NEW APPLICATIONS OF PATTERN RECOGNITION TO POWER SYSTEM SECURITY

A THESIS SUBMITTED TO

THE FACULTY OF GRADUATE STUDIES

IN PARTIAL FULFILLMENT

OF THE REQUIREMENTS FOR THE DEGREE

DOCTOR OF PHILOSOPHY

by

ELSAYED ABD-ELALIEM MOHAMED

B.Sc., Ain-Shams University, Egypt, 1977

M.Sc., Ain-Shams University, Egypt, 1983

DEPARTMENT OF ELECTRICAL ENGINEERING

THE UNIVERSITY OF MANITOBA

WINNIPEG, CANADA

ISBN 0-315-37141-2

NEW APPLICATIONS OF PATTERN RECOGNITION TO POWER SYSTEM SECURITY

BY

ELSAYED ABD-ELALIEM MOHAMED

A thesis 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

ABSTRACT

Modern power systems complexity increases daily as new interconnections are made. For this reason, energy control centers are developed and being implemented in order to provide real-time security assessment and control for these systems. The prediction of system anomalies is therefore essential.

A new prediction system based on pattern recognition techniques is designed. This system is to predict generator self-excitation of the Manitoba Hydro Northern power system. Converter station blocking, sometimes followed by machine trips, is the contingency behind that condition. A corrective algorithm based on sensitivity analysis is developed to provide operators with suggested actions in order to improve system security.

Secondly, a prediction system is designed to predict dynamic overvoltages due to converter blocking (load rejection). Overvoltage corrective actions are derived using the corrective algorithm. A least squares based algorithm is developed in order to estimate the post-contingency dynamic overvoltages.

Thirdly, a multi-class prediction system approach is investigated. The approach is new, it has some advantages and at the same time it possess some limitations. It was applied to the Northern system in order to predict self-excitation as well as dynamic overvoltages.

Finally, a new pattern recognition based voltage contingency analysis algorithm is developed. Using this algorithm a discriminant hyperplane is defined as a performance index. Using this index, a screening and ranking of voltage outages (contingencies) can be efficiently achieved. The speed of computation involved is slightly lower, in the case of on-line design, when compared to other methods. The off-line design approach is efficient and reliable but it has an excessive computational burden. The algorithm and other methods are investigated on a sample system and on the Northern power system as well.

ACKNOWLEDGEMENTS

I wish to express my deep appreciation to Professor G. W. Swift, for suggesting this topic as well as his counsel, guidance, patience and encouragement during the course of this study.

The author is indebted to Mr. F. A. Jost, and the research committee of Manitoba Hydro for the financial aid and encouragement during the different stages of this project. The staff of system operating department at Hydro, Mr. Haywood, Dr. Chand, are also acknowledged for useful discussions and the assistance for using their digital simulation and computational facilities.

It is my great honour to have Prof. G. Heydt as my external examiner; Prof. M. Z. Tarnawecky, Prof. A. Gole, and Prof. F. Zeiler as my committee of examiners. Thanks are due to them for their valuable time, advice, and comments. Thanks, also, are due to Prof. E. Shwedyk and Prof. Pawlak for their valuable discussions and comments on using pattern recognition techniques.

The support from the National Sciences and Engineering Research Council is also acknowledged.

Finally, I am particularly grateful to my parents, my wife Hala, and daughter Hadiel for their unwavering support and encouragement over the course of this project.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

CHAPTER 1

INTRODUCTION

1.1 Power System Operation

Electric power systems consist mainly of generation, transmission, and distribution systems. The distribution systems must supply the consumption centers, or load centers, with a continuous service of satisfactory quality. Therefore, this means that the power system should operate in such a way to provide the customer with a service of the quality requirements. These requirements mean that both the voltage and frequency at the load centers must be within certain limits.

In order to satisfy the customer's demands, the electrical transmission network should operate within certain constraints that are acceptable for the system equipment operation. This imposes new constraints on the planning and operation of power system. In general, all the necessary conditions required for normal operation of a power system could be expressed in two sets of constraints : the *equality constraints* called the load constraints, and the *inequality constraints* called the operating constraints. The load constraints set the conditions which satisfy the requirements that the load demands will be met by the system, while the operating constraints impose maximum or minimum operating limits (e.g. voltage limits,loading limits, ..) on variables associated with the component parts of the system.

The conditions of operation of a power system can be categorized into three operating states : *normal* , *emergency* , and *restorative* [1,2]. A system is in the normal state when the load and operating constraints are satisfied. The system may be operating within the space of all feasible normal operating states. This space is completely defined by the load constraints and the operating constraints. A system is in the emergency state when the operating constraints are not satisfied. For example, when an equipment loading limit is exceeded or when the voltage limit at a bus is

violated. A system is in the restorative state when the loading constraints are violated. This means a condition of either a partial or total shutdown.

1.2 Power System Security Assessment

Modern interconnected power systems demand a high degree of security for normal operation. This demand is due to the fact that some contingencies can have a catastrophic results on these systems.

Security assessment is a new approach to power system operation, helping the system operator to detect conditions that may lead to a failure or deterioration in the quality of the power supply before it occurs. The security analysis could be considered as two parts [3], the first is to determine whether the system is in the secure state or not. The second is to determine what corrective action should be taken when the system is insecure.

Three types of security assessment can be identified, the *steady state* , the *transient state* , and the *dynamic state* . The steady state security assessment examines the steady state response of the system under credible outage conditions. Each contingency in the steady state security analysis causes transients which can result in very undesirable conditions, such as loss of synchronism of generators. To prevent such conditions fast corrective action must be taken which requires transient security assessment. The dynamic security assessment pertains to system response of the order of a few minutes.

The results of a security assessment study must be presented to the power system operator in such a way so that he may be able to take necessary corrective action effectively [3]. Such information may be presented in terms of simple indices and may be a part of the power system control centers duties.

1.3 Motivation and Literature Review

The main objective of this project has been guided materially by Manitoba Hydro. Several discussions at Hydro were behind the need of an on-line *prediction* system to assess the *dynamic security* of the Northern system against generator *self −excitation* and system *dynamic overvoltages* , due to the blocking of converter stations (dc load rejection). Also, a prediction (security) index that tells the operator how far the system is from having self-excitation was required.

Referring to literature, it was found that the first study on load rejection overvoltages was done by de Mello [4] using analog simulation. It was concluded that generator self-excitation can occur more readily with hydro-generators than with steam-generators, because of the higher overspeeds obtained after load rejection as well as the associated extensive transmission systems.

Dandeno [5] has investigated the effect of dynamic overvoltages, produced on the 500 KV line of Ontario Hydro due to load rejection on hydro-generators, on the system design e.g. insulation coordination, relaying schemes, etc. It was reported also that when self-excitation occurs, the voltage regulator becomes unstable and can increase system overvoltages.

De Mello then reported [6] that the lack of negative field current capability, on the static excitation systems, causes self-excitation when the direct axis reactance (X_d) of the machine is greater than the capacitive reactance seen at the machine terminals (X_c). An extremely rapid rise in voltage (2 p.u./sec) is recorded during self-excitation. Moreover, he has recommended, as a special overvoltage protection requirement, the study with modern simulation tools which permit prediction of conditions under load rejection.

Gole in [7] is the first to investigate the effect of converter station blocking on the excitation system of the Manitoba Hydro Northern system. The study revealed that the most severe case of overvoltage occurs, due to machine trips, when a small number of machines are feeding a capacitive load (ac filters). Also, it was indicated

that the same overvoltage conditions could happen, but less severe, in the case of converter blocking (dc load rejection) when a few machines are left connected to large filter banks.

Pattern recognition techniques have been applied to power systems as a fast on-line assessment method. Several attempts have been made to apply pattern recognition to load forecasting, steady state security assessment [1,2,8], and transient security assessment [2,9-21].

Pang [1,2] has applied pattern recognition to steady state and transient state security assessment. He has used single and repetitive ranking for feature selection, least squares and optimal search for classifier design. He has obtained good simulation results on the CIGRE-225 KV system.

Tokyo Electric power company in Japan was the first utility to have experimented with using the pattern recognition approach for on-line fast transient stability assessment. The possibility of implementing a pattern recognition approach as a part of the security monitoring package has been reported in [9,10,16]. It was concluded that these advanced techniques are feasible not only for on-line security assessment, but also for off-line analysis in power system planning and operation.

Gupta [11,18] has used the linear transformation to extract required features. Also, he was the first to report on the application of PDM (Polynomial Discriminant Method) for the classifier design.

Hakimmashadi and Heydt [13,19] have recommended the use of transient measurements as features for transient stability assessment. They have investigated the application of: Bayes rule, PDM, and Nearest Neighbor for the classifier design. They have concluded that the Bayes design is the fastest approach, while the Nearest Neighbor has proved to be the most accurate. Moreover, they proved that pattern recognition is a fast, and serious candidate for power system on-line security assessment compared to other methods such as Liapunov and numerical integration.

Recent research [14] has reported that pattern recognition techniques are suitable for power system security, either in off-line mode or on-line mode. A data generation algorithm has been developed and the selection of variables related to generators as primary variables has been recommended.

New York Power Pool [8] have introduced a prediction system using pattern recognition for steady state voltage security analysis. They have used the intuitive features and Nearest Neighbor design. The system proved to be efficient and reliable and it is being implemented.

Finally, Mokhtari in [21] has applied generalized square distance and K-th Nearest Neighbor (K-NN) classifier on the large BPA 345-230 KV system for a transient security assessment. He has introduced new features (post fault as well as pre-fault features). Also, new feature selection and performance estimation methods have been investigated.

Modern energy control centers [22-26] utilized in electric utility control and monitor schemes are required to provide real-time security assessments. Steady state assessments are usually evaluated for a large number of actual and anticipated contingencies. Therefore, as the number of contingencies increase, the solution time will be a burden to the overall real-time control function. Under these circumstances, the selection of the most important (critical) contingencies to perform a detailed contingency analysis, becomes more desirable and is gaining a lot of utilities attention [27-37]. The selection process of these contingencies is called Automatic Contingency Selection (ACS).

Ejebe [28] has introduced the first ACS algorithm. Line and generator outages are ranked according to their severity as reflected in voltage level degradation and circuit overloads. The algorithm uses Tellegen's theorem to generate the sensitivities of a system-wide performance index with respect to outages.

The application of decision theory to ACS has been reported, as a new approach, by Fischl [31,36]. The approach is a theoretical one which allows

development of a method for finding the performance index as a volume maximization problem. The use of Bayes criterion has been recommended since it offers an additional flexibility to the classification problem.

Nara [27] has presented a new ACS algorithm for contingency selection concerning voltage security analysis. This algorithm proves to be efficient both in contingency selection and ranking. The algorithm is based on the definition of a performance index as a second order vector norm in the voltage space. Performing a dc load flow for each contingency, the algorithm can be used to select and rank critical contingencies.

An investigation of the current algorithms [27-37] indicated inefficient and unreliable performance due in part to a lack of one or more of the following considerations:

1- Ranking process is not efficient due to misclassified cases.

2- Generator VAR limits are ignored.

3- Bus current injections are assumed constant.

4- Screening criterion is not efficiently achieved.

5- Execution time at the beginning of sensitivity analysis is not that low.

6- Most algorithms do not perform reliably.

7- Single and multiple outages are not considered.

Due to the important considerations discussed above, a new approach to the voltage contingency analysis, screening and ranking, is developed. The approach is based on pattern recognition techniques in order to come up with the appropriate performance index design. Once the performance index is designed, off-line or on-line, then using a dc load flow to derive the post-contingency conditions, a screening and ranking of the most critical cases can be achieved.

1.4 Thesis Outlines

The main objective of this project has been discussed. Chapter 2 describes the basic configuration of a prediction system based on pattern recognition techniques. Pattern acquisition, pattern preprocessing, feature vector identification, predictor design, and performance evaluation have been discussed.

Chapter 3 deals with the load rejection overvoltages on the Northern system. Self-excitation, dynamic overvoltages, and normal conditions are explained. Effects of machine tripping, filter tripping are also investigated on the Northern system.

In Chapter 4 a prediction system is designed for self-excitation conditions due to load rejection and double contingency. A corrective action algorithm is developed.

Chapter 5 concerns with the prediction of dynamic overvoltages due to load rejection. A prediction system is designed and a least squares method is used for the estimation of the post-contingency voltages.

Chapter 6 treats the load rejection operating problem as a multi-class pattern recognition problem. A multi-class prediction system is described, advantages and limitations of this scheme are outlined.

A new approach to the voltage contingency analysis is described in chapter 7. Pattern recognition techniques are applied in this approach. A comparison with other current algorithms are included.

Finally, the general conclusions and the future recommendations are summarized.

CHAPTER 2

PATTERN RECOGNITION BASED PREDICTION SYSTEM DESIGN

2.1 Introduction

Pattern recognition, which is a part of the larger topics *Artificial Intelligence* and *Knowledge Based* methods, has been used for a wide variety of applications [1,2,8-21]. The main advantage of the pattern recognition method is that a given new pattern can be quickly identified as belonging to a known class of patterns. The intelligence to make this *classification* comes from extensive off-line computation.

Pattern recognition principles and methodologies have influenced the course of technological development in almost every knowledge-based field. In many fields, pattern recognition is an effective candidate for problem solving, capable of producing dramatic results. Pattern recognition has been very much a product of today's computer technology. *Automatic programming*, *parallel processing*, *color graphics*, *high speed computation*, and *microprocessors*, have all benefited from the concepts and methods of pattern recognition applications.

A *prediction* system is one which uses pattern recognition techniques to predict the class membership of a given pattern. This chapter considers the different phases required for the design process of the prediction system. Several algorithms are employed to come up with such design as well as to evaluate its performance.

2.2 Prediction System Configuration

The configuration of the prediction system can be divided into four phases, as shown in Fig. 2.1:

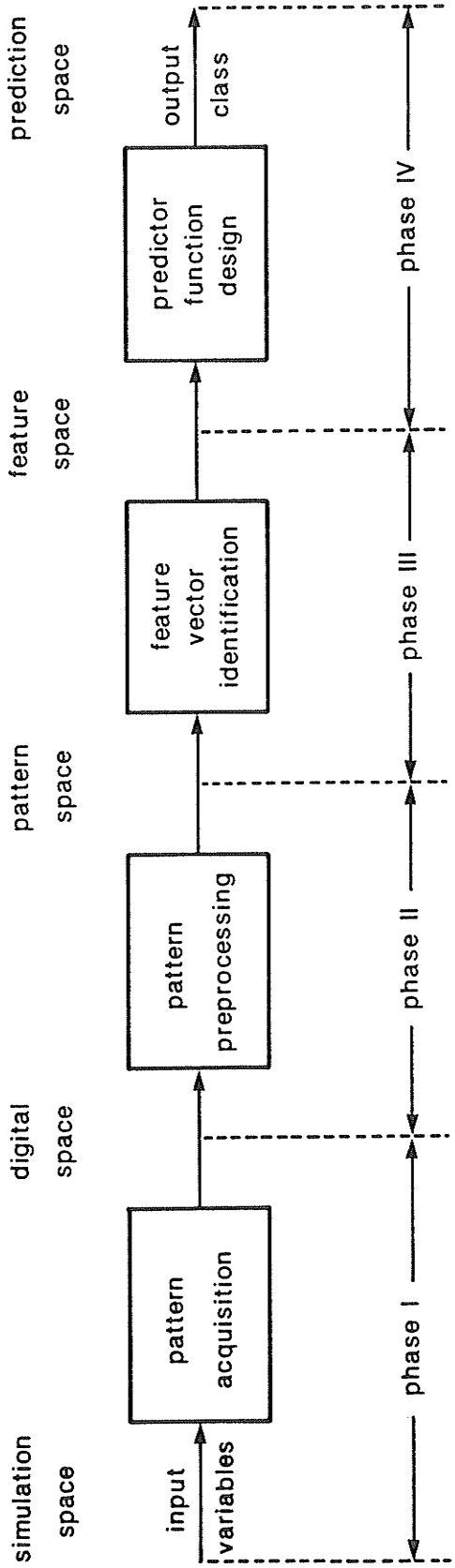


Fig. 2.1 Design process of a pattern recognition based prediction system.

(1) pattern acquisition;

(2) pattern preprocessing;

(3) feature vector identification;

(4) predictor function design.

In the *pattern acquisition* phase, digital patterns generated from a system simulation model, are gathered in a format suitable for further computer processing.

The measured patterns are then used as the input to the second phase, *pattern preprocessing*, and grouped into a set of characteristic patterns as output. The general intent of this process is simply to learn about the patterns, determine the different pattern classes that might exist in the patterns, and organize the learned knowledge into a form most efficient for further processing.

In the third phase, *feature vector* identification, some techniques are applied to the output patterns from the second phase, in order to derive the required feature vector.

The fourth phase in a typical prediction system is a *predictor function* design. This predictor is in the form of a set of *discriminant functions* required to provide the system with the appropriate decisions. Its purpose is to predict, based on the information obtained, what class would be assigned to a given unknown pattern.

2.3 Pattern Acquisition Phase

As stated before, pattern acquisition is the process of generating and converting a pattern from its main source into a form acceptable to the digital computation system for further processing. Also as explained before, the design of a prediction system is based on a set of patterns which represent typical operating conditions of the system under study.

Pattern acquisition is to obtain a *pattern set* which represents the characteristics of different classes which are to be classified or predicted. It should be emphasized, that all the consecutive analyses are based on these patterns and on the

assumption that all the information required for recognition is contained in the pattern set.

A *pattern* is a vector of variables, these variables should represent different characteristics of the object under consideration. The task of a prediction system is the identification of the common characteristics of a class of patterns and on the basis of this information, to classify a new pattern. The successful identification of these patterns is significantly dependent on the chosen pattern variables. In fact, the selection of these variables sets the lower bound for *recognition error* [13] and therefore the selection of low quality variables could weaken the applicability of a prediction system. In general, the selection of these variables require an intimate *knowledge* of the system under consideration and the specific application of the pattern recognition methodology. Pattern variables should assess the properties of each class and should fulfill three main requirements [14], namely:

- *discrimination* between different classes,
- *reliability* ; i.e. variables should take on similar values for all patterns of the same class,
- *independence* , i.e. variables should be uncorrelated with each other, although they might be combined to reduce noise sensitivity.

It is very essential to have a representative and adequate pattern set for a meaningful solution to the overall problem. The patterns should include all the information pertinent to the recognition process. As has been shown, there are statistical reasons behind most of the methodologies in pattern recognition. Therefore, the problem arises of how many patterns are needed for an adequate prediction system design. It has been shown [38,39] that the error rate of *misrecognition* on the design patterns is an estimate of the *asymptotic* error rate of the prediction system. The amount of this bias is a function of the ratio of the number of patterns per class to the number of features used. Foley [38] has shown that a ratio of greater than three will be acceptable for an accurate prediction system design.

2.4 Pattern Preprocessing Phase

This stage could be considered as a preparation stage for the next processing stages. Three activities can be involved under this stage, namely:

- pattern normalization,
- noise filtering, and
- estimation of measurements failure.

Usually, the system measurements are in different units, therefore it is necessary to normalize the input patterns in order to speed up the numerical convergence of the design process. A study [12] has shown that zero mean and unit variance is the optimal normalization algorithm.

In the field implementation of a prediction system, two problems may arise: one is due to the measurement errors which implies that additive noise is superimposed on the ideal representation of patterns, the other problem arises when one or more measurements are not available for some reason. The problem of measurement errors is *relevant* in the situation when the design of a prediction system is based on a design set of patterns generated some way different than that presented to the actual system in the field operation. Researchers [13,19] have concluded that measurement errors affect the performance optimality of the prediction system and increase the recognition error probability. However, if the pattern generating process, i.e. simulation model of the studied system, is of an acceptable accuracy for the representation of the real system characteristics for specific application, it will lead to a reasonable optimal prediction scheme.

When a designed prediction system is in the field implementation, it may happen that one or more measurements fail for some practical reasons. Thus, the recognition capability may drop to a very low level which of course will affect the overall performance of the prediction system. Costa [14] has outlined different approaches in order to overcome this problem. One of these approaches, the less demanding one, is to replace the missing measurements by their *historical* average values or by their

last values if suitable.

2.5 Feature Vector Identification Phase

The problem of concern in this stage is how to identify or extract from a set of variables $Y = (y_1, y_2, \ldots, y_v)^T$ those that have the best *discriminatory* power, called *features* or *key variables*, where v is the number of pattern variables. These features should permit the prediction system to distinguish between patterns belonging to different classes.

The number of variables obtainable for a pattern can be very large in any pattern recognition problem. Therefore, it is desirable to extract a small number of these variables from the initial set to be the features. Usually *intuition* and knowledge of the problem being studied guide the listing of potentially useful variables to be considered. However, it is a very difficult task to identify the best variables required to represent the system state.

With feature identification we aim to achieve three objectives [40-46], namely:

(i) to select with appropriate methods the most useful information from the *pattern vector* Y and to present it in the form of a *feature vector* $X = [x_1, x_2, \ldots, x_f]^T$ of lower dimensionality $f < v$, where f is the number of selected features;

(ii) to remove any redundant and irrelevant information which may have a detrimental effect on the prediction system; and

(iii) to rearrange the variables in terms of their discriminatory power in order to provide the consecutive design stages with the most informative variable to be considered.

For this stage, the following algorithms are recommended for the application of concern:

(a) Variable Separability Measure Algorithm (VSM)

In this algorithm a function *VSM* is used as a criterion in the identification of a feature [1,2] since it provides a measure of the *separability* between any two classes *j& k* for each variable. This function could be written as:

$$VSM_{jk}(i) = ABS\ [\ (M_j(i) - M_k(i)\ )\ /\ (SD_j(i) + SD_k(i)\ )\ ] \qquad (2.1)$$

$$i = 1, 2, ..., v\ .$$

where:

$M_j(i)$: mean of variable i in class j,

$M_k(i)$: mean of variable i in class k,

$SD_j(i)$: standard deviation of variable i in class j,

$SD_k(i)$: standard deviation of variable i in class k.

ABS: absolute value.

The algorithm procedures could be summarized as follows, see Appendix (A):

1- identify the variable with the largest class separability measure *VSM* ;

2- discard the highly correlated variables, i.e. variables of *correlation coef f icients* $\geq$ 0.9, and

3- repeat the above procedure till the required features are obtained.

(b) Intuitive Identification Method

In this method [8,47], usually intuition and knowledge of the problem being studied guide the listing of potentially useful variables to be considered as features. In general, features to be selected by this method are based on the availability of:

1- telemetering measurements;

2- intuition gained from experience;

3- correlation coefficients obtained from multi-variable regression analysis; and

4- engineering judgment .

(c) Stepwise Identification Algorithm

In this algorithm, as in the case of *stepwise regression* analysis [48], the identification of features to be used in the final analysis is often not known in advance. A stepwise analysis is a sequence of simple analysis that moves from one analysis to the next by adding , and sometimes deleting, a feature variable at each step. The most commonly used method [49] of selecting variables to be entered or deleted is based on the ratio of the *within class generalized dispersion* to the *total class generalized dispersion* for the selected variables. The within class generalized dispersion is simply the determinant of the within group sum of cross-product matrix $W(X)$ for the feature vector $X = (x_1, x_2, ..., x_f)^T$. As such it may be viewed as a generalization of Fisher's ratio,

$$R(X) = |W(X)|/|T(X)| \tag{2.2}$$

where : $T(X)$ is the total class sum of cross-product matrix.

The *selection measure* SM , Appendix (B), for the entry of variable y_j into the set X is given by :

$$SM_j = (a/b)\,[1/MI_j - 1] \tag{2.3}$$

and for the case of removing of the feature x_i from the set $X = (x_1, x_2, ..., x_i, ..., x_f)^T$, the *removing measure* could be written as :

$$RM_i = ((a + 1)/b)\,[MI_i - 1] \tag{2.4}$$

where MI_k *multiplicative increment* in Fisher's ratio $R(X)$ resulting from the entry of the variable y_k or the deletion of the feature x_k , and a & b are constants for a given case.

The algorithm rules can be summarized as follows:

1- not to remove a feature if its removing measure *RM* value is greater than or equal to a specified limit (threshold).

2- not to enter or select a variable if its selection measure *SM* is below a specified limit.

3- not to select a variable if its tolerance value is below a specified limit.

Tolerance for the identification of a feature variable is one minus the square of its within group multiple correlation with the currently entered variables. This algorithm is a very powerful technique since it has the advantage of the *backtracking* phenomena.

2.6 Predictor Function Design Phase

Having chosen the pattern features, this stage is to design a surface in the feature space which separates the patterns of one class from those of the other. Therefore, this design is equivalent to fitting a surface, called *predictor or classifier* in pattern recognition terminology, to correctly classify patterns from different classes. Depending on the information contained in the feature vectors and the distribution of patterns in the feature space, the predictor function can be a simple or a very complicated expression.

As stated before, the aim of a prediction system is to identify the underlying characteristics which are common to a class of patterns. The identification of these characteristics enables one to correctly predict the class of a new pattern.

Pattern recognition problems can be divided into two groups: *supervised* and *unsupervised*. In the supervised group, there is a *prior classification knowledge* available about design patterns. On the other hand, the unsupervised problem has no prior information about the classification of design patterns. Power system operating problems can be considered to be of the supervised group from the point of view of pattern recognition theory.

In pattern recognition theory, many techniques are available in order to design the prediction surface [40-46]. However, the predictor function must be as simple as possible mathematically, and must consists of as small a number of features as is allowable. Also, the prediction system should be easy to implement in the practical field. Therefore, the following design schemes are developed in order to achieve the aforementioned characteristics.

2.6.1 Hyperplane Predictor Algorithm

In this design the prediction surface is constructed from discriminant functions, the forms of which are known to be linear. The design patterns are to be used for the estimation of the predictor parameters. Also, this design is suitable for situations where no knowledge of the forms of underlying probability density distributions are required, and in this sense it can be considered as a *nonparametric* approach.

Linear discriminant functions have a variety of pleasant properties from an analytical point of view. They can be optimal if the underlying distributions are cooperative. Even, when they are not optimal, one might be willing to sacrifice some performance to gain the advantage of simplicity. Linear discriminant functions are relatively easy to compute, and a predictor of fixed structure is an attractive candidate for implementation on a digital model.

2.6.1.1 Two-Class Prediction Problem

Consider a *two−class* problem in which a predictor or discriminant function, which is a linear combination of the feature-vector components X, can be written as:

$$\begin{aligned} d(X) &= w_1 x_1 + w_2 x_2 + \cdots + w_f x_f + w_0 \\ &= W^T X + w_0 \end{aligned} \tag{2.5}$$

The f-dimensional vector $W = [\, w_1, w_2, \ldots, w_f \,]^T$ is called the *weight vector* and w_0 is the *threshold* weight. Also, the f-dimensional vector $X = [\, x_1, x_2, \ldots, x_f \,]^T$ is

called the feature vector.

The *prediction rule* corresponding to the predictor function $d(X)$ is to assign the pattern X to class 1 if $d(X) > 0$ and to class 2 if $d(X) \leq 0$. The prediction boundary is defined by the equation:

$$d(X) = 0 \tag{2.6}$$

As $d(X)$ is a linear function of X therefore the prediction surface is a linear surface i.e. a *hyperplane* in f-dimensional space given by:

$$W^T X = -w_0 \tag{2.7}$$

From the geometry of hyperplanes, a very important property can be derived ,Fig. 2.2. On this diagram, let us assume X_1 and X_2 be two patterns on the hyperplane and U be the unit vector normal to the hyperplane at pattern X_1. As X_1, X_2 are both on the hyperplane, we can write

$$W^T X_1 + w_0 = W^T X_2 + w_0 = 0$$

$$W^T (X_1 - X_2) = 0 \tag{2.8}$$

From Eq. (2.8), we can conclude that W is normal to the hyperplane and

$$U = W / \| W \|,$$

$$\| W \| = [\sum_{i=1}^{f} w_i^2]^{1/2} \tag{2.9}$$

Therefore, the vector W defines the orientation of the hyperplane.

The normal *Euclidean* distance from the origin to the hyperplane is given by :

$$X_1^T U = X_1^T \frac{W}{\| W \|} = -\frac{w_0}{\| W \|} \tag{2.10}$$

Therefore, with an appropriate normalization, the threshold weight w_0 defines the

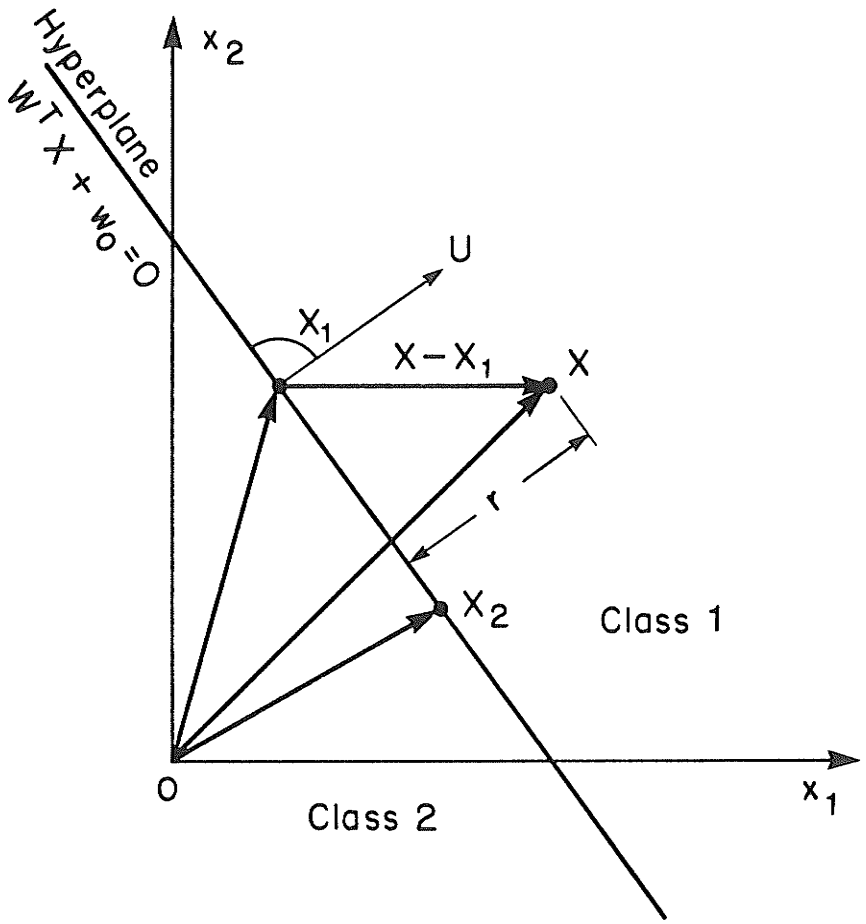


Fig. 2.2 The geometry of hyperplanes.

location of the hyperplane. Now, the Euclidean distance r from X to the hyperplane can be written as :

$$r = U^T (X - X_1) = \frac{W^T}{\| W \|} (X - X_1)$$

$$= \frac{W^T X + w_0}{\| W \|}$$

$$= d(X) / \| W \| \qquad (2.11)$$

i.e. the predictor function $d(X)$ gives an algebraic measure of the distance from X to the hyperplane.

To summarize, a linear predictor function divides the feature space by a hyperplane surface. The orientation of the surface is determined by the weight vector W and the location of the surface is determined by the threshold weight w_0 .

2.6.1.2 Multi-class Prediction Problem

The case of a *multi−class* problem can be treated in the same manner exactly as the case of the two-class problem. Many of the concepts discussed in the two-class case can be extended to the case of multi-class problem. The multi-class case offers some more possibilities and on the other hand it raises some additional difficulties.

One way of treating a c class problem is to convert it into c *two−class subproblems*. Each subproblem amounts to discriminate the patterns belonging to class i from those which do not belong. Although this approach is intuitively satisfying , it has the disadvantage that some future patterns may be found with undefined decisions or belong to several classes up to $c-1$ simultaneously, see Fig. 2.3 for a three class problem.

A second possibility is to convert the c-class problem into $c-1$ two-class subproblems exactly as described in the previous approach. This way this approach has

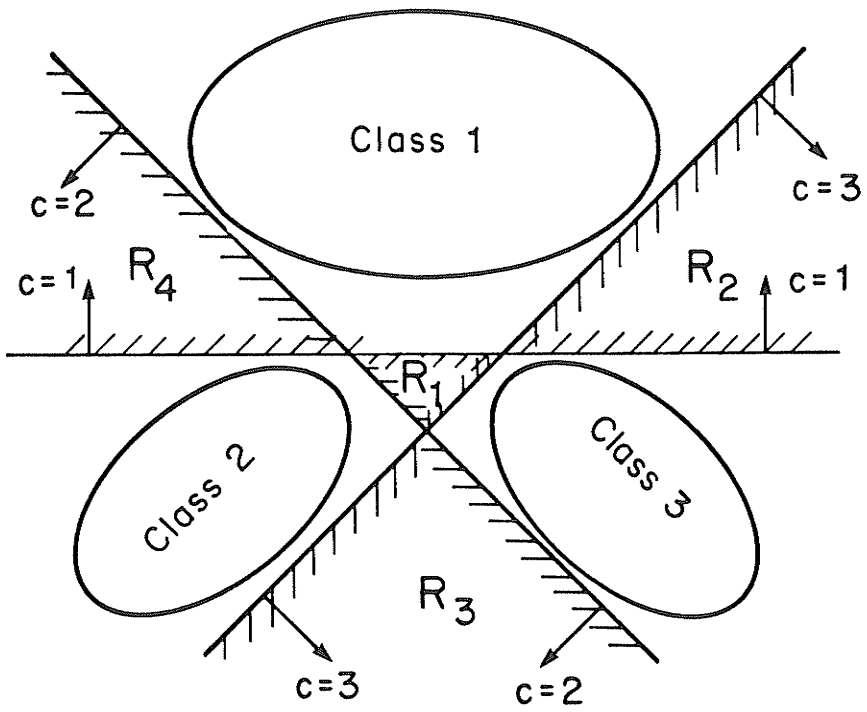


Fig. 2.3 Three two-class subproblems approach.

