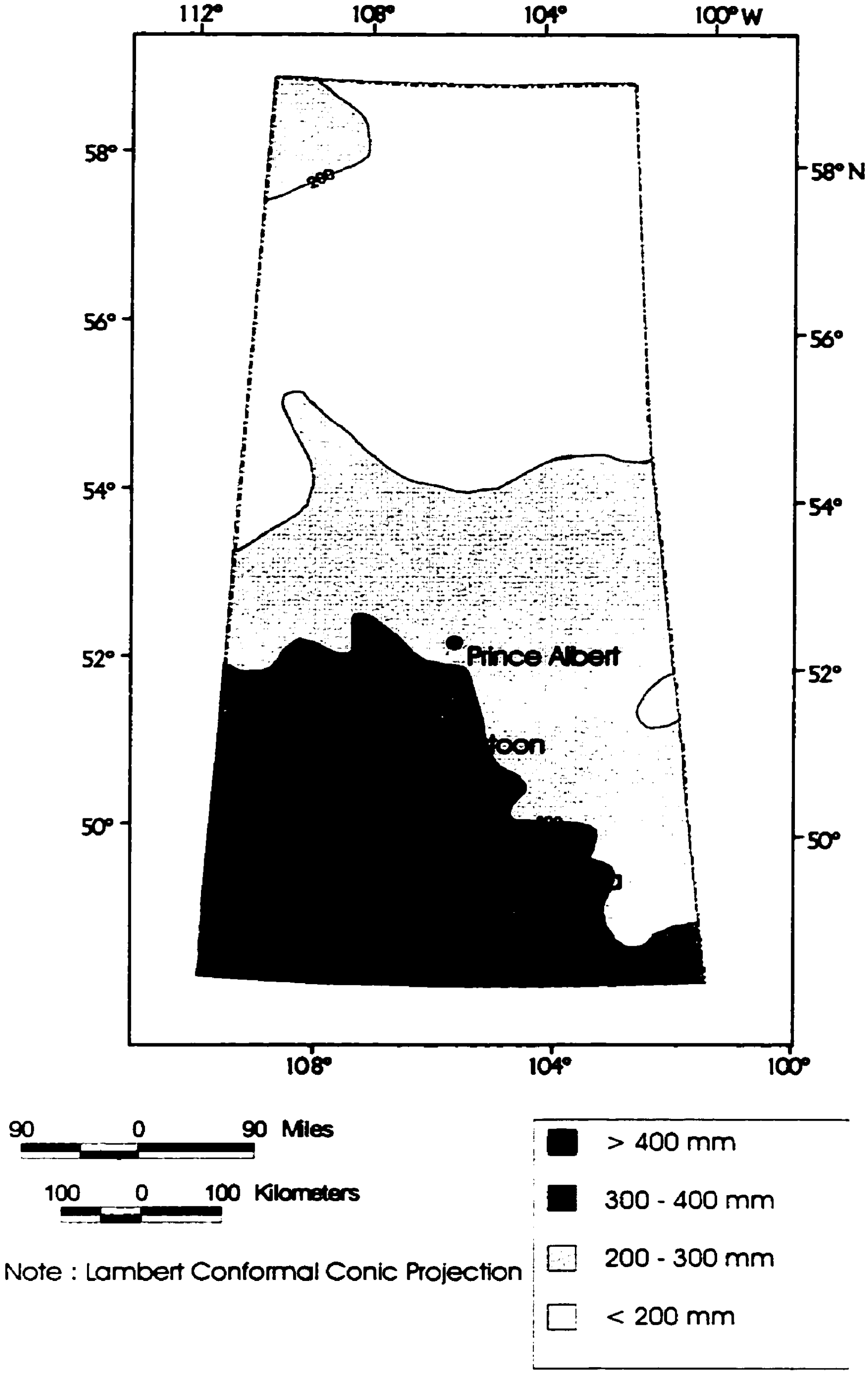


Figure 2.11 Seasonal water deficits at 50% probability, May - Sept. 1931-60

extensive droughts typify the climate of this area. The resulting mixed-grass vegetation includes what are often referred to as "short grass" (blue grama grass and sedge) and "mild to tall grass" (wheatgrasses, June grass, needle-and-thread, and porcupine grass), along with pasture sage and moss phlox. Northward and eastward from the mixed grassland, moisture deficits are less severe and droughts are less prolonged. Mid-grasses dominate these areas, along with an increase in the extent of shrublands, aspen grove woodlands, and wetlands.

The smooth climatic and vegetation zonation that extends from southwest to northeast across the southern part of the province is interrupted by a prominent uplands. High elevation result in a climate that is cooler and more moist than the surrounding dry grasslands.

Many small wetland areas, or sloughs, occur throughout the Prairie ecozone. In the more humid part, sloughs tend to be more permanent, the water is relatively fresh, and they are ringed by willows and trembling aspen; in drier parts of the ecozone, however, the sloughs are less permanent and more saline, and the transition from the wetland to the grassland. Most freshwater wetlands are characterized by emergent vegetation such as sedges, bulrushes, cattails and red grasses on their margins. In the open water , submerged growth of pondweeds, yellow water crowfeet and greater bladderwort may be present. Saline lands do not have a marginal ring of willows, but rather have shorelines heavily encrusted with white salts and usually bare of vegetation except for a few salt-tolerant plants like red samphire. Salt-tolerant grasses, such as seaside arrow-grass and alkali grass, grow at the margin of salt crust.

2.8.4 Selection of crop districts

Droughts are generally well reflected in the types of vegetation and moisture deficit conditions. Based on these criteria Saskatchewan Agriculture and Food (1997) divided the province into four drought-based agro-climatic regions: (i) a region where drought is a frequent hazard, (ii) a region where drought occasionally occurs and poses a problem, (iii) a region where drought may occur but does not usually pose a problem, and (iv) a region where rainfall is usually adequate for crop production These regions are shown in Fig. 2.12.

The inter-regional variation in agroclimatic conditions contributes to the inter-district variation in the yield characteristics (average yield and coefficient of variation (CV) in yield) as shown in Table 2.3. It would be interesting to examine how the difference in ecological and agroclimatic conditions affects the process of improving drought prediction. It was therefore decided to test the hypotheses on a few crop districts that represent varying agroclimatic conditions and for which the required data were available. The crop districts in Saskatchewan are demarcated based on rural municipality-boundaries for administrative purposes.

The CV was considered as a means to gauge the inter-annual fluctuation in the yield data pertaining to a district. Though the wheat yield data (in digital form) were available for the 1975-96 period for all of the crop districts in Saskatchewan, the climatic data were not available for the same period for all weather stations in Saskatchewan. While the monthly weather data were available for most stations,

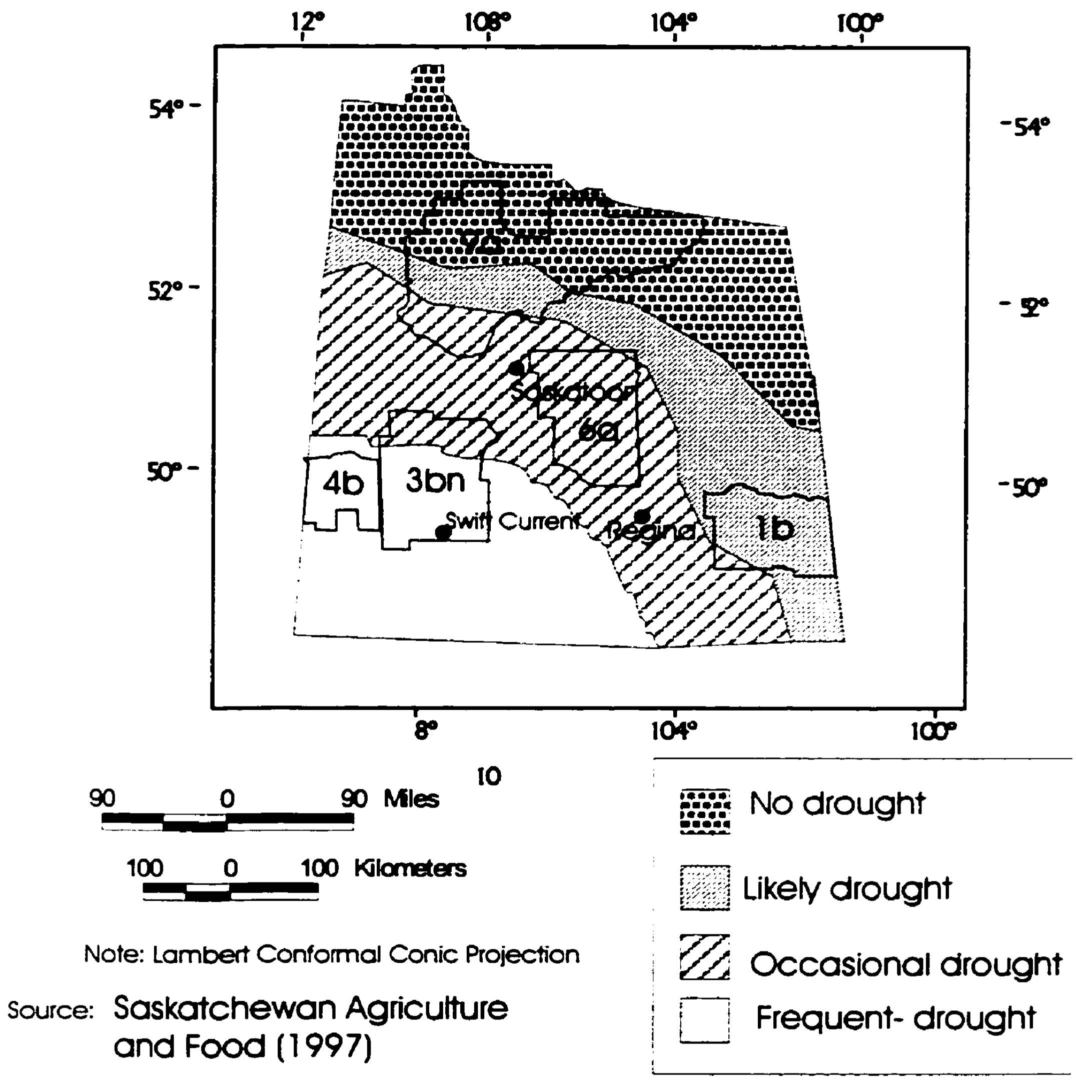


Figure 2.12 Drought-based agro-climatic zones of Saskatchewan

Table 2.3 Spring wheat yield variation in Saskatchewan during 1975-96

Crop district	Min. (t/ha)	Max. (t/ha)	Mean (t/ha)	Stdev. (t/ha)	CV
1a	0.93	2.33	1.70	0.36	0.21
1b	1.08	2.51	1.80	0.39	0.22
2a	0.72	2.38	1.62	0.46	0.28
2b	0.95	2.53	1.91	0.38	0.20
3an	0.52	2.25	1.70	0.44	0.26
3as	0.59	2.13	1.67	0.44	0.26
3bn	0.5	2.23	1.77	0.43	0.24
3bs	0.38	2.36	1.69	0.46	0.27
4a	0.4	2.22	1.58	0.43	0.27
4b	0.44	2.42	1.65	0.56	0.34
5a	0.95	2.39	1.87	0.35	0.19
5b	1.17	2.41	1.95	0.32	0.16
6a	0.61	2.18	1.78	0.36	0.20
6b	0.48	2.45	1.85	0.43	0.23
7a	0.6	2.62	1.92	0.44	0.23
7b	1.21	2.55	1.97	0.34	0.17
8a	0.94	2.8	2.03	0.48	0.24
8b	0.87	2.7	2.05	0.4	0.20
9a	1.3	2.62	2.00	0.29	0.15
9b	1.27	2.45	1.97	0.28	0.14

[Source:The above basic statistical parameters have been derived from the wheat yield data collected from Statistics Canada]

daily weather data were available only for a limited number of stations for the 1975-96 period. To ensure an equal length (i.e., 1975-96) of yield and weather data required to test the hypotheses, and to represent the districts experiencing low, medium, and high degrees of yield fluctuation, five crop districts in Saskatchewan were selected for the study. The districts selected were 1b, 3bn, 4b, 6a, and 9a (Fig. 2.13). Since demarcation of crop districts is not based on natural boundaries, ecological heterogeneity within a crop district is expected. Soil, water-deficit, and vegetation conditions for the selected district are presented in Figs. 2.14, 2.15, and 2.16, respectively, and a synthesis of their characteristics along with the description on average climate parameters, is provided in Table 2.4.

From Fig. 2.14, it can be observed that while districts 3bn, 4b, and 6a possess only one type of soil (brown, brown, and dark brown, respectively), districts 1b and 9a have more than one type of soil (black and dark brown for 1b, and bark brown, black, and dark gray for district 9a). Also, a significant variation in water deficits is

Table 2.4 Some agroclimatic characteristics of selected districts in Saskatchewan

District	Soil Type	Vegetation Type	Avg. Water Deficit (mm, approx.)	Temperature and Precipitation (average of 1974-96 monthly data)				
				During Summer (April-September)		During Winter (October-March)		Annual Precip. (mm)
				Avg. Temp. (°C)	Total Precip. (mm)	Avg. Temp. (°C)	Total Precip. (mm)	
1-b	Black	Aspen Parkland	250	12.6	320.2	-8.2	108.5	428.7
3-bn	Brown	Short- and Mixed-Grass Prairie	400	12.8	273.9	-4.0	95.1	369.0
4-b	Brown	Aspen Parkland and Short-Grass Prairie	> 400	14.4	224.6	-5.5	72.6	297.2
6-a	Dark Brown	Aspen Parkland	350	13.0	308.8	-7.6	100.1	408.9
9-a	Dark Gray	Boreal and Mixed Forest	< 200	12.4	303.5	-9.5	96.6	400.1

[Source for monthly weather data: Statistics Canada; Information on soil, vegetation, and water deficits have been extracted from Fig. 2.6, 2.9, and 2.11, respectively.]