$R_1 - R_4$: undefined decision regions.

the advantages of a lesser number of required decisions and a reduced region of multiclass decisions.

A third approach is to convert the c-class problem into $c(c-1)/2$ two-class subproblems. This approach is called the *pairwise* predictor functions. The prediction rule in this case is to assign the unknown pattern to the class that reach $c-1$ positive answers. In addition to being costly, this approach also has the disadvantage that some future decisions may be undefined, Fig. 2.4.

All approaches discussed so far have the disadvantages of undefined or ambiguous decisions. These disadvantages can be overcome by defining c linear predictor functions as:

$$d_i(X) = w_{i1}x_1 + w_{i2}x_2 + ... + w_{if}x_f + w_{i0}$$

$$= W_i^T X + w_{i0}, \quad i = 1, 2, ..., c \tag{2.12}$$

The prediction rule in this case, is to assign pattern X to class i if $d_i(X) > d_j(X)$ for all $j \neq i$, Fig. 2.5. This way, the space is partitioned into c prediction regions R_i given by :

$$R_i = [X / d_i(X) > d_j(X); j = 1, 2, ..., c \; ; j \neq i \;] \tag{2.13}$$

Each region R_i may be shown, Fig. 2.5, to be a *convex polyhedron* with at most $c-1$ portions of the hyperplanes defined by:

$$d_i(X) = d_j(X) \; ; j = 1, 2, ..., c \; ; j \neq i \tag{2.14}$$

The predictor scheme in this case is given by Fig. 2.6 .

2.6.2 Deterministic Learning Algorithms

As stated before, the predictor function has a weighting vector W and a threshold weight w_0. The process of determining these weights is called *learning*. In the pattern recognition literatures, there are many learning algorithms, Appendix (C), however the algorithms selected here are superior since they are based on

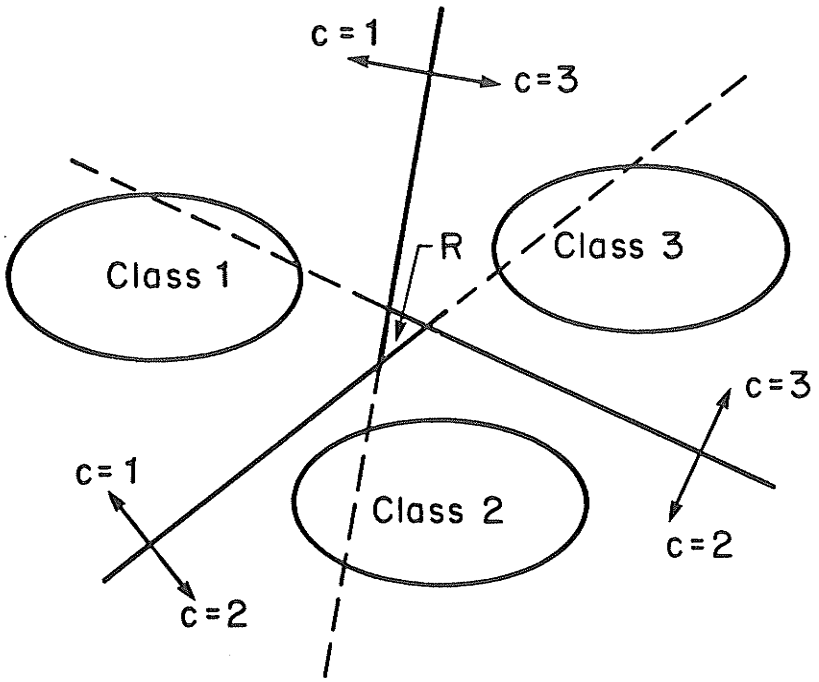


Fig. 2.4 Pairwise predictor approach.

R : undefined decision region.

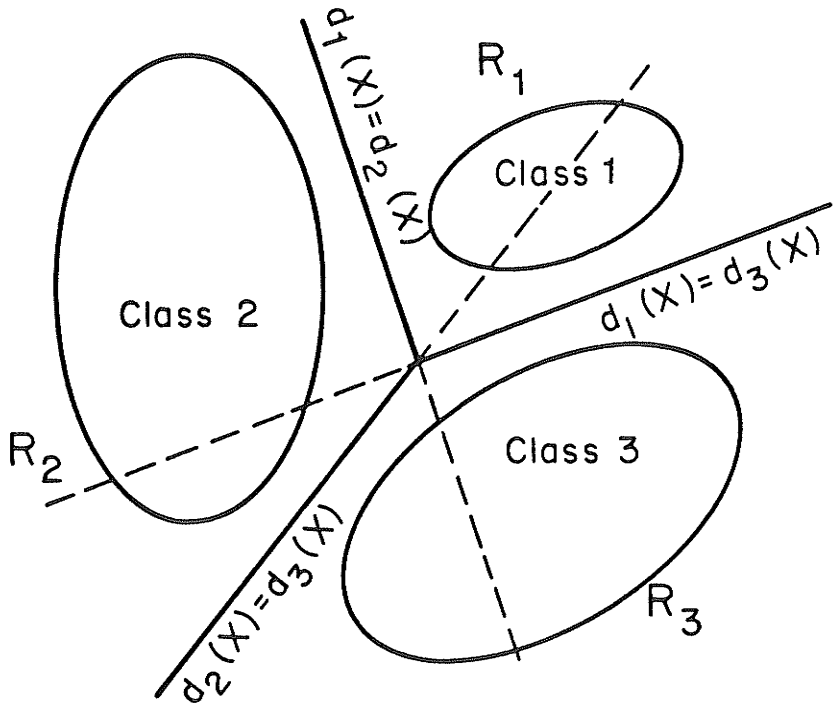


Fig. 2.5 Three linear predictor functions approach.

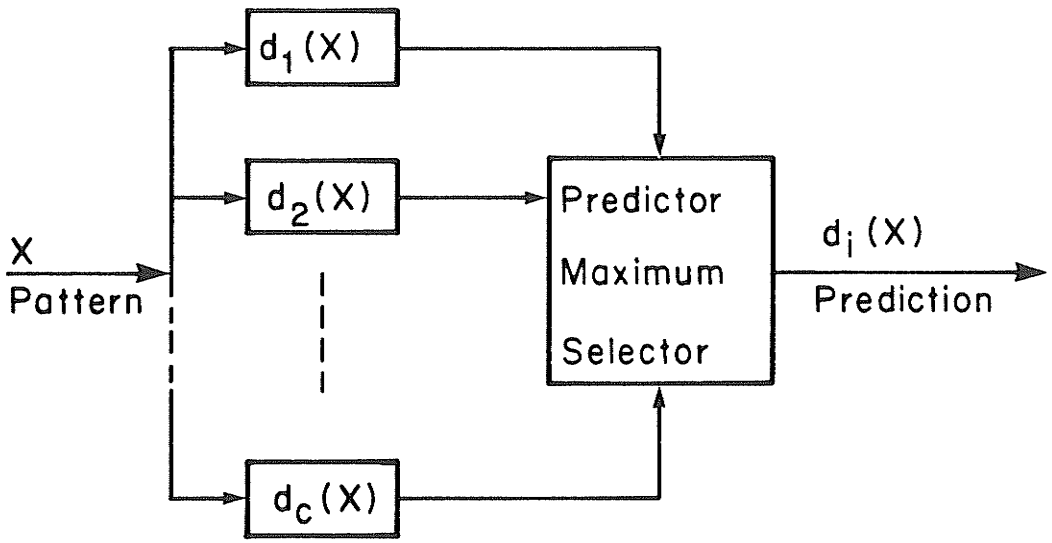


Fig. 2.6 Multiclass predictor scheme.

nonparametric estimation of the probability density function for each class to be classified.

(a) Linear Learning Algorithm

This algorithm is advantageous since it is based on nonparametric estimation of the probability density function $f\ (X/i)$, in addition to the application of *Bayes* decision rule for classification. One way to estimate the class density function taking into consideration the contribution of each pattern X_j is given by [50]:

$$f\ (X/i) = \frac{1}{[\ N_i\ (2\pi)^{f/2}\ \sigma^f\ ]} \cdot \sum_{j=1}^{N_i} \exp[-(X - X_{ij})^T\ (X - X_{ij})/\ 2\sigma^2] \qquad (2.15)$$

where :

X_{ij} : pattern j in class i ,

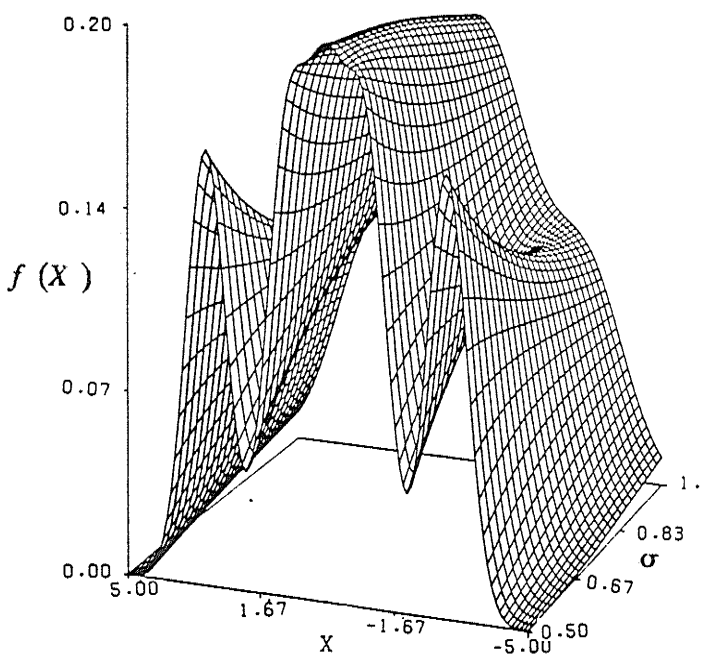
N_i : number of patterns in class i ,

σ : smoothing factor.

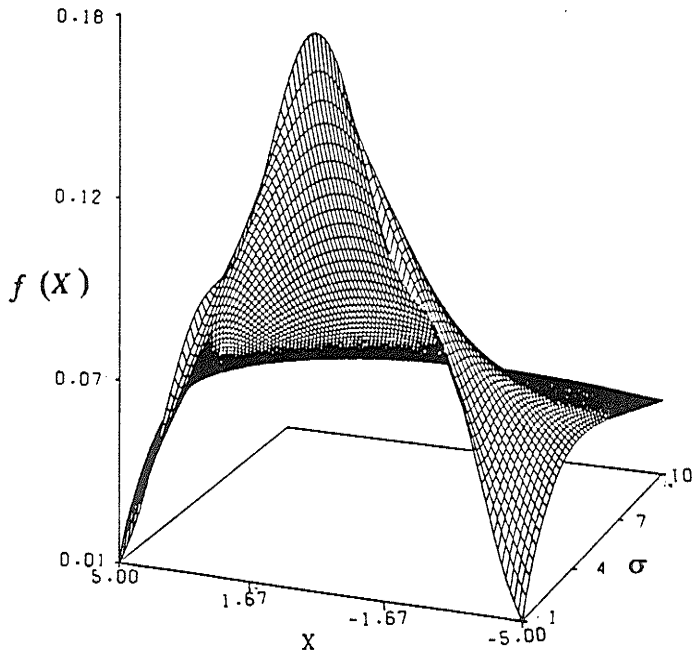
The parameter σ dictates the smoothness of the estimate of the density function. When σ is small the estimated density approximates the true distribution of patterns in the design set more closely. As σ is increased the estimated density is smoothed, Fig. 2.7, and can therefore be represented by lower order terms. Hence, for a given number of training patterns, σ should be at least large enough to provide smoothing between adjacent patterns. It may even be increased above this minimum to limit the order of function representing the decision boundary.

Equation (2.15) can be used directly with the Bayes decision rule. However, Specht [50] proposed a simplified approximation to this equation using Taylor's expansion series. As explained in Appendix (C), the final form of the weighting vector W_i, for the case of design patterns X_i from class i of *prior probability* p_i, can be estimated from design patterns during the design stage as:

$$w_{ik} = \frac{1}{\sigma^2} \cdot [\frac{p_i}{N_i} \sum_{j=1}^{N_i} x_{ijk} \cdot B_{ij}] \qquad (2.16)$$



(a)



(b)

Fig. 2.7 Effect of the smoothing factor σ on the density function.

(a) small values of σ ($0.5 \leq \sigma \leq 1.0$)

(b) large values of σ ($1.0 \leq \sigma \leq 10.0$)

where: $k = 1, 2, ..., f$, and

$$w_{i0} = [\frac{p_i}{N_i} \cdot \sum_{j=1}^{N_i} B_{ij}] \tag{2.17}$$

$$i = 1, 2, \cdots, c.$$

where:

$X_{ij} = [x_{ij1}, x_{ij2}, ..., x_{ijf}]^T$ j-th feature vector from class i.

$W_i = [w_{i1}, w_{i2}, ..., w_{if}]^T$ weighting vector for class i.

$$B_{ij} = \exp(\frac{-0.5\, X_{ij}^T X_{ij}}{\sigma^2})$$

As an example, consider a two-class prediction system, where the predictor function is given as:

$$d(X) = d_1(X) - d_2(X)$$

$$= W^T X + w_0 \tag{2.18}$$

where $W^T = (W_1^T - W_2^T)$, $w_0 = (w_{10} - w_{20})$. With some mathematical manipulation we can have the final form of the weighting vectors as :

$$w_k = \frac{1}{N_1}[\sum_{j=1}^{N_1} x_{1jk} \cdot exp(\frac{-X_{1j}^T X_{1j}}{2\sigma^2})]$$

$$- \frac{PR}{N_2}[\sum_{j=1}^{N_2} x_{2jk} \cdot exp(\frac{-X_{2j}^T X_{2j}}{2\sigma^2})] \tag{2.19}$$

$$k = 1, 2, ..., f .$$

$$w_0 = \frac{1}{N_1}[\sum_{j=1}^{N_1} exp(\frac{-X_{1j}^T X_{1j}}{2\sigma^2})]$$

$$- \frac{PR}{N_2}[\sum_{j=1}^{N_2} exp(\frac{-X_{2j}^T X_{2j}}{2\sigma^2})] \tag{2.20}$$

where PR is the *prior probability ratio*.

(b) Stepwise Learning Algorithm

Having chosen the stepwise pattern features, this stage is to build a surface in the feature space which separates patterns of one class from those of the other. Therefore, this design is called the *stepwise predictor*. The stepwise development of the predictor functions is dependent almost exclusively on the current values of the status matrices. These matrices are equal to the within and total sums of cross-products. They are updated from one step to another be means of the sweeping operation [48,49]. When the stepping is complete, or when the number of features selected is equal to the one specified, the predictor function weighting vectors for class i are computed as:

$$W_i = (N - c)\, W(X)^{-1} M_i \tag{2.21}$$

and the corresponding predictor function thresholds are given by:

$$w_{i0} = \ln p_i - 0.5\,(N - c)\, M_i^T\, W(X)^{-1}\, M_i \tag{2.22}$$

$$i = 1, 2, \ldots, c\,.$$

where p_i is the prior probability for class i, and M_i is the features mean vector for class i.

As an example, consider a two-class prediction system, where the predictor function is given by Eq. (2.18) and substituting with W_1, W_2, w_{10}, and w_{20} from Eqns. (2.21) and (2.22), the weighting vector can be obtained as:

$$W = (N - 2)\, W(X)^{-1}\, (M_1 - M_{2)} \tag{2.23}$$

and the predictor constant is given by:

$$w_0 = \ln(\, 1/PR\,) - 0.5(N - 2)[\, M_1^T\, W(X)^{-1}\, M_1 - M_2^T\, W(X)^{-1}\, M_2\,] \tag{2.24}$$

2.7 Prediction System Performance Evaluation

As far as a performance evaluation methods are concerned, there are two approaches; on-line and off-line. For the *on −line* method, the prediction system is installed on the actual system, and its parameters are adjusted in accordance with experience. This method is probably only practical for "fine tuning" of the prediction system, Fig. 2.8.

The *off −line* approach uses an accurate computer model to generate test patterns and estimate the probability of misprediction for each test pattern.

In this study we will deal with the second approach and incorporate error estimation methods. The performance evaluation given by these methods could be written as:

$$Performance = 100.\ (1.0 - 1/N_t \sum_{j=1}^{N_t} [\ 1 - \delta_j\ ]\ ) \qquad (2.25)$$

where : δ_j is the *Kronecker* delta function, i.e. $\delta_j = 1$ in case of correct prediction, $\delta_j = 0$ in wrong prediction cases, and N_t is the number of patterns in the test set.

Generally speaking, the last step in a pattern prediction system design is to evaluate the future classification performance of this system. Several estimates exist in the literatures [40-46]. However, only some of them are used here, Appendix (D), as : 1- Resubstitution Estimate, 2- Hold-out Estimate, 3- Leave-one-out Estimate, and 4- Rotation Estimate.

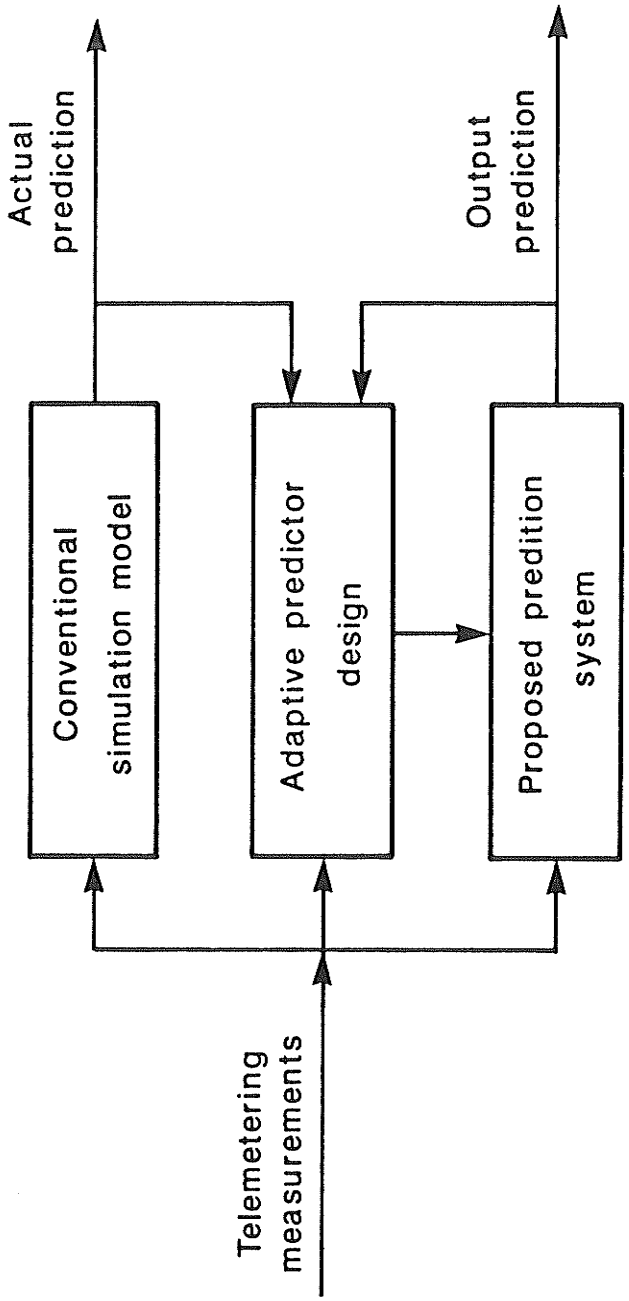


Fig. 2.8 On-line prediction system performance evaluation.

CHAPTER 3

LOAD REJECTION OVERVOLTAGES ON NORTHERN SYSTEM

3.1 Introduction

In power system applications involving synchronous generators connected to long distance ac transmission systems or connected to dc systems with large filter banks close to the machines, serious overvoltage conditions can arise following major system *contingencies*. Examples of such contingencies are ac load rejection or blocking of converter stations.

These overvoltages have been reported before [4,6] using analog and digital simulations. They are a function of the ac transmission system charging capacitance or the amount of filter capacitance for ac/dc systems, the number of machines connected, the over-speed characteristic of the turbine-generator set, and the excitation system response. The problem of load rejection overvoltages on hydro-generator systems can be avoided by eliminating or at least minimizing the possibility of generator *self −excitation*, and controlling system *dynamic overvoltages* to acceptable levels.

The generating stations in power systems could be either steam-turbine or hydro-turbine units. Historically, hydro-turbine and steam-turbine units have been faced with overvoltages and generator self-excitation when long unloaded EHV transmission systems are left connected to the generators [4]. Generator self-excitation has been a common concern on hydro systems due to the high over-speed condition on load rejections [4]. In practice, for steam-turbine generator units load rejection can be tolerated with little or no significant damage to the unit, and the unit can continue to run on a self-supporting basis, for periods of up to several hours, without incurring major operating difficulties [51].

Load rejection overvoltages are considered in this chapter for the Manitoba Hydro Northern system. Existing preventive and protective schemes are outlined.

3.2 Load Rejection Overvoltage Conditions

3.2.1 Introduction

As shown in Fig. 3.1, the Manitoba Hydro Northern system on the Nelson River consists of two interconnected hydro generating stations (Kettle and Long Spruce) feeding power to Southern Manitoba through two dc bipoles: bipole 1 (BP1) and bipole 2 (BP2). AC filter banks are connected at Radisson and Henday converter stations and are of a 229.5 & 500 MVARs capacity respectively. These filters are to filter out the harmonic currents of the 5-th, 7-th, 11-th, 13-th, and higher order.

The system under study, simplified in Fig. 3.2, is only a part of the Northern system in order to identify the problem very well and to avoid excessive computations at this stage. This system consists of Kettle generating station connected to Radisson through K-lines (lines K1 to K7 in Fig. 3.1). At Radisson there are ac filter banks connected to the ac side of bipole 1 converter stations, to filter out the generated current harmonics. In such a system dc load rejections caused by the blocking of Radisson converter stations, can lead to situations where certain combinations of over-speeding machines are left connected to filter banks. These conditions can give rise to serious overvoltages, especially when the filter VARs are excessive relative to the number of generating units remaining connected.

The digital *simulation model* used for this study has been developed by Manitoba Hydro [52]. The hydraulic turbine generator unit is represented by: the governor-turbine system, the synchronous machine equations, and the excitation system. Appendix (E), shows a block diagram for a salient pole machine; a representation of static exciter without negative field current capability and voltage regulator; a block diagram of the hydro turbine governor; and a detailed nonlinear hydraulic system simulation incorporated with the governor.

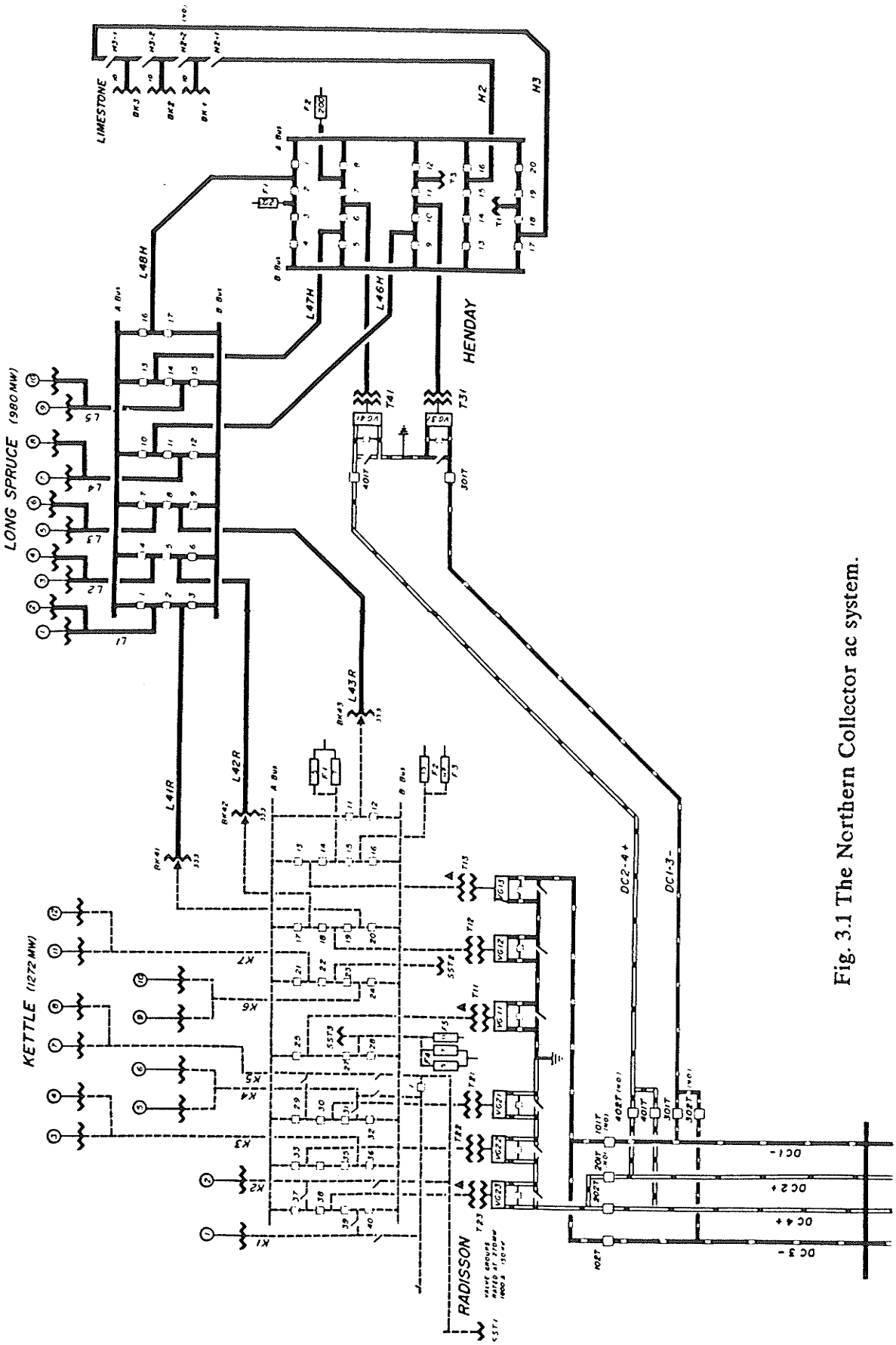


Fig. 3.1 The Northern Collector ac system.

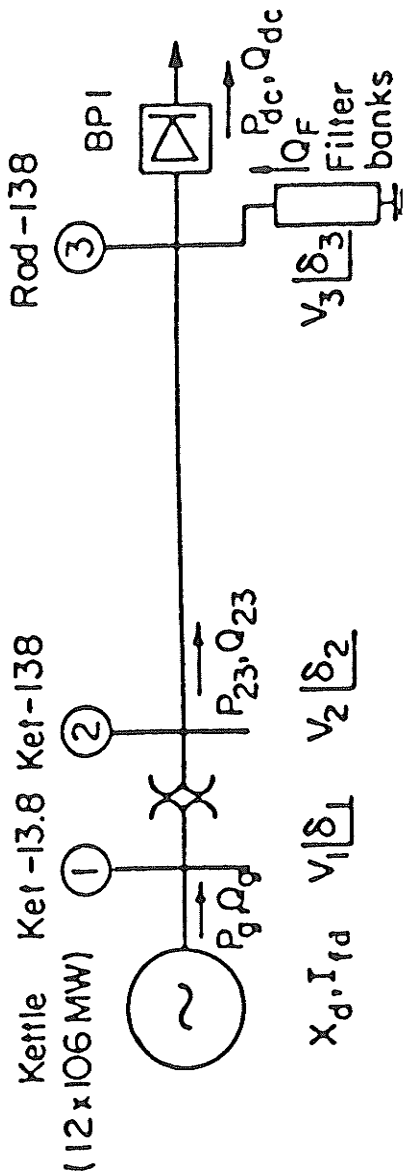


Fig. 3.2 The power system under study.

3.2.2 No Self-Excitation Condition

For the Northern system simplified in Fig. 3.2, hydro generating plants as Kettle are feeding a dc system known as bipole 1 (BP1).

Figure 3.3 shows a typical pattern of *no self −excitation* results. In this pattern, 6 units at Kettle, 229.5 MVARs filters load at Radisson, and 600 MW of 0.86 power factor lagging dc load on BP1, are considered. These results show that load rejection overvoltages are all characterized by an instantaneous rise at the instant of rejection (0.1 sec), then followed by a temporary voltage limiting by regulator control and ending with a more gradual rise. Also, it can be seen that the generator dynamic voltage rise is about 20%, while the steady state rise is about 1%. On the other hand, the corresponding Radisson overvoltages are about 28% and 10% respectively. An increase in the generator speed with about 25 Hz can be observed.

3.2.3 Dynamic Overvoltage Condition

In some situations [53], involving a few machines connected to a particular filter configuration, full load rejection does not indicate imminent self-excitation, but nevertheless, can lead to poor *voltage regulation* in the dynamic and/or steady state periods following the load rejection. For example, it is possible for the bus voltage to rise to approximately 1.4 p.u. during the dynamic period immediately following the load rejection, and to settle several seconds later at more than 1.2 p.u. Either of these could be considered abnormally high, even though self-excitation has not occurred.

3.2.4 Self-Excitation Condition

For the system under study, Fig. 3.2, *hydro −generating* plants are feeding a *dc system* (BP1). Self-excitation usually occurs during an *over −f requency* condition (up to 80 Hz). It occurs following *resonance* between the machine direct axis reactance X_d and the capacitive reactance seen at the machine terminals X_c ,

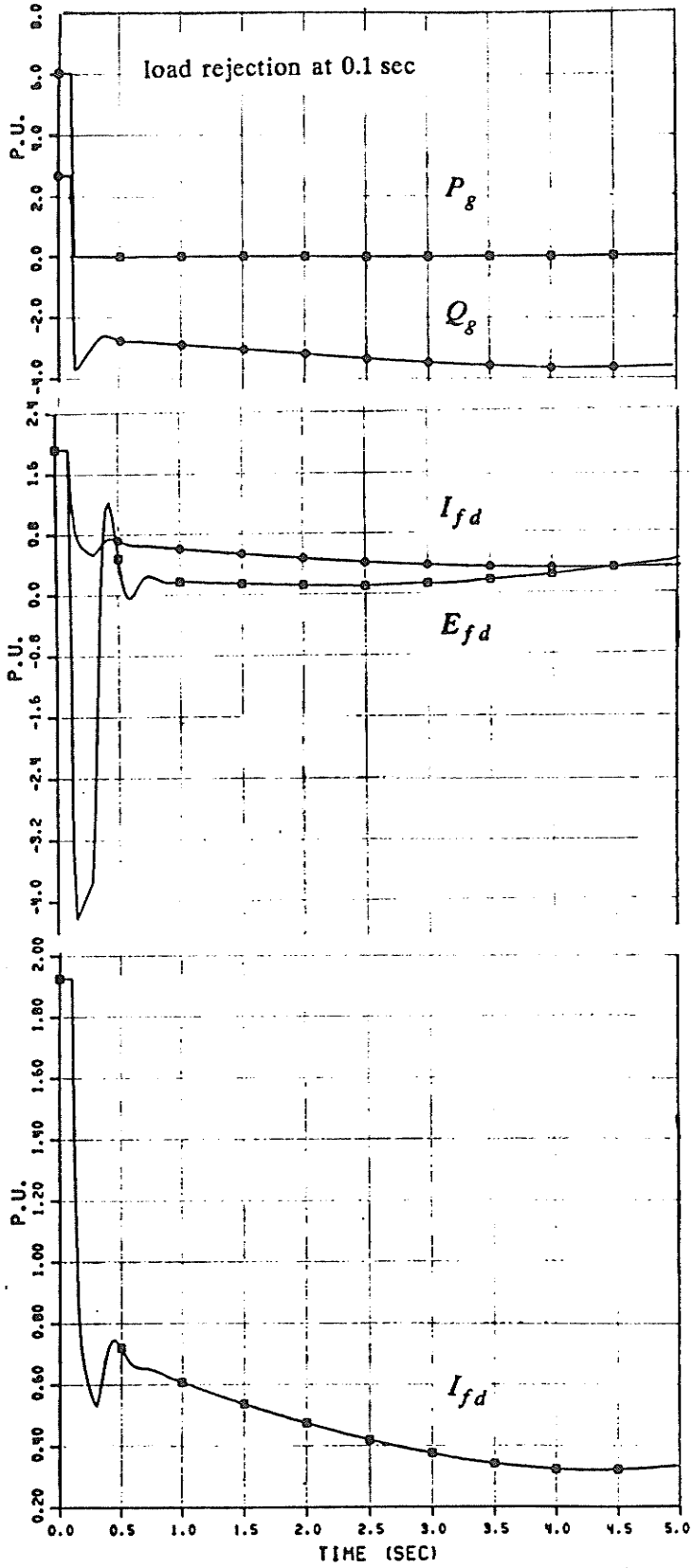


Fig. 3.3 DC load rejection results for no self-excitation condition.

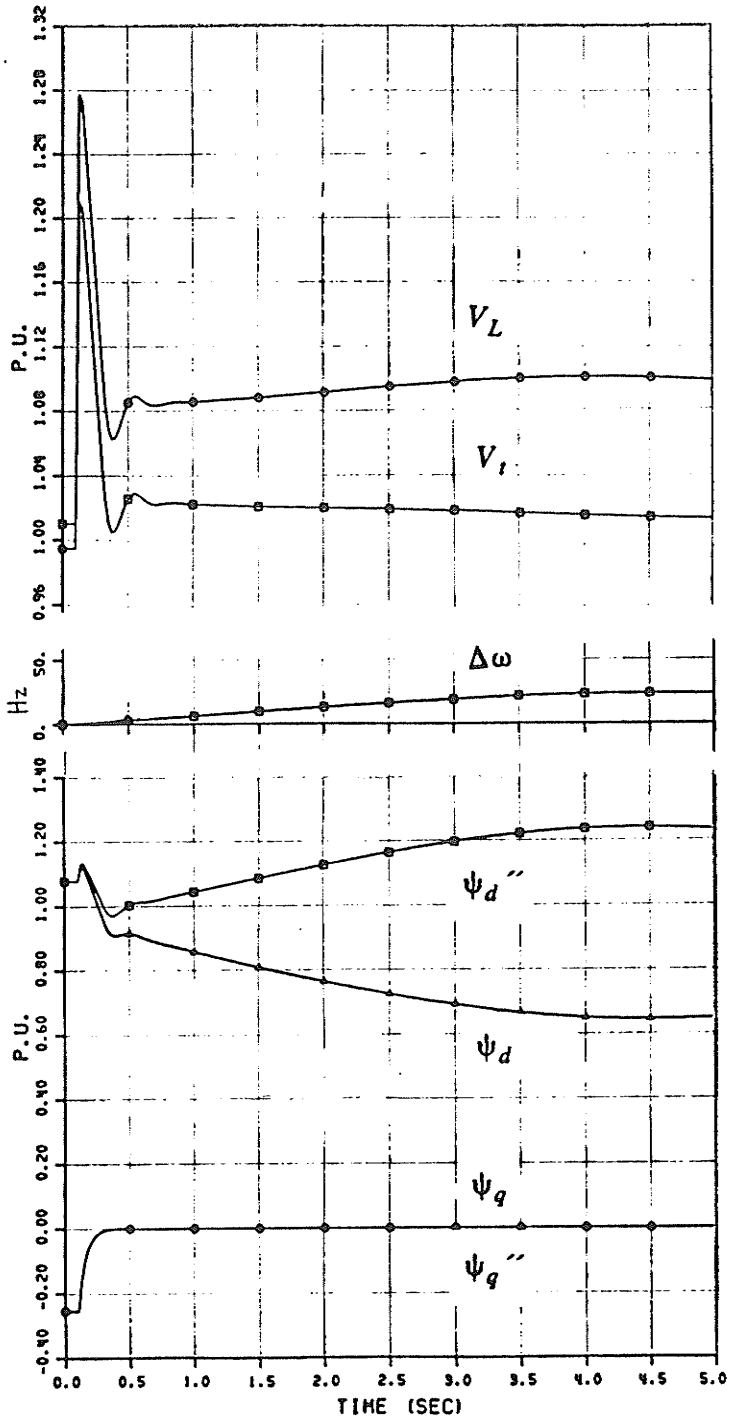


Fig. 3.3 (continued).

particularly for machines equipped with static excitation systems without negative field current capability. The effects of having or not having negative field current capability in the excitation system have been treated in some detail [6].

A pattern of results for self-excitation conditions is illustrated in Fig. 3.4. This pattern is for the same configuration given in the case of no self-excitation except that the number of Kettle units is 4 in this case. The converter station, blocked at 0.1 second, produces the field current and voltage responses given in Fig. 3.4. The generator field current becomes negative at approximately 2.5 seconds, at the same time the direct axis flux and voltages start to reverse directions.

During the condition of self-excitation, the stator current due to the capacitive load tends to excite the machines. When this happens, the exciter and field circuit have no control over the machine terminal voltage and it can rise dramatically, with a rate of about 1 - 2 p.u./sec, to an extremely high level, Fig. 3.4. At the same time, the field current is forced rapidly to zero causing very high *voltage stresses* across the field, which can damage the excitation system as explained by Gole [7]. Damage to an exciter of the Kettle generating plant has been reported.

In summary, the problem of load rejection overvoltages can be considered as a problem of machine/filter interaction following load rejection.

3.2.5 Effect of Generator Units Trip

The design of Kettle and Long Spruce involved terminating two units per high voltage breaker position. This is a concern at Kettle where the existing K-line protection is set to trip at 172 MVA into the generating station. With two units of Kettle per K-line, the possibility of losing the line following a dc load rejection exists. The machine VAR intake during the dynamic period following a rejection is important due to the loss of field and loss of excitation relays at Kettle. Comparing the results of Fig. 3.3 with those of Fig. 3.4, the effect of a two units trip at Kettle is clear.

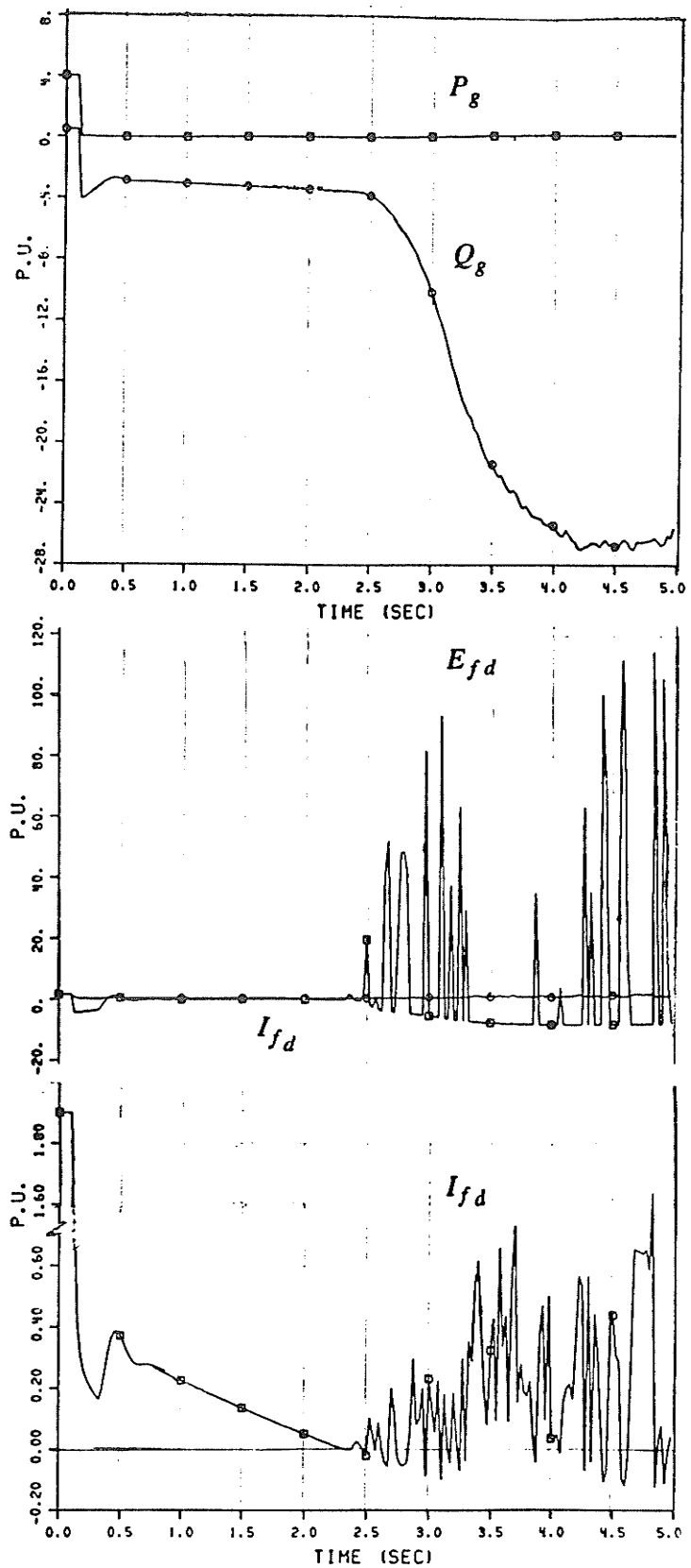


Fig. 3.4 DC load rejection results for self-excitation condition.

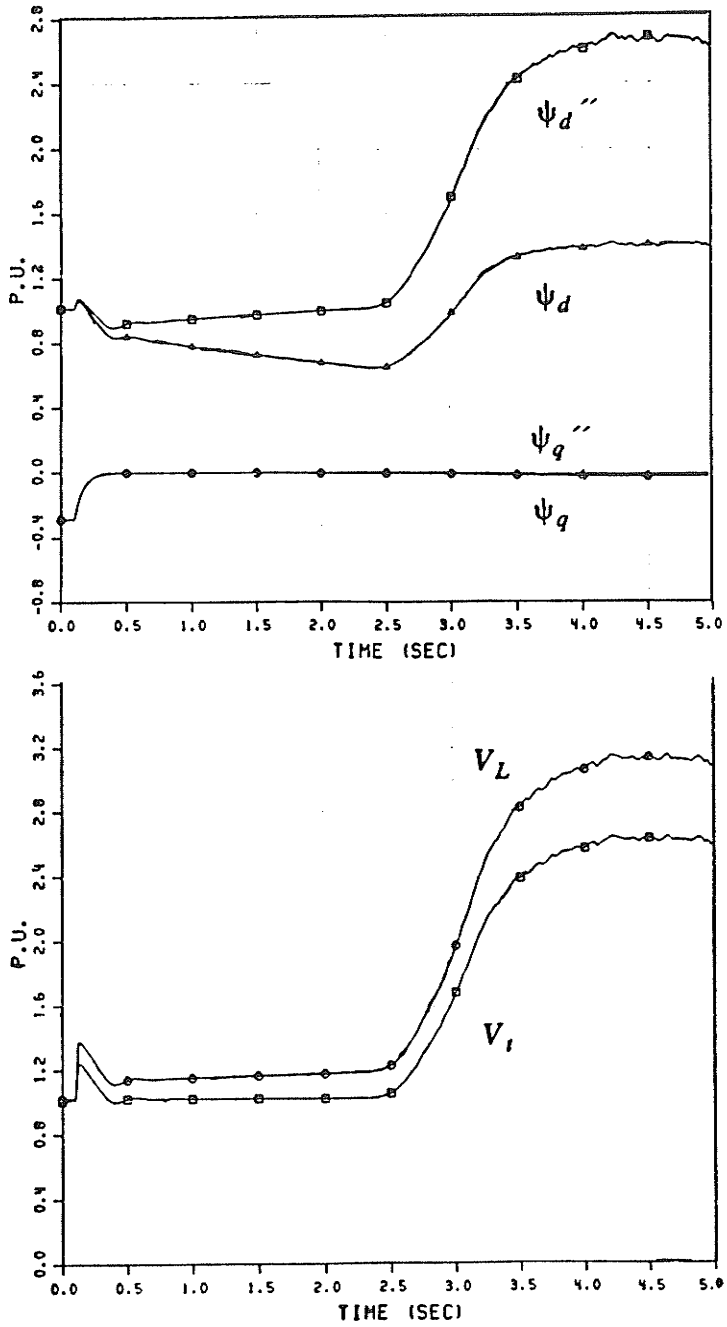


Fig. 3.4 (continued).

3.2.6 Effect of Filter Bank Trips

The effect of filter tripping can be explained by Fig. 3.5. It shows a comparison between the results of two patterns: the first pattern is self-excitation condition, and the second is a no self-excitation pattern. The self-excitation pattern in this case includes 12 units at Kettle, 600 MVARs of filter banks, and 1200 MW of 0.86 power factor BP1 load. On the other hand, a tripping of 20 MVARs of filters at 1.0 second and for the same system configuration as that for the first pattern, produces no self-excitation condition for the second pattern. From these results, it is clear to see the effect of filter tripping on the rate of rise of the generator and load terminal voltages i.e. the *filter tripping* affects the load rejection overvoltage conditions in the *corrective* action direction.

3.3 Existing Preventive and Protective Measures

DeMello [6] concluded that the excitation system design can be an important means for solving system overvoltage problems. However, Dandeno [5] mentioned that for a hydro system it may be very costly to design an excitation system such that self-excitation cannot occur at *overspeeds*. A protective method based on the *monitoring* of machine field current has been proposed. This proposal can solve major problems but for some cases it can indicate a problem when the overvoltage is not serious.

Another protective measure [6] suggests that overvoltage relays should drop reactive sources (filters) when safe voltage levels are exceeded. It is difficult to avoid *overstressing* the breakers because the voltage can rise very fast.

Manitoba Hydro has implemented a *look −up* table method [53]. The strategy is prevention via configuration operating restrictions. It is based on maintaining the machine/filter combinations within the operating restrictions and therefore involves a very high number of computation runs. For example, Table 3.1 presents the relation between the minimum number of machines corresponding to different filter

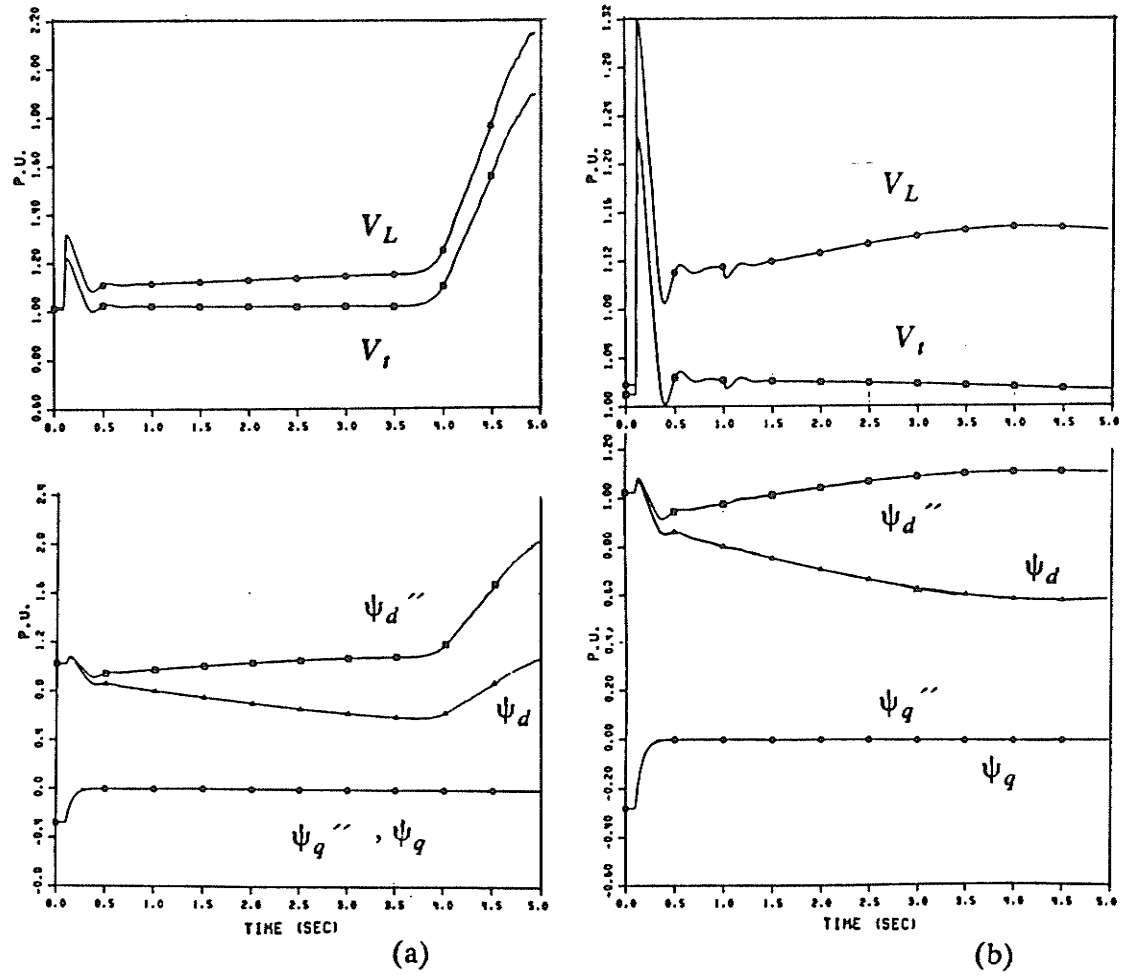


Fig. 3.5 DC load rejection results.

(a) self-excitation condition

(b) no self-excitation due to tripping of the excess filter banks

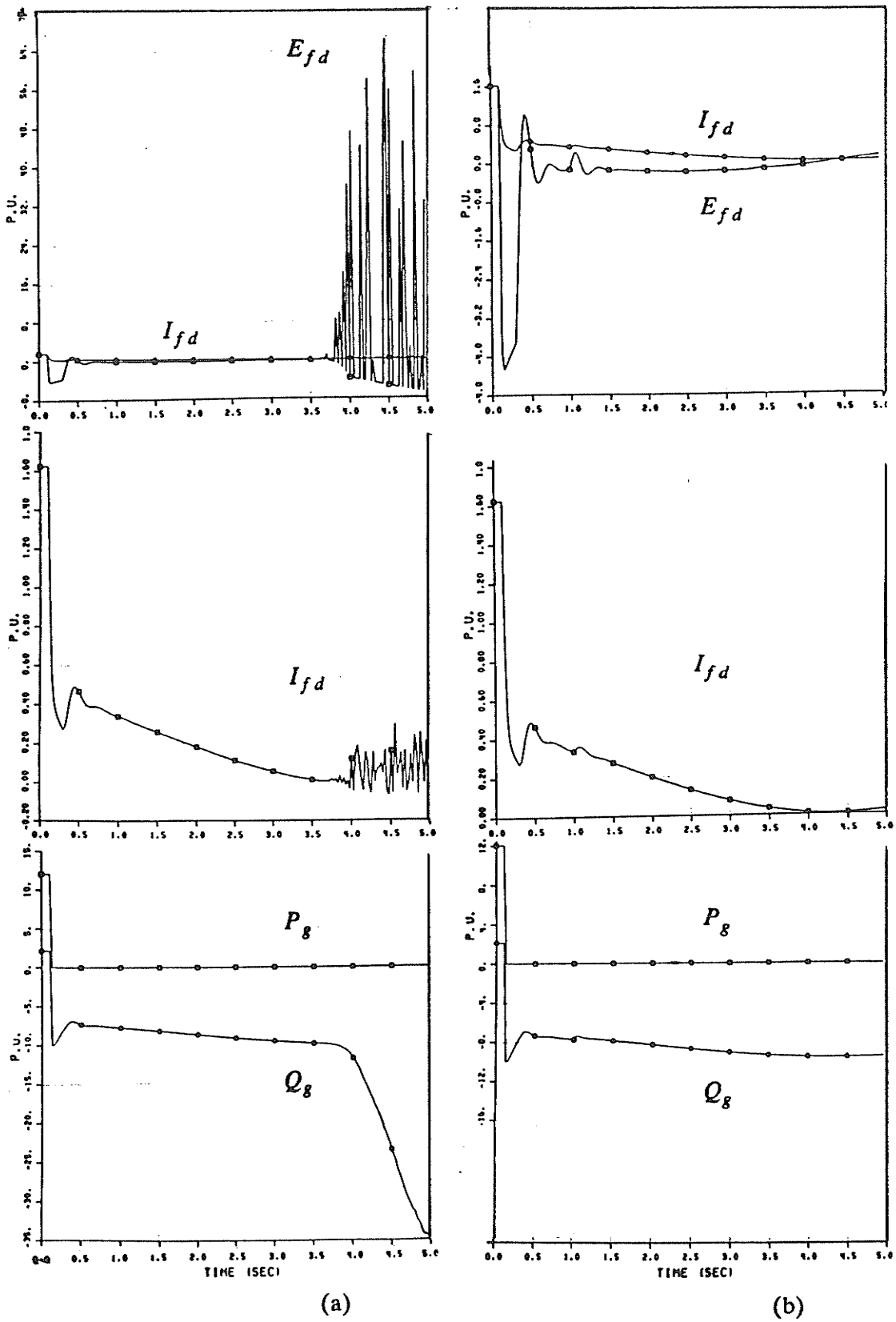
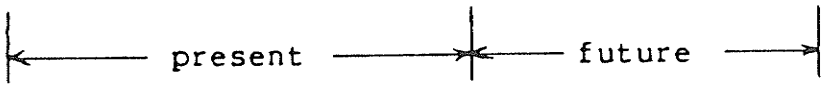


Fig. 3.5 (continued).

Table 3.1 The minimum number of machines allowed for no self-excitation

Filter (MVARs)	79	100	150	229.5	300	500	580
No. of machines	2	2	4	5	7	11	12



configurations, required to prevent self-excitation at rated loading conditions. The "boundary" is not that simple for the wide variety of possible operating conditions: hence a procedure such as the one described in this thesis is highly recommended.

Another approach is to use a complete digital computer system model to assess the current operating configuration. In this scheme the evaluation is based on transient stability and load flow simulations which takes about 50.6 seconds cpu time for one case on the Prime computers. Since more than one "case" needs to be run, i.e. more than one anticipated fault condition for a particular operating condition, this method is too slow and too expensive.

3.4 Proposed Approach

Any protection scheme for load rejection overvoltages (system dynamic overvoltages and/or generator self-excitation) must be capable of offering protection for the following operating situations [53]:

(1) Provide protection in the event of loss of machines. The loss of two units is the most probable situation.

(2) Provide protection for the case of a complete bipole block or a large dc reduction.

On the other hand, the successful protection scheme should have the following qualifications:

(1) It should reduce or eliminate the large number of machine/filter operating restrictions.

(2) It should only take action as required and only affect the generating station/filters in danger.

(3) It should be easily extended as additional generation and dc facilities are added.

(4) It should not interfere with existing machine protection.

Looking for a protection scheme to operate in the above situations and possess all required qualifications, all protection schemes mentioned in the above section have drawbacks. The approach suggested here is to use a *prediction* system, as recommended by de Mello in [6], that will assess the current machine/filter configuration using a "fast calculation" based on a few key measured variables. The design is based on the use of *pattern recognition techniques*. It will help the power system operator to detect after receiving system informations, within 10 msec prediction time on AMDAHL computer, conditions that may lead to the aforementioned abnormal conditions before they occur.

Pattern recognition has a wide variety of applications [2,8,9,54-59]; however the application discussed here is *unique* in both the problem that being addressed and the manner in which the technique is applied.

3.5 Conclusions

The Manitoba Hydro Northern system has experienced generator self-excitation as well as system dynamic overvoltages, due to converter station blocking (dc load rejection). These conditions are found to be dependent on the system configuration i.e. machine/filter combinations. It was also explained that the lesser the number of machines, the more severe the condition will be.

Investigating the current protection schemes, it was concluded that they are unreliable, and therefore the suggesting prediction scheme was recommended. The reduction of filter MVARs was proved to be a suitable corrective action.

CHAPTER 4

SELF-EXCITATION: PREDICTION SYSTEM DESIGN

4.1 Introduction

As discussed before, power system load rejection may result in generator self-excitation and/or system dynamic overvoltage depending on the machine/filter combinations. Therefore, in the field implementation there are two prediction system schemes: the *self −excitation prediction* system and the *dynamic overvoltage prediction* system. Each of these systems is considered as a two-class prediction problem e.g. the normal class and the abnormal class.

This chapter concerns the self-excitation prediction system design. The main purpose of this system is to assess the current machine/filter configuration using a fast calculation method based on the available knowledge contained in design patterns.

The application of pattern recognition techniques, described in chapter 2, is illustrated here to design the required prediction system for the simplified Manitoba Hydro Northern system. A detailed description of the self-excitation prediction system, its performance evaluation, and the effect of failure in communication channels are provided. Also, the idea behind the development of an algorithm to iteratively simplify the prediction system structure is presented. Finally, a *corrective measure* is proposed in order to prevent self-excitation.

By using pattern recognition techniques, including the establishment of the appropriate features and *predictor* design, a scheme is suggested to aid power system operators in anticipating self-excitation operating problems and to evaluate the required action in a very short time compared with other methods.

4.2 Creation of Characteristic Patterns

This stage is to create a set of patterns called *pattern set*. The pattern set should be representative of the whole range of power system operating conditions. Otherwise, the prediction system obtained may not be able to correctly predict future patterns.

The characteristic pattern set for a particular class i in the variable space may be defined as:

$$Y_i = \begin{bmatrix} y_{i11} & y_{i12} & \cdots & y_{i1N_i} \\ y_{i21} & y_{i22} & \cdots & y_{i2N_i} \\ \cdots & \cdots & \cdots & \cdots \\ y_{iv1} & y_{iv2} & \cdots & y_{ivN_i} \end{bmatrix} \tag{4.1}$$

which could be rewritten in the following format:

$$Y_i = (y_{ijk} \;/\; j = 1, 2, ..., v\,;\, k = 1, 2, ..., N_i) \tag{4.2}$$

where v is the number of *pattern variables* and N_i is the number of patterns in class i $(i = 1, 2)$.

In order to avoid the large off-line computational burden involved to create the pattern set, a simplified version of the Manitoba Hydro Northern system is studied (see Fig. 3.2 in chapter 3). Additionally, a number of self-excitation relevant system configurations are considered, in order to test the proposed prediction system without using excessive computation at this stage.

For an isolated hydro generating plant as Kettle feeding a dc load as bipole 1, self-excitation prediction *assessment* involves new kinds of *contingencies*. As mentioned in the previous chapters, the prediction system should be capable of providing self-excitation protection for the most probable contingencies. For this reason, the complete bipole block and the loss of one or two Kettle units following the dc load rejection are considered as the system operating contingencies of major interest.

Any power system pattern can be fully described by a set of variables. The number of variables comprising a pattern for self-excitation studies will depend on the detail to which the system is modelled, as well as the size of the system. Nevertheless, a great number of variables is always needed and the problem of high dimensionality can be avoided using feature identification techniques.

During the pattern set creation, it was found that the following variables are the most *informative* for the prediction process of self-excitation patterns:

P_{dc}, Q_{dc} = active and reactive load into a converter station

Q_F = reactive generation of filter banks at a bus

X_d = machine direct axis synchronous reactance at a bus

I_{fd} = machine field current at a bus

V, δ = voltage magnitude and angle at a bus respectively

P_g, Q_g = active and reactive power generation at a bus respectively

P_l, Q_l = active and reactive line flows respectively

Using the Manitoba Hydro self-excitation digital simulation, discussed in some detail in chapter 3, a load rejection pattern set of 124 patterns of 12 variables each is created. These patterns are then classified into no self-excitation (90 patterns) and self-excitation (34 patterns). Out of these 90 patterns, a *feasible* pattern set of 88 patterns is constituted for the case of *double contingency* i.e. the loss of one or two units following load rejection. The last pattern set is again classified into 51 no self-excitation and 37 self-excitation patterns.

4.3 Identification of Required Features

Pattern sets having been created, this stage is to identify the required features. By applying the *principal component* method or *KL−transformation*, Appendix (A), to these pattern sets in the variable space, the required number of features sufficient for a prediction process can be estimated as shown in Table 4.1.

Table 4.1 Principal components results

Feature	LR-predictor		DC-predictor	
no.	e.v.	sum %	e.v.	sum %
1	401.31	94.66	91.81	88.5
2	17.46	98.78	7.47	95.7
3	4.72	99.89	3.99	99.60
4	0.46	99.99	0.42	99.99

LR: load rejection contingency

DC: double contingency

e.v.: eigen-value

Then applying feature identification techniques, discussed in chapter 2, to a pattern set given by Eq. (4.1) will produce the corresponding *feature space* described as:

$$X_i = \begin{bmatrix} x_{i11} & x_{i12} & \cdots & x_{i1N_i} \\ x_{i21} & x_{i22} & \cdots & x_{i2N_i} \\ \cdots & \cdots & \cdots & \cdots \\ x_{if1} & x_{if2} & \cdots & x_{ifN_i} \end{bmatrix} \tag{4.3}$$

rewritten in the form:

$$X_i = (x_{ijk} \ / \ j = 1, 2, ..., f \ ; k = 1, 2, ..., N_i) \tag{4.4}$$

where f is the number of required features i.e. $f = 4$.

Table 4.2 presents the features identified using the variable separability measure (VSM) and intuitive techniques for the load rejection contingency as well as the double contingency.

The *stepwise* feature identification algorithm is then applied to the available pattern sets and the results obtained are provided in Table 4.3. These results show that at each step there is one feature to be added to or removed from the current features until the stepping is complete. The selection and removing threshold used is 0.001.

4.4 Predictor Design and Performance Evaluation

Having identified the required features, we can design the predictor and evaluate its performance in predicting self-excitation conditions. The design process is really based on the *recognition* knowledge contained in the features and on the distribution of the pattern set in the feature space. For this design the following two schemes are employed: *Linear prediction* and *Stepwise prediction* .

Table 4.2 Features identification results

Feature no.	LR-predictor		DC-predictor		Intuitive features
	feature	VSM	feature	VSM	
1	V_3	0.9501	Q_{23}	1.3561	X_d
2	Q_{23}	0.9187	V_3	1.1913	Q_F
3	P_g	0.3908	X_d	0.5187	Q_g
4	X_d	0.3807	Q_F	0.4513	I_{fd}

Table 4.3 Stepwise features identification results

Step no.	LR-predictor		DC-predictor	
	feature	SM	feature	SM
1	V_3	89.0420	Q_{23}	115.7065
2	I_{fd}	11.9342	Q_g	22.8578
3	δ_3	3.2814	X_d	23.2773
4	Q_{23}	.9693	V_3	4.1879

4.4.1 Linear prediction scheme

In this approach the Resubstitution and Hold-out design methods are involved along with the linear *learning* algorithm (section 2.6.2). This scheme deals with the features identified using the variable *separability* measure and the intuitive algorithms (section 2.5), see Fig. 4.1.

(a) The Resubstitution design

In this method the predictor scheme is designed using the pattern set. Figure 4.2 shows the design procedure for this case. This design is a function of two parameters, σ and PR, the *smoothing factor* σ dictates the smoothness of the estimated density function, while the *prior probability ratio* PR adjusts the design so as to have a desired accuracy in the prediction of patterns from a particular class.

The design results, shown in Fig. 4.3, explain the selection process of these two parameters. The selection criterion involved here is based on the effectiveness of the overall prediction accuracy. These results declare that both designs start with accuracy around 30 % and builds up till it saturates to around 80 % using VSM features and 90 % using the intuitive features.

The design parameter PR starts flat and settles for few times then builds up till it saturates at 1.45 for VSM and 1.15 for intuitive features. The appropriate designs are selected at $\sigma = 11$ and $\sigma = 6$ using VSM and intuitive features respectively. The results obtained in Fig. 4.3 prove that the intuitive features are more informative than the VSM features for the case of the load rejection predictor scheme.

The selected design of the predictor hyperplane d for the case of *load rejection* and using VSM features is obtained as:

$$d = W^T X + w_0$$

where

$$W = [\, -0.4543,\ 1.3928,\ -0.0721,\ -0.2614 \,]^T \ , \ w_0 = -0.4183$$

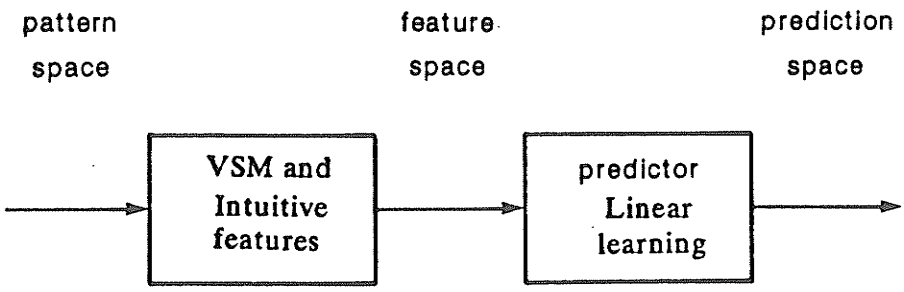


Fig. 4.1 Linear prediction scheme.