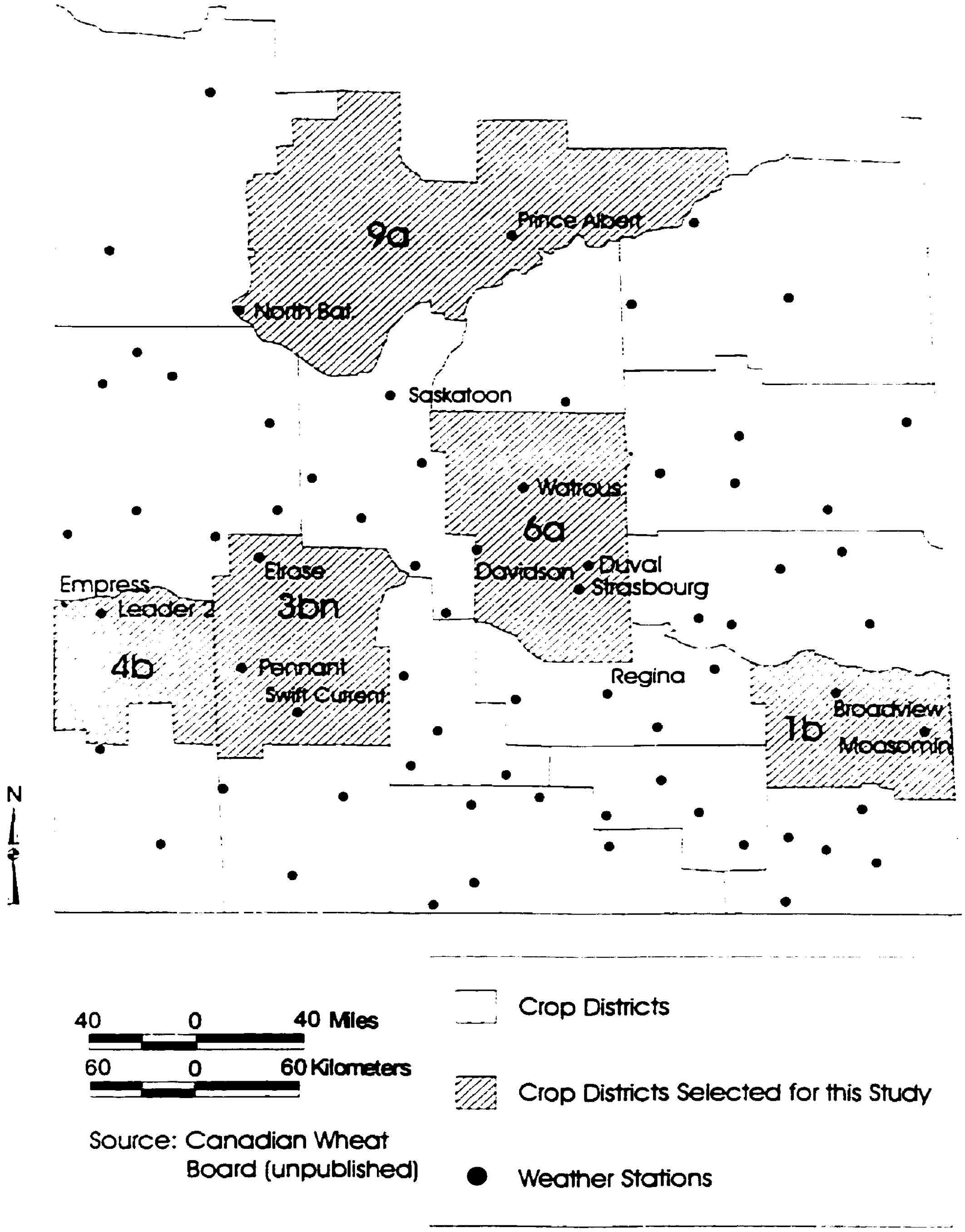


Figure 2.13 Crop districts and weather stations in Saskatchewan

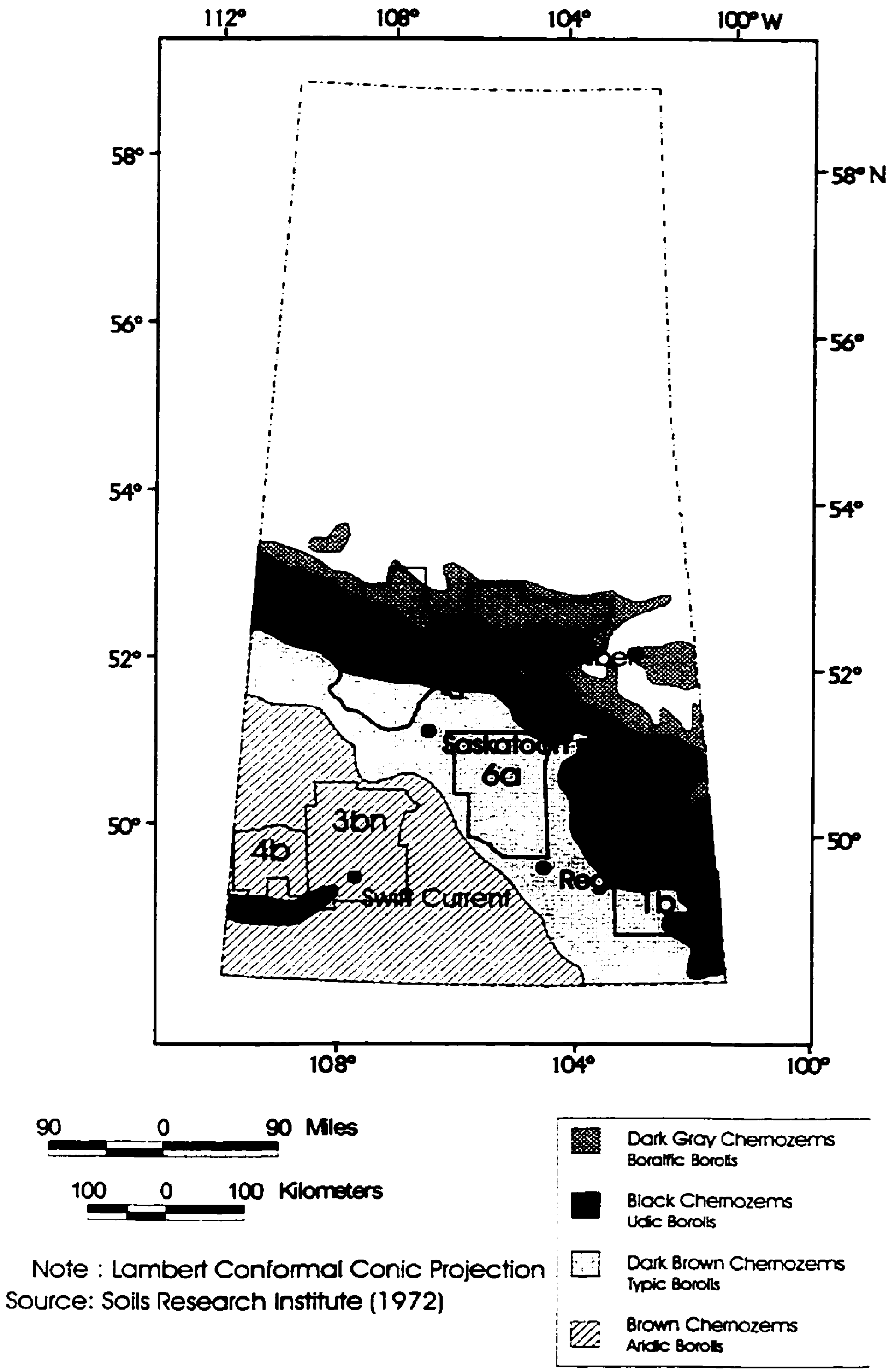


Figure 2.14 Soil types in selected districts in Saskatchewan

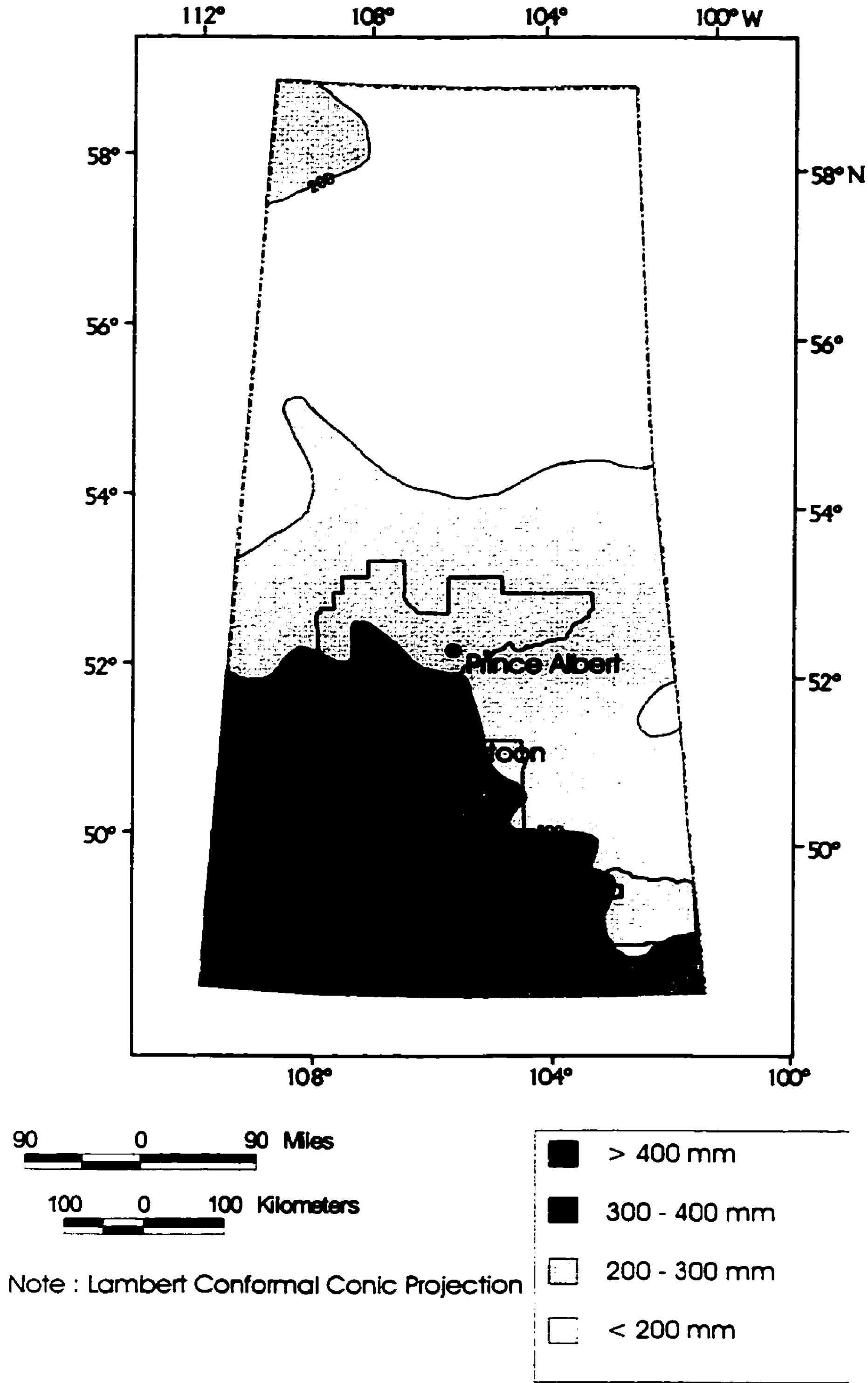


Figure 2.15 Seasonal water deficits at 50% probability (May - Sept. 1931-60) in selected districts in Saskatchewan

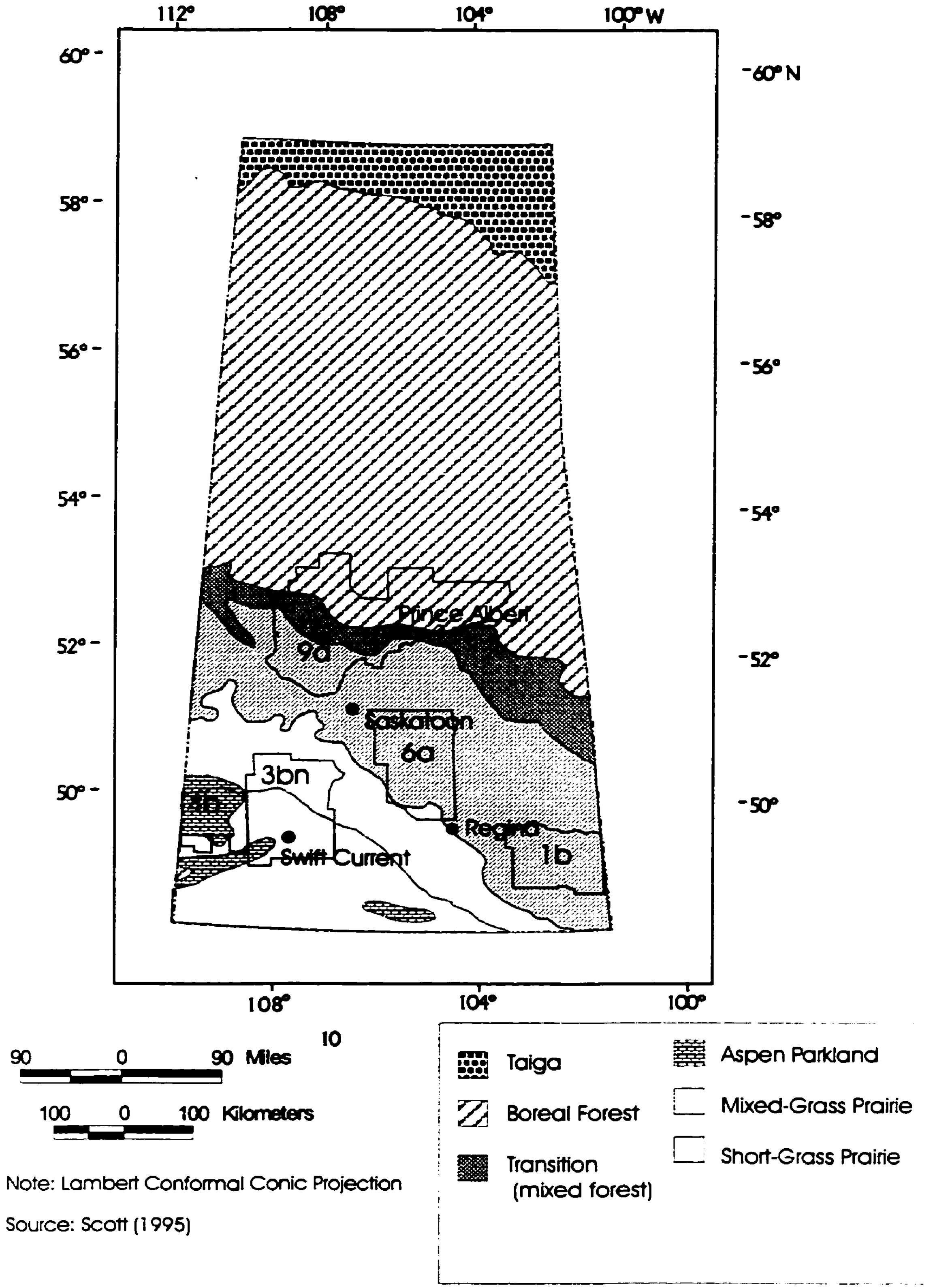


Figure 2.16 Vegetation types in selected districts in Saskatchewan

found in the selected districts. Water deficits are minimum (< 200 mm) for districts 1b and 9a, and maximum for district 4b (> 400mm). Two water deficit zones rather equally divide district 3bn. In the upper half of district 3bn, water deficit is more precarious (> 400 mm) compared to the lower half of the district where it ranges from 300 mm to 400 mm. District 6a experiences moderate water deficit (300 - 400 mm).

Soil types and water deficit (as result of temperature and precipitation trends) in an area largely determine the vegetation types of the area. In districts 1b, 4b, and 6a, single vegetation type (Aspen Parkland) dominates, whereas district 3bn has a combination of Short-Grass and Mixed-Grass Prairie (Fig. 2.15). Upper portion of district 9a is covered with Boreal Forest, middle with Transition Forest, and southern with Asper Parkland.

It can be seen from Table 2.4 that summer as well as annual precipitation are lowest in district 4b followed by district 3bn. Both annual and summer precipitation (428.7 mm and 320.2 mm, respectively) are highest for district 1b, while districts 6a and 9a experience relatively lower amount of precipitation (about 400 mm annually and 300 mm during the summer).

Overall, in terms of intradistrict variation, district 9a appeared to have the more heterogeneous characteristics relative to others. Similarly, a higher degree of diversity was recognized for district 3bn; other districts were less heterogeneous.

2.8.5 Data procurement and derivation

The following data pertaining to the selected districts were collected to test the hypotheses.

a. Crop yield data: Wheat yields (kg/ton) for the period from 1975 to 1996 were collected in digital form from the Canadian Wheat Board; the original source of this data was Statistics Canada.

b. Climate data: Monthly data (maximum temperature, minimum temperature, average temperature, and total precipitation) for the period from 1974 to 1996, and daily data (maximum temperature, minimum temperature, average temperature, and total precipitation) for the period from 1975 to 1996 were obtained from the Canadian Wheat Board.

For data collected from the selected weather stations, Table 2.5 was prepared to examine the spatial variability in weather data within a district. It can be observed from Table 2.5 that, the correlation (determined by R^2) among precipitation records at weather stations within a district is lower than that among temperature records. This is to say that spatial variability in precipitation records is greater than that in temperature records. This can be attributed to a greater degree of within-district variation in vegetation conditions for districts 3bn and 9a than for district 6a.

Table 2.5 Collinearity in data recorded at different weather stations within selected districts

Crop District	Weather Station(s)	Coefficient of Determination (R^2) among Data for Weather Stations within District	
		Average Monthly Temperature (1974-96)	Total Monthly Precipitation (1974-96)
3bn	Elrose, Penant, Swift Current	0.99	0.67
6a	Duval, Watrous	0.97	0.75
9a	Prince Albert, North Battleford	0.99	0.60

c. Satellite data: The NOAA/AVHRR-based NDVI data pertaining to the Prairies were available only since 1987. Such data were collected from the Spatial Analysis and Geomatics Application Division, Statistics Canada, Ottawa, for the period from 1987 to 1996. The NDVI data refer to the NDVI values generated from the weekly composited images that were produced by the Manitoba Centre for Remote Sensing (MCRS) following a procedure developed by Robertson et al. (1992). The NDVI data generated at the MCRS are rescaled and standardized by Statistics Canada (1996) to allow for a better inter-annual comparison of the NDVI data. The standardized data were employed in the present study.

After procuring the wheat yield, climate, and NDVI data, the hypotheses were tested. The next chapter tests the first hypothesis.

CHAPTER THREE

PATTERN RECOGNITION IN DROUGHT PREDICTION

3.1 Hypothesis I and Methodology

The goal of this chapter is to test the first hypothesis which states: "Pattern recognition techniques can be applied to predict drought effectively on a short-term basis."

In Section 2.3.2 it was explained how a pattern recognition technique, namely, the Error Correction (EC) procedure, could be utilized in predicting whether or not a predefined drought would occur. In order to use the EC procedure, a quantitative definition of drought is required. One of the common ways of defining drought quantitatively is to: (i) choose the long-term average of the yield of the main crop as a reference yield, (ii) compute the threshold yield based on an assumed percent value, as follows:

$$\text{Threshold yield} = (1\text{- percent value})/100^*(\text{average yield}), \qquad [3.1]$$

(iii) compare the yield in a given year with the threshold yield, and (iv) mark the yield as drought if the yield is less than the threshold yield. Kumar (1993) and Kumar and Panu (1997) used this definition in classifying drought severities in an arid region.

Using the drought definition described above, annual yield data in past years can be grouped into two categories: (i) drought, and (ii) nondrought. This categorization can be observed in Table 3.1 which was prepared to apply the EC procedure on 1975-91 data, leaving 1992-96 data for testing.

Should a different threshold yield be chosen to define drought, it is likely that for a given set of yield data the number of drought (or nondrought) cases will change. As a result, a different solution vector, W (if it exists), is likely to emerge. It should be noted here that the process of determining W is an iterative one, and it is possible that W may not exist for the case pertaining to a specific percent value which was used to define the threshold yield. In an attempt to find W, a change in either the threshold yield or the elements of a drought vector, or both, is made. Subsequently, the EC procedure is repeated. The objective is to search for W, which can linearly separate drought patterns from nondrought patterns. If W is identified it paves the way for application of the EC procedure to drought prediction.

In order to enhance the probability of finding W, a number of percent values were chosen to define drought and various elements were selected to vectorize it. The district 3bn, one of the five selected districts, was selected for application of the EC procedure because drought is a frequent hazard in this district and a sufficient number of drought vectors could be formed for this district.

The initial value utilized in order to define drought was 10 percent. Additional values were obtained with an interval of 5 percent. However, only four values (10, 25, 45 and 70 percent) were found to be relevant, because a change from one percent value to the next one (among the four values) caused a change in the ratio of the drought cases to nondrought cases. Other intermittent values (i.e., 15, 20, 30, 35, 40, 50, 55, 60, 65, 75...) did not provide a new case to be tested.

Table 3.1 Yield vectorization for district 3bn, Saskatchewan

Year	Yield (t/ha)	% Deviation From Mean Yield	Yield Category	Elements of Yield Vector		
				Coeff. of Var. in Temp.	Coeff. of Var. in Precip.	Augmenting element
1975	1.67	-2.77	ND	0.286	0.059	1
1976	2.21	28.67	ND	0.170	0.198	1
1977	2.04	18.77	ND	0.139	0.177	1
1978	1.71	-0.44	ND	0.174	0.092	1
1979	1.53	-10.92	D	-0.302	-0.096	-1
1980	1.54	-10.34	D	-0.117	-0.161	-1
1981	1.80	4.80	ND	0.260	0.160	1
1982	2.21	28.67	ND	0.268	0.137	1
1983	1.88	9.46	ND	0.308	0.200	1
1984	1.28	-25.48	D	-0.276	-0.154	-1
1985	0.91	-47.02	D	-0.188	-0.080	-1
1986	2.08	21.10	ND	0.166	0.127	1
1987	1.86	8.29	ND	0.142	0.124	1
1988	0.50	-70.89	D	-0.139	-0.098	-1
1989	1.67	-2.77	ND	0.229	0.091	1
1990	2.08	21.10	ND	0.223	0.152	1
1991	2.23	29.83	ND	0.243	0.132	1

D = Drought category, ND = Nondrought category
Note: Elements with negative signs represent drought

Following the selection of a definition for drought, attempts were made to vectorize it. In order to vectorize a drought, the yield influencing factors were studied. However, in the Prairies, weather is the primary factor responsible for drought occurrence (Walker, 1989). Further, emphasizing the significance of temperature on wheat production in the Canadian Prairies, McKay (1983) observed that "A 1 ^{0}C drop in the mean annual temperature accompanied by a 9-to-15 day reduction in the length of the growing season could prove critical for wheat production."

In this study, the temperature and precipitation data have therefore been considered as basic parameters (Table 3.2) to derive variables that could be used as elements of yield vectors. The 12 variables which were derived are: (i) Win_Tavg (average temperature during winter, from November to March), (ii) Win_P (total precipitation during winter, from November to March, (iii) May_T (average temperature in May), (iv) May_P (total precipitation in May), (v) Sum_Tavg (average temperature during summer, from May to August), (vi) Sum_P (total precipitation during summer, from May to August), (vii) Sum1_Tavg (average temperature from May to July), (viii) Sum1_P (total precipitation from May to July), (ix) CV_T (coefficient of variation in temperature during summer, from May to August), (x) CV_P (coefficient of variation in precipitation during summer, from May to August), (xi) CV1_T (coefficient of variation in temperature from May to July), and (xii) CV1_P (coefficient of variation in precipitation from May to July).

Subsequently, variables were grouped into six pairs: (i) Win-T and Win_P, (ii) May_T and May_P, (iii) Sum_T and Sum_P, (iv) Sum1_T and Sum1_P, (v) CV_T and CV_P, and (vi) CV1_T and CV1_P. The paired variables were then used as elements of yield vectors.