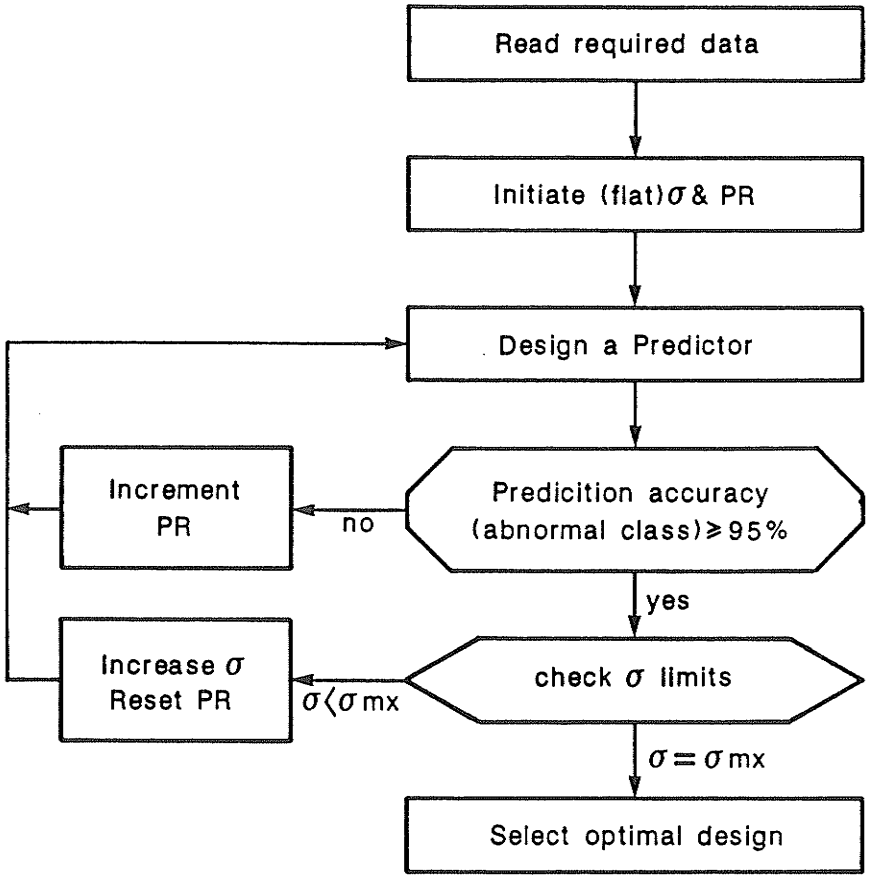
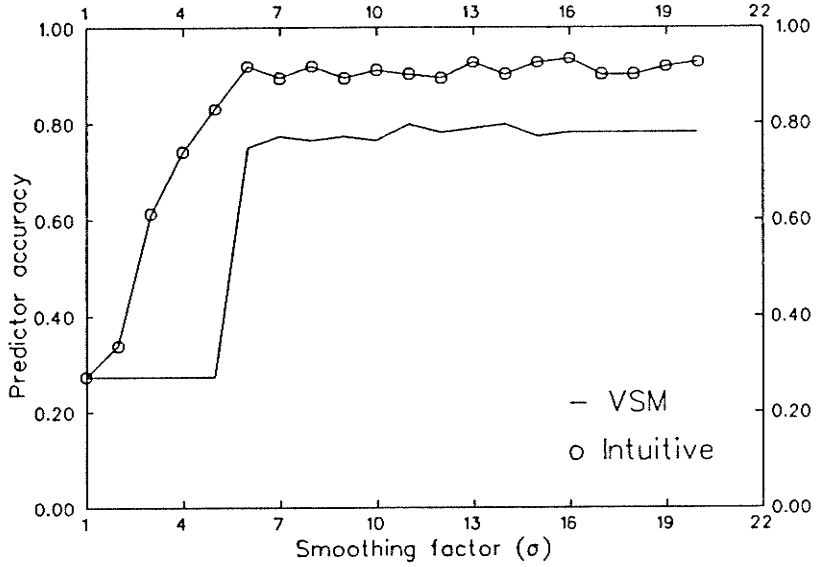
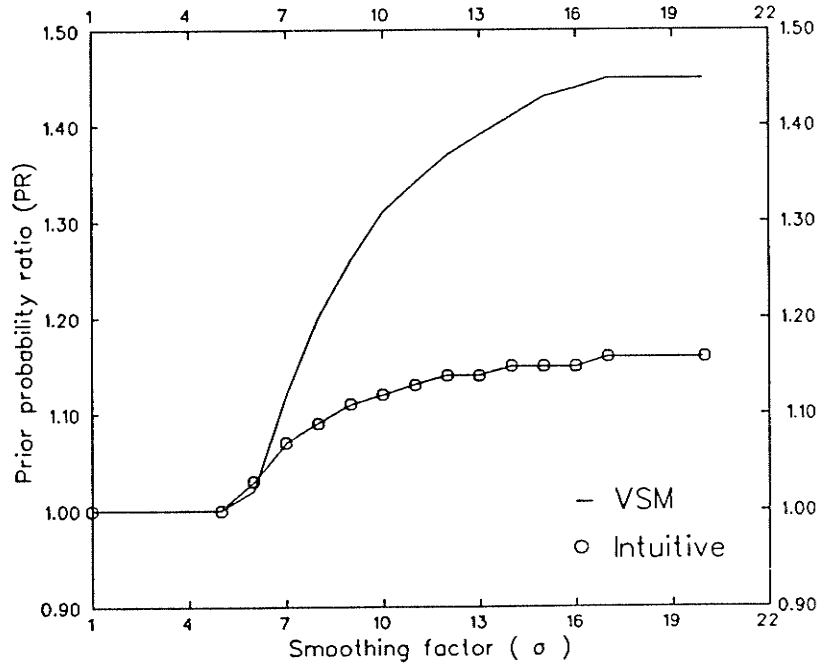


Fig. 4.2 Predictor design algorithm.



(a) predictor accuracy



(b) optimum pairs of σ & PR

Fig. 4.3 Resubstitution predictor design due to load rejection.

$$X = [\, V_3, Q_{23}, P_g, X_d \,]^T \tag{4.5}$$

and for the *intuitive* features:

$$W = [\, -0.1649, -0.6042, 1.5471, 0.0616 \,]^T \;, w_0 = -0.1451$$

$$X = [\, X_d, Q_F, Q_g, I_{fd} \,]^T \tag{4.6}$$

where W is the weighting vector, w_0 is the threshold weight, and X is the feature vector. From these designs, Eqns. (4.5) and (4.6), it can be seen that the hyperplane surface is inversely proportional to V_3, P_g, X_d, and Q_F. For example, if X_d increases (meaning fewer machines) then the predictor hyperplane will tend towards the self-excitation patterns, which is logical. On the other hand, the predictor surface is directly proportional to Q_{23}, Q_g , and I_{fd} which agrees with the physical interpretation of self-excitation conditions.

Similarly, Figure 4.4 presents the design results obtained for the selection of design parameters, σ and PR, for the double contingency condition. In this case the prediction accuracy builds up faster than the case of load rejection, and reaches about 90 %. It must be noted here that the design *philosophy* is to keep the prediction of self-excitation patterns greater than or equal to 95 %. The preferred designs in this case are selected at $\sigma = 4$ and $\sigma = 5$. On the other hand, the PR parameter starts flat then builds up to a saturation levels of 1.3 and 1.4 for both feature cases.

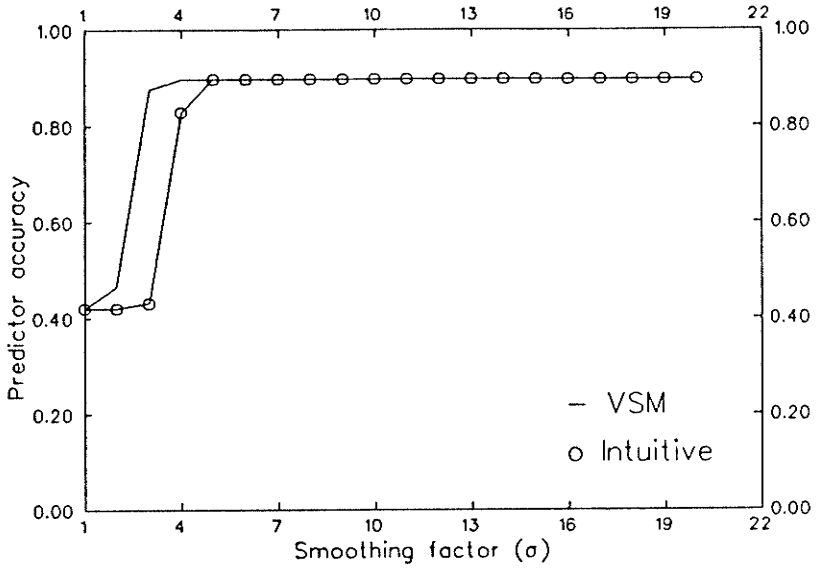
The hyperplane design for the double contingency case and using VSM features can be derived as:

$$W = [\, 1.1268, -0.2203, -0.0909, -0.9064 \,]^T \;, w_0 = -0.2057$$

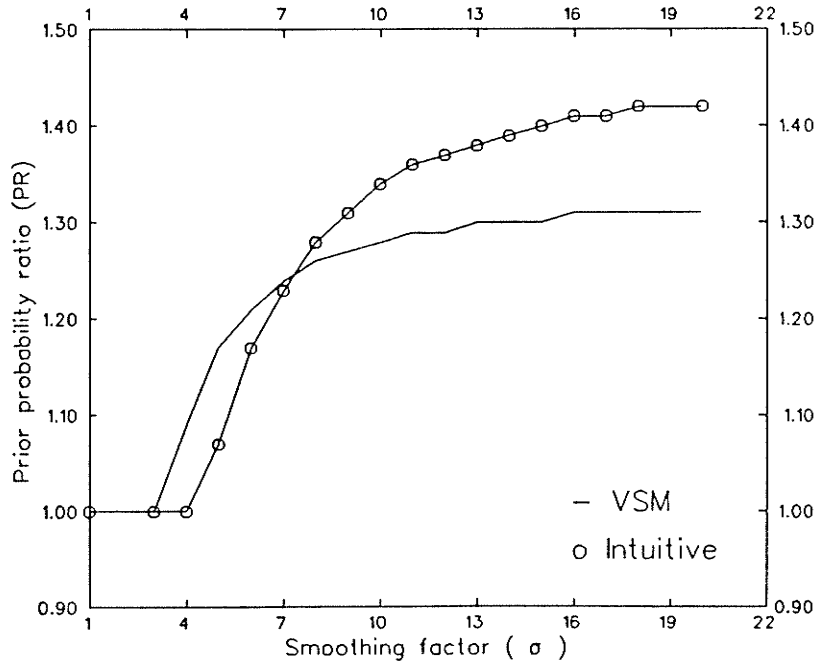
$$X = [\, Q_{23}, V_3, X_d, Q_F \,]^T \tag{4.7}$$

and for the intuitive features case is obtained as:

$$W = [\, -0.0926, -1.0635, 1.1053, -0.2479 \,]^T \;, w_0 = -0.2393$$



(a) predictor accuracy



(b) optimum pairs of σ & PR

Fig. 4.4 Resubstitution predictor design due to double contingency.

$$X = [\, X_d , Q_F , Q_g , I_{fd} \,]^T \tag{4.8}$$

Table 4.4 presents the performance evaluation results using the above described design procedure for the first class (no self-excitation), the second class (self-excitation), and for the overall problem. These results confirm that an overall prediction accuracy of greater than 90% can be achieved for the load rejection case and around 90% for the double contingency case.

(b) **The Hold-out design**

The design procedure involved in this method is based on the *partition* of the available pattern set into a *design* set and a *test* set. The design set is to be used for the design of the predictor hyperplane, while the test set is to be used for the evaluation of the predictor performance.

The partition method developed here is based on the idea of selecting patterns which are near to the hyperplane separating surface to construct the design set, Fig. 4.5. This method has the advantage of getting the highest predictor accuracy using the smallest number of design patterns compared to the case of using patterns gathered at random [9,10].

The Hold-out design algorithm developed is shown in Fig. 4.6. This algorithm presents different steps which are required to come up with the appropriate design. These steps could be summarized as follows:

(1) Design an approximate predictor using an initial design set.

(2) Update the design set using the nearest patterns to the separating surface.

(3) Redesign the predictor surface using the new design set.

(4) Calculate the predictor accuracy. If it is satisfactory terminate this process, otherwise proceed to step (2).

The results obtained using this algorithm, given in Fig. 4.7, explain the effect of selecting the design set on the achieved predictor performance. These results are

Table 4.4 Resubstitution design performance evaluation

Class type	LR-predictor		DC-predictor	
	VSM	Intuitive	VSM	Intuitive
Class 1	72.22	90.0	84.31	84.31
Class 2	100.0	97.06	97.3	97.3
Overall	79.84	91.94	89.77	89.77

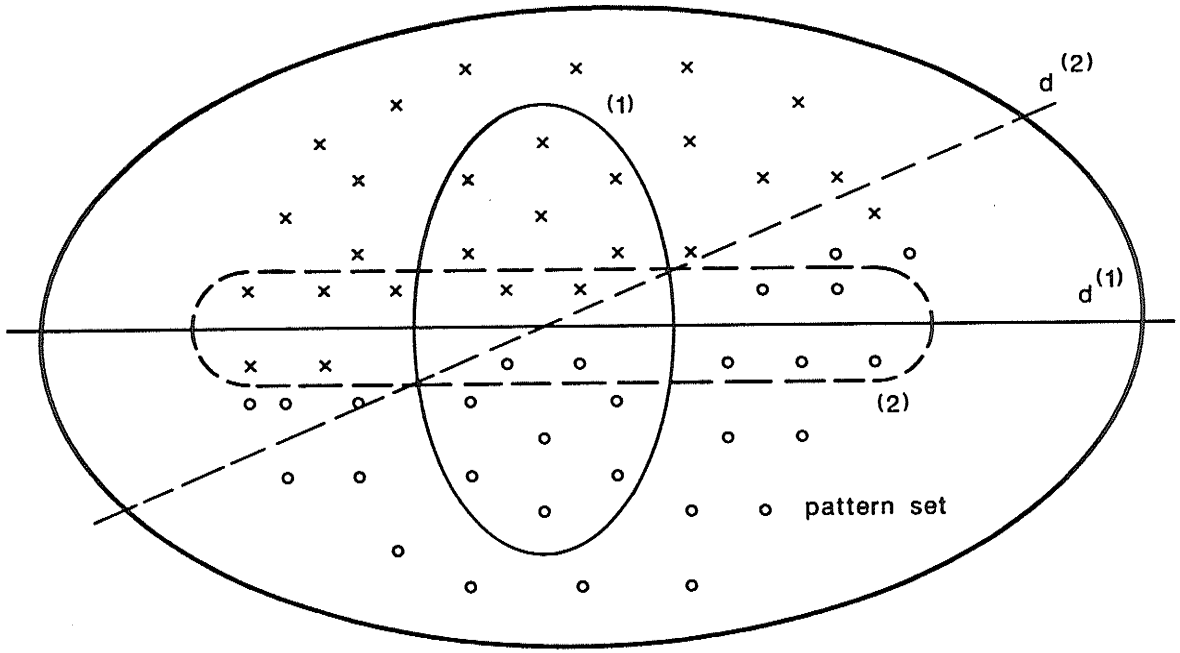


Fig. 4.5 Partitioning of the pattern set.

$d^{(1)}$: initial predictor design corresponding to design set (1)

$d^{(2)}$: updating predictor design using updating design set (2)

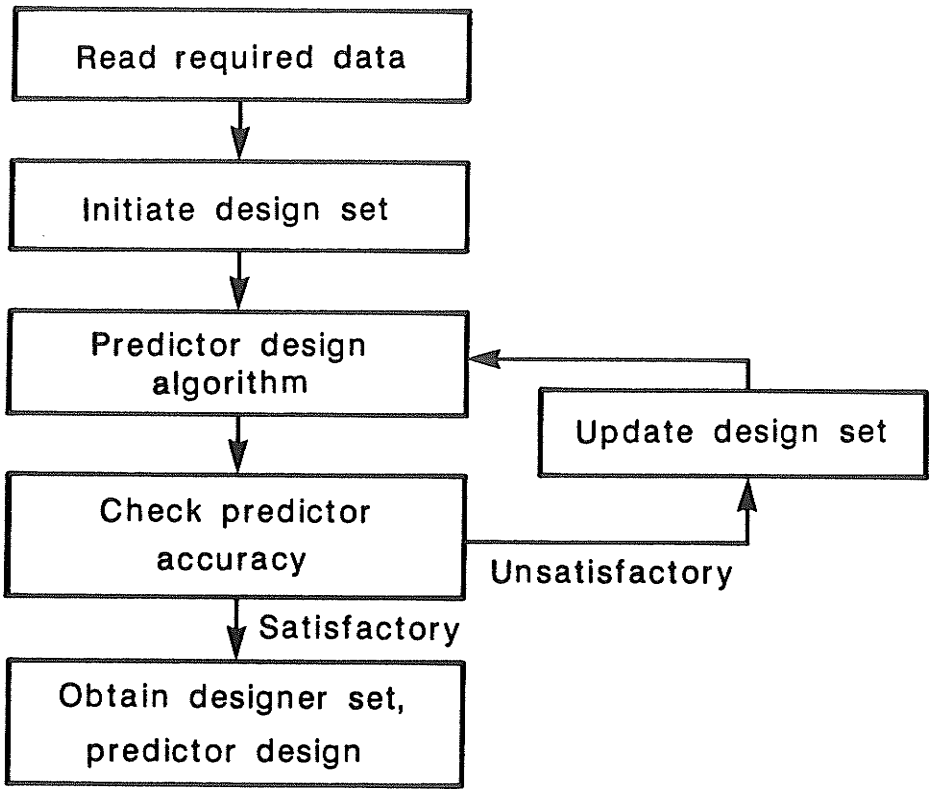


Fig. 4.6 Hold-out predictor design algorithm.

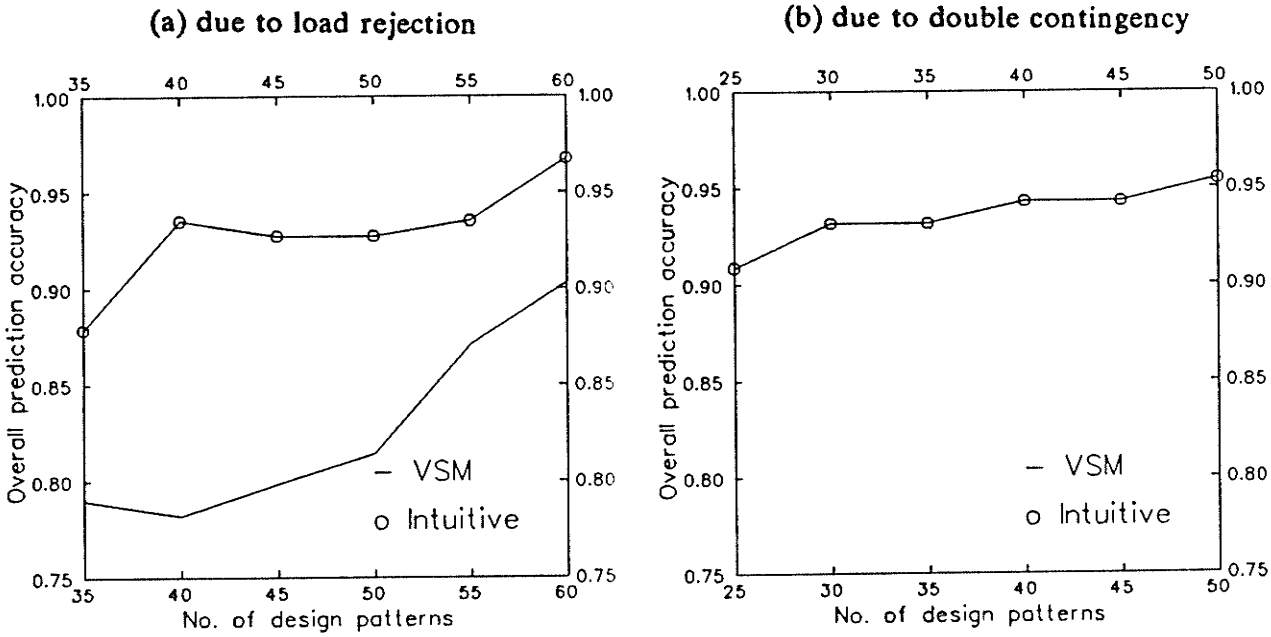


Fig. 4.7 Selection of design set.

obtained for the case of load rejection as well as for the double contingency case. It can be concluded that the final designs could be derived after only 6 iterations.

Each of the final designs obtained before, is a function of the design parameters (σ and PR). Figure 4.8 shows the results obtained regarding the selection of these parameters, using the VSM and intuitive features for the case of load rejection contingency. The optimal design obtained using the VSM features can be written as:

$$W = [\ -0.0209, 0.6098, -0.0622, -0.0209\ ]^T\ ,\ w_0 = -0.0089 \qquad (4.9)$$

and that obtained using the intuitive features is given by:

$$W = [\ -0.1059, -0.3769, 0.9344, 0.0669\ ]^T\ ,\ w_0 = -0.0608 \qquad (4.10)$$

The results shown in Fig. 4.9 explain the effect of the design parameters on the design process for the condition of double contingency. It can be observed that the intuitive features are quite effective and informative than the VSM features in this design. Additionally, the predictor hyperplane designs are selected at a prior probability ratio PR of 1.0 and 1.01, and at a smoothing factor σ of 3.0 and 4.0, in the case of VSM and intuitive features respectively.

The selected hyperplane equations are obtained as:

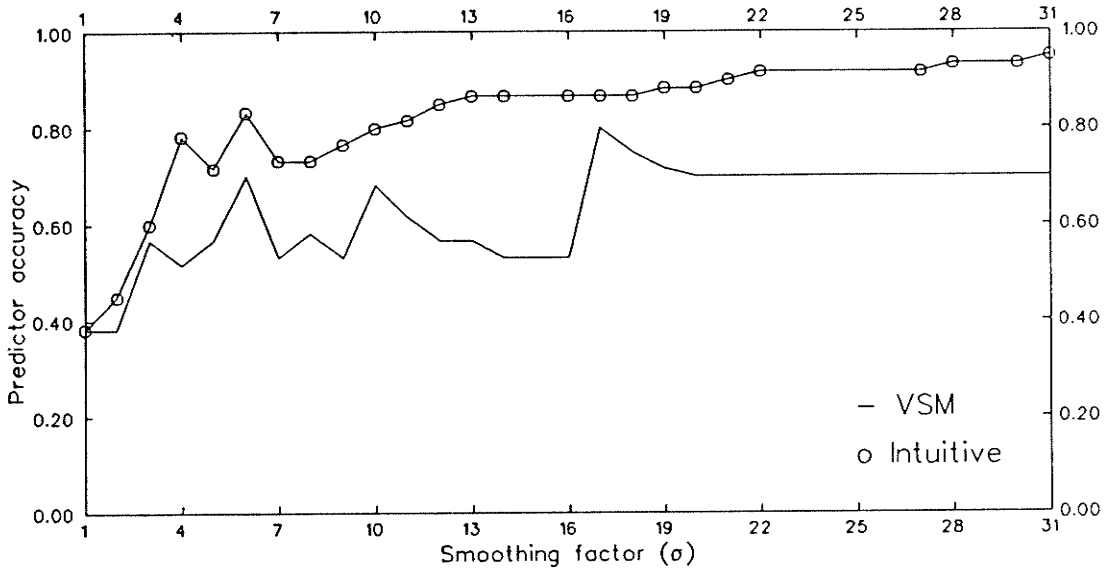
$$W = [\ 0.5548, -0.2052, -0.1063, -0.3004\ ]^T\ ,\ w_0 = -0.1978 \qquad (4.11)$$

for the VSM features, and

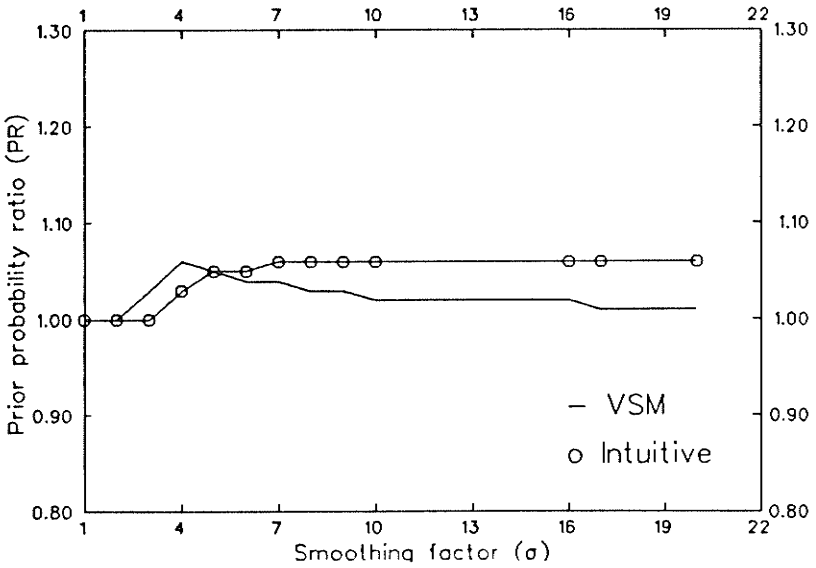
$$W = [\ -0.1087, -0.3761, 0.6386, -0.2903\ ]^T\ ,\ w_0 = -0.2074 \qquad (4.12)$$

for the intuitive features.

The performance evaluation using the Hold-out method is presented by Table 4.5. The results obtained indicate that the lowest design accuracy is about 80 % and the highest is about 95 %. On the other hand, the test results were as low as 98.44 % prediction accuracy. Generally speaking, the overall prediction accuracy is greater

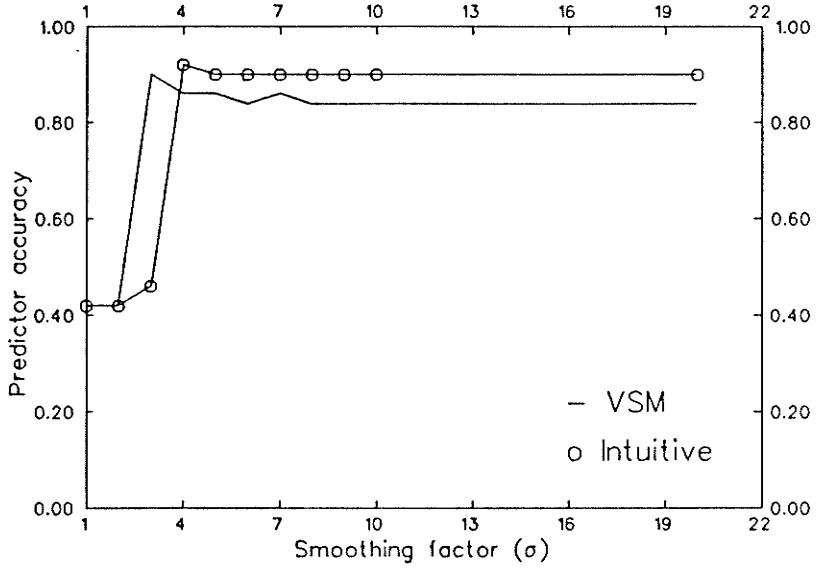


(a) predictor accuracy

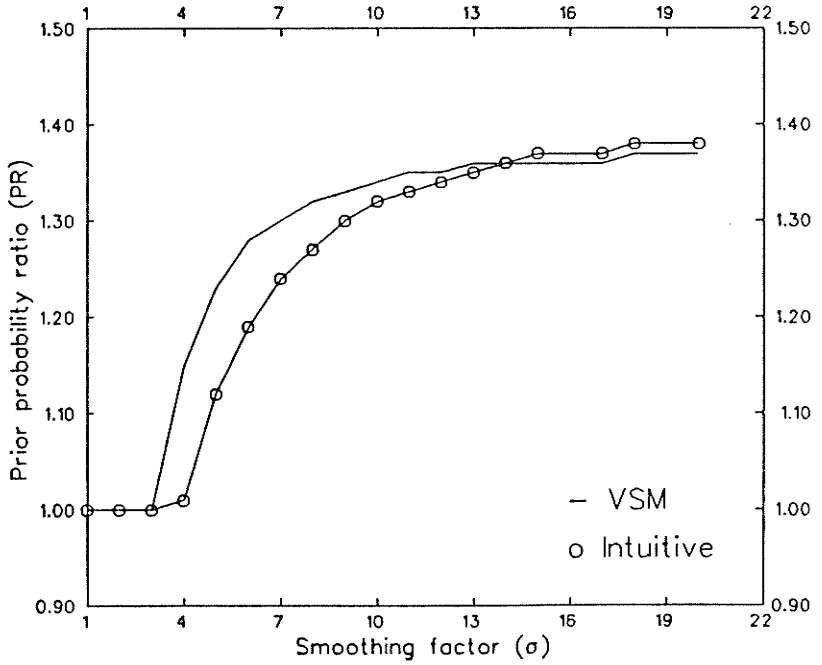


(b) optimum pairs of σ & PR

Fig. 4.8 Hold-out predictor design due to load rejection.



(a) predictor accuracy



(b) optimum pairs of σ & PR

Fig. 4.9 Hold-out predictor design due to double contingency.

Table 4.5 Hold-out design performance evaluation

Class type	LR-predictor		DC-predictor	
	VSM	Intuitive	VSM	Intuitive
Class 1	67.57	91.89	82.76	86.21
Class 2	100.0	100.0	100.0	100.0
Total	80.00	95.00	90.00	92.00

(a) design stage

Class type	LR-predictor		DC-predictor	
	VSM	Intuitive	VSM	Intuitive
Class 1	100.0	100.0	100.0	100.0
Class 2	100.0	90.91	100.0	100.0
Total	100.0	98.44	100.0	100.0
Overall	90.32	96.77	94.32	95.46

(b) testing stage

than 90 %, and the intuitive features are proved to be more powerful and informative than the VSM features.

The diagram shown in Fig. 4.10 explains the predictor design *spectrum* for the available pattern set. Self-excitation patterns (abnormal patterns) have a prediction index which varies from -2.0 to near zero. On the other hand, no-self-excitation patterns have the range from near zero to 8.0 which could be divided into three levels: weak, average, and strong. Therefore, the level of a given new pattern can be simply identified and how close this pattern is to other levels. In other words, the relative degree by which a pattern belongs to a certain level can be evaluated.

4.4.2 Stepwise prediction scheme

This is another design scheme where a stepwise learning algorithm and stepwise features are employed in order to come up with the required prediction scheme. Figure 4.11 presents a block diagram for this scheme. The prediction of self-excitation due to the load rejection as well as the double contingency is considered in the system under design.

The stepwise learning algorithm is essentially based on the available pattern set and the distribution of these patterns in the feature space. This algorithm is a function of the prior probability ratio (PR). This parameter PR affects the constant term or threshold weight w_0 and it does not affect the weighting vector W. By changing this parameter the hyperplane constant term changes which means that the hyperplane moves up or down in parallel by increasing or decreasing the parameter value.

The results obtained, shown in Fig. 4.12, explain the hyperplane predictor design and its relation with the prior probability of class no. 1 (p_1) for different numbers of features (f) (only f =1 and f =4 are shown). These results prove that the highest design overall accuracy occurs at $p_1 = 0.6$. It should be emphasized that if $p_2 > p_1$, the prediction will be *biased* towards the second class patterns, and on the contrary, if $p_1 > p_2$ the first class patterns will be biased. Therefore, a strong

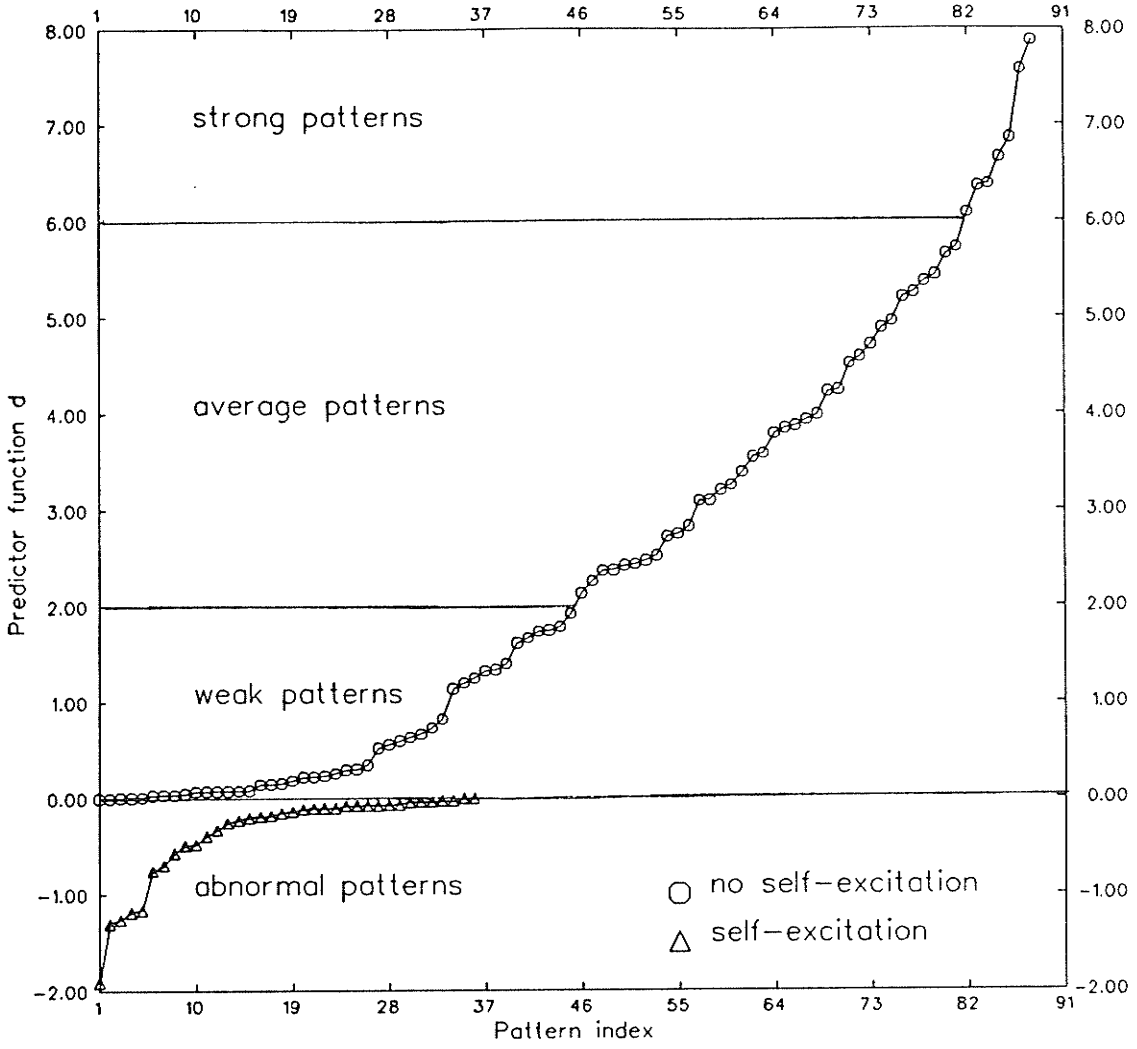


Fig. 4.10 Predictor design spectrum.