Table 3.2 Some basic variables used to characterize yield (or drought) vectors for district 3bn, Saskatchewan

Year	Yield (t/ha)	%Deviation from the mean yield	Variables									
			T_{avg}* (^{0}C)	P_{avg}** (mm)	Average Temp. (^{0}C)				Total Precip. (mm)			
					May	June	July	Aug.	May	June	July	Aug.
1975	1.67	-2.77	-7.4	119.9	10.1	14.7	20.6	14.8	43.0	60.7	57.9	36.6
1976	2.21	28.67	-5.4	108.9	13.4	15.0	18.8	19.1	51.0	130.0	16.0	48.2
1977	2.04	18.77	-4.0	58.6	13.3	16.9	18.0	14.4	76.2	30.8	115.7	22.4
1978	1.71	-0.44	-8.9	140.5	12.2	17.1	18.5	17.5	34.5	49.7	53.9	21.7
1979	1.53	-10.92	-10.4	110.2	9.0	16.4	19.7	18.6	39.0	64.5	36.5	27.1
1980	1.54	-10.34	-5.1	68.1	14.1	16.2	18.4	15.0	66.3	94.4	12.4	42.9
1981	1.80	4.80	-3.0	85.7	12.1	13.0	18.6	20.6	79.8	103.8	26.3	29.3
1982	2.21	28.67	-8.4	109.4	9.0	15.6	17.7	17.0	102.8	32.1	121.2	51.2
1983	1.88	9.46	-4.6	117.4	9.6	16.1	19.1	21.3	121.7	27.1	45.8	28.9
1984	1.28	-25.48	-5.2	78.0	10.6	16.1	20.1	20.7	32.2	74.5	35.3	17.0
1985	0.91	-47.02	-7.1	103.4	13.8	13.7	20.1	16.9	20.7	16.1	34.1	26.6
1986	2.08	21.10	-5.6	84.4	12.2	17.2	17.3	18.1	68.9	54.7	97.3	22.8
1987	1.86	8.29	-2.1	69.3	14.1	18.6	18.3	14.8	72.6	31.9	22.6	57.1
1988	0.50	-70.89	-3.4	52.6	15.6	21.6	20.1	17.9	30.1	63.8	28.5	44.9
1989	1.67	-2.77	-7.1	107.4	11.5	16.4	20.5	18.1	32.4	90.2	77.3	74.6
1990	2.08	21.10	-3.9	109.3	10.8	16.5	17.8	18.7	102.3	62.8	52.2	15.9
1991	2.23	29.83	-5.2	132.7	11.4	16.0	18.7	20.8	89.5	152.3	71.7	41.4
1992	1.96		-2.3	73.0	11.2	15.4	16.4	16.1	72.2	34.6	41.8	49.0
1993	2.21		-5.4	108.7	12.5	14.3	15.6	16.3	86.9	80.6	15.2	112.7
1994	1.83		-6.9	116.7	12.2	15.7	18.9	18.1	29.3	76.2	65.8	37.4

Note: *Average Temp. during Nov.-March, **Average Total Precip. during Nov.-March [Sources: Environment Canada for climatic data, and Statistics Canada for yield data]

The four percent values (10, 25, 45, and 70 percent) and the six types of yield vectors resulted in a total of 24 cases. A computer program was developed and employed to determine if, for any combination of percent value (out of the four values) and a pair of variables (out of the six pairs), it was possible to discover a solution vector, W.

3.2 Results and Discussion

In the computer program which implemented the EC procedure, an iteration limit of 2000 was set. It was found that the solution vector, W, did not exist in any of the 24 cases considered. Hence, it was not possible to predict drought using the EC procedure in the present case. Nevertheless, the above analysis has corroborated the complex nature of droughts, as cautioned by Yevjevich et al. (1978). It should be noted that the EC procedure seeks only linear classification of drought vectors from nondrought ones. It may be possible that a nonlinear classification could separate the two categories of drought for which other pattern recognition techniques, such as nearest-neighbourhood classification (Duda and Hart, 1973), could be successful. Also, additional variables derived from the temperature and precipitation parameters could further enhance the possibility of finding a solution vector, W.

In this chapter, the EC procedure of pattern recognition was attempted to predict drought on a short-term basis. However, prediction of drought on a long-term basis is also required for economic planning and drought preparedness. This can be achieved using time series analysis as discussed in the following chapter.

CHAPTER FOUR

YIELD FORECASTING USING TIME SERIES ANALYSIS

4.1 Hypothesis II and Methodology

The second hypothesis states, "Long-term estimates of wheat yield can be obtained using time series analysis. Different types of time series models may best fit the yield series in different crop districts, disputing the current practice of only using a linear trend model for the Prairie region as a whole."

In order to test the above hypothesis, a detailed time series analysis was performed on yield data of the past years. The time series plots of yields in the selected districts are shown in Chapter 2 (Fig. 2.3). These yield series were modelled using various time series techniques. The first step in modelling a series using time series analysis is to test whether a given series is stationary or nonstationary. To test this, the Unit Root Test (Augmented Dickey-Fuller test), as explained in Section 2.4.1, was performed. Table 4.1 contains the τ statistics computed using SHAZAM software for different yield series for the selected crop districts. The yield series with varying lengths of yield data (19, 20, and 21 observations) are considered for each district. This was attempted to forecast yield for three different years as explained later in this section. The τ statistics have been computed first at a differencing level 0. If the computed τ statistic (absolute value) is found to be greater than the critical value (i.e., - 3.13 at a chosen 10 percent confidence level which is a default value in

SHAZAM software), the series is termed as stationary. It can be observed from Table 4.1 that series 1b is stationary, and the yield series for the remaining districts are nonstationary at the differencing level 0. The nonstationary series are transformed to the stationary series for the reasons described in Section 2.4.1. The higher differencing levels are chosen until the transformed series become stationary. It can be seen from Table 4.1 that all the yield series have become stationary after the differencing level of 3 or lower.

Table 4.1 The τ statistics for yield series for the selected crop districts (Asymptotic critical value at 10 percent confidence level = -3.13)

Number of Observations in Yield Series	Differencing Level	The τ Statistic for Crop District				
		1b	3bn	4b	6a	9a
19 (1975-1993)	0	-3.40	-1.61	-2.46	-1.87	-1.17
	1		-4.15	-3.55	-2.82	-3.01
	2				-4.53	-3.75
20 (1975-1994)	0	3.58	-2.16	-2.75	-1.87	-2.55
	1		-3.91	-3.47	-2.96	-2.43
	2				-4.22	-3.13
	3					-4.1
21 (1975-1995)	0	-3.73	-2.23	-2.58	-2.09	-2.57
	1		-4.08	-3.76	-3.03	-2.56
	2				-4.46	-3.73

The difference in stationarity among the original yield series (i.e., before any transformation) demonstrates that a single time series technique cannot be applied to all of the series under study. This is because, as explained in Section 2.4, the time series techniques are stationarity-specific. Ideally, the possibility of fitting a trend to the stationary series can be ruled out because a trend may fit to a nonstationary series but not to a stationary series in which the mean and variance remain constant over time. To verify which technique will best model a series, six forecasting techniques were used: (i) linear trend, (ii) quadratic trend, (iii) simple exponential smoothing, (iv) double exponential smoothing, (v) double moving averaging, and (vi) ARIMA. Tables 4.2 and 4.3 present coefficients of the models that were found to fit the yield series for the selected districts. With applications of these models, yields were forecasted for 1994, 1995, and 1996, using the 1975-1993, 1975-1994, and 1975-1995 yield data, respectively.

4.2 Evaluation of the Forecasting Techniques

Various models were developed for each selected district and yields were forecasted. The forecasted and reported yields are shown in Table 4.2. The Mean Absolute Percent Error (MAPE), as defined by Equation 4.1, was used to evaluate the forecasting performance of a model.

$$\text{MAPE} = \frac{\sum_{i=1}^{n} | Y_{est.i} - Y_i | / Y_i * 100}{n} \qquad [4.1]$$

where $Y_{est.i}$ is the estimated (i.e., forecasted) yield for year i, Y is reported yield, and n is the number of observations in the yield data used for the model development.

Table 4.2 The coefficients of the time series models that fitted the yield series for the selected districts

Crop District	Yield Data Range	Coefficients							
		Trend analysis					Exponential smoothing		
		Linear		Quadratic			Simple	Double	
		β_0	β_1	β_0	β_1	β_2	α	α	β
1b	1975-93	1.652	0.0118	1.8212	-0.037	0.0024	0.049	1.1677	0.0620
	1975-94	1.664	0.0099	1.777	-0.021	0.0015	-0.014	1.1256	0.0246
	1975-95	1.654	0.0113	1.776	-0.020	0.0014	-0.0029	1.1745	0.0413
3bn	1975-93	1.723	0.0033	2.194	-0.131	0.0067	0.0451	0.1762	0.1740
	1975-94	1.719	0.0039	2.129	-0.108	0.0053	0.0779	0.4495	0.1038
	1975-95	1.711	0.0049	2.082	-0.091	0.0044	0.0063	0.1571	0.1817
4b	1975-93	1.676	-0.009	2.141	-0.143	0.0066	0.4835	0.5388	0.1617
	1975-94	1.646	-0.005	2.116	-0.134	0.0061	0.2681	0.7902	0.0094
	1975-95	1.561	0.006	2.176	-0.155	0.0073	0.4631	0.5241	0.0100
6a	1975-93	1.815	-0.005	2.024	-0.066	0.0030	0.0574	0.1187	0.2428
	1975-94	1.797	-0.003	2.019	-0.064	0.0029	0.0180	0.7165	0.0061
	1975-95	1.787	-0.002	2.001	-0.006	0.0025		0.5025	0.0711
9a	1975-93	1.851	0.011	1.950	-0.017	0.0014	0.0935	0.3416	0.1463
	1975-94	1.856	0.0103	1.927	-0.009	0.0009	0.0713	0.335	0.0159
	1975-95	1.856	0.0097	1.913	-0.004	0.0006	0.0158	0.2442	0.1511

Table 4.3 Estimated coefficients of the ARIMA models developed for the selected districts

Yield Data Range	Crop District	Model Specifications	Estimated coefficients							
			AR(1)	AR(2)	AR(3)	MA(1)	MA(2)	MA(3)	MA(4)	Constant
1975-93	1-b	ARIMA (2,0,2)	0.272	-0.194		0.077	0.818			1.607
	3-bn	ARIMA (2,1,2)	-0.053	-0.376		0.556	0.327			0.009
	6-a	ARIMA(3,2,0)	-0.818	-0.734	-0.715					0.021
	9-a	ARIMA (3,2,0)	-0.826	-0.881	-0.957					0.026
1975-94	1-b	ARIMA(2,0,2)	0.318	-1.001		0.284	-0.873			3.018
	3-bn	ARIMA(2,1,2)	0.259	-0.443		1.035	0.235			0.020
	6-a	ARIMA(3,2,0)	-0.805	-0.718	-0.695					0.003
	9-a	ARIMA(3,3,1)	-0.979	-0.999	-0.965	0.905				-0.026
1975-95	1-b	ARIMA(2,0,2)	0.323	-1.001		0.302	-0.920			2.974
	3-bn	ARIMA(3,1,2)	0.462	-0.491	0.257	1.140	0.064			0.002
	6-a	ARIMA(3,2,1)	-0.387	-0.516	0.557	1.098				-0.012
	9-a	ARIMA(3,2,4)	-0.942	-0.851	-0.609	0.693	0.391	0.366	-0.539	0.004

Table 4.4 Evaluating performance of various forecasting techniques on the selected crop districts

Crop District	Year	Reported Yield (t/ha)	Forecasted Yield (t/ha)					
			Trend Analysis		Exponential Smoothing		Moving Average	ARIMA
			linear	Quadratic	simple	double		
1b	1994	1.76	1.89	2.06	1.81	1.93	1.85	1.21
	1995	1.98	1.87	1.99	1.72	1.67	1.75	2.21
	1996	2.24	1.90	2.03	1.73	1.74	1.87	1.86
		MAPE	9.37	8.98	12.91	15.88	11.08	19.94
3bn	1994	1.83	1.79	2.26	1.80	1.74	2.09	1.95
	1995	1.88	1.80	2.21	1.80	2.11	2.02	1.76
	1996	1.74	1.82	2.19	1.81	1.86	1.86	1.78
		MAPE	3.68	22.3	3.31	8.02	9.52	5.08
4b	1994	1.78	1.48	1.94	2.01	1.94	2.20	
	1995	2.42	1.53	2.00	1.75	2.15	2.01	
	1996	2.21	1.69	2.31	2.14	2.20	2.10	
		MAPE	25.72	10.29	14.59	6.87	15.17	
6a	1994	1.87	1.70	1.91	1.78	1.70	1.91	2.14
	1995	1.83	1.73	1.95	1.80	1.87	1.86	2.21
	1996	2.14	1.74	1.96	1.80	1.86	1.85	1.83
		MAPE	11.08	5.70	7.45	8.12	5.78	16.56
9a	1994	2.02	2.07	2.17	1.98	2.04	1.92	2.80
	1995	2.04	2.07	2.14	1.99	2.11	2.11	1.94
	1996	2.62	2.08	2.13	2.00	2.10	2.03	1.89
		MAPE	8.19	10.34	9.37	8.09	10.47	23.79

The values of the MAPE for each district are shown in Table 4.4. On the basis of the MAPE, one can decide which technique has provided the best forecast. It can be seen from Table 4.4 that the quadratic trend, simple exponential smoothing, double exponential smoothing, quadratic trend, and double exponential smoothing have produced the best forecasts for the districts 1b, 3bn, 4b, 6a, and 9a, respectively., The linear trend technique was found to be the best for none of the districts.

Although the time series techniques have oeen selected on the basis of the MAPE, their selection cannot be generalized as the selection procedure depends not only on the yield data but also on the deviation of the yield (to be forecasted) from the series average. Furthermore, the forecasting performance of a model tends to improve if the value to be forecasted does not disturb the series characteristics which were identified before the model was developed. It is noteworthy here that the process of characterization of time series achieves a greater reliability if the sample size is larger (say, 30 or more).

In this chapter, attempts were made to generate long-term yield estimates by modelling the past yield data using time series techniques. These estimates are available before wheat is sown. Based on the forecasted yields, it can be estimated whether a drought of certain intensity may occur or not. Drought intensity can be gauged on the basis of the percent reduction in the forecasted yield from the long-term average of the yield in a district. An average yield for the entire Prairies can also be determined from the forecasted yields and the corresponding harvested area for the Prairie districts.