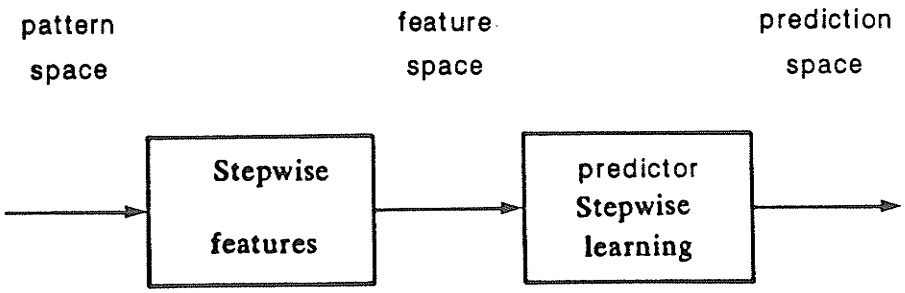
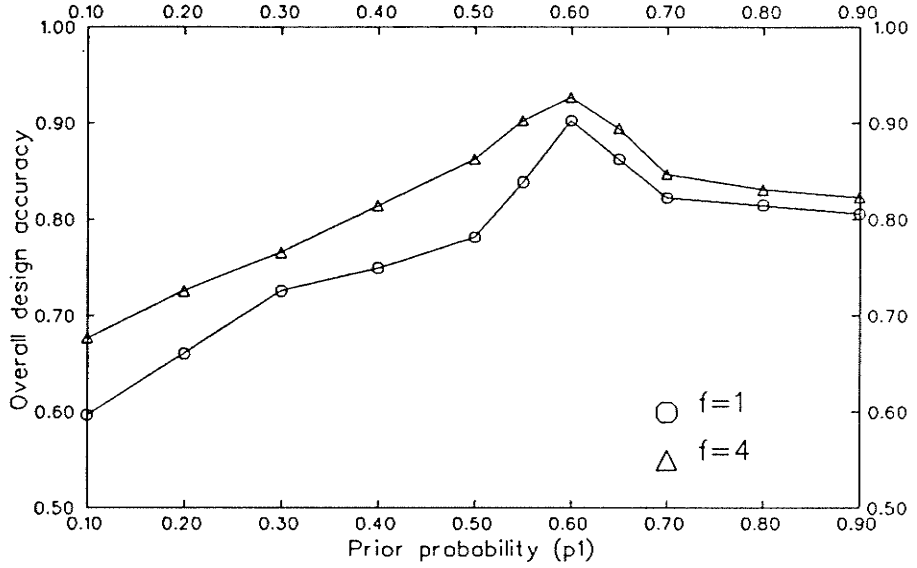
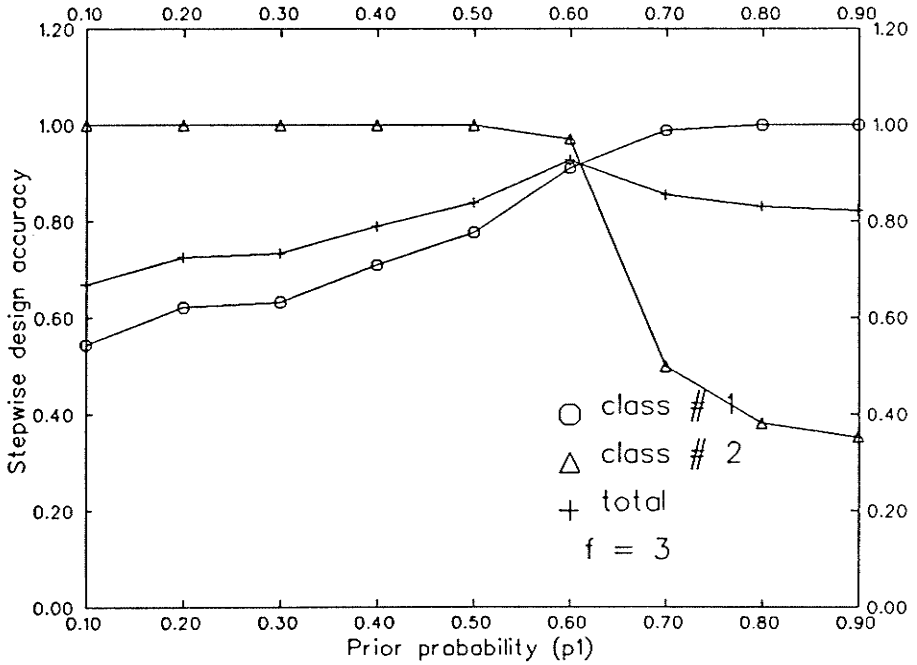


Fig. 4.11 Stepwise prediction scheme.



(a) effect of number of features



(b) effect of prior probability

Fig. 4.12 Stepwise predictor design results.

conclusion from these results can be drawn. It is that the prior probability ratio (p_2/p_1) could be estimated to a value around 1.0. This estimation looks reasonable for most situations where this design scheme is evolved.

Table 4.6 records the results obtained during the design stage as well as the testing stage. In the design stage, the whole pattern set is used to derive the design scheme, while in the testing stage, the *leave −one −out* estimate is evolved in order to evaluate the performance of the prediction scheme. It can be proved that the highest prediction accuracies are of 92.7, 94.3 %, and are obtained when $f = 3$.

4.5 Effect of Telemetering Communications Failure

Failure of the *communication* channels is a practical possibility. Therefore it is necessary to study the effect of the failure on the performance of the prediction system and to identify the most informative channels. This is in order to avoid a *degradation* in the prediction system accuracy and keep the accuracy of the predictor as high as possible.

As discussed in chapter 2, there is another approach in order to avoid a large drop in the prediction system performance in case of communications failure. This is: either to use the last sample of the failed channel, or to use the average value of recent samples of the failed channel.

Assuming that one channel failure at a time is the most likely situation. Table 4.7 presents the results obtained for the linear prediction scheme (section 4.4.1) in case of communications failure. From these results, it can be proved that the failure of X_d channel or sometimes V_3 channel had no effect (or occasionally a slight effect) on the prediction system performance. On the other hand, the failure of Q_{23} channel or Q_F or Q_g channels had a drastic effect and the prediction performance dropped to a very low level.

The results obtained in Table 4.8 explain the effect of communications failure on the stepwise prediction scheme performance. They indicate that all the channels

Table 4.6 Stepwise prediction performance evaluation

No. of features	Design accuracy			Test accuracy		
	Class 1	Class 2	Total	Class 1	Class 2	Total
1	87.8	97.1	90.3	87.8	97.1	90.3
2	94.4	79.4	90.3	94.4	73.5	88.7
3	91.1	97.1	92.7	90.0	94.1	91.1
4	93.3	91.2	92.7	90.0	88.2	89.5

(a) load rejection

No. of features	Design accuracy			Test accuracy		
	Class 1	Class 2	Total	Class 1	Class 2	Total
1	80.4	94.6	86.4	80.4	94.6	86.4
2	88.2	100.	93.2	86.3	100.	92.0
3	92.2	97.3	94.3	90.2	97.3	93.2
4	90.2	97.3	93.2	90.2	97.3	93.2

(b) double contingency

Table 4.7 Effect of telemetering communications failure on the linear prediction scheme

Channel failed	Resubstitution design			Hold-out design		
	Class 1	Class 2	Total	Class 1	Class 2	Total
V_3	87.78	73.53	83.87	92.22	94.12	92.74
Q_{23}	0.0	100.0	27.42	0.0	100.0	27.42
P_g	84.44	67.65	79.84	100.0	35.29	82.26
X_d	72.22	97.06	79.03	88.89	100.0	91.94

(a) load rejection: VSM features

Channel failed	Resubstitution design			Hold-out design		
	Class 1	Class 2	Total	Class 1	Class 2	Total
Q_{23}	0.0	100.0	42.05	0.0	100.0	42.05
V_3	86.27	94.59	89.77	94.12	94.59	94.32
X_d	84.31	97.3	89.77	92.16	97.3	94.32
Q_F	100.0	10.81	62.5	100.0	37.84	73.86

(b) double contingency: VSM features

Table 4.7 (continued).

Channel failed	Resubstitution design			Hold-out design		
	Class 1	Class 2	Total	Class 1	Class 2	Total
X_d	93.33	94.12	93.55	97.78	91.18	95.97
Q_F	100.0	29.41	80.65	100.0	29.41	80.65
Q_g	0.0	100.0	27.42	0.0	100.0	27.42
I_{fd}	82.22	100.0	87.10	81.11	100.0	86.29

(c) load rejection: Intuitive features

Channel failed	Resubstitution design			Hold-out design		
	Class 1	Class 2	Total	Class 1	Class 2	Total
X_d	84.31	97.30	89.77	92.16	100.0	95.45
Q_F	100.0	5.41	60.22	100.0	32.43	71.59
Q_g	0.0	100.0	42.05	0.0	100.0	42.05
I_{fd}	90.2	91.89	90.91	96.08	78.38	88.64

(d) double contingency : Intuitive features

Table 4.8 Effect of telemetering communications failure on the stepwise prediction scheme

Channel failed	LR-Prediction		
	Class 1	Class 2	Total
V_3	100.0	0.0	72.58
I_{fd}	100.0	0.0	72.58
δ_3	100.0	8.82	75.0

(a) load rejection

Channel failed	DC-Prediction		
	Class 1	Class 2	Total
Q_{23}	0.0	100.0	42.05
Q_g	100.0	2.70	59.09
X_d	100.0	37.84	73.86

(b) double contingency

are very informative and highly important for the prediction system.

As a summary, it can be concluded that the X_d channel is the least significant channel for the linear prediction scheme.

4.6 Reduction of Prediction System Structure

The designed predictor hyperplane (dimension) is a function of identified features. Therefore, the *operating prediction structure*, Fig. 4.13, has a number of *telemetering* channels equivalent to the number of predictor features. These communication channels may be costly. For this reason, there is a need for an algorithm to reduce the number of required channels, which of course will also simplify the prediction system structure.

The following procedure has been developed in order to reduce the required number of features without affecting the overall system performance appreciably:

1- Design the predictor hyperplane with the features identified.

2- Avoid the feature with the smallest *discriminatory* power from the features used in step 1.

3- Redesign the predictor with the reduced features obtained in step 2.

4- Check the prediction system accuracy: if it is decreased terminate this procedure, if not return to step 2.

Table 4.9 presents the results obtained using this procedure, from which it is clear that the prediction system structure can be reduced by one channel without any significant effect on the overall performance.

4.7 Development of Self-Excitation Corrective Action

If a system state is predicted to be abnormal, then a fast corrective action is needed in order to improve system security. As to the required actions, several methods are available [1,9]. For example, a change of the power system configuration, or a modification of the distribution of the active or reactive injections. An

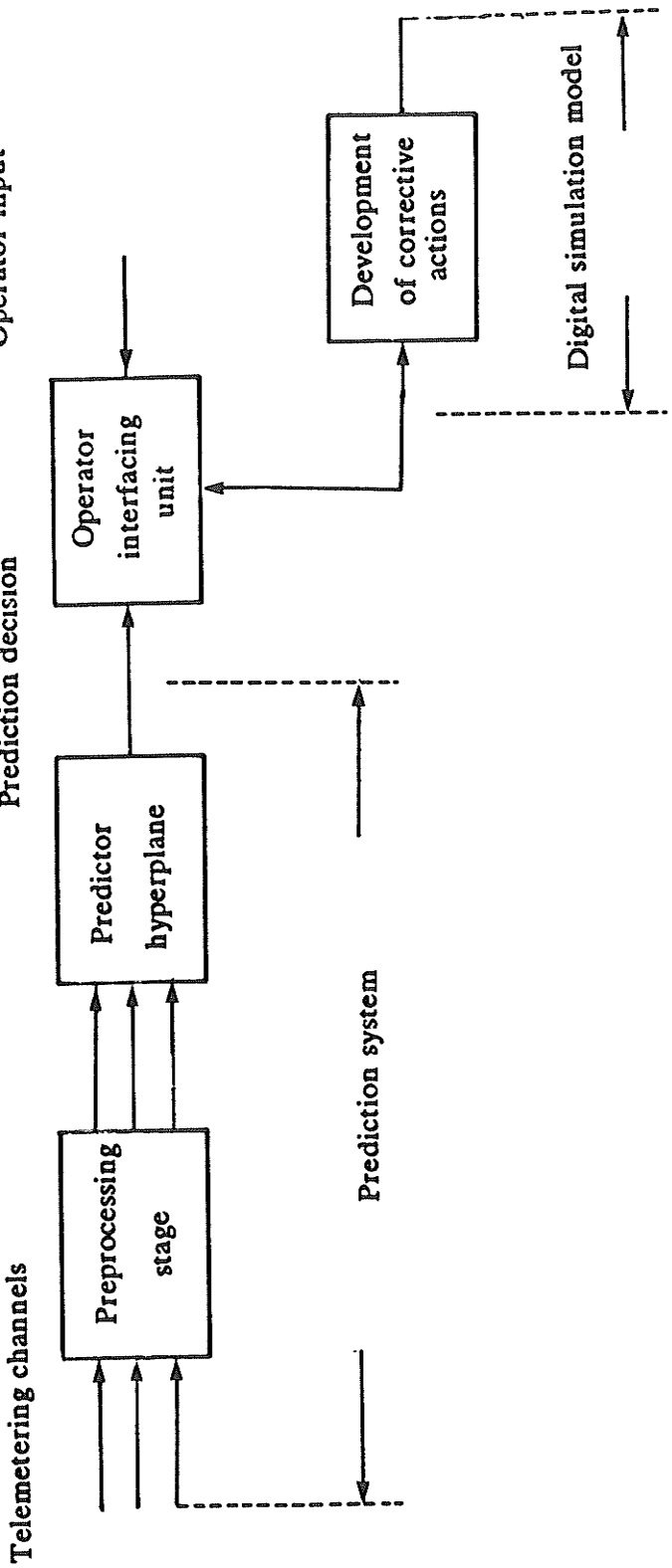


Fig. 4.13 Operating prediction system structure.

Table 4.9 Reduction of prediction system structure

No. of features	Resubs. design		Hold-out design	
	VSM	Intuitive	VSM	Intuitive
4	79.84	91.94	90.32	96.77
3	79.84	93.55	91.94	96.77
2	75.00	68.55	75.00	87.10

(a) load rejection

No. of features	Resubs. design		Hold-out design	
	VSM	Intuitive	VSM	Intuitive
4	89.77	89.77	94.32	95.46
3	89.77	89.77	93.18	95.46
2	87.50	88.64	90.91	89.77

(a) double contingency

alternative approach would be to employ a load flow simulation constrained by the predictor function.

Regarding the self-excitation problem which is basically a violation in the reactive flow distribution throughout the system, the corrective action suggested is the adjustment in the reactive power injections from the reactive sources (filter banks). A corrective algorithm is developed in order to correct the system *alert* (insecure) state into a *preferred* (secure) state. This algorithm is based on the *sensitivity* of the predictor design in terms of the selected features.

To provide the system operator with the suggested appropriate action, the following algorithm, Fig. 4.14, is developed as:

(1) Derive the required predictor design d using the appropriate *learning* scheme. For example, assume the predictor design is a function of all features as:

$$d = F(x_1, x_2, ..., x_f) \tag{4.13}$$

(2) Formulate the predictor sensitivity w.r.t. all features involved in the design scheme as:

$$\Delta d = \frac{\partial F}{\partial x_1}\Delta x_1 + \frac{\partial F}{\partial x_2}\Delta x_2 + ... + \frac{\partial F}{\partial x_f}\Delta x_f \tag{4.14}$$

(3) Select one of the *feasible* features, means all features under the operator control, to be the *control feature* by which the operator will adjust system conditions.

(4) Re-evaluate the prediction sensitivity in terms of the control feature taking into consideration other features contribution.

$$\Delta d = [\ \frac{\partial F}{\partial x_1}\ \frac{\partial x_1}{\partial x_i} + \frac{\partial F}{\partial x_2}\ \frac{\partial x_2}{\partial x_i} + ... + \frac{\partial F}{\partial x_f}\ \frac{\partial x_f}{\partial x_i}\ ]\ \Delta x_i$$

$$= h\ \Delta x_i \tag{4.15}$$

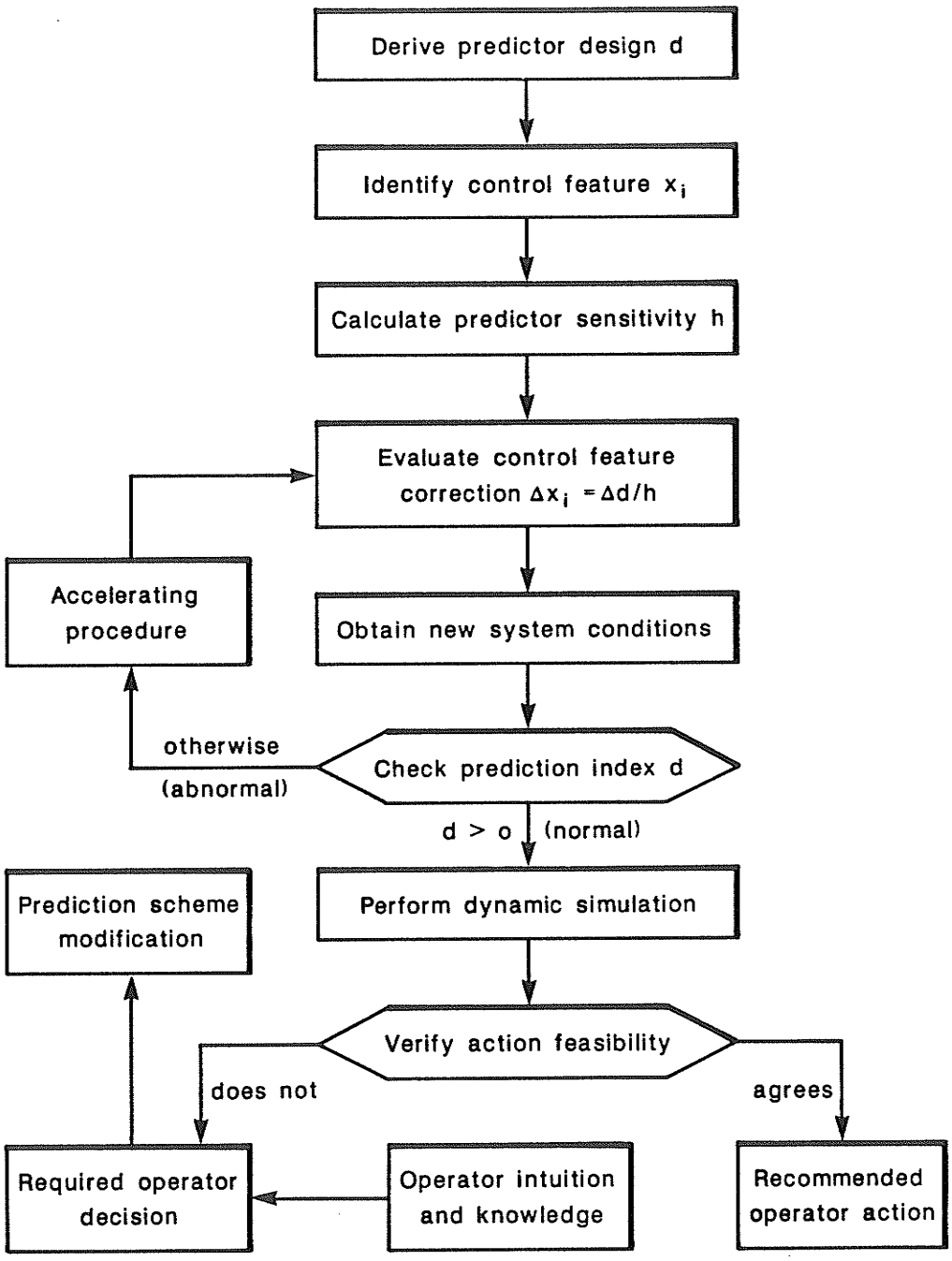


Fig. 4.14 Flow chart of the corrective algorithm.

where x_i is the control feature and h is a constant.

(5) Determine the amount of change in the control feature Δx_i just required to bring the system into a normal state as:

$$\Delta x_i = \Delta d \ / \ h \tag{4.16}$$

(6) Using a fast load flow simulation (DC load flow), obtain the new system conditions in general and the design features in particular.

(7) Calculate the prediction index d and verify the required constraints i.e. if $d > 0$ proceed to next step, otherwise return to step (5) since the system is in the abnormal (alert) state.

(8) Confirm the action feasibility using system dynamic simulation. If it agrees then operator may take action. In case it does not agree, which may happen in a very rare situations due to a misprediction by the prediction scheme involved, operator has to recommend for a modification in the design of the prediction system and to decide an action based on his own knowledge and field experience as explained in Fig. 4.14 .

In case of excessive modifications, an accelerating procedure could be employed as:

$$\Delta x_i = a \ \Delta d \ / \ h$$

$$a = \frac{|d^{(0)}|}{(d^{(1)} - d^{(0)})} \tag{4.17}$$

where a is the acceleration coefficient and $d^{(0)}$ & $d^{(1)}$ is the prediction index after zero, one iteration respectively.

4.7.1 Corrective Action of Linear Prediction Scheme

With respect to the application of this algorithm to the problem of concern, the following steps are considered:

(1) Consider a predictor design as for example the scheme using Hold-out method with the intuitive features involved as given by Eq. (4.10).

(2) Select the control feature as in this case the reactive power injection of filter banks Q_F.

(3) Assume that the change in filter banks MVARs ΔQ_F is followed by an equal and opposite change in generator MVARs ΔQ_g i.e. $\Delta Q_F = -\Delta Q_g$.

(4) Evaluate the predictor sensitivity in terms of the control feature as: $\Delta d = h\ \Delta Q_F$, $h = -1.3113$ in this case.

A sample of the results obtained using the proceeding algorithm is presented in Table 4.10a. It can be observed that only one iteration step is required for most of cases and two iterations for one case in order to come up with the required correction, without the use of any accelerating process.

4.7.2 Corrective Action of Stepwise Prediction Scheme

Consider, in this case, the stepwise prediction scheme designed for the double contingency condition as discussed before in section 4.4.2 . This design can be written as:

$$d = W^T X + w_0\ ,\ W^T = [w_1, w_2, w_3]\ ,\ X^T = [X_d, Q_g, Q_{23}] \qquad (4.18)$$

The predictor sensitivity could be written as:

$$\Delta d = w_2\ \Delta Q_g + w_3\ \Delta Q_{23} \qquad (4.19)$$

and assume that approximately $\Delta Q_g = \Delta Q_{23} = -\Delta Q_F$ then substitute back in Eq. (4.19), therefore, we can write:

Table 4.10 Self-Excitation corrective actions

Initial predic. index $d^{(0)}$	Initial filter MVARs $Q_F^{(0)}$	Required filter MVARs $Q_F^{(k)}$	Final predic. index $d^{(k)}$	No. of steps k	APPROVED OPERATOR ACTION $Q_F^{(k)} - Q_F^{(0)}$
-0.4776	75.00	38.58	0.1005	1	-36.42
-0.1786	105.00	91.38	0.0301	1	-13.62
-0.0694	150.00	144.71	0.0106	1	-5.29
-0.0775	200.00	194.09	0.0116	1	-5.91
-0.3915	229.50	194.13	0.0112	2	-35.37
-1.1571	229.50	141.26	0.2094	1	-88.24
-0.0265	245.00	242.98	0.0042	1	-2.02
-0.1125	300.00	291.42	0.0173	1	-8.58
-0.1489	400.00	388.64	0.0227	1	-11.36
-0.1861	500.00	485.81	0.0271	1	-14.19
-0.2239	600.00	582.92	0.0315	1	-17.08

(a) linear prediction scheme

$$\Delta d = -(w_2 + w_3)\,\Delta Q_F = h\,\Delta Q_F \qquad (4.20)$$

where h is equal to -0.16724 in this design scheme.

From Eq. (4.20) the correction in the control feature can be determined in order to enhance the system conditions. Following the corrective algorithm consequence including the accelerating procedure, a sample of results are obtained and presented in Table 4.10b . It can be proved that the developed corrective algorithm is very effective and helpful to the power system operators providing them with appropriate actions required to improve system security.

4.8 Conclusions

This chapter provides a new application using pattern recognition techniques. An efficient prediction system based on these techniques is designed. It rapidly predicts self-excitation conditions so that the operator can take action necessary to improve the system state.

The results obtained for the studied power system confirm the high performance of the pattern-recognition based prediction system presented. The intuitive identification of features was very effective for this application. The discriminant hyperplane used for the predictor design was proven applicable for the prediction of self-excitation problems. The use of this design method permits a recursive adaptation to new system conditions. The developed prediction-system structure-reduction method was very useful. The developed corrective algorithm proved its effectiveness for the application of concern.

The main conclusion in this chapter is that pattern recognition techniques can be applied to the prediction of power system self-excitation operating problems. These techniques are very attractive candidates for on-line prediction schemes: speed, accuracy, and ease of implementation can be accomplished.

Table 4.10 (continued).

Initial predic. index $d^{(0)}$	Initial filter MVARs $Q_F^{(0)}$	Required filter MVARs $Q_F^{(k)}$	Final predic. index $d^{(k)}$	No. of steps k	APPROVED OPERATOR ACTION $Q_F^{(k)} - Q_F^{(0)}$
-0.0817	150.00	85.95	0.0019	2	-64.05
-0.1534	195.00	87.00	0.0004	2	-108.00
-0.0668	250.00	204.22	0.0009	2	-45.78
-0.0947	350.00	285.19	0.0015	2	-64.81
-0.1429	450.00	355.51	0.0006	2	-94.49
-0.1982	550.00	417.96	0.0011	3	-132.04

(b) stepwise prediction scheme

CHAPTER 5

LOAD REJECTION DYNAMIC OVERVOLTAGES: PREDICTION SYSTEM DESIGN

5.1 Introduction

As discussed before in chapter 3, load rejection may lead to dynamic overvoltage problems. This situation can arise in cases particularly where self-excitation has not occurred, depending on the number of machines in service relative to the number of filter banks. Dynamic overvoltages are considered unacceptable condition in power systems. They have many side-effects on the operation of hydro power stations [5].

An *overvoltage prediction* system is to be designed in this chapter. Linear and stepwise prediction schemes are considered in the design process. Communications failure and its effect on the performance of the prediction system is discussed. Reduction of the prediction system structure is treated. Finally, a corrective action algorithm is developed.

A *Load rejection* predictor scheme is shown in Fig. 5.1. In this figure two predictor schemes are connected in series: the self-excitation predictor and the *overvoltage predictor*. The self-excitation predictor predicts self-excitation states, and the overvoltage predictor predicts overvoltage conditions for all no-self-excitation states.

5.2 Dynamic Overvoltage Pattern Set

Radisson Converter Station blocking (dc load rejection) is the most likely contingency for the Northern system. Therefore, the first step is to create, due to that contingency, a dynamic overvoltage pattern set sufficient for the prediction system design. In order to do so, the load rejection pattern set (124 patterns) created before

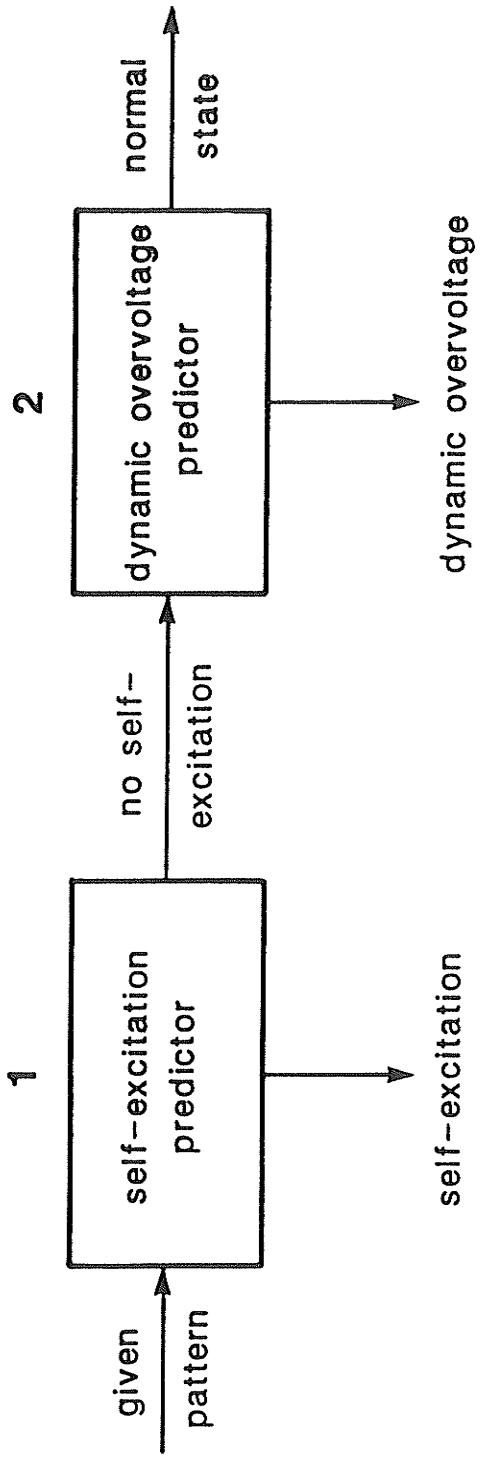


Fig. 5.1 Load rejection predictor scheme.

in Chapter 4, is used again, but in this case *excluding* all self-excitation cases. In other words, the overvoltage pattern set is constructed from all no-self-excitation conditions (90 patterns) created due to load rejection.

As shown in Fig. 5.2, the dynamic overvoltage "limit" is taken here at 1.26 p.u. This results in an overvoltage pattern set having 50 normal patterns and 40 abnormal patterns.

5.3 Dynamic Overvoltage Features

As mentioned before, the *principal component* method is applied here again to the *voltage pattern set* in order to determine the sufficient number of features needed for the prediction process. Table 5.1 presents the results obtained in this case. These results confirm that only four features (for a tolerance of 1%) are required to provide the prediction system with the appropriate information.

In order to select the most informative variables to be considered as features, the feature identification techniques of Section 2.4 are employed. Table 5.2 includes the features obtained in this case.

5.4 Overvoltage Predictor Design and Evaluation

The pattern set having been created and the informative features identified, the next stage is to design the *overvoltage predictor* and evaluate its performance. In this design process, two schemes are considered: the linear prediction scheme and the stepwise prediction scheme.

5.4.1 Linear prediction scheme

In this approach the Resubstitution and Hold-out design methods are involved, along with the VSM and intuitive feature approaches.

Regarding the design process using the Resubstitution method, the results shown in Fig. 5.3 show the selection process of the two *design parameters* σ and

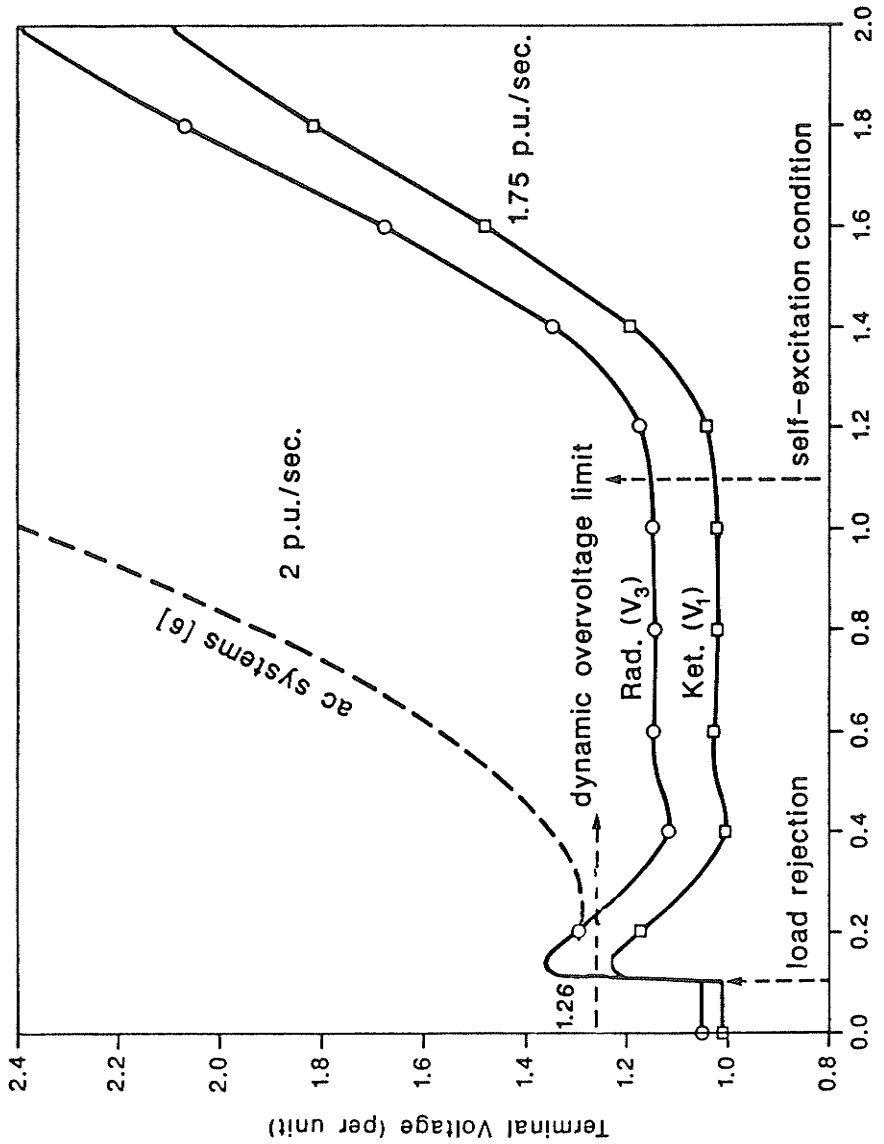


Fig. 5.2 Load rejection overvoltage results.