The estimation of wheat yields much before wheat is even sown is of great significance to the export marketing strategy. However, these estimates tend to be coarse as they do not depend on the data which directly affects the yield (i.e., weather data over the growing season). Yield estimates using climatic data over the growing season are currently being produced using the Western Canada Wheat Yield model based on the monthly weather data. However, crop-physiological considerations suggest that the use of daily data is likely to improve yield estimation. The third hypothesis which concerns monthly versus daily data is tested in the following chapter.

CHAPTER FIVE

EFFECT OF TEMPORAL RESOLUTION OF CLIMATE DATA ON YIELD ESTIMATION

The WCWY model currently being used by the CWB employs only monthly temperature and precipitation data to obtain short-term yield estimates for the entire Prairies. It was discussed in Section 2.5.4 that the WCWY model estimated the daily data (temperature and precipitation) from their monthly averages. The estimated daily data are then utilized to compute the daily drought index for each weather station. However, the estimated daily data for a weather station remain constant throughout a month. This constant value could be significantly different from the daily data actually recorded at the weather station. It was therefore considered appropriate to determine if the performance of the model is improved by the direct use of the recorded daily data. In other words, by testing the third hypothesis, the effect of the temporal resolution of data (monthly versus daily) on the predictive capability of the model was investigated.

5.1 Hypothesis III and Methodology

The third hypothesis states: "Direct use of daily climate data, instead of monthly data, can lead to an improvement in short-term yield estimates."

The execution of the task required to test the above hypothesis called for an overhaul of the computer program which is currently being used to obtain the short-term yield estimates. The original program written in Pascal was obtained from the Canadian Wheat Board and converted to C++ for the present study. The program was able to compute drought indices only for the past 12 years (Bullock, 1996). For example, if the program is run after the cropping season has passed (i.e., after August) in 1996, the drought indices produced by the program will include the period from 1985 to 1996, thus totalling 12 years. However, this limited period of 12 years was not considered sufficient for the present study which is based on a longer span of data (1975 to 1996, a total of 22 years). The longer span of data was preferred in order to enhance the reliability of any statistical inference drawn from the data analysis.

For computing drought indices for the required number of years (i.e., 22) and also to access daily weather data, the computer program was modified. Two separate programs were written in C++: one that computes drought indices on the basis of monthly data, and the other that computes drought indices on the basis of daily data. These computations were performed for the weather stations falling within the selected districts. Although the total number of such weather stations was 13 (2, 3, 2, 4, and 2 for the districts 1b, 3bn, 4b, 6a, and 9a, respectively), the daily data (since 1975) were available only for nine stations. It may be recalled that the availability of the daily data was a factor affecting the selection of the districts (Section 2.8.3). The stations for which the drought indices were computed using monthly and daily data were: Broadview (district 1b); Elrose, Pennant, and Swift Current (district 3bn);

Empress (district 4b); Duval and Watrous (district 6a); and Prince Albert and North Battleford (district 9a). Empress, though falling just outside the boundary of district 4b, was selected as the data were not complete for the stations falling within the district boundary.

5.2 Monthly versus Daily Data

Tables 5.1. and 5.2 present values of the drought indices that were computed using the modified computer programs, accessing the monthly and daily data, respectively, for all nine of the weather stations. The model using daily data has been referred to as the modified model. To evaluate how the use of daily data has affected the performance of the model, an average drought index was determined for the district from a drought index (or indices) for weather station(s) within the district (Table 5.3).

The average drought index for a district was regressed against the district yield using 1975-1996 data. Since the purpose of developing a regression model is only for prediction, the coefficient of determination, r^2, can be chosen as a measure of performance of the model (Gujarati, 1995). In addition to r^2, the Durban Watson (DW) statistic, and a coefficient (i.e., α) reflecting autocorrelation in the residuals (difference between the observed yields and the estimated yields) are also considered as factors that can be used to evaluate performance of a regression model. The values of r^2, DW statistic, and α are shown in Table 5.4.

It can be observed from Table 5.4 that the direct use of daily data has significantly increased the values of r^2 in all of the five districts except for district 9a which did not exhibit a relationship between yield and drought index ($r^2 = 0.00$). Non-existence of a relationship between the yield and drought index in district 9a where soil moisture supply is usually adequate leads to the conclusion that the concept of drought index for estimating yield is not appropriate for the areas with adequate soil moisture. Secondly, it can also be inferred that the direct use of daily

Table 5.1 Drought indices derived from monthly data for weather stations within the selected districts

Year	Districts and Weather Stations								
	1b	3bn			4b	6a		9a	
	road-view	Elrose	Pennant	Swift Current	Empress	Duval	atrou	Prince Albert	North attleford
1975	57.2	58.1	54.5	58.3	55.0	57.2	57.7	60.0	58.5
1976	60.6	55.2	58.5	60.6	48.5	58.1	59.8	60.6	59.9
1977	57.8	58.8	59.6	60.0	41.9	59.0	59.5	43.6	61.4
1978	56.8	55.2	56.5	56.4	55.6	58.0	61.9	54.9	58.4
1979	51.4	48.7	52.1	51.4	50.0	48.8	52.8	58.3	55.7
1980	58.1	52.4	59.6	57.7	50.2	51.0	57.7	61.1	59.9
1981	54.6	56.3	56.6	60.2	56.8	56.3	58.7	57.6	59.1
1982	60.7	58.1	58.9	61.8	55.8	60.1	61.8	60.3	60.7
1983	51.1	54.2	55.0	54.1	52.6	53.1	54.3	54.0	56.3
1984	50.2	52.5	51.4	47.5	50.6	48.1	54.0	56.1	56.8
1985	53.8	49.3	48.6	49.9	46.5	59.6	58.6	50.7	57.7
1986	60.1	57.3	57.1	61.3	51.3	60.8	61.8	58.0	57.8
1987	61.3	53.8	50.9	61.8	47.5	59.0	60.1	62.0	60.3
1988	56.0	43.7	44.7	50.7	42.6	43.9	48.1	55.1	59.3
1989	53.8	48.0	55.6	57.1	45.0	52.5	50.2	55.0	56.4
1990	61.2	57.6	59.3	60.7	47.5	59.5	56.9	61.4	60.8
1991	59.2	56.3	56.7	58.2	54.4	56.1	56.7	57.7	53.4
1992	47.0	58.4	61.9	51.0	56.7	51.2	52.0	58.5	56.8
1993	54.4	56.6	54.2	49.9	57.6	53.3	51.3	49.5	49.7
1994	58.0	54.7	57.6	60.9	56.0	61.2	61.4	58.5	59.8
1995	56.1	58.4	57.4	58.1	55.3	57.7	60.2	57.8	55.0
1996	57.5	54.0	52.8	55.2	51.0	55.9	55.7	57.6	58.3

[Source for monthly data: Environment Canada] data has enhanced the predictive power of the model in the remaining districts which are drought prone to varying extent. Therefore, the modified model can be treated as an improvement over the original WCWY model for obtaining the short-term wheat yield estimates.

Table 5.2 Drought indices derived from daily data for weather stations within the selected districts

Year	District and Weather Station(s)								
	1b	3bn			4b	6a		9a	
	Broadvie	Elros	Pennan	Swift Current	Empress	Duva	Watrous	Prince Albert	North Bat.
1975	54.9	55.9	51.5	57.6	52.3	55.8	56.2	59.9	59.3
1976	60.4	57.1	58.4	59.4	50.2	58.5	59.2	60.5	58.5
1977	59.1	59.7	59.7	61.9	42.1	60.9	61.2	58.1	61.4
1978	59.9	54.3	53.2	53.7	54.7	57.7	58.8	61.5	55.0
1979	50.3	48.7	53.5	52.4	48.0	51.1	52.5	57.4	56.3
1980	57.2	54.9	60.7	57.0	49.9	54.4	58.2	60.1	59.6
1981	53.1	55.9	54.2	58.1	55.1	56.6	57.5	55.1	58.3
1982	55.6	57.3	57.6	60.0	54.5	57.8	58.5	59.3	59.6
1983	53.8	54.8	55.2	55.5	53.4	54.4	55.1	54.6	55.7
1984	45.9	49.0	47.6	44.4	45.4	47.1	51.8	55.1	55.3
1985	61.4	45.8	45.1	46.1	39.9	59.0	61.6	60.0	60.9
1986	60.0	56.1	57.0	60.2	50.9	58.3	59.5	60.6	61.9
1987	60.8	54.0	50.9	61.0	46.7	59.7	60.7	61.2	61.6
1988	55.3	41.3	43.0	50.5	38.5	46.1	48.0	57.6	60.0
1989	51.9	46.4	55.8	56.1	43.0	53.8	49.6	55.1	56.6
1990	58.8	55.8	56.3	57.7	45.7	57.1	55.8	58.0	58.1
1991	59.8	57.7	58.0	59.5	56.6	57.7	57.8	57.7	54.3
1992	59.2	58.2	58.9	61.7	56.8	61.9	62.0	61.6	60.7
1993	59.9	61.2	60.5	56.9	61.8	60.5	60.2	60.3	58.4
1994	61.6	54.3	55.7	58.0	54.1	59.1	59.5	60.8	58.3
1995	58.3	58.8	57.6	59.9	54.8	58.5	59.5	58.7	59.4
1996	58.5	57.1	53.7	58.4	51.2	57.8	58.0	58.6	60.3

[Source of daily data: Environment Canada]

Table 5.3 Average drought indices for the selected districts using monthly and daily data

Year	Yield (t/ha) in District					Using Monthly Data					Using Daily Data				
	1a	3bn	4b	6a	9a	1a	3bn	4b	6a	9a	1a	3bn	4b	6a	9a
1975	1.53	1.67	1.82	1.66	1.83	57.2	57.0	55.0	57.5	59.3	54.9	55.0	52.3	56.0	59.6
1976	1.69	2.21	1.86	2.16	2.08	60.6	58.1	48.5	59.0	60.3	60.4	58.3	50.2	58.9	59.5
1977	2.28	2.04	1.37	2.05	2.01	57.8	59.5	41.9	59.3	52.5	59.1	60.4	42.1	61.1	59.8
1978	2.25	1.71	1.61	1.98	1.91	56.8	56.0	55.6	60.0	56.7	59.9	53.7	54.7	58.3	58.3
1979	1.08	1.53	1.77	1.37	1.61	51.4	50.7	50.0	50.8	57.0	50.3	51.5	48.0	51.8	56.9
1980	1.37	1.54	1.61	1.37	1.92	58.1	56.6	50.2	54.4	60.5	57.2	57.5	49.9	56.3	59.9
1981	1.90	1.80	1.84	1.78	1.92	54.6	57.7	56.8	57.5	58.4	53.1	56.1	55.1	57.1	56.7
1982	1.74	2.21	2.13	2.18	2.00	60.7	59.6	55.8	61.0	60.5	55.6	58.3	54.5	58.2	59.5
1983	1.58	1.88	1.94	1.89	1.90	51.1	54.4	52.6	53.7	55.2	53.8	55.2	53.4	54.8	55.2
1984	1.26	1.28	1.22	1.33	1.84	50.2	50.5	50.6	51.1	56.5	45.9	47.0	45.4	49.5	55.2
1985	1.78	0.91	0.83	1.90	2.20	53.8	49.3	46.5	59.1	54.2	61.4	45.7	39.9	60.3	60.5
1986	2.31	2.08	1.92	2.00	1.97	60.1	58.6	51.3	61.3	57.9	60.0	57.8	50.9	58.9	61.3
1987	1.93	1.86	1.50	1.93	1.91	61.3	55.5	47.5	59.6	61.2	60.8	55.3	46.7	60.2	61.4
1988	1.25	0.50	0.44	0.61	1.30	56.0	46.4	42.6	46.0	57.2	55.3	44.9	38.5	47.1	58.8
1989	1.36	1.67	0.93	1.47	2.15	53.8	53.6	45.0	51.4	55.7	51.9	52.8	43.0	51.7	55.9
1990	2.51	2.08	0.56	1.89	2.58	61.2	59.2	47.5	58.2	61.1	58.8	56.6	45.7	56.5	58.1
1991	2.10	2.23	2.23	1.97	2.31	59.2	57.1	54.4	56.4	55.6	59.8	58.4	56.6	57.8	56.0
1992	1.96	1.96	2.16	1.98	1.63	47.0	57.1	56.7	51.6	57.7	59.2	59.6	56.8	62.0	61.2
1993	1.74	2.21	2.23	1.84	2.20	54.4	53.6	57.6	52.3	49.6	59.9	59.5	61.8	60.4	59.4
1994	1.76	1.83	1.78	1.87	2.02	58.0	57.7	56.0	61.3	59.2	61.6	56.0	54.1	59.3	59.6
1995	1.98	1.88	2.42	1.83	2.04	56.1	58.0	55.3	59.0	56.4	58.3	58.8	54.8	59.0	59.1
1996	2.24	1.74	2.21	2.14	2.62	57.5	54.0	51.0	55.8	58.0	58.5	56.4	51.2	57.9	59.5

[Source for monthly and daily data: Environment Canada]

Table 5.4 Yield models based on monthly and daily data for selected districts

Crop district	Yield Model							
	Using monthly data (original model)				Using daily data (modified model)			
	Regression equation	r^2	Durban Watson - statistic	Autocor. Coeff., α	Regression equation	r^2	Durban Watson Statistic	Autocor. Coeff., α
1b	Y = -0.7557 + 0.0455 D	0.20	1.65	0.13	Y = -1.7960 + 0.0630 D	0.43	1.995	-0.04
3bn	Y = -3.8281 + 0.1008 D	0.71	2.38	-0.23	Y = -3.0889 + 0.0879 D	0.82	2.16	-0.10
4b	Y = -2.9189 + 0.0891 D	0.57	1.97	-0.05	Y = -2.3390 + 0.0794 D	0.73	1.79	0.04
6a	Y = -1.9208 + 0.0659 D	0.58	1.49	0.17	Y = -2.7257 + 0.0792 D	0.71	1.62	0.14
9a	Y = 2.2364 – 0.0041 D	0.00	1.65	0.06	Y = 2.2220 - 0.0038 D	0.00	1.66	0.06

Note: Y represents the annual wheat yield of district (t/ha), and D denotes the average drought index for district using 1975-96 data.