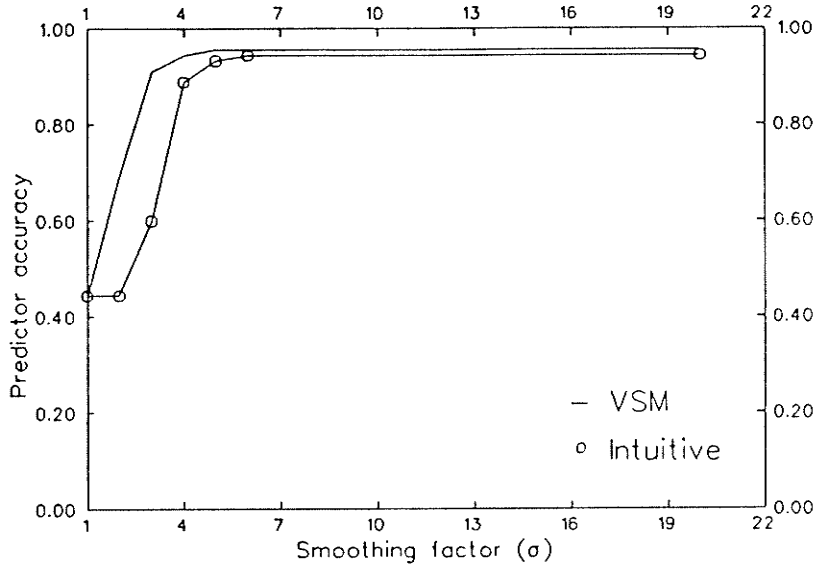
Table 5.1 Principal components results

Feature no.	1	2	3	4
e.v.	218.6780	16.0525	4.4473	0.4396
sum %	91.26	97.96	99.81	99.99

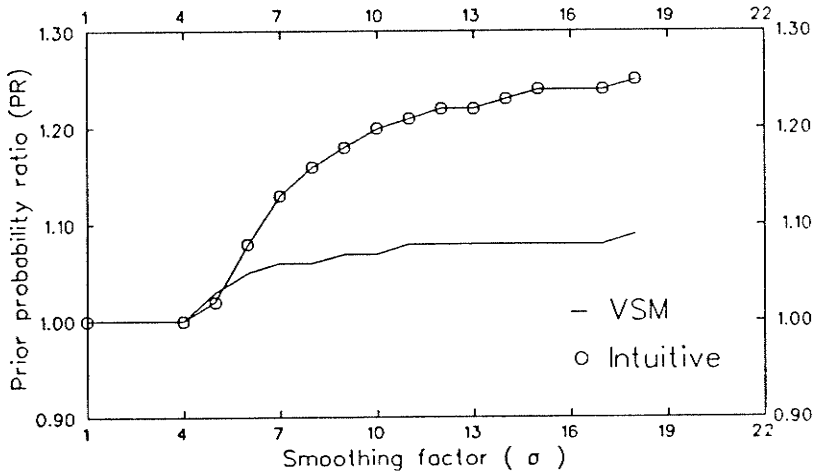
e.v.: eigen-value

Table 5.2 Dynamic overvoltage features

Feature no.	VSM method		Stepwise method		Intuitive features
	feature	VSM	feature	SM	
1	V_3	1.2954	V_3	133.7678	X_d
2	Q_{23}	1.0369	δ_2	79.8336	Q_F
3	Q_F	0.6128	I_{fd}	12.0086	Q_g
4	X_d	0.3362	Q_{23}	3.7632	I_{fd}



(a) predictor accuracy



(b) optimum pairs of σ & PR

Fig. 5.3 Resubstitution overvoltage predictor design.

PR. The best designs are selected at $\sigma = 5.0$ and PR = 1.03 for the case of using the VSM features , and at $\sigma = 6.0$ and PR = 1.08 using the intuitive features.

The selected design of the predictor hyperplane for this case and using VSM features is obtained as:

$$d = W^T X + w_0$$

where:

$$W = [\, -0.1138,\ 1.4888,\ -1.0326,\ -0.0908\,]^T \ , \ w_0 = -0.0956$$

$$X = [\, V_3,\ Q_{23},\ Q_F\ ,\ X_d\,]^T \tag{5.1}$$

and for the intuitive features is given by:

$$W = [\, -0.1039,\ -1.2818,\ 1.3719,\ -0.0932\,]^T \ , \ w_0 = -0.1824$$

$$X = [\, X_d\ ,\ Q_F\ ,\ Q_g\ ,\ I_{fd}\,]^T \tag{5.2}$$

The evaluation results shown in Table 5.3 indicate the performance of the prediction system designed herein. These results verify that the overvoltage prediction accuracy can be as low as 94.44% and as high as 95.56%. Also, it can be observed that the VSM features are more powerful and informative than the intuitive features for the application of concern.

Now considering the predictor design using in this case the Hold-out method and taking into account that the VSM features are involved as well as the intuitive features, Table 5.4 presents the results derived due to the *partitioning* of the pattern set into design and test sets. These results indicate that only 35 patterns are good enough for the design set. Also, the VSM features are more effective and superior than the intuitive ones. Moreover, a 95.56% overall prediction accuracy is achieved with only 2 iterations to come up with the accurate design set.

Table 5.3 Evaluation of OV-Prediction system performance using Resubstitution method

Feature type	Prediction Accuracy		
	class 1	class 2	overall
VSM	96.00	95.00	95.56
Intuitive	94.00	95.00	94.44

Table 5.4 Selection of the OV-design set

No. of patterns	VSM features			Intuitive features		
	design	test	overall	design	test	overall
30	93.33	95.00	94.44	90.00	95.00	93.33
35	91.43	98.18	95.56	82.86	100.0	93.33
40	85.00	100.0	93.33	75.00	100.0	88.89

OV : means overvoltage

Using the achieved design set, Figure 5.4 provides the design results obtained. From these results the selected design schemes in this case are given as:

$$W = [\ -0.0447, 0.5396, -0.3711, -0.1040\ ]^T\ ,\ w_0 = -0.0347 \qquad (5.3)$$

using the VSM features, and

$$W = [\ -0.1277, -0.4895, 0.4558, -0.0259\ ]^T\ ,\ w_0 = -0.0567 \qquad (5.4)$$

for the case of intuitive features.

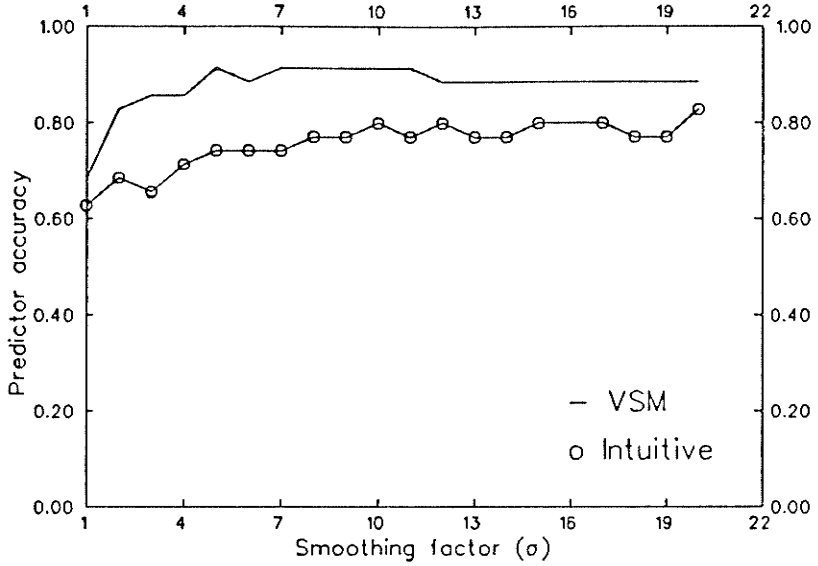
Table 5.5 shows the results achieved for the overvoltage prediction system evaluation. These results confirm that the highest accuracy (95.56%) achieved by the prediction system presented and on the application of concern, using the VSM features method.

5.4.2 Stepwise prediction scheme

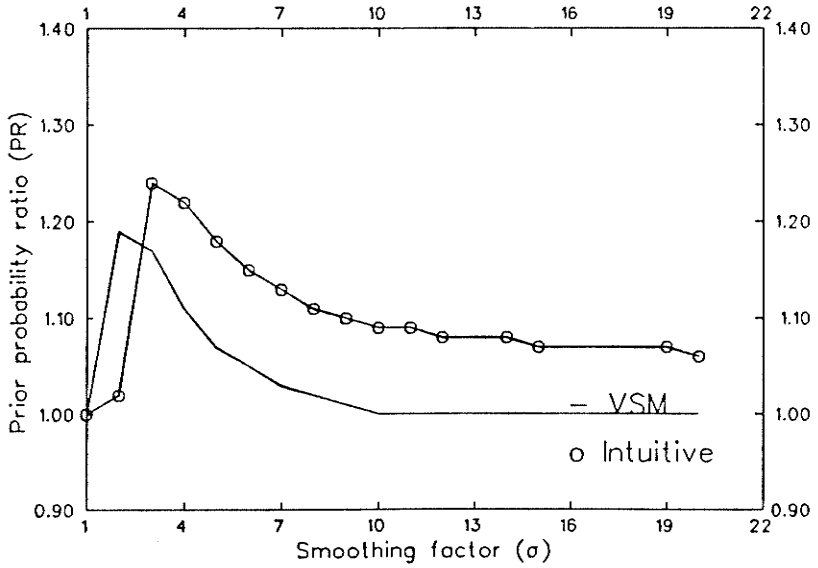
In this design scheme, the stepwise features are involved as well as the stepwise learning algorithm. Also, using this scheme, the *dependence* on the number of features or steps (f) and on the *prior probability* of different classes, could be derived from the results presented in Fig. 5.5. A perfect prediction accuracy is obtained at $p_1 = 0.6$. The optimal design selected is given by:

$$W = [\ -0.9996, 0.0086, -0.0288, -0.00087\ ]^T\ ,\ w_0 = 1.0969 \qquad (5.5)$$

The results presented by Fig. 5.6 show the prediction accuracy of a particular class of patterns and its prior probability factor. Patterns of class 1 are completely recognized when p_1 is greater than 0.6, and on the other hand, patterns of class 2 are also completely recognized when p_2 is greater than 0.2 . In addition, it can be observed that both classes are of 100% prediction in this case for the range of $0.6 \leq p_1 \leq 0.8$.



(a) predictor accuracy



(b) optimum pairs of σ & PR

Fig. 5.4 Hold-out overvoltage predictor design.

Table 5.5 Evaluation of OV-Prediction system performance using Hold-out method

Feature type	Design Accuracy			Testing Accuracy			Overall Accuracy
	class 1	class 2	total	class 1	class 2	total	
VSM	81.82	95.83	91.43	100.0	93.75	98.18	95.56
Intuitive	61.54	95.45	82.86	100.0	100.0	100.0	93.33

Table 5.6 Evaluation of OV-prediction system performance (stepwise scheme)

No. of features	Design Accuracy			Testing Accuracy		
	Class 1	Class 2	Total	Class 1	Class 2	Total
1	88.0	90.0	88.9	88.0	90.0	88.9
2	98.0	100.0	98.9	98.0	97.5	97.8
3	98.0	100.0	98.9	98.0	97.5	97.8
4	100.0	100.0	100.0	98.0	97.5	97.8

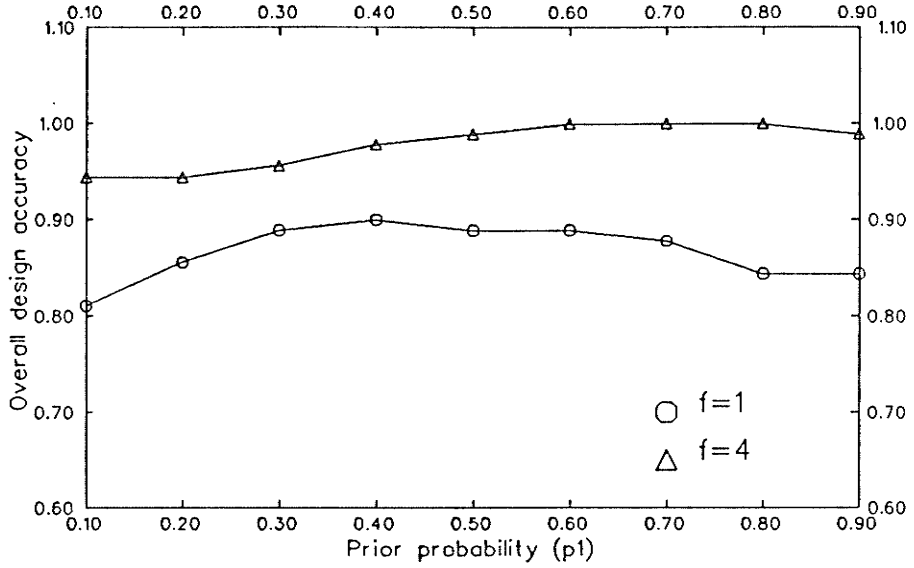


Fig. 5.5 Stepwise overvoltage predictor design (effect of no. of features).

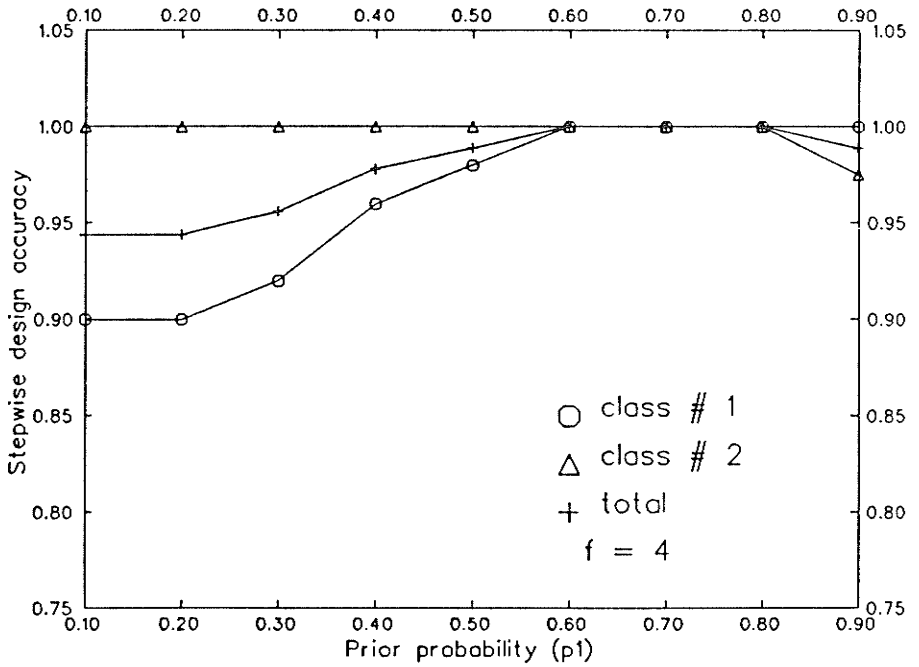


Fig. 5.6 Stepwise overvoltage predictor design (effect of prior probability).

The performance evaluation of the prediction scheme under consideration in the design stage is reported in Table 5.6. The system evaluation in the testing stage is also verified. From these results, a prediction accuracy of 100% in the design stage ($f = 4$) is achieved and a 97.8% evaluation accuracy can be confirmed.

In *conclusion*, the highest prediction accuracy of a particular class of patterns, for a two-class problem, can be achieved when its prior probability is greater than 0.5

The stepwise predictor *spectrum*, shown in Fig. 5.7, interprets the *prediction output* (index) and its change for the available patterns. It extracts weak, average, and strong patterns according to the prediction output value given to each pattern in the pattern set. In addition, it displays the range of the prediction system presented for this application.

5.5 Prediction of System Overvoltages

As discussed before, the main function of the overvoltage prediction system is mainly to predict overvoltage problems. This is of course in order to avoid any overvoltage stresses on the power system elements. Another purpose based on the predictor outputs (prediction indices) could be the *estimation* of system overvoltages. The estimation of these overvoltages is very valuable for the procedure of identification of overvoltage busses, which in turn will guide the process of taking corrective actions.

The procedure developed herein is mainly based on the *regression* of the system voltage deviations and the predictor outputs. Using the *least squares* method, *post −contingency* system voltages on all busses can be evaluated. In fact, the basic idea behind this algorithm is that the voltage deviations are a function of the predictor outputs. We assume that this relation is in a quadratic format:

$$\Delta V = a_0 + a_1 d + a_2 d^2 \tag{5.6}$$

where a_0, a_1& a_2 are constants for each bus. These constants are to be determined

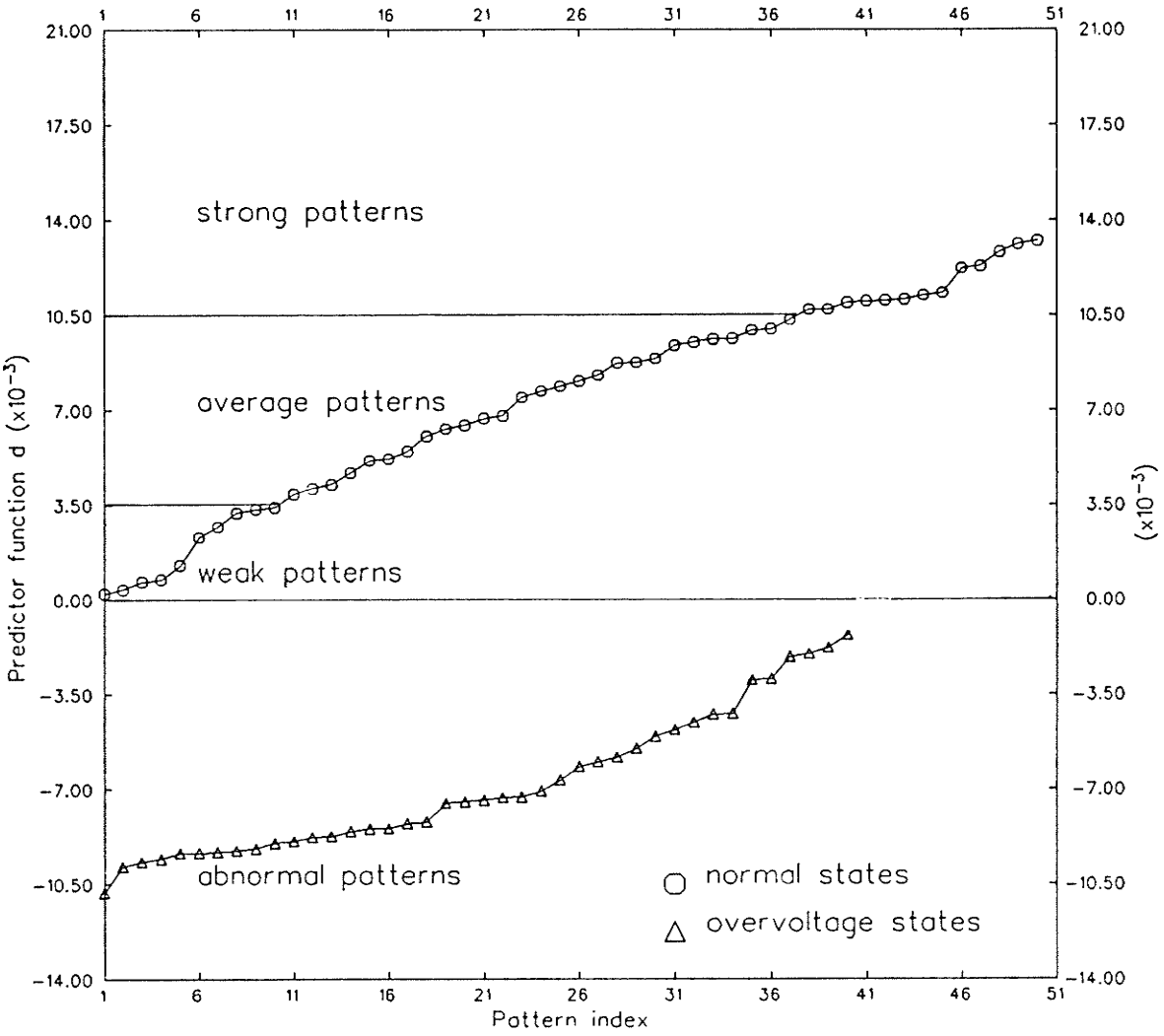


Fig. 5.7 Overvoltage Predictor spectrum.

using a least squares technique, taking into account that the voltage deviation ΔV at each bus and predictor output d are available for all cases in the pattern set. Then, knowing the *pre−contingency* conditions V_M, the post-contingency voltages V_P could be estimated.

The algorithm procedure starts by calculating the a's constants at each bus. For example, consider any particular bus and evaluate the voltage deviations and the corresponding predictor outputs for all patterns in the pattern set. Therefore, the following form can be obtained as:

$$V_{Pi} - V_{Mi} = a_0 + a_1 d_{(i)} + a_2 d_{(i)}^2 \tag{5.7}$$

$$i = 1, 2, ..., n$$

In matrix form this equation could be rewritten as:

$$[V_P - V_M] = [D] \,.\, [X] \tag{5.8}$$

where:

n = is the number of patterns in the voltage pattern set.

$$[V_P - V_M] = [V_{P1} - V_{M1}, V_{P2} - V_{M2}, ..., V_{Pn} - V_{Mn}]^T$$

$$[D] = \begin{bmatrix} 1 & d_{(1)} & d_{(1)}^2 \\ 1 & d_{(2)} & d_{(2)}^2 \\ .... & .... & .... \\ 1 & d_{(n)} & d_{(n)}^2 \end{bmatrix}$$

$$[X] = [a_0, a_1, a_2]^T$$

From eqn. (5.8), there are 3 unknowns and n observations with $n > 3$. Therefore, the solution of this equation can be derived using the least squares technique. The obtained solutions for the system under study and using the available states in the pattern set are given by:

$$[X]^T = [0.1754, -0.8289, -53.9512], \quad \text{for } \textit{voltage } V_1$$

$$= [0.2514, -1.2477, -77.5191] , \quad \text{for } \textit{voltage} \quad V_2$$

$$= [0.2590, -1.2115, -80.0605] , \quad \text{for } \textit{voltage} \quad V_3 \qquad (5.9)$$

Figure 5.8 shows the relation between the voltage deviation ΔV and the predictor output d at bus 1, bus 2, and bus 3 using the least squares solution. This solution has a maximum error of 1.5% for the case of bus 1, 2.0% for bus 2, and 2.1% for bus 3. These errors are all in the *pessimistic* direction, that is the predicted voltages are usually worse than the actual values. This is thought to be due to the predictor *mis-ranking* process of some states in the pattern set. In addition, the mean absolute error is equal to 0.71% in predicting voltage V_1, 0.98% in predicting voltage V_2 , and 1.02% in predicting voltage V_3.

In order to evaluate the developed algorithm, a set of 20 *historic* testing states of different conditions is used. Table 5.7 presents the results achieved using this algorithm. It can be seen that the results are promising and they indicate that this algorithm can be used effectively, and it has a potential for further *extension*. The maximum error recorded is about 0.3% in predicting voltage at bus 1 , about 0.5% in predicting voltage at bus 2, and about 0.7% in predicting voltage at bus 3.

The algorithm presented here has many advantages compared to the one presented by Mc Clelland [8]:

1- There is no need to store all patterns in the pattern set, therefore, there is no a storage problem.

2- The larger computation part is to be done off-line, hence, the computation time required for real time implementation is very small compared to that presented in Ref.[8].

3- The algorithm is simple and has no weighting coefficients to be calculated by trial and error as in [8].

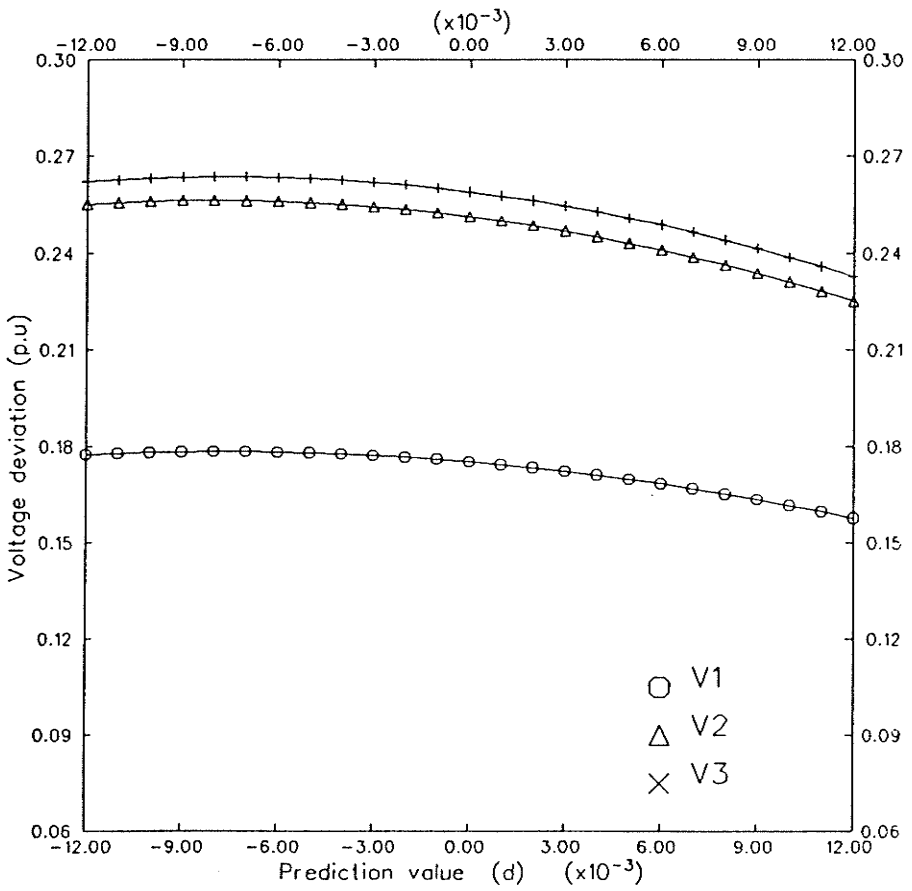


Fig. 5.8 Least square voltage solution.

Table 5.7 Prediction of system overvoltages

Case	V_M^1	V_P^1	V_P^{1*}	Error %
1	1.0100	1.1898	1.1883	-0.13
2	1.0100	1.1894	1.1884	-0.08
3	1.0100	1.1892	1.1883	-0.07
4	1.0100	1.1897	1.1883	-0.12
5	1.0100	1.1893	1.1884	-0.07
6	1.0100	1.1897	1.1884	-0.11
7	1.0100	1.1894	1.1885	-0.08
8	1.0100	1.1896	1.1881	-0.12
9	1.0100	1.1892	1.1883	-0.07
10	1.0100	1.1891	1.1882	-0.07
11	1.0100	1.1883	1.1879	-0.03
12	1.0100	1.1850	1.1866	0.13
13	1.0100	1.1838	1.1877	0.33
14	1.0100	1.1883	1.1883	0.0
15	1.0100	1.1890	1.1883	-0.06
16	1.0100	1.1890	1.1884	-0.05
17	1.0100	1.1895	1.1884	-0.09
18	1.0100	1.1889	1.1858	-0.26
19	1.0100	1.1893	1.1882	-0.09
20	1.0100	1.1893	1.1882	-0.09

V_M^1 : pre-contingency voltage at bus 1
V_P^1 : post-contingency voltage at bus 1
V_P^{1*}: predicted voltage at bus 1

Table 5.7 (continued)

Case	V_M^2	V_P^2	V_P^{2*}	Error %
1	1.0200	1.2774	1.2761	-0.11
2	1.0190	1.2760	1.2752	-0.06
3	1.0190	1.2753	1.2751	-0.01
4	1.0200	1.2774	1.2761	-0.10
5	1.0190	1.2757	1.2752	-0.04
6	1.0200	1.2773	1.2762	-0.09
7	1.0190	1.2762	1.2753	-0.07
8	1.0203	1.2773	1.2762	-0.09
9	1.0194	1.2758	1.2755	-0.02
10	1.0204	1.2767	1.2764	-0.03
11	1.0207	1.2758	1.2763	-0.04
12	1.0277	1.2783	1.2815	0.25
13	1.0251	1.2739	1.2804	0.51
14	1.0207	1.2757	1.2768	0.09
15	1.0196	1.2757	1.2757	0.0
16	1.0191	1.2753	1.2753	0.0
17	1.0190	1.2760	1.2753	-0.05
18	1.0205	1.2765	1.2733	-0.25
19	1.0204	1.2770	1.2764	-0.05
20	1.0197	1.2763	1.2756	-0.05

$V_{M_2}^2$: pre-contingency voltage at bus 2
$V_{P_2}^2$: post-contingency voltage at bus 2
V_P^{2*}: predicted voltage at bus 2

Table 5.7 (continued)

Case	V_M^3	V_P^3	V_P^{3*}	Error %
1	1.0180	1.2839	1.2811	-0.22
2	1.0170	1.2824	1.2803	-0.16
3	1.0170	1.2816	1.2802	-0.11
4	1.0180	1.2837	1.2812	-0.19
5	1.0170	1.2818	1.2803	-0.12
6	1.0180	1.2834	1.2812	-0.17
7	1.0170	1.2822	1.2804	-0.14
8	1.0186	1.2829	1.2815	-0.11
9	1.0176	1.2814	1.2807	-0.05
10	1.0189	1.2818	1.2819	0.01
11	1.0195	1.2796	1.2821	0.19
12	1.0270	1.2825	1.2875	0.39
13	1.0242	1.2779	1.2864	0.66
14	1.0195	1.2795	1.2826	0.24
15	1.0180	1.2810	1.2811	0.01
16	1.0173	1.2809	1.2806	-0.02
17	1.0178	1.2829	1.2811	-0.14
18	1.0190	1.2812	1.2784	-0.22
19	1.0188	1.2823	1.2817	-0.04
20	1.0179	1.2819	1.2808	-0.08

V_M^3 : pre-contingency voltage at bus 3
V_P^3 : post-contingency voltage at bus 3
V_P^{3*}: predicted voltage at bus 3

4- The developed algorithm considers different system configurations and loading conditions and can be extended to include all likely system contingencies.

5- The extended version can be installed in a real time computer system as an implementation of a fast dynamic voltage prediction, which could be automatically executed every 5 minutes.

5.6 Effect of Telemetering Communications Failure

The prediction system functions cannot be fulfilled without a supporting *communication system* (information system). The principle function of this communication system is to determine the operating conditions of the power system and provide the necessary information inputs to the prediction system. In addition, the communication system will provide the *link* between the power system and the system operator or *dispatcher*, giving him information on request.

The required information represents the measurements, and measurements include data obtained from the power system and from the environment. The environment also in turn represents all kind of external factors which affect the electrical conditions of the system. The problems of measurement and their hardware and communication aspects will have a *decisive* influence on the structure of the *control center* in general and the prediction system in particular [60].

The main purpose of the communication system is to provide the *information paths* from local stations to the control center and the *control paths* back from the control center. Therefore, the failure of any of these paths (channels) may have a large effect on the performance of the prediction system. In order to avoid such a problem, we must either reduce the probability of failure or improve the system performance despite the failure. The first approach, to be considered here, requires identification of those highly informative channels. For this case Table 5.8 presents the system performance due to the lack of each channel one at a time. From these results, the most powerful channels can be identified as well as the least

Table 5.8 Communications failure and the OV-Prediction system performance

Channel failed	Resubstitution design			Hold-out design		
	Class 1	Class 2	Total	Class 1	Class 2	Total
V_3	96.00	92.50	94.44	96.00	92.50	94.44
Q_{23}	0.0	100.0	44.44	0.0	100.0	44.44
Q_F	100.0	5.0	57.78	100.0	7.50	58.89
X_d	96.00	95.00	95.56	96.0	95.00	95.56

(a) VSM features

Channel failed	Resubstitution design			Hold-out design		
	Class 1	Class 2	Total	Class 1	Class 2	Total
X_d	94.00	95.00	94.44	90.00	95.00	92.22
Q_F	100.0	5.00	57.78	100.0	7.50	58.89
Q_g	0.0	100.0	44.44	0.0	100.0	44.44
I_{fd}	94.00	90.00	92.22	92.0	95.00	93.33

(b) Intuitive features

Table 5.8 (continued)

Channel failed	Stepwise design		
	Class 1	Class 2	Total
V_3	100.0	0.0	55.56
δ_2	100.0	0.0	55.56
I_{fd}	100.0	0.0	55.56
Q_{23}	100.0	95.00	97.78

(c) Stepwise features

significant channels.

5.7 Reduction of Overvoltage Prediction System Structure

As discussed in section 4.6, an attractive algorithm is developed in order to reduce or simplify the prediction system structure. Since the main purpose of the communication system is to provide the prediction system with the required information, it is more economical to use the minimum number of required channels.

The developed algorithm is applied in this case to the dynamic overvoltage prediction system and the results are reported in Table 5.9. These results confirm the possibility of reducing the number of channels (predictor features) by one channel. The overall prediction system performance should not be significantly affected by this process.

5.8 Development of Overvoltage Corrective Action

Using the developed *corrective* algorithm discussed in section 4.7, a *preventive* measure or improving measure can be achieved. The dynamic overvoltage problem, the problem of concern in this chapter, has a strong correlation with the reactive sources in the power network. Therefore, the adjustment of the reactive power sources is the control feature to be considered in this application.

Considering the linear prediction scheme, see section 4.7.1, designed for this application and following the same sequences of the corrective algorithm, a sample of results, Table 5.10, can be obtained. It can be seen that only one iteration step is needed in order to develop the required corrections.

Table 5.9 Reduction of OV-prediction system structure

No. of features	Resubs. design		Hold-out design	
	VSM	Intuitive	VSM	Intuitive
4	95.56	94.44	95.56	93.33
3	95.56	94.44	95.56	93.33
2	76.67	73.33	44.44	62.22

Table 5.10 Overvoltage corrective actions

Initial predic. index $d^{(0)}$	Initial filter MVARs $Q_F^{(0)}$	Required filter MVARs $Q_F^{(1)}$	Final predic. index $d^{(1)}$	APPROVED OPERATOR ACTION $Q_F^{(1)} - Q_F^{(0)}$
-1.2489	100.00	52.94	0.0735	-47.06
-0.5304	150.00	130.01	0.0363	-19.99
-1.3036	250.00	200.87	0.1003	-49.13
-0.6469	300.00	275.62	0.0507	-24.38
-2.8646	450.00	342.05	0.2360	-107.95
-3.6485	550.00	412.51	0.3069	-137.49

5.9 Conclusions

A new application based on pattern recognition techniques is discussed in this chapter. A successful overvoltage prediction system based on these techniques is designed. It predicts overvoltage conditions in a very short time so that power system dispatcher can decide and take actions as required in the appropriate time in order to improve the system conditions.

The results provided in this chapter confirm a very highly performed prediction system. The VSM features method was very effective compared to the intuitive method for this application of concern. A perfect prediction of overvoltage problems was achieved using the stepwise hyperplane design. An overvoltage correction is developed using the corrective algorithm.

A general conclusion can be drawn: pattern recognition techniques can be applied to the prediction of power system dynamic overvoltage operating problems.

CHAPTER 6

MULTI-CLASS PREDICTION SYSTEM DESIGN

6.1 Introduction

Power system dynamic overvoltages as well as generator self-excitation have been treated so far as a two-class problem in chapters 4 and 5. This chapter concerns the *multi−class* approach i.e. the class prediction of a given pattern among patterns belonging to more than two classes e.g. normal, dynamic overvoltage, and self-excitation patterns due to power system load rejections.