The above analysis has established the significance of the daily data in improving the performance of the WCWY model. To further improve the performance of the model, as discussed in Section 2.6, satellite data-based variables can be incorporated into the model. It would therefore be appropriate to study the relationship between the NOAA/AVHRR-based NDVI data and the yield. Accordingly, an appropriate NDVI-based variable could be derived. The derived variable could then be incorporated into the modified model to determine if the performance of the model is improved. This notion is appropriately expressed in hypothesis IV which is tested in the following chapter.

CHAPTER SIX

DEVELOPMENT OF A HYBRID MODEL

In Section 2.6, the utility of the NOAA/AVHRR data in monitoring vegetation conditions was defined. It was further noted that the NDVI derived from the NOAA data is widely used to monitor crop conditions over large areas. Realising the significance of the NDVI data, the fourth hypothesis was constructed.

6.1 Hypothesis IV and Methodology

The fourth hypothesis states, " Incorporation of the NDVI data into the current wheat yield model (i.e., the Western Canada Wheat Yield model) can improve the performance of the model in obtaining short-term yield estimates."

In order to test the above hypothesis, the NDVI data were acquired from Statistics Canada for the period of 1987 to 1996; prior to 1987, the NDVI data were not available. The NDVI data refer to the NDVI values of the weekly composites (i.e., the temporal resolution of the NDVI data was one week). In order to facilitate the combined use of satellite and climatic data into a hybrid model, it was considered appropriate to convert the weekly NDVI data into the daily data just by retaining weekly values for every day in the corresponding week. In order to select the most suitable NDVI-based variable for inclusion into the hybrid model, the relationships of various derivatives of the NDVI data with yield were studied.

6.2 Derivation of the NDVI-based Variables

Variation in the NDVI data during the period from sowing to harvest (in other words, the NDVI profile) is considered to be related to fluctuation in yield. Simple parameters capable of characterizing the NDVI profile are the mean and standard deviation. Therefore, two variables – average NDVI (AvgNDVI) and standard deviation in the NDVI data (StdevNDVI) – were determined from the daily NDVI data for the entire growing period (May 1 to August 31).

In addition, the maximum value of NDVI data during the growing period (MaxNDVI) was computed as a variable because the maximum NDVI reflects the best vegetative condition of crop, which may be related to yield. Furthermore, the NDVI profile corresponding to the year of maximum yield was considered as a reference profile. The coefficient of correlation between the reference profile and the NDVI profile in a given year was computed. This coefficient of correlation was considered as a measure of deviation from the maximum yield and was referred to as CoefCor_NDVI_max_yld. However, the yield is influenced significantly by the crop conditions during the heading stage which usually falls during July in the case of wheat over the Prairies (Bullock, 1995). This led to the formation of two additional variables from NDVI data during July: average NDVI during July (i.e., AvgNDVI_Jul) and standard deviation in NDVI data during July (i.e., StdevNDVI_Jul).

The commencement and termination of the heading stage depends on the planting date, maximum and minimum daily temperatures, and the photoperiod. Using these parameters, Robertson (1968) developed a biometeorological time scale model to estimate duration of various phenological stages (Section 2.6.1). Robertson's model was applied on the selected districts, using 1987-96 data. The planting dates were collected from the Winnipeg Climate Centre (Raddatz, 1996). Photoperiod was estimated for the locations of the selected weather stations using a computer program (Robertson and Rusello, 1968). The original program was written in Fortran, but was converted to C++ for the present study. Another computer program in C++ was developed to apply the Robertson's model, using the coefficients provided in Chapter 2 (Table 2.2), to estimate the beginning and the end of various phenological stages for all nine of the weather stations selected for the study. Following the estimation of timing and duration of the heading stage, the corresponding NDVI data were identified for each weather station. Subsequently, an additional variable -- namely, avgNDVI_heading s -- was generated for each district. The variables thus derived from the NDVI data are shown in Table 6.1.

6.3 Development of the Hybrid Model

To determine which variable should be incorporated into the modified model, a regression analysis was performed between yield and each of the NDVI-based variables, using 1987-1996 data. Table 6.1 presents values of coefficient of determination, r^2, which is considered as a means to measure significance of a variable. Since the length of the NDVI data is limited (only 10 years), it was not possible to develop the regression model on the initial few years of data and test the model over the remaining years. However, under the present situation of limited satellite data, the derived variables were compared on the basis of r^2.

Table 6.1 Coefficient of determination, r^2, pertaining to linear regression model developed between the reported yield and various NDVI-based variables

Crop District	Coeff. of Determination, r^2						
	Average NDVI during May-Aug., AvgNDVI	Average NDVI during July, AvgNDVI_Jul	Maximum NDVI during May-Aug., MaxNDVI	Std. Dev. in NDVI during May-Aug., Stdev NDVI	Std. Dev. in NDVI during July, Stdev NDVI_Jul	Coeff. of Corr. between NDVI profile and reference profile, CoefCor NDVI_max_yld	Average NDVI during heading stage, AvgNDVI_heading
1-b	0.46 (11.98)	0.62 (10.91)	0.45 (12.97)	0.43 (13.45)	0.06 (17.06)	0.46 (14.09)	0.79 (8.28)
3-bn	0.62 (18.02)	0.21 (30.63)	0.40 (26.25)	0.24 (30.75)	0.31 (27.65)	0.06 (33.60)	0.46 (24.25)
4-b	0.61 (26.38)	0.61 (32.03)	0.54 (32.37)	0.13 (60.17)	0.18 (54.21)	0.59 (39.59)	0.78 (25.48)
6-a	0.76 (13.6)	0.63 (16.0)	0.74 (12.89)	0.52 (19.53)	0.26 (23.53)	0.24 (24.11)	0.78 (11.57)
9-a	0.00 (15.9)	0.29 (15.4)	0.38 (14.1)	0.38 (12.8)	0.01 (16.4)	0.31 (13.3)	0.37 (14.19)

Note: Values inside brackets are the corresponding MAPEs.

From Table 6.1 it can be observed that, the r^2 was highest for districts 1b, 4b, and 6a while using AvgNDVI_heading. Further, use of AvgNDVI gave highest r^2 for district 3bn, and MaxNDVI or StdevNDVI for district 9a. It was therefore decided to select AvgNDVI for inclusion into the modified model (developed in Section 5.1) for the district 3bn, and the AvgNDVI_heading for the remaining four districts; the value of r^2 using AvgNDVI_heading was very close to the highest for district 9a. The models thus developed with inclusion of the NDVI-based variable are referred to as hybrid models. These models were developed using the dataset presented in Table 6.3.

Table 6.2 Hybrid models for selected districts

Crop District	Hybrid Model	R^2
1b	Y = - 11.2116 + 0.002437 D + 0.0008659 N	0.79
3bn	Y = -13.4582 + 0.0815 D + 0.0008385 N	0.96
4b	Y = - 9.9967 + 0.0496 D + 0.000523 N	0.83
6a	Y = - 8.3463 + 0.05219 D + 0.0004933 N	0.95
9a	Y = - 4.0661 - 0.03405 D + 0.0005369 N	0.39

Note: Y = yield (t/ha), N = Average NDVI during May-Aug. i.e., AvgNDVI value (for district 3bn) and averge NDVI during the heading stage i.e., AvgNDVI_heading, for the remaining districts, and D = drought index (based on the daily data).

Table 6.3 Dataset used to develop hybrid models for the selected districts

Year	Yld-1b	yld-3bn	yld-4b	yld-6a	yld-9a	di-d-1b	di-d-3bn	di-d-4b	di-d-6a	di-d-9a	av-nd-1b	av-nd-3bn	av-nd-4b	av-nd-6a	av-nd-9a	av-nd-hs-1a	av-nd-hs-3bn	av-nd-hs-4b	av-nd-hs-6a	av-nd-hs-9a
1987	1.93	1.86	1.5	1.93	1.91	60.8	55.3	46.7	60.2	61.4	13902	12778	12430	13407	14107	15098	13202	12737	14622	15142
1988	1.25	0.5	0.44	0.61	1.3	55.3	44.9	38.5	47.05	58.8	13485	12302	11936	12744	13925	14661	12955	12189	13352	14693
1989	1.36	1.67	0.93	1.47	2.15	51.9	52.8	43.0	51.7	55.85	13359	12846	12353	12931	13881	14200	14072	12868	14200	15539
1990	2.51	2.08	0.56	1.89	2.58	58.8	56.6	45.7	56.45	58.05	13955	13038	12245	13267	14189	15596	14214	12713	14753	15815
1991	2.1	2.23	2.23	1.97	2.31	59.8	58.4	56.6	57.75	56	13924	12965	12712	13350	14094	15040	14127	13821	14564	15388
1992	1.96	1.96	2.16	1.98	1.63	59.2	59.6	56.8	61.95	61.15	14081	12801	12582	13387	13984	15053	13718	13313	14334	15156
1993	1.74	2.21	2.23	1.84	2.2	59.9	59.5	61.8	60.35	59.35	13641	12690	12552	13084	13700	14595	13767	13459	14063	14660
1994	1.76	1.83	1.78	1.87	2.02	61.6	56	54.1	59.3	59.55	13853	12856	12571	13286	13788	14979	13958	13503	14738	15251
1995	1.98	1.88	2.42	1.83	2.04	58.3	58.8	54.8	59	59.05	13589	12592	12380	13076	13409	15157	13847	13404	14636	14583
1996	2.24	1.74	2.21	2.14	2.62	58.5	56.4	51.2	57.9	59.45	13679	12766	12346	13279	13761	15194	14120	13377	14993	15487

Note: Yld = wheat yield (t/ha), di-d = drought index using daily data, av-nd = average NDVI over the growing period considered (May 1-August 31), av-nd-hs = average NDVI during heading stage.

6.4 Evaluation of the Hybrid Model

Table 6.4 presents the comparison between the modified model and the hybrid model on the basis of r^2. It is confirmed from Table 6.4 that the predictive power of the hybrid model is greater than that of the modified model. Nonetheless, it was not possible to reliably test the hybrid model, as the sample size was very small. With such a small sample size, it was not feasible to select some of the observations for the model development and leave the remaining observations for model testing, while ensuring adequate reliability.

Table 6.4 Comparison of various models on the basis of coefficient of determination, r^2

Crop District	Modified model			Hybrid model		
	r^2	Durban-Watson Statistic	Autocor. Coeff., $\propto$	r^2	Durban-Watson Statistic	AutoCor. Coeff., $\propto$
1b	0.43	1.995	-0.04	0.79	1.38	0.17
3bn	0.82	2.16	-0.10	0.96	2.92	-0.57
4b	0.73	1.79	0.04	0.83	1.27	0.20
6a	0.71	1.62	0.14	0.95	1.49	0.07
9a	0.00	1.66	0.06	0.39	1.77	-0.02

CHAPTER SEVEN

CONCLUSIONS AND RECOMMENDATIONS

As stated in Section 2.7.1, the objectives of the present study were: (i) to develop an improved method for obtaining short-term and long-term wheat yield estimates for the Canadian Prairies, and (ii) to improve understanding of factors contributing to drought prediction based on yield prediction for the Canadian Prairies. In order to meet these objectives the hypotheses presented below were formulated and then tested on five crop districts of Saskatchewan (districts 1b, 3bn, 4b, 6a, and 9a). Saskatchewan was selected for the study because it produces the largest amount of wheat compared to the other two Prairie provinces. The selected districts were located in varying agroclimatic conditions and were associated with varying degrees of fluctuation in wheat yield, thus enhancing the scope of the testing. The four hypotheses of the study were:

I. Pattern recognition techniques can be applied to predict drought effectively on a short-term basis;

II. Long-term estimates of wheat yield can be obtained using time series analysis. Different types of time series models may best fit the yield series in different crop districts, disputing the current practice of only using a linear trend model for the Prairie region as a whole;

III. Direct use of daily climatic data, instead of monthly data, can lead to an improvement in short-term yield estimates; and

IV. Incorporation of the NDVI data into the current wheat yield model (i.e., the Western Canada Wheat Yield model) can improve the performance of the model in obtaining short-term yield estimates.

The above hypotheses were tested leading to the following conclusions.