The design of the required prediction scheme, called a *linear prediction machine*, is described. This approach has never appeared in the literature so far. It offers more *capabilities* and at the same time it has some *limitations*. One of its main advantages could be the reduction in the telemetering communications compared to the two-class approach. Moreover, this scheme could be much more efficient and of less computation time (during the design process). However, one of the disadvantages is the reduction in performance when a telemetering channel fails, in comparison to the two-class approach. The estimation of prior probability could make the design of multi-class system a difficult task.

This algorithm could be useful and it has some variety of applications, e.g. it could be used for the on-line assessment of both steady state and transient security problems at the same time [2,8].

6.2 Configuration of Multi-Class System

As discussed in chapter 2, the multi-class prediction system configuration is very similar to that of the two-class case, except that the predictor design scheme is different in this case.

As stated before, power system load rejection may result in a generator self-excitation and/or a system dynamic overvoltage. Therefore, a situation of similar conditions could be treated as a three-class recognition problem e.g. normal class, overvoltage class & self-excitation class, as explained in Fig. 6.1.

In order to come up with the required multi-class prediction scheme, it is required to follow the same design phases as described in chapter 2.

Regarding the pattern generation phase, all load rejection patterns (124 of them) are classified using the employed dynamic simulation model into normal patterns (50), overvoltage patterns (40) & self-excitation patterns (34).

On the other hand, the feature vector is identified using the stepwise algorithm. The features obtained in this concern are:

$$V_3, \delta_3, Q_F, \delta_2, Q_{23}, I_{fd}, P_{23}, X_d, P_g$$

6.3 Multi-class Predictor Design and Evaluation

Following the same procedure as that for the stepwise algorithm, the multi-class design scheme can be derived. The design process in this case is a function of four variables: f , $p_1, p_2 \& p_3$ where

f is the required number of features

$p_1, p_2 \& p_3$ are the prior probability of class 1 , class 2 and class 3 respectively.

Utilizing a numerical experimentation approach, the optimal design scheme can be defined with the main *objective* of achieving the maximum overall prediction accuracy. Figure 6.2 presents the results obtained under these conditions. The prediction accuracy obtained when p_1 and f are changing and assuming that $p_2 = p_3$, are given by Fig. 6.2a. Similarly, the prediction results obtained considering the change in f and p_2, f and p_3 are shown in Fig. 6.2b and in Fig. 6.2c respectively.

The results provided during the design have come to the optimal number of features and the optimal combination of prior probabilities, these are:

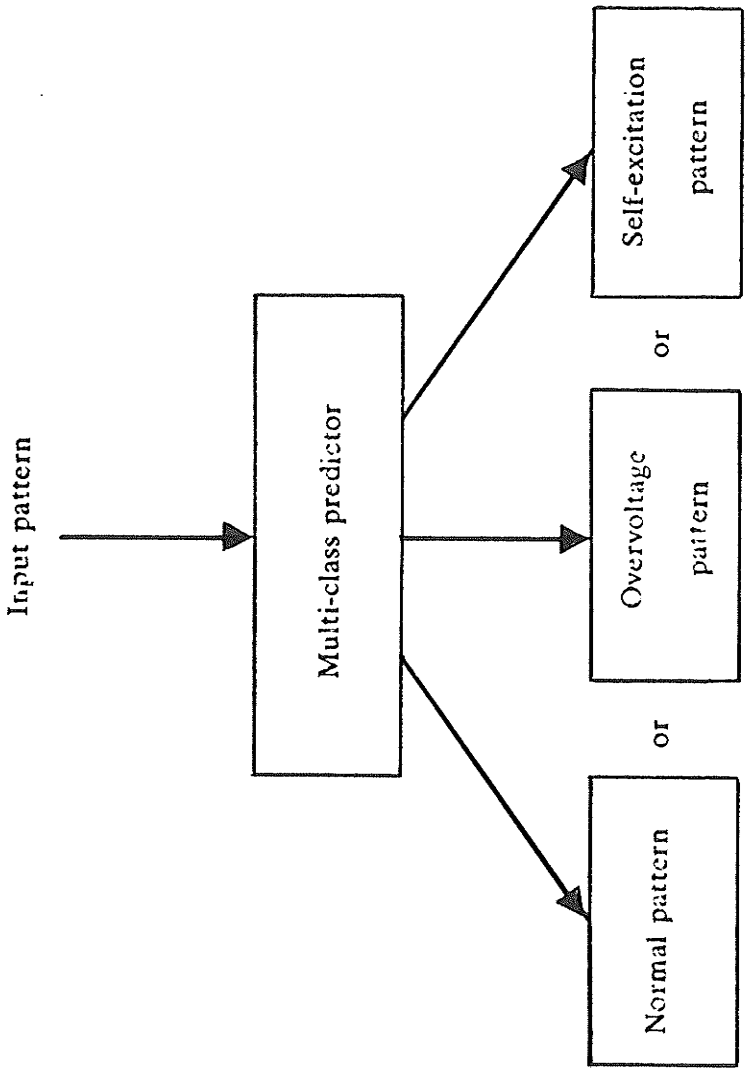


Fig. 6.1 Multi-class predictor outputs.

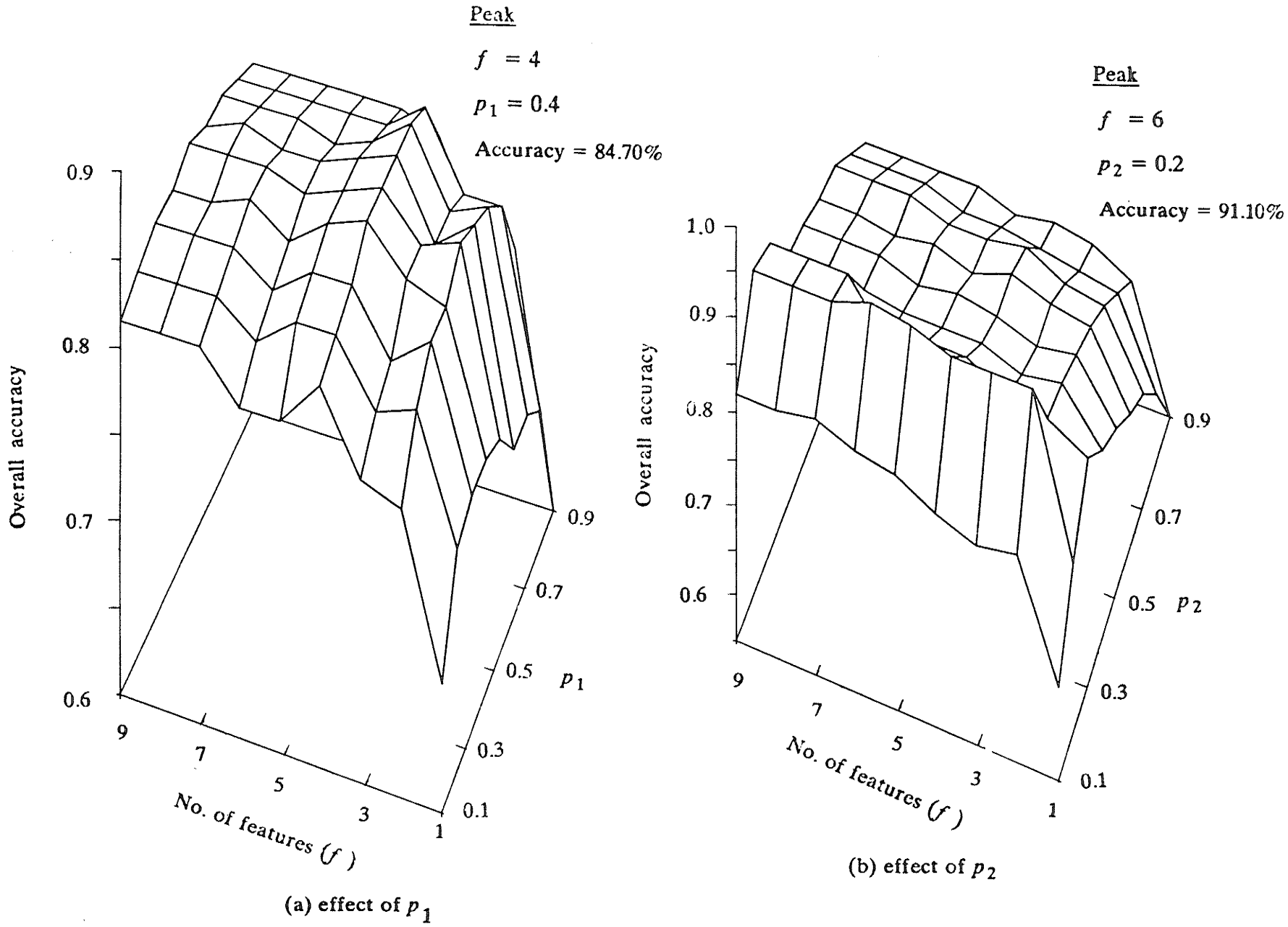
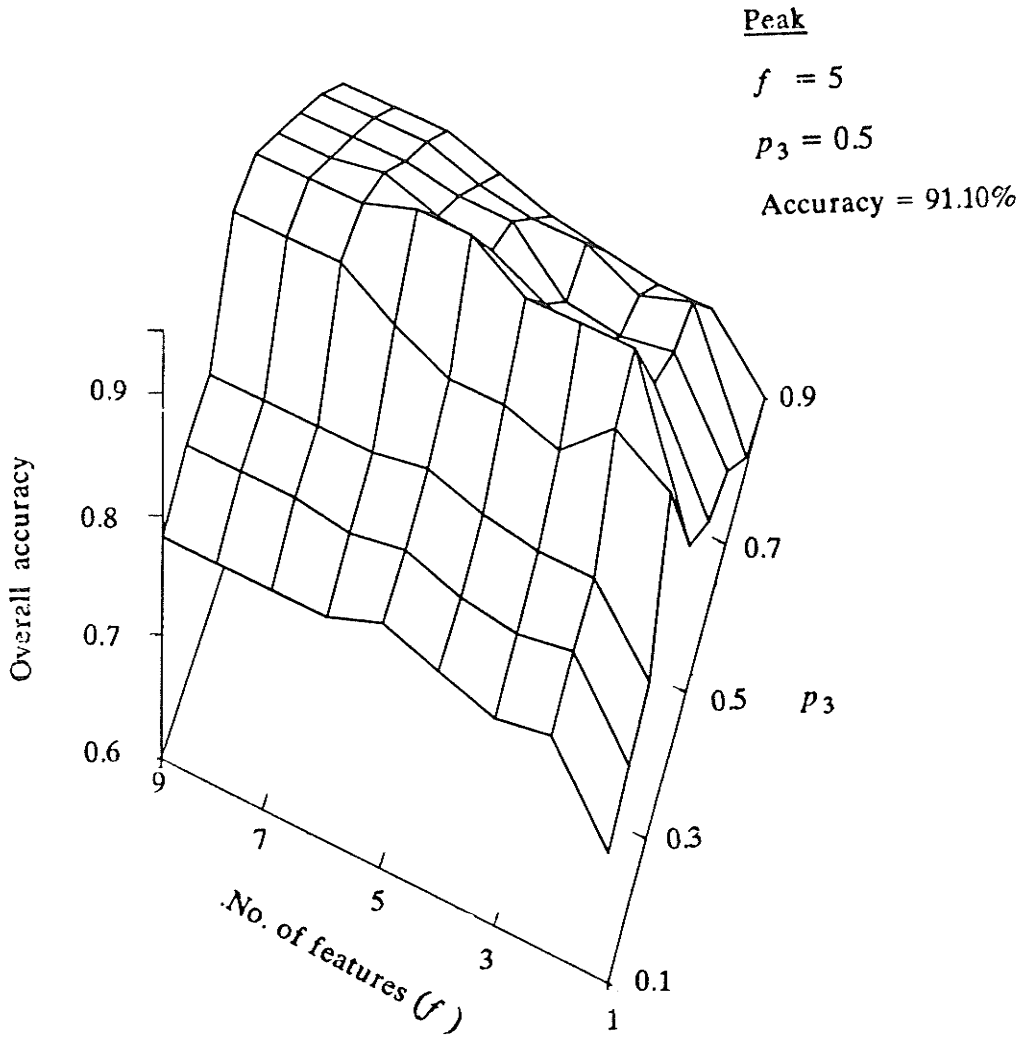


Fig. 6.2 Multi-class predictor design.



(c) effect of p_3

Fig. 6.2 (continued).

$f = 5, p_1 = p_2 = 0.25, p_3 = 0.50$. The predictor maximum accuracy reached is 91.1%.

In order to investigate the effect of number of features and the prior probability on the three-class prediction accuracy, Fig. 6.3 is provided. In Fig. 6.3a the number of features is fixed at 5. The insignificant effect of prior probability p_3 on the prediction of class 1 (normal patterns) can be verified, while the high *dependence* of class 2 (overvoltage patterns) and class 3 (self-excitation patterns) on p_3 can also be observed. In addition, Fig. 6.3b explains the relation between the number of features and the prediction accuracy of each class, for the best probability conditions.

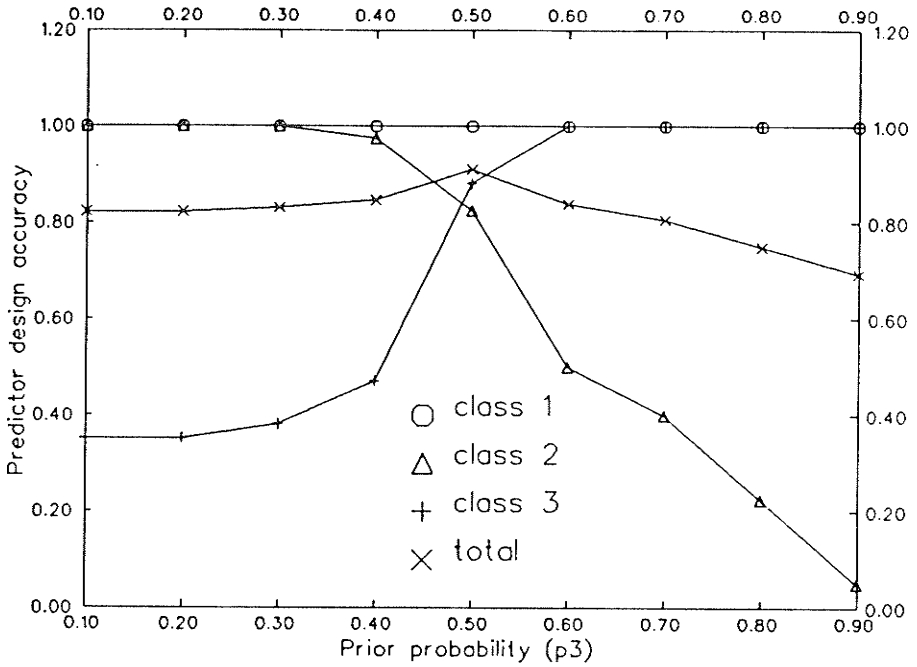
Regarding the evaluation process for the prediction system of concern , the leave-one-out method is used. The evaluation results are important in providing us with the degree of *conf idence* by which we can rely on the achieved design scheme, see Table 6.1 .

It must be emphasized here that the total CPU computation (design) time for a two-class problem is about 0.86 sec (AMDAHL), while it is about 0.96 sec for a three-class case.

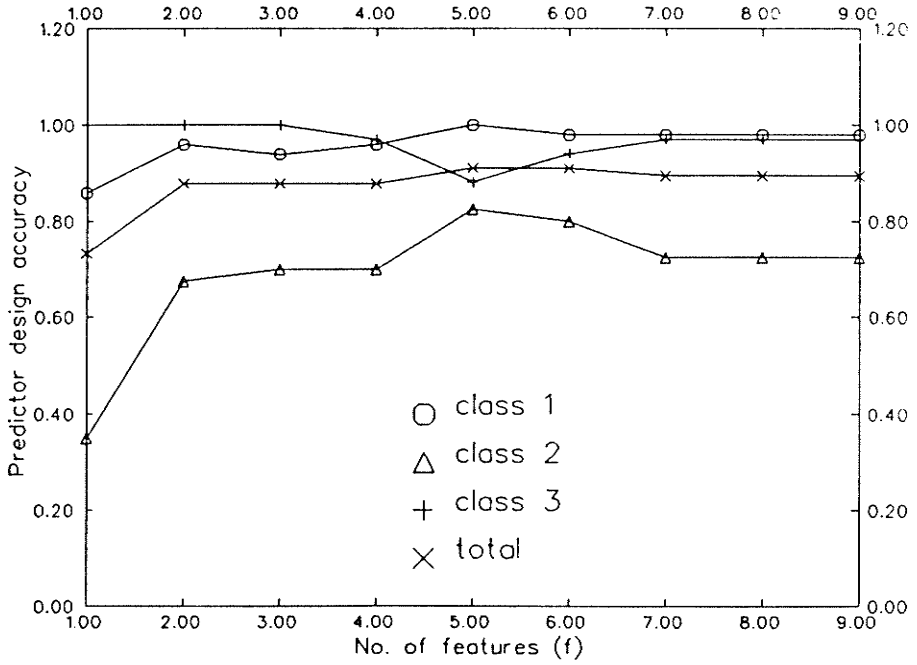
6.4 Multi-class System Performance and Effect of Communications Failure

The performance of a multi-class prediction system is very dependent on the communication channels. This is due to the fact that these channels are involved in the process of recognition of more than two classes at the same time. Therefore, the multi-class system performance will be highly affected by the degree of *availability* of communications.

Table 6.2 shows the behavior of the prediction system of concern due to the failure of telemetering channels. It can be seen that the prediction of class 1 is very sensitive to δ_3 channel, while the prediction of class 2 is highly dependent on Q_F, V_3& δ_2 channels. In addition, the failure of Q_F channel has no significant



(a) effect of prior probability ($f = 5$)



(b) effect of number of features ($p_1 = p_2 = 0.25$, $p_3 = 0.5$)

Fig. 6.3 Prediction accuracy of multi-class scheme.

Table 6.1 Evaluation of multi-class system performance

No. of features	Design accuracy			
	Class 1	Class 2	Class 3	Total
1	86.0	35.0	100.0	73.4
2	96.0	67.5	100.0	87.9
3	94.0	70.0	100.0	87.9
4	96.0	70.0	97.1	87.9
5	100.0	82.5	88.2	91.1

Table 6.1 (continued)

No. of features	Testing accuracy			
	Class 1	Class 2	Class 3	Total
1	86.0	35.0	100.0	73.4
2	96.0	65.0	100.0	87.1
3	94.0	62.5	85.3	81.5
4	96.0	57.5	85.3	80.6
5	98.0	70.0	61.8	79.0

Table 6.2 Multi-class prediction performance due to communication failure

Channel failed	Prediction accuracy			
	Class 1	Class 2	Class 3	Total
V_3	100.00	0.00	0.00	40.32
δ_3	0.00	100.00	0.00	32.26
Q_F	100.00	17.50	94.12	71.77
δ_2	100.00	0.00	0.00	40.32
Q_{23}	78.00	97.50	55.88	78.23

influence on the prediction process of class 3 .

The results provided by Table 6.2 can lead us to a general conclusion: the overall performance of the multi-class prediction system developed has dropped to a low level (78.23%) due to a failure in communications i.e. the system performance is highly degraded. Therefore, it is considered as one of the limitations concerning the application of multi-class approach.

6.5 Multi-class Canonical Transformation

In this section a transformation algorithm [49] is applied in order to derive a set of *canonical* functions (CF). For a multi-class system (c classes), there are $c-1$ canonical functions e.g. for a three-class system there two canonical functions. These functions are to be determined for each pattern of the whole pattern set.

The algorithm involved here is based on the solution of a generalized eigen-value problem:

$$T(X)\, v_i = \lambda_i\, W(X)\, v_i$$

$$v_i{}^T\, W(X)\, v_i = \delta_{ij} \qquad (6.1)$$

$$i = 1, 2, ..., f$$

where:

$T(X)$, $W(X)$: are the total and within groups sum of cross product matrices respectively

$[\lambda_1, \lambda_2, ..., \lambda_f]$: are the eigen-values in a descending order

$[v_1, v_2, ..., v_f]$: are the corresponding eigen-vectors

δ_{ij} : is the Kronecker delta function

$X = [x_1, x_2, ..., x_f]^T$: is the feature vector

Solving Eq. (6.1) then using the v_i solution to define the canonical functions as:

$$CF_i(X) = X\, v_i \;\; ; \;\; i = 1, 2, ..., c-1 \qquad (6.2)$$

The plot shown in Fig. 6.4, provides the scattering of the three classes of concern. This plot is obtained from the two canonical functions (variables) determined for each pattern. The letters A, B & C are to indicate patterns from class 1, class 2 & class 3 respectively. On the other hand, the numbers 1, 2 & 3 are to represent each group mean , while the asterisks denote overlap between different groups.

6.6 Conclusions

A multi-class prediction system is designed and applied to a three-class prediction problem. The multi-class approach has many advantages and it posses some limitations. One of the advantages is the communication reduction which may simplify the structure of the communication system required. On the other hand, the design of an optimal scheme is not an easy task and it has some difficulty. Also, the required classification time for the multi-class approach may be longer or equal to that required for the two-class approach.

In general, for power system applications the two-class approach has proved to be applicable and for the purpose of prediction it is highly recommended.

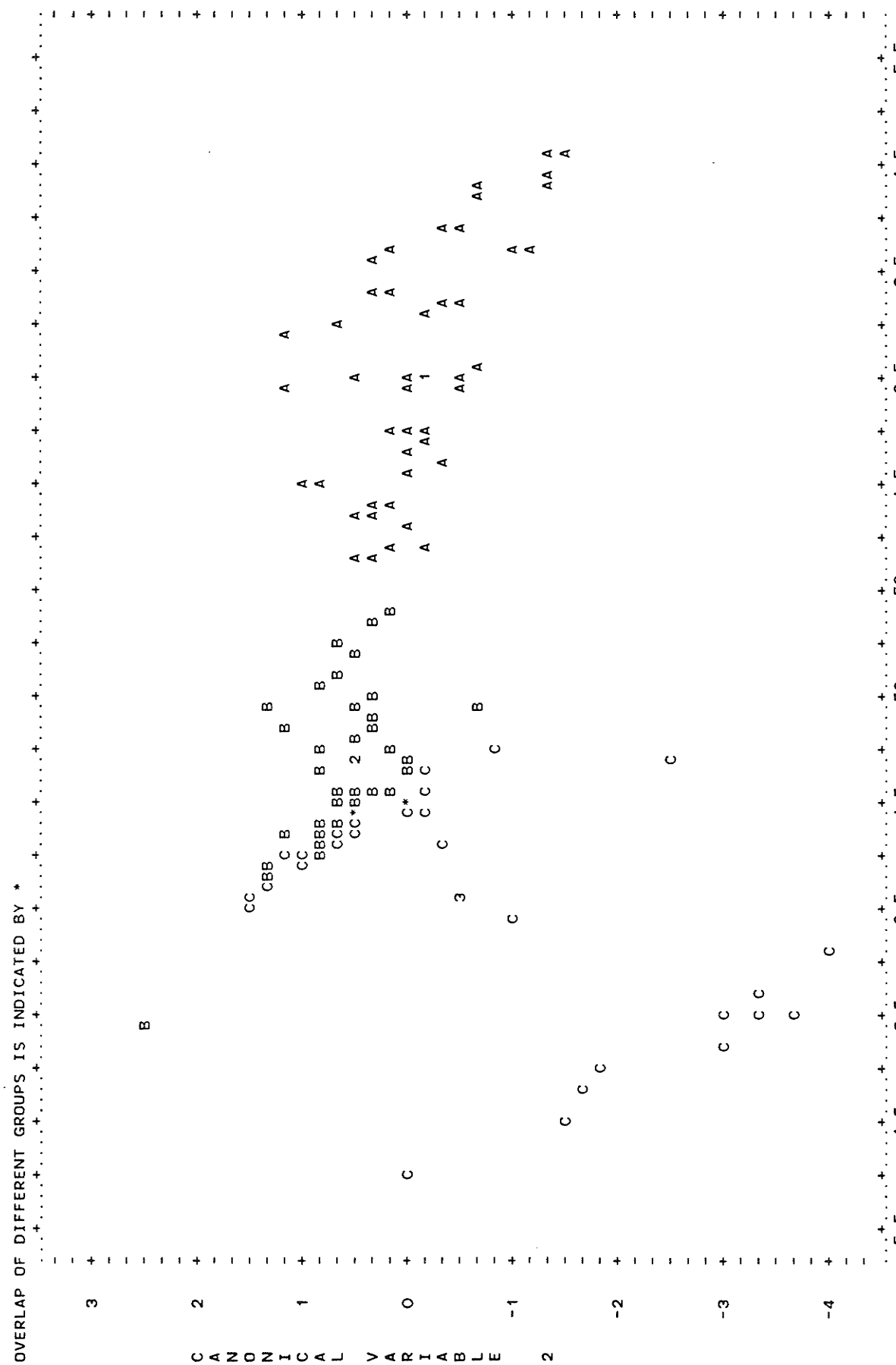


Fig. 6.4 Multi-class scattering diagram.

CHAPTER 7

VOLTAGE CONTINGENCY ANALYSIS DUE TO SYSTEM OUTAGES

7.1 Introduction

In the last few chapters (4,5 & 6), we were concerned with the application of pattern recognition techniques to the prediction of load rejection self-excitation conditions and dynamic overvoltages as well. Therefore, this application was for a *dynamic* power system operating problems. This chapter deals with the normal, (*steady state*), operating condition. For this condition power system on-line *contingency analysis* is of major interest (for system planning and operation) at *control centers* .

A new pattern recognition based algorithm for power system voltage contingency analysis, is introduced in this chapter. The algorithm is based on the design of a *contingency selector* (CS), or *discriminant hyperplane* in pattern recognition terminology. The design of such a CS is based on a training knowledge. Using the CS a *screening* and *ranking* of voltage problem contingencies can be obtained. Further processing on the finalized list of contingencies can be performed in order to develop, if called for, the required corrective actions.

7.2 Voltage Contingency Analysis and Methodologies

Since modern power systems are getting large and more extensively interconnected, the task of contingency analysis and control is becoming difficult for the system dispatcher. For this reason, on-line contingency analysis is gaining much attention, and rightfully so, as new energy control centers are developed and applied [22-26].

In the methodology of contingency analysis, overload and voltage contingency analysis have been separately considered [27-37]. The overload contingency analysis (P/δ) has been efficiently and reliably developed, and is being implemented in new control centers [61]. On the other hand, voltage contingency analysis (Q/V) is a very difficult problem due to its nonlinear behavior. An efficient and reliable algorithm has not been developed yet [62].

The direct approach to contingency analysis means that at a given system condition, AC power flows for all likely contingencies (hundreds), single and multiple, is to be performed for every assigned period, usually in the range of 2 - 3 minutes. Therefore, a large computation burden is required in order to analyze all these contingencies, which is difficult to overcome. As a result, a contingency selection methodology was introduced. The idea behind these approaches is to select only those contingencies considered critical and then perform AC power flows for these cases (see Fig. 7.1).

Several papers have proposed CS techniques. However, CS techniques can be grouped into two different groups: a *sensitivity* analysis based group [28-37,61], and a *fast load flow* based group [27]. Sensitivity techniques are based on performing the base case AC load flow and using the sensitivity matrices to predict system changes due to outages. On the other hand, the fast load flow group is based on the evaluation of DC or fast decoupled power flow for each assigned contingency (outage); then by use of a predefined performance index, a short list of critical contingencies can be produced for further analysis. It must be noted here that, sensitivity analysis is less efficient than the fast load flow algorithms but more reliable [62].

Currently CS algorithms are judged in terms of their *Capture Rate* (CR) and the *False-Alarm Rate* (FR), which are defined by [31]:

$$CR = 1 - NM / NNC , \quad FR = NFA / NC \tag{7.1}$$

where NM is the number of misses out of NNC the number of noncritical

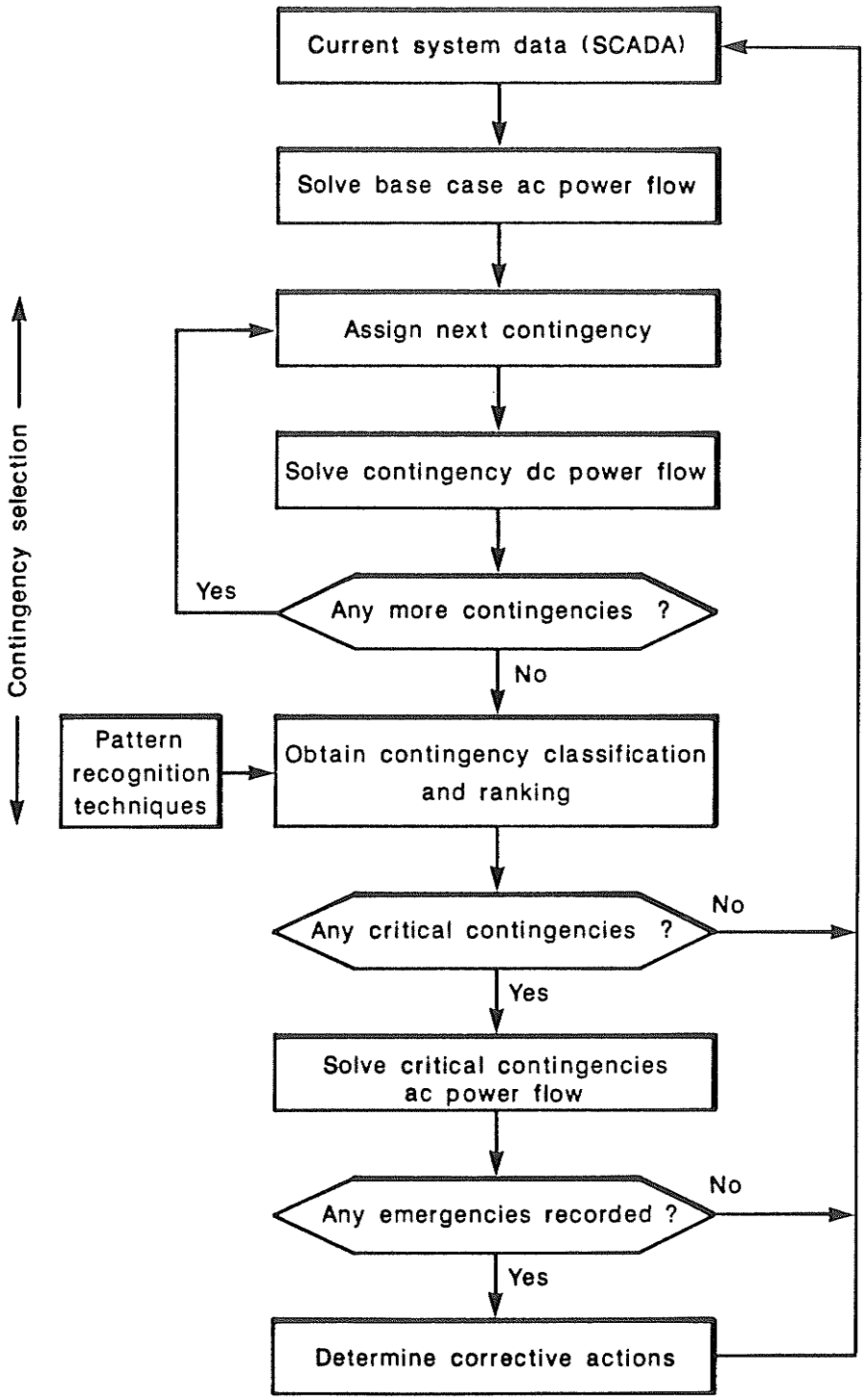


Fig. 7.1 Methodology of contingency selection and analysis.

contingencies and NFA is the number of false alarms out of NC the number of critical contingencies. It must be noted that $Miss$ means classifying a critical contingency as being noncritical and $False-Alarm$ means classifying a noncritical contingency as being critical.

If there are no misses and false-alarms, then $CR=1, FR=0$, and the CS algorithm correctly predicts the impact of every contingency. Thus the effectiveness of the CS algorithm should be evaluated in terms of both CR & FR.

7.2.1 Nara Selection Algorithm [27]

This algorithm is a new concept in formulating a performance index for voltage contingency selection. It is a DC load flow based algorithm. The index is a second order vector norm in the voltage space. Two types of voltage limits are defined: *alarm limits* and *security limits*. Contingency cases of indices greater than 1.0 are considered insecure cases.

This algorithm has many advantages compared to others [28-37]. These are:

(1) It considers the magnitude of voltage violation as: (a) maximum magnitude ; (b) combined magnitude; and (c) average of voltage violations.

(2) It takes into account the number of voltage violations.

(3) It is a function of the system load level.

(4) It is a function of the distance between voltage violations and their limits.

(6) It is based on the relative importance of each voltage violation.