7.1 Conclusions

1. Using the Error Correction procedure of pattern recognition, it was not possible to find a solution vector that would linearly separate drought patterns from nondrought patterns. Hence, qualitative drought prediction was not successful for district 3bn. This corroborates the complex nature of agricultural drought. Nevertheless, there could exist a nonlinear classifier, which might be discovered using nonlinear classification techniques. These techniques were not attempted in the present study.

2. Time series analysis succeeded in providing long-term estimates of wheat yield. However, different models were found to be suitable in different districts selected. The quadratic trend, simple exponential smoothing, double exponential

smoothing, quadratic trend, and double exponential smoothing techniques of time series analysis produced the best forecast for districts 1b, 3bn, 4b, 6a, and 9a, respectively. This was at variance with what is presently practiced (i.e., linear trend only) to obtain long-term forecast of wheat yield. It would therefore be more appropriate to model yield series at the district level and to forecast its yield for the year ahead. From the district-level forecasts average yield for the entire Prairies could be estimated taking into account the district level yields and the corresponding harvested areas.

For a time series analysis, the length of the series plays a significant role. The longer the series, the better the forecast. In the present study, the wheat yield data were available only since 1975. With a longer yield series, one could expect more reliable findings.

3. Direct use of daily climatic data (temperature and precipitation), instead of the estimated daily data according to the Western Canada Wheat Yield model, has significantly improved the predictive power of the model. The coefficient of determination, r^2, obtained as a result of regression between yield and drought index, increased from 0.20, 0.71, 0.57, and 0.58 to 0.43, 0.82, 0.73, and 0.71, for districts, 1b, 3bn, 4b, and 6a, respectively. However, for district 9a, where the soil-water availability is adequate for crop production, the value of r^2 was 0.00 in both cases (related to monthly and daily data). Therefore, the concept of estimating yield on the basis of a drought index (using either monthly or daily data) does not hold well for a district which is not drought prone. For such a

district other methods of yield estimation could be explored.

Furthermore, the climatic data (monthly and daily) which were used to compute drought indices were available only for a limited number of weather stations. The total number of weather stations (with desired length of data) existing within the selected districts was 13, but the daily data (since 1975) were available only for nine stations, which were used in the present study. If the data were available for additional stations, more reliability in the regression models could be achieved.

4. The use of NDVI data further increased the predictive power of the yield model. Two NDVI-based variables were found to be the most suitable for inclusion into the yield models. Except for district 3bn, average NDVI during the heading stage identified using a biometeorological time scale, was found to be the most suitable variable for inclusion into the yield model. For district 3bn, average NDVI over the entire growing period (i.e, May 1-August 31) was the best predictor. When these variables were included along with the drought index (computed using recorded daily data) as explanatory variables against the yield (dependent variable), the values of r^2 were 0.79, 0.96, 0.78, 0.95, and 0.39 for districts 1b, 3bn, 4b, 6a, and 9a, respectively, as opposed to 0.20, 0.71, 0.57, 0.58, and 0.00 from the existing Western Canada Wheat Yield model.

Conclusively, the predictive powers of the hybrid models were found to be significantly higher than those of the current model. More reliability would likely be achieved with a longer study period.

On the basis of the conclusions drawn from testing the hypotheses of the present study, it can be summarized that, by using hybrid models, wheat yields can be estimated with greater accuracy. Furthermore, drought can be predicted where the estimated yields are found to be significantly lower than the long-term averages of wheat yields. Nevertheless, an appropriate threshold for the drought-defining reduction of wheat yield needs to be set to best suit the needs of the marketing strategies of agencies such as the Canadian Wheat Board.

7.2 Recommendations for Further Research

The following is suggested for further research on drought prediction for the Canadian Prairies.

1. Nonlinear classification techniques of pattern recognition (e.g., nearest neighbourhood) should be attempted to examine the possibility of separating drought from nondrought events for qualitative prediction of drought.

2. Daily temperature and precipitation series should be modelled separately using appropriate time series techniques. Subsequently the daily data (temperature and precipitation) can be forecasted for the next growing season. Utilizing the forecasted daily data, drought index and then wheat yield can be estimated.

3. As the NOAA/AVHRR data are not available under cloudy conditions, it would be prudent to make use of microwave data (for example, European Remote Sensing Satellite, RADARSAT, or Indian Remote Sensing Satellite-C) which are available under cloudy conditions, for wheat yield estimation in significantly cloudy growing seasons. Microwave data have the capability of supplying information on crop condition and hence final yield by estimating various parameters such as the leaf area index, total biomass, and crop canopy (Ulaby and Bush, 1976; Le Toan et al., 1984; Daughtry et al., 1991; Major et al., 1992).

REFERENCES

Aasae, J.K., and F.H. Siddoway. 1981. Spring wheat yield estimates from spectral reflectance measurements. *IEEE Transactions on Geosciences and Remote Sensing* 19(2):78-84.

Aber, J.D., and J.M. Melillo 1991. *Terrestrial Ecosystems.* Chicago: Saunders College Publishing.

Abraham B., and J. Ledolter. 1983. *Statistical Methods for Forecasting.* New York: John Wiley and Sons.

Acton, D.F., G.A. Padbury, and C.T. Stushnoff. 1998. *The Ecosystems of Saskatchewan.* Saskatchewan Enviornmental Resources Management, Canadian Prairies Research Centre, Regina, Saskatchewan.

Ahlrichs, J.S., and M.E. Bauer. 1983. Relation of agronomic and multispectral reflectance characteristics of spring wheat canopies. *Agronomy Journal* 75: 987-993.

Asrar, G., E.T. Kanemasu, and M. Yoshida. 1985. Estimates of leaf area index from spectral reflectance of wheat under different cultural practices and solar angle. *Remote Sensing of Environment* 17:1-11.

Awumah, K., C. Booy, and L.M. Lye. 1990. Accumulated basin storage as an indicator of agricultural drought. *Canadian Water Resources Journal* 15(2):105-124.

Bair, W. 1977. Crop weather models and their use in yield assessments. *WMO Technical Note No. 151*. Geneva: World Meteorological Organization.

Bair, W. 1982. Agroclimatic modeling - an overview. *Agroclimatic Information for Development: Reviving the Grain Revolution*. [D. F. Cusack (ed.)]. Boulder, Colorado: Westview. 57-82.

Barnett, T. L., and D. R. Thompson. 1982. The use of large-area spectral data in wheat yield estimation. *Remote Sensing of Environment* 12: 509-518.

Bird, P.M., and D.J. Rapport 1986. *State of the Environment Report for Canada.* Ottawa: Minister of Supplies and Services.

Bonsal, B.R., and A.K. Chakravarti 1993. Teleconnections between north Pacific SST anomalies and growing season extended dry spells on the Canadian Prairies. *International Journal of Climatology* 13:865-878.

Booy, C., and L.M. Lye. 1986. Accumulated basin storage as a factor in the correlation structure of annual peak flows in the Red River. *Canadian Journal of Civil Engineering* 13(3):365-374.

Borchert, J.R. 1971. The dust bowl in the 1970's. *Annals of American Assoc. Geog.* 61(1):1-22.

Box, G.E.P., and G.M. Jenkins. 1976. *Time Series Analysis: Forecasting and Control.* San Fransisco: Holden-Day.

Brooks, R.H. 1972. Flight from disaster: drought perception as a force in migration from Ceara, Brazil. Doctoral dissertation (unpublished). Boulder, Colorado: Universty of Colorado.

Brooks, R. H. 1973. Differential perception of drought in northeast Brazil. *Proceedings of the Association of American Geographers* 5:31-35.

Brown, R.J.E. 1967. Map – Permafrost in Canada. Geological Survey of Canada, map 1246A.

Brown, R.J. 1986. An overview of remote sensing agricultural applications in North America: past present and future. *Proc. International Geoscience and Remote Sensing Symposium, Zurich,* Sept. 8-11, 733-737.

Brown, R.J. , M. Bernier, G. Fedosejevs, and L. Skretkowicz. 1982. NOAA /AVHRR crop condition monitoring. *Symposium Photogrammetric and Remote Sensing,* Toulouse, France. Sept. 13-17.

Bullock, P.R. 1992. Operational estimates of western Canadian grain production using NOAA/AVHRR LAC data. *Canadian Journal of Remote Sensing* 8(1):23-28.

Bullock, P.R. 1995. Personal communication. July, 1995. Winnipeg, Manitoba, Canada: Weather and Crop Surveillance division, Canadian Wheat Board.

Bullock, P.R. 1996. Personal Communication. March, 1996. Winnipeg, Manitoba, Canada: Weather and Crop Surveillance division, Canadian Wheat Board.

Burton, I., R. Kates, and G. White (eds.). 1978. *The Environment as Hazard.* New York: Oxford University Press.

Campbell, C. A., R.P. Zenter, and P.J. Johnson. 1988. Effect of crop rotation and fertilization on the quantitative relationship between spring wheat yield and moisture use in southern Saskatchewan, *Canadian Journal of Soil Science* 68:1-16.

Chow, V.T. 1964. *Handbook of Applied Hydrology*. New York: McGraw-Hill.

Clevers, J.G.P.W. 1989. The application of weighted infra red vegetation index for estimating leaf area index by correcting for soil moisture. *Remote Sensing of Environment* 29:25-37.

Curtis, L.F. 1978. Remote sensing systems for monitoring crops and vegetation. *Progress in Physical Geography* 2:55-79.

Daughtry, C.S.T., K.J. Ranson, and L.L. Biehl. 1991. C-band backscattering from corn canopies. *International Journal of Remote Sensing* 12(5):1097-1109.

Diaz, R.A., A.D. Matthias, and R.J. Hanks. 1983. Evapotranspiration and yield estimation of spring wheat from canopy temperature. *Agronomy Journal* 75:805-810.

Doorenbos, J., and W.O. Pruitt. 1977. *Crop Water Requirements*. FAO Irrigation and Drainage Paper 24. Rome, Italy: Food and Agriculture Organisation.

Dracup, J.A., K.S. Lee, and E.G. Paulson, Jr. 1980. On definitions of droughts. *Water Resources Research* 16(2):297-302.

Duda, R.O., and P.E. Hart. 1973. *Pattern Recognition and Scene Analysis*. New York: John Wiley and Sons.

Emanuel, W.R., H.H. Shugart, and M.P. Stevensen 1985. Climate change and broad scale distribution of terrestrial ecosystem complexes. *Climate Change* 7:29-43.

Environment Canada. 1989. *Water Shortages in Canadian Prairies*. Regina, Saskatchewan: Inland Waters Directorate, Western and Northern Region.

Garcia, R.V. 1981. *Nature Pleads Not Guilty*. Oxford: Pergamon Press.

Garnett, E.R., and M.L. Khandekar 1992. The impact of large-scale atmospheric circulations and anomalies on Indian monsoon droughts and floods and on world grain yield — a statistical analysis. *Agricultural and Forest Meteorology* 61:113-128.

Garnett, E.R., M.L. Khandekar, and J.C. Babb 1998. On the utility of ENSO and PNA Indices for long-lead forecasting of summer weather over the crop growing region of the Canadian Prairies. *Theor. Appl.Climatol.* 60:37-45.

Gillette, H.P. 1950. A creeping drought underway. *Water and Sewage Works*, March issue:104-105.

Goward, S. N., B. Markham, D.G. Dye, W. Dulaney, and J. Yang. 1991. Normalized difference vegetation index measurements from the advanced very high resolution radiometer. *Remote Sensing of Environment* 35: 257-277.

Goward, S. N., C. J. Tucker, and D. G. Dye. 1985. North American vegetation patterns observed with the NOAA-7 advanced very high resolution radiometer. *Vegetation* 64: 3-14.

Grace, B. and D.L. Johnson. 1985. The drought of 1984 in southern Alberta: its severity and effects. *Canadian Water Resources Journal* 10(2):29-38.

Great Britain Meteorological Office 1951. *The Meteorological Glossary*. New York: Chemical Publishing Co.

Gujarati, D.N. 1995. *Basic Econometrics*. New York: McGraw Hill.

Gutman, G. G. 1987. The derivation of vegetation indices from AVHRR data. *International Journal of Remote Sensing* 8(8):1235-1243.

Gutman G.G. 1991. Vegetation indices from AVHRR - An update and future prospects. *Remote Sensing of Environment* 35:121-136.

Hall, A.L. 1978. *Drought and Irrigation in the Northeast Brazil.* Cambridge, London:Cambridge University Press. 22-23.

Hanke J.E. and A.G. Reitsch. 1995. *Business Forecasting*. Fifth ed. Englewood Cliff, New Jersey: Prentice Hall.

Hatfield, J. L. 1983. Remote sensing estimators of potential and actual crop yield. *Remote Sensing of Environment* 13:301-311.

Heathcote, R.L. 1969. Drought in Australia: A problem of perception. *Geographical Review* 59(2):175-194.

Heathcote, R.L. 1974. Droughts in south Australia. *Natural Hazards: Local, National, Global.* [White, G.F.(ed.)]. London: Oxford University Press.

Hewitt, K. 1983. Interpreting the role of hazards in agriculture. *Interpretations of Calamity from View Point of Human Ecology*.[Hewitt, K. (ed.)].Boston: Allen and Unwin Inc. 123-139.

Holben, B. N. 1986. Characteristics of maximum-value composite images from temporal AVHRR data. *International Journal of Remote Sensing* 7(11): 1417-1434.

Idso, S.B., J.L Hatfield, R.J. Raginato, and R.D. Jackson. 1978. Wheat yield estimation by albedo measurement. *Remote Sensing of Environment* 7:273-276.

Jackson, E.L., and T. Mukerjee. 1974. Human adjustment to the earthquake hazard to San Francisco, California. *Natural Hazards: Local, National, Global.* [White. G.F.(ed.)]. New York: Oxford University Press. 160-174.

Kates, R. W. 1967. The perception of storm hazard on the shores of magalopolis. *Environmental Perception and Behaviour* [D. Loenthal (ed.)]. Research Paper No. 109, Department of Geography. Chicago: University of Chicago. 60-74.

Kates, R.W., J.H. Ausubel, and M. Berberian (eds.).1985. *Climate Impact Assessment*. Toronto: John Wiley and Sons.

Kelly, P.M. and P.B. Wright. 1978. The European drought of 1975-76 and its climatic context. *Progress in Physical Geography* 2:237-263.

Kirkby, A.V. 1974. Individual and community responses to rainfall variability in Oaxaco, Mexico. *Natural Hazards: Local, National, Global.* [White, G.F.(ed.)]. New York: Oxford University Press. 119-128.

Knox, J.L., and R.G. Lawford. 1990. The relationship between Canadian Prairie dry and wet months and circulation anomalies in the mid-troposphere. *Atmosphere-Oceans* 28(2):189-215.

Krishnan, A. 1979. Definitions of drought and factors relevant to specifications of agricultural and hydrological droughts. *Proc. Int. Symp. on hydrological aspects of droughts.* Vol I. New Delhi, India: Indian Institute of Technology. 67-102.

Kumar, V. 1988. Satellite remote sensing and management of agricultural droughts in India. *Proc. National Symp. Remote Sensing in Rural Development*, Indian Society of Remote Sensing, Nov. 17-19. Hisar, India: Haryana Agricultural University. 7-15.

Kumar, V. 1993. A predictive assessment of agricultural drought using climatic and remotely sensed data. Master of Engineering dissertation (unpublished), University of Jodhpur, Jodhpur, Rajasthan, India.

Kumar, V. 1998. An early warning system for agricultural drought in an arid region using limited data. *Journal of Arid Environments* 40(2): 199-209.

Kumar, V. and U. Panu 1994. On application of pattern recognition in drought classification. *Proc. Annual Conference, Can. Soc. Civil Engineers*, Winnipeg, June1- 4.

Kumar, V., and U. Panu. 1997. Predictive assessment of severity of agricultural droughts based on agro-climatic factors. *Journal of American Water Resources Association* 33(6):1255-1264.

Kumar, V., C.E. Haque, and M. Pawlak. 1998. Determining agricultural drought using pattern recognition. *Prairie Perspectives – Geographical Essays*. [Lehr, J.C., and H.J. Selwood (eds.)]. Winnipeg: Department of Geography, University of Winnipeg.

Laycock, A.H. 1964. *Water Deficiency Patterns in the Prairie Provinces*. Prairie Provinces Water Board, Regina, Saskatchewan, Rep. No. 8:34

Laycock, A.H. 1967. *Water Deficiency and Surplus Patterns in the Prairie Provinces*. Prairie Provinces Water Board, Regina, Saskatchewan, Rep. No. 13:185.

Le Toan, T., A. Lopes, and M. Huet. 1984. On the relationships between radar backscattering coefficient and vegetation canopy characteristics. *Proc. International Geoscience and Remote Sensing Symposium*, Strasburg, France, Aug. 27-30:155-160.

Liverman, D.M. 1990. Drought impacts in Mexico: climate, agriculture, technology, and land tenure in Sonora, and Puebla. *Annals Amer. Assoc. Geog.* 80: 49-72.

Lockwood, J.G. 1986. The causes of drought with particular reference to the Sahel. *Progress in Physical Geography* 10(1): 111-119.

Longley, R.W. 1972. *The Climate of the Prairie Provinces.* Climatological Studies No. 13, Environment Canada, Toronto.

Major, D.J., M.J. Hill, W.D. Willms, B.Brisco, R.J. Brown. 1992. Radar backscatter from a short grass prairie. Proc. International Geoscience and Remote Sensing symposium, Houston, Texas, May 26-29:1683-1685.

Mass, S.J. 1988. Use of remotely sensed information in agricultural crop growth models. *Ecological Modelling* 41: 247-268.

McKay, G.A. 1983. Climatic hazards and Canadian wheat trade. *Interpretations of Calamity from the View Point of Human Ecology.* [Hewitt, K. (ed.)]. Boston: Allen and Unwin Inc. 220-228.

Morgan, R. 1985. The development and application of a drought early warning system in Botswana. *Disasters* (9): 44-50.

Oliver-Smith, A. 1982. Here there is a life: the social and cultural dynamics of successful resistance to resettlement in post-disaster Peru. *Involuntary Migration and Ressetlement: the Problems and Responses of Dislocated People.* [Hansen, A. and A. Oliver-Smith (eds.)]. Boulder, Colorado: Westview Press. 85-104.

Palmer, W.C. 1965. *Meteorological Drought.* Weather Bureau. Washington., D.C: U.S. Dept. of Commerce, Paper No. 45: 1-58.

Parry, M.L., T.R. Carter, and N.T. Konijn (eds.). 1988. *Impact of climatic variations on agriculture Vol. 1, Assessment in Cool temperature and cold regions.* London: Kluwer Academic Publishers.

Prankatz, A. 1983. *Forecasting with Univariate Box-Jenkins Models.* New York: John Wiley and Sons.

Prouit, N., N. Dalezios, M. Manore, J. Tracey, R. Brown, and G. Walker. 1986. *Agricultural drought: precipitation deficiency and NOAA/AVHRR indices comparison. Proc. Canadian Hydrology Symp.* June 3-6. Regina, Saskatchewan: 91-100.

Raddatz, R.L., C.F. Shaykewich, and P.R. Bullock. 1994. Prairie crop yield estimates from modelled phenological development and water use. *Can. J. Plant Sci.* 74: 429-436.

Raddatz, R. 1996. Biometeorological time scale computations. Winnipeg Climate Centre, Environment Canada (unpublished).

Rao, A.R., and G. Padmanabhan. 1984. Analysis and Modeling of Palmer's Drought Index Series. *Journal of Hydrology* 68:211-229.

Robertson, B., A. Ericson, J. Friedel, B. Guindon, T. Fisher, R. Brown, P. Teillet, M. D'Iorio, J. Cihlar, and A. Sanz. 1992. GEOCOMP - a NOAA/AVHRR geocoding and geocompositing system. *Proc. Int. Soc. Photogram. Rem. Sen.*, Washington, D.C.:223-228.

Robertson, G.W. 1968. A biometeorological time scale for a cereal crop involving day and night temperatures and photoperiod. *Int. J. Biometeor.* 12 (3):191-223.

Robertson, G.W. (ed.) 1983. *Guidelines on Crop Weather Models.* Task Force on crop weather models. Geneva: World Meteorological Organization.

Robertson, G.W. and D.A. Russelo. 1968. Agrometeorological estimator for estimating time when the sun is at any elevation, time elapsed between the same elevations in the morning and afternoon, and hourly and daily values of solar energy, *Q. Ag. Met. Tech. Bull.* 14, Agrometeorology Section, Agriculture Canada, Ottawa, 22 pp.

Rosenberg, N.J. (ed.) 1980. *North American Droughts.* Westview Press, Boulder, Colorado.

Saarinen, T.F. 1966. *Perception of Drought Hazard on the Great Plains.* Department of Geography, Research Paper No. 106. Chicago: University of Chicago.

Sakamoto, C.M. 1978. The Z-index as a variable for crop yield estimation. *Agricultural Meteorology* 19: 305-313.

Sakamoto, C.M. 1981. Climate crop regression yield model: An appraisal. *Application of Remote Sensing to Agricultural Production Forecasting*. [Berg, A. (ed.)]. Balkema, Rotterdam.131-138.

Saskatchewan Agriculture and Food. 1997. **http://www.agr.gov.sk.ca/saf** .

[accessed June 12].

Scott, G.A.J. 1995. *Canada's Vegetation: A World Perspective.* Montreal: McGill-Queen's University Press.

Sen, Z. 1990. Critical drought analysis by second order markov chain. *Journal of Hydrology* 120:183-202.

Shukla, J. 1985. *Predictability - Issues in Atmospheric and Oceanic Modelling.* [S. Manabe (ed.)]. New York: Academic Press. 87-122.

Sinha, C.P., U.C. Chaube, and R.P. Saxena. 1992. A critical analysis of drought indices for different regions. *Proc. World Congress on Natural Hazards*, New Delhi, January 10-14.

Slabbers, P. J., and F.X. Dunin. 1981. Wheat yield estimation in northwest Iran. *Agricultural Water Management* 3 : 291-304.

Smart, G.M. 1983. Drought analysis and soil moisture prediction. *Journal of Irrig. Drain. Div., ASCE* 109(2):251-261.

Statistics Canada. 1996. *Report - Geocomp Re-Processing project: Processing Methodlogy and Summary Report.* Ottawa: Statistics Canada, Spatial Analysis and Geomatics Applications, Agriculture Division. July, 1996.

Statistics Canada. 1998. *Estimates of production of principal crops*. Field Crop Reporting Series No. 8. Ottawa: Statistics Canada.

Tarpley, J.D., S.R. Schneider, and R.L. Money. 1984. Global vegetation indices from the NOAA-7 meteorological satellite. *Journal of Climate and Applied Meteorology* 23:491-495.

Thompson, D.R., and O.A. Wehmanen. 1979. Using Landsat digital data to detect moisture stress. *Photogrammetric Engineering and Remote Sensing* 45(2): 201-207.

Tucker, C.J. 1979. Red and photographic infrared linear combinations for monitoring vegetation. *Remote Sensing of Environment* 8:127-150.

Tucker, C. J., B.N. Holben, J. H. Elgin Jr., and J. E. III McMurtrey. 1980. Relationship of spectral data to grain yield variation. *Photogrammetric Engineering and Remote Sensing* 46(5): 657-666.

Tucker, C.J., J.A. Gatlin, and S.R. Schnieder. 1984. Monitoring vegetation in the Nile Delta with NOAA-6 and NOAA-7 AVHRR imagery. *Photogrammetric Engineering and Remote Sensing* 50(1):53-61.

Ulaby, F.T., and T.F. Bush. 1976. Monitoring wheat growth with radar. *Photogrammetric Engineering and Remote Sensing* 42(4):557-568.

van Diepen C.A. and T van der Wall. 1996. Crop growth monitoring and yield forecasting at regional and national scale. *Proc. Workshop for Central and Eastern Europe on Agrometeorological Models: Theory and Applications, The MARS Project*. [Dallemend and P. Vossen, (eds.)] Ispra, Italy, November 21-25, 1994. Luxembourg: European Commission. 143-157.

Walker, G.K. 1989. Model for operational forecasting of western Canada wheat yield. *Agricultural and Forest Meteorology* 44:339-351.

Walker, G.T. and E.W. Bliss. 1932. World Weather. V. *Mem. Royal Meteorological Society* 4:53-84.

Wallace, J.M. and D.S. Gutzler 1981. Teleconnections in the geopotential height field during the northern hemisphere winter. *Monthly Weather Review*, 109:784-812.

Warrick, R.A. 1983. Drought in the US Great Plains: shifting social consequences. *Interpretations of Calamity from view point of human ecology*. [Hewitt, K. (ed)]. Boston: Allen and Unwin Inc.:67-82.

Wiegand, C. L., A. J., Richardson, and E.T. Kanemasu, 1979. Leaf area estimates for wheat from LANDSAT and their implications for evapotranspiration and crop modeling. *Agronomy Journal* 71: 336-342.

Wiegand, C.L., and A.J. Richardson. 1983. Leaf area, light interception, and yield estimates from spectral components analysis. *Agronomy Journal* 76: 543-548.

Wiegand, C. L., and A. J. Richardson. 1990. Use of spectral vegetation indices to infer leaf area, evapotranspiration and yield: I. Rationale. *Agronomy Journal* 82: 623-629.

Wiegand, C. L., A. J. Richardson, D.E. Escobar, and A.H. Gerbermann. 1991. Vegetation indicies in crop assessments. *Remote Sensing of Environment* 35:105-119.

Wilhite, D.A., and M.H. Glantz. 1987. Understanding the drought phenomenon. *Planning for Drought: Toward a Reduction of Societal Vulnerability*. [Wilhite, D.A., W.E. Easterling, and D.A. Wood (eds.)]. Boulder: Westview Press. 11-27.

Wisner, B. and P. Mbithi. 1974. Drought in Eastern Kenya: nutritional status and farmer activity. *Natural Hazards: Local, National, Global*. [White, G. (ed)]. New York: Oxford University Press. 87-97.

World Meteorological Organization. 1985. *Report of the WMO/UNEP/ICSU-SCOPE, expert meeting on the reliability of crop climate models for assessing the impact of climate change and variability, WCP-90.* Geneva: World Meteorological Organization.

Yevjevich, V., W.A. Hall, and J.D. Salas. 1978 (ed.). *Drought Research Needs.* Fort Collins, Colorado: Water Resources Publications.