This algorithm can be explained by defining the following terms (see Fig. 7.2):

$V^{(i)}$: voltage magnitude at bus i

$FL^{(i)}(L), FU^{(i)}(L)$: lower and upper voltage alarm limits at load level L

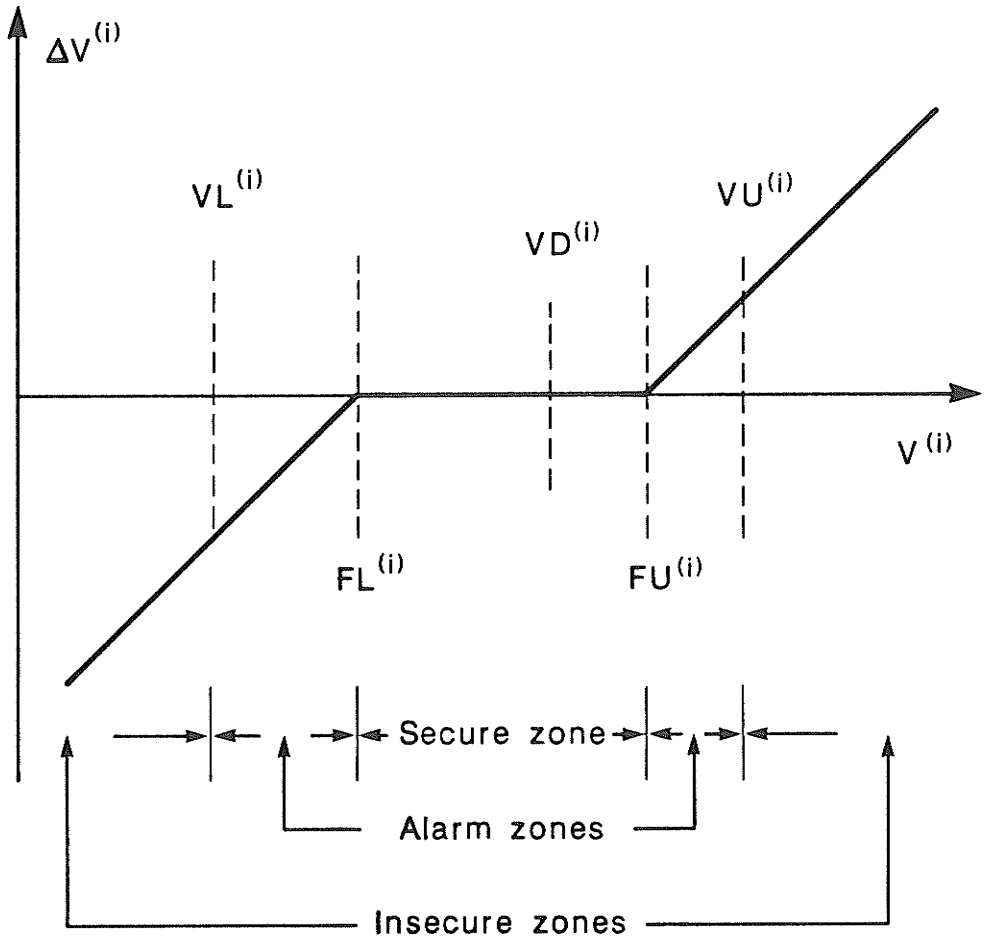


Fig. 7.2 Secure, alarm, and insecure voltage zones.

$VL^{(i)}(L)$, $VU^{(i)}(L)$: lower and upper voltage security limits at load level L

$VD^{(i)}(L)$: desired voltage at bus i and load level L

$\Delta V^{(i)}$: voltage deviation beyond alarm limits

NB : number of system busses.

The voltage deviation $\Delta V^{(i)}$ can be defined as:

$$\begin{aligned}\Delta V^{(i)} &= V^{(i)} - FU^{(i)}(L) \quad ; \text{ if } V^{(i)} > FU^{(i)}(L) \\ &= FL^{(i)}(L) - V^{(i)} \quad ; \text{ if } V^{(i)} < FL^{(i)}(L) \\ &= 0 \qquad\qquad ; \text{ if } FL^{(i)}(L) \le V^{(i)} \le FU^{(i)}(L)\end{aligned} \tag{7.2}$$

Rewriting $\Delta V^{(i)}$ into two terms DU , DL (corresponding to two limits) as defined by

$$\begin{aligned}DU^{(i)}(L) &= [V^{(i)} - FU^{(i)}(L)] \quad ; \text{if } V^{(i)} > FU^{(i)}(L) \\ &= 0 \qquad\qquad ; \text{if } V^{(i)} \le FU^{(i)}(L)\end{aligned} \tag{7.3}$$

and

$$\begin{aligned}DL^{(i)}(L) &= [FL^{(i)}(L) - V^{(i)}] \quad ; \text{if } V^{(i)} < FL^{(i)}(L) \\ &= 0 \qquad\qquad ; \text{if } V^{(i)} \ge FL^{(i)}(L)\end{aligned} \tag{7.4}$$

The performance index PIV_1 can then be defined as:

$$PIV_1 = \left[\sum_{i=1}^{NB} \left(AU^{(i)}\, DU^{(i)}(L) \right)^4 + \sum_{i=1}^{NB} \left(AL^{(i)}\, DL^{(i)}(L) \right)^4 \right]^{1/4} \tag{7.5}$$

which describes a *hyper −ellipse* in the voltage space. The terms $AU\&AL$ are defined by

$$\begin{aligned}AU^{(i)} &= 1.0 \,/\, [VU^{(i)}(L) - FU^{(i)}(L)] \\ AL^{(i)} &= 1.0 \,/\, [FL^{(i)}(L) - VL^{(i)}(L)]\end{aligned} \tag{7.6}$$

From the definition of the performance index, PIV_1, the system voltage conditions are classified as:

secure state if $PIV_1 = 0$

alarm state if $0 < PIV_1 \leq 1$

insecure state if $PIV_1 > 1$

The algorithm presents an accurate *screening* and *ranking* voltage index. At the same time, it has some disadvantages: 1- the computation time needed is larger than the sensitivity based approach for on-line applications; 2- it generates bonafide as well as pseudo voltage violations.

7.2.2 Ejebe Selection Algorithm [28]

This algorithm is a *sensitivity* based approach. It presents a methodology developed for ranking line and generator outages according to the severity of their effects on bus voltage profile. It employs *Tellegen*'s theorem to derive the sensitivities of a performance index due to system outages. By ordering these sensitivities a contingency ranking can be obtained. Full AC load flow must be carried out for critical contingencies. The set of critical contingencies is determined by simply running load flow from one contingency to another starting at the top of the list and screening out cases that do not give problems.

The voltage constraints at the load busses are usually expressed in terms of a high and low limit. The high limit is imposed by the maximum voltage value of the system, and the low limit is a value below which the load can no longer be supplied with the required quality of supply. In light of these constraints, the voltage performance index must be defined in such a way to reflect the severity of out-of-limit voltage values. Therefore, the performance index defined by this algorithm is chosen to quantify system deficiency due to out-of-limit bus voltages. It is defined by:

$$PIV_2 = \sum_{i=1}^{NB} 0.5 \left[\left(V^{(i)} - VD^{(i)}\right) / \Delta VL^{(i)} \right]^2 \qquad (7.7)$$

where:

$VD^{(i)}$: desired (specified) voltage magnitude at bus i

$\Delta VL^{(i)}$: voltage deviation limit at bus i

The voltage deviation $\Delta VL^{(i)}$ represents the voltage limit. Outside this limit voltages are unacceptable and yield a high value of index PIV_2. On the other hand, when voltage deviations are within $\Delta VL^{(i)}$ the system voltages will not be a problem and the index in this case will be small. Generally speaking, this voltage index measures the severity of the out-of-limit bus voltages, for a set of contingencies, it provides a direct means of ranking the relative severity of these contingencies on the system voltage profile.

This ranking algorithm has the advantage of reducing the computation time required for on-line applications purpose, compared to the Nara approach. On the other hand, it has some drawbacks: 1- it is not an efficient ranking algorithm; 2- the screening criterion involved needs to be enhanced.

7.3 New Pattern Recognition Based Voltage Contingency Analysis

7.3.1 Introduction

This section is to present a CS algorithm based on pattern recognition techniques, to be used for voltage contingency analysis. The automatic contingency analysis problem is concerned with the developing of an efficient algorithm which can rank and classify contingencies in terms of their impact on system performance.

None of the available contingency selection methods can recognize all critical contingencies. Some severe contingencies may be omitted and some that are not severe may be included. In many cases, an increase in accuracy can be obtained only at the price of a decrease in execution speed. Moreover, some of the selection algorithms [27,28] have the desirable property that each bus voltage has its own reference voltage. Unfortunately, this feature results in computational difficulties which

prevent an efficient calculation of the change in index when a circuit is dropped, for sensitivity methods.

In the last few sections we have discussed two different approaches currently used to come up with the CS algorithm. This section is to investigate a new approach to the definition of CS (performance index). The basic idea behind this approach is the design of a discriminant hyperplane which can classify contingencies into critical and noncritical. Using this hyperplane a measure of contingency impact on the system can be determined.

7.3.2 Discriminant Hyperplane Selection Algorithm

The application of *decision theory* to the contingency selection, which is a binary decision problem, has been reported [31,36]. On the other hand, the application of pattern recognition techniques has been recommended by Fischl [36], and as far as we know this application has never been appeared on the literature so far.

Contingency selection is to select critical contingencies among all different contingencies. In other words, it is required to recognize critical contingencies from noncritical ones. Therefore, the pattern recognition discriminant hyperplane can be used as a performance index to classify those critical and noncritical contingencies. In other words, using the discriminant hyperplane a screening process between critical and noncritical cases can be achieved.

Moreover, with respect to the ranking criterion, the discriminant hyperplane represents the "boundary" surface between different kinds of contingencies. By calculating the *separation* (distance) of each contingency from this surface a ranked list of contingencies can be derived. As discussed in Chapter 2, it was concluded that the value of the discriminant hyperplane gives a measure of the separation for each contingency. As a conclusion, the function of a performance index or a CS algorithm is exactly the same as the function of the discriminant hyperplane from the point of view of screening and ranking of contingencies.

The performance index format using the discriminant hyperplane is given by:

$$PIV_3 = \sum_{i=1}^{f} w_i \, \Delta VF^{(i)} + w_0$$

$$= W^T \, \Delta VF + w_0 \tag{7.8}$$

where:

$W^T = [w_1, w_2, ..., w_f]$ is the weighting vector,

$\Delta VF = [\Delta VF^{(1)}, \Delta VF^{(2)}, ..., \Delta VF^{(f)}]^T$ is the voltage deviation feature vector,

$\Delta VP = [\Delta VP^{(1)}, \Delta VP^{(2)}, ..., \Delta VP^{(NB)}]^T$ is the voltage deviation vector,

$\Delta VP^{(i)} = 100.0 \, ABS \, [V^{(i)} - VD^{(i)}]$ is the absolute voltage deviation in per cent at bus i ,

f is the number of voltage features.

The voltage features ΔVF are to be identified from ΔVP using the stepwise features algorithm. Also, the weighting vector W and threshold w_0 are to be determined using the stepwise discriminant method, discussed in Chapter 2.

As discussed in Chapter 2, the stepwise hyperplane design is a function of two parameters f & PR (prior probability ratio). In order to come up with the optimal design, the two parameters must be selected in such a way that the screening and ranking processes should be efficiently achieved. The parameter f affects the ranking and screening processes, while PR affects only the screening accuracy.

By changing PR, the threshold w_0 changes and the hyperplane moves into another parallel surface. Therefore, the ranking process will not change. Assuming an equal prior probability criteria for all contingencies (PR =1.0), the whole design will rely only on the number of features f . Then, using a numerical experimentation approach, f can be selected such that the ranking process can be efficiently achieved. Finally, by controlling PR, an optimal design efficient in ranking and screening, can be determined.

7.3.2 Mahalanobis Distance Selection Algorithm

In this section another approach is discussed. This approach is really based on the determination of separability measure between different contingencies and the group mean of noncritical contingencies, see Fig. 7.3 . It can be seen that the larger the separation is, the more critical the contingency will be, and vice versa.

The statistical separability measure selected here is called the *Mahalanobis* $-D^2$ (MH-D*D) in pattern recognition theory. MH-D*D is a measure of the distance between each contingency and each group mean. Therefore, a ranking process could be developed using the MH-D*D method. On the other hand, this method does not provide a screening process and it has been used here only to provide a ranking process. However, the screening criteria could be achieved by monitoring of bus voltages.

The performance index in this case is defined to be the MH-D*D, which can be given as (for case j in group i from the mean of group k) [48,49]:

$$PIV_4 = (N - 2) \sum_{r=1}^{f} \sum_{s=1}^{f} (\Delta VF_{ijr} - \Delta \overline{VF}_{kr}) a_{rs} (\Delta VF_{ijs} - \Delta \overline{VF}_{ks}) \quad (7.9)$$

$$i \ \& \ k = 1, 2 \quad ; j = 1, 2, ..., N_i$$

where

N_i is the number of cases in group i ,

ΔVF_{ijr} is the value of voltage deviation feature r in the case j of the group i ,

$\Delta \overline{VF}_{ir}$ is the value of group i mean for voltage deviation feature r ,

$A = (a_{rs})$ is the covariance matrix, see Appendix (A),

N is the total number of cases in all groups.

The voltage deviation features ΔVF are to be selected again using the stepwise selection algorithm, see Chapter 2 for more detail. The number of features f is to be selected using a numerical experimentation method. This algorithm could be considered relatively an efficient approach.

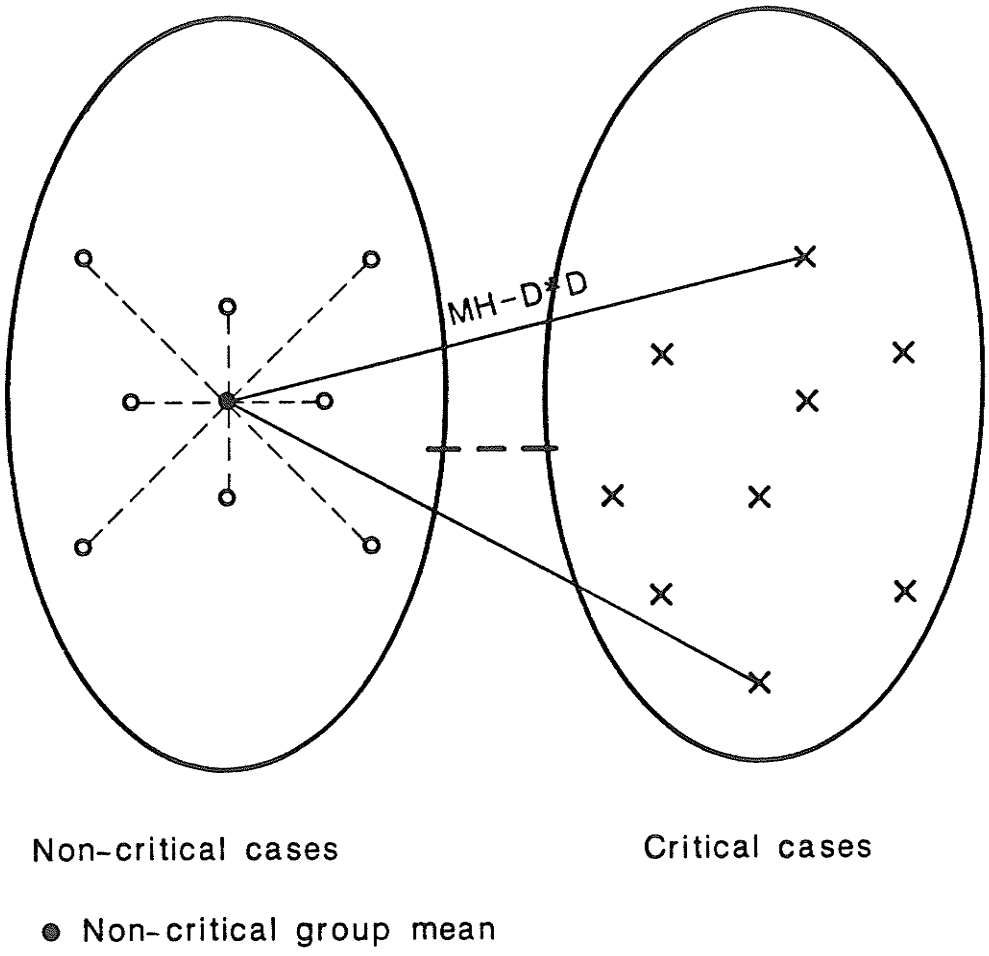


Fig. 7.3 Mahalanobis separability measure between contingencies.

7.4 Numerical Application Results and Discussions

The evaluation of the above mentioned selection algorithms is to be examined through this section. Two power systems are to be tested: a 7-bus sample system, and an 8-bus real system (Northern system). The 7-bus sample system (see Fig. 7.4) data are given in Table 7.1, while the Northern system (Fig. 7.5) data are listed by Table 7.2 .

The MVA base used here is 100 MVA. Also we assume:

$$\Delta VL^{(i)} = \pm 0.03\, VD^{(i)} ,$$

$$FU^{(i)} = 1.03\, VD^{(i)} , \quad FL^{(i)} = 0.97\, VD^{(i)} ,$$

$$VU^{(i)} = 1.05\, VD^{(i)} , \quad VL^{(i)} = 0.95\, VD^{(i)} ,$$

$$VD^{(i)} = 1.0 \quad (sample\ system),$$

$$VD^{(i)} = base\ case\ values\ (real\ system) ,$$

$$i = 1, 2, ..., NB$$

The power system will be in an alarm state if the bus voltage $V^{(i)} / VD^{(i)}$ is greater than 1.03 or less than 0.97 p.u. . Moreover, the system will be in the insecure state if $V^{(i)}$ is greater than $1.05\, VD^{(i)}$ or less than $0.95\, VD^{(i)}$ for any given loading condition. Three loading conditions are used to demonstrate the effectiveness of the proposed performance indices. The capture rate CR and the false alarm rate FR are to be determined.

The value of the discriminant performance index (PIV_3) indicates the state of the system, summarized as follows:

$$voltage\ secure\ state \quad ; \text{if } PIV_3 \geq 0.0$$

$$voltage\ insecure\ state \quad ; \text{if } PIV_3 < 0.0$$

Therefore, the ranking process using PIV_3 starts with the lowest value as the most critical contingency and the second higher value as the second critical contingency

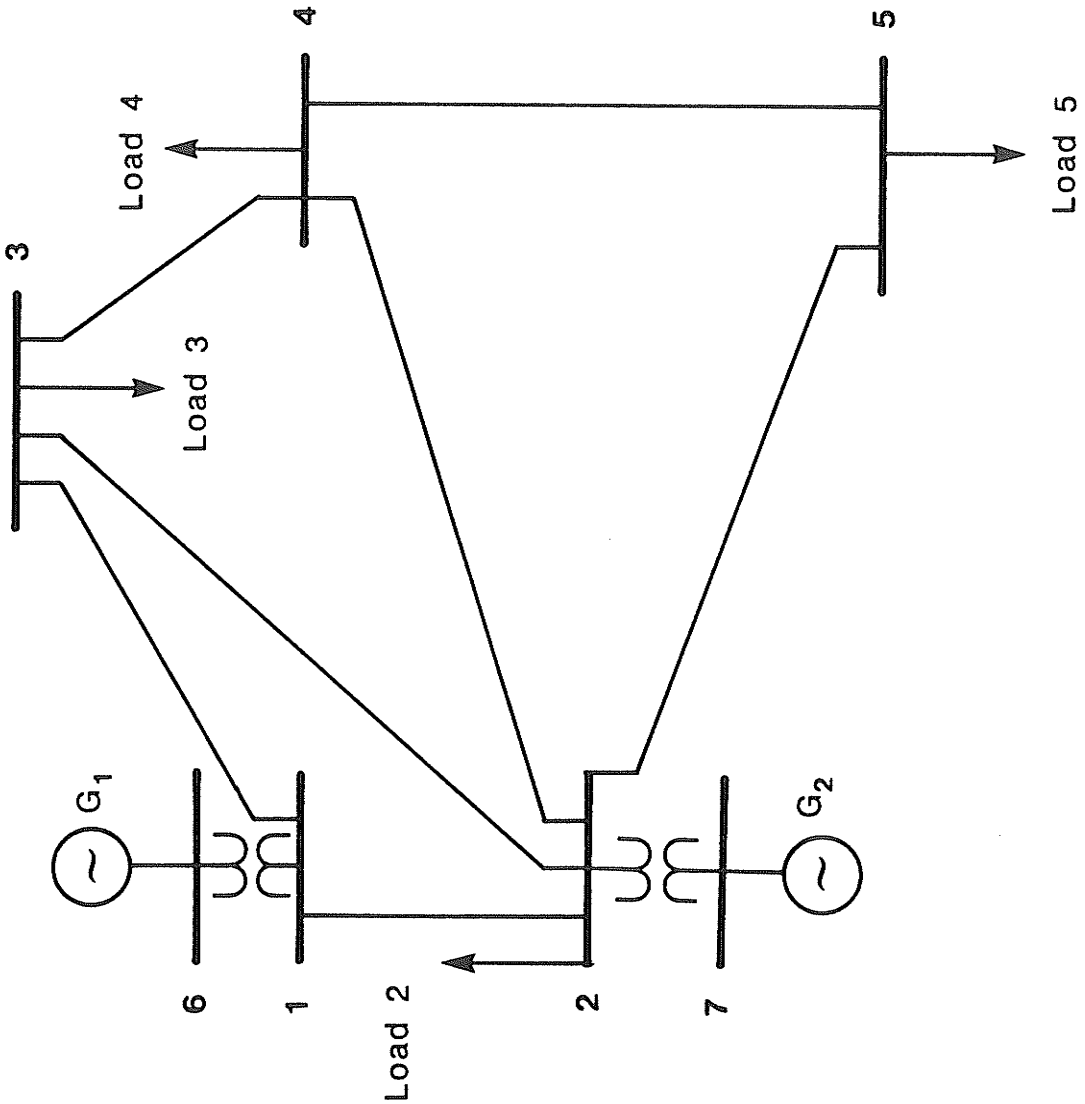


Fig. 7.4 7-bus sample power system.

Table 7.1 Sample system data

(a) Line data

From bus	To bus	R (%)	X (%)	Total MVAR
1	2	0.6	6.0	6.0
1	3	0.8	8.0	5.0
2	3	1.8	18.0	5.0
2	4	1.8	18.0	4.0
2	5	1.2	12.0	3.0
3	4	0.3	3.0	2.0
4	5	1.2	12.0	5.0
6	1	0.4	4.0	
7	2	1.3	13.0	

(b) Bus data

Bus #	Rated voltage	Generation				Light load		Middle load		Heavy load	
		MW	MVAR	Qmax	Qmin	MW	MVAR	MW	MVAR	MW	MVAR
1	1.0										
2	1.0					10	-7	20	10	20	10
3	1.0					30	-5	45	15	55	20
4	1.0					20	-5	40	5	50	10
5	1.0					30	-9	60	10	70	15
6	1.0		swing	bus							
7	1.0	40	30	40	-15						

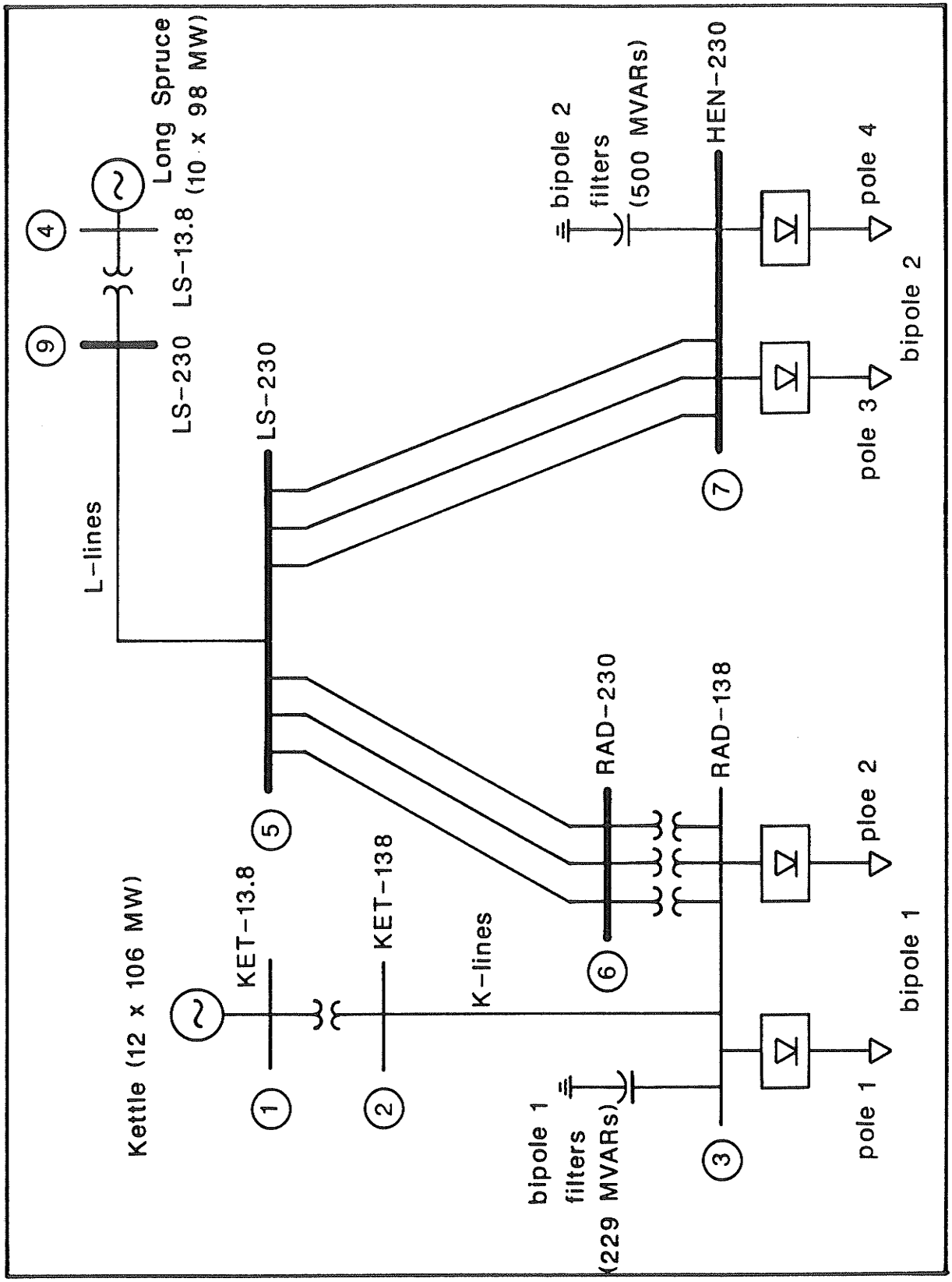


Fig. 7.5 Northern power system.

Table 7.2 Northern system data

(a) Line data

From bus	To bus	No. cts	R (%)	X (%)	Total MVAR	Tap p.u.
2	1	12	.329	8.570	0.0	1.025
2	3	7	.066	0.593	0.147	1.0
5	6	3	.180	1.850	3.610	1.0
9	4	10	.179	6.920	0.0	1.050
5	9	5	.000	0.100	0.050	1.0
6	3	3	.059	3.750	0.0	1.0
5	7	3	.220	2.320	4.490	1.0

(b) Bus data

Bus #	Rated voltage	Generation MW	MVAR	Qmax	Qmin	Light load MW	MVAR	Middle load MW	MVAR	Heavy load MW	MVAR
1	1.01		swing bus								
2	1.0										
3	1.0					9.0	5.4	12.0	7.2	13.2	7.9
4	1.03	9.77		9.99	-9.99						
5	1.0										
6	1.0										
7	1.0					7.5	4.5	10.0	6.0	11.0	6.6
9	1.0										

* All values are in per unit.

and so on in an increasing order.

The calculation results obtained using the selection algorithms of concern are illustrated in Table 7.3. The design of the discriminant index involved here is based on an equal prior probability approach. Also, the selection of two features was considered enough for this design, and for the system under study.

The timing analysis of the discriminant algorithm is essential at this stage. It can be explained as follows: if nv represent the number of variables and N is the number of cases, then the number of required operations can be estimated by [49]

$nv\ (nv\ +6)\ N$ for sums of cross products,

$3\ nv^2\ (nv\ +1)$ for selection of features (stepping),

$2\ nv\ (nv\ +1)$ for the discriminant design

For example, the on-line CPU time needed to come up with the discriminant index is about 0.4 sec on the AMDAHL machine, for the sample system. Therefore, the proposed algorithm has a small time delay compared to others [27,28].

We believe that for a reliable contingency selection, an accurate index must be considered even that the computing speed is sacrificed. With the ever increasing speed of modern computers and various new techniques such as parallel processing, computing speed will be less a problem [27].

The results obtained using Nara and Ejebe algorithms are included in Table 7.3 in order to compare them with the results obtained by the developed algorithms. It must be noted that the AC load flow technique is used for all methods and only single contingencies (outages) are considered. It can be seen that the ranking of severe cases is almost the same for all methods, even though the respective values of the performance indices are different. A value of CR equals 1 and FR equals 0 are noticed which prove the effectiveness of the proposed algorithms.

Table 7.4 shows the results obtained using the discussed algorithms for the case of the Northern system. Single as well as multiple contingencies are considered in this situation. Two design approaches are employed with the indices PIV_3 & PIV_4,

Table 7.3 Voltage contingency analysis for sample system

Outage	Nara [58]		Ejebe [59]		Discriminant*		MH-D*D*		Voltage deviations
	PIV_1	Rank	PIV_2	Rank	PIV_3	Rank	PIV_4	Rank	outside limits (p.u.)
L12	0.4757	1	2.8367	2	-0.2715	1	30.8	1	0.0094, 0.0032, 0.0041
G2	0.3785	2	3.0623	1	-0.2126	2	22.8	2	0.0075, 0.0033, 0.0003
L13	0.2749	3	2.3057	3	-0.1071	3	11.6	3	0.0041, 0.0044, 0.0040
L25	0.2050	4	1.4389	5	-0.1071	4	11.6	4	0.0041
L34	0.1796	5	1.7287	4	-0.0667	5	8.3	5	0.0028, 0.0032
LD3			1.3953	6	0.0728	6	1.1	7	
LD4			1.2629	7	0.1101	7	0.3	8	
L24			1.1184	8	0.1597	8	0.0	12	
L23			1.0273	9	0.1845	10	0.2	9	
LD2			1.0137	10	0.1907	11	0.2	10	
LD5			0.9835	11	0.2745	12	2.5	6	
L45			0.9822	12	0.1752	9	0.1	11	
CR	1		1		1				
FR	0		0		0				

(a) Light loading

L12 : outage of line between bus 1 and bus 2

LD2 : outage of load at bus 2

G2 : outage of generator at bus 2

* = on-line design

Table 7.3 (continued)

Outage	Nara [58]		Ejebe [59]		Discriminant*		MH-D*D *		Voltage deviations
	PIV_1	Rank	PIV_2	Rank	PIV_3	Rank	PIV_4	Rank	outside limits (p.u.)
L13	1.6163	1	5.6777	1	-0.3642	2	44.8	3	-.0231, -.0242, -.0261 -.0032
L25	1.5502	2	4.0478	3	-0.3929	1	48.9	1	-.0309, -.0103, -.0057
G2	0.7707	3	4.1707	2	-0.2966	3	45.4	2	-.0032, -.0147, -.0091 -.0072, -.0032
L45	0.6500	4	2.1940	6	-0.2422	7	29.3	7	-.0130
L12	0.6219	5	3.0248	4	-0.2925	4	35.3	4	-.0121, -.0067, -.0047 -.0006
L34	0.4669	6	2.2781	5	-0.2207	8	27.0	8	-.0086, -.0068
L24	0.2378	7	2.0507	7	-0.2638	6	31.8	6	-.0047, -.0022
L23	0.1900	8	1.9742	8	-0.2709	5	32.7	5	-.0038, -.0005
LD2			0.8602	9	0.1309	9	2.3	9	
LD4			0.6579	10	0.1884	10	0.8	12	
LD3			0.3095	11	0.4109	12	2.0	10	
LD5			0.2175	12	0.3749	11	1.1	11	
CR	1		1		1				
FR	0		0		0				

(b) Middle loading

Table 7.3 (continued)

Outage	Nara [58]		Ejebe [59]		Discriminant*		MH-D*D*		Voltage deviations
	PIV_1	Rank	PIV_2	Rank	PIV_3	Rank	PIV_4	Rank	outside limits (p.u.)
L13	3.7609	1	14.4378	1	-5.5958	2	13.0	2	-.0544, -.0571, -.0593 -.0225
L25	3.3521	2	10.0804	3	-6.8928	1	17.4	1	-.0654, -.0344, -.0267 -.0031
G2	2.4259	3	10.7290	2	-4.0629	3	8.6	3	-.0228, -.0414, -.0336 -.0298, -.0228, -.0019
L12	2.0301	4	7.4361	4	-3.3673	4	6.9	4	-.0355, -.0286, -.0249 -.0166
L34	1.9249	5	6.0450	5	-3.0372	6	6.1	6	-.0327, -.0320, -.0079 -.0054
L45	1.6977	6	5.1129	6	-3.1433	5	6.3	5	-.0336, -.0138, -.0116 -.0030
L24	1.4132	7	5.0916	7	-2.0349	7	4.1	7	-.0242, -.0215, -.0169 -.0005
L23	1.2999	8	4.8294	8	-1.8345	8	3.8	8	-.0225, -.0185, -.0170 -.0005
LD2	0.6066	9	2.7589	9	-0.5728	9	1.9	9	-.0118, -.0066, -.0044
LD4			1.6372	10	0.2997	10	1.0	10	
LD3			1.0937	11	0.7478	11	0.6	11	
LD5			0.7907	12	2.5754	12	0.0	12	
CR	1		1		1				
FR	0		0		0				

(c) Heavy loading

Table 7.4 Voltage contingency analysis for Northern system

Outage	Nara [27]		Ejebe [28]		Discriminant*		MH-D*D*		Discriminant**		MH-D*D**		Voltage deviations
	PIV_1	Rank	PIV_2	Rank	PIV_3	Rank	PIV_4	Rank	PIV_3	Rank	PIV_4	Rank	outside limits (p.u.)
LST	4.9091	1	25.3942	1	-2.8851	1	29.5	1	-18.7677	1	56.2	1	-.0102, -.0156, -.0798 -.0575, -.0502, -.0800
HFT	2.0789	2	4.7668	2	-0.9461	2	8.8	3	-6.3848	2	14.5	2	-.0032, -.0081, -.0023
BP2	1.5906	3	3.1045	3	-0.5446	4	6.6	4	-3.7668	3	9.5	3	0.0365
3L56	0.8388	4	2.2412	4	-0.1839	5	2.6	5	-0.5779	5	3.5	5	-.0186
BP1	0.3658	5	1.8142	5	-0.5988	3	9.2	2	-2.8815	4	9.2	4	0.0028, 0.0076
LD3	0.2301	6	0.9706	6	-0.0789	6	1.8	8	-0.5085	6	2.5	6	0.0053
RLD			0.6237	7	0.0281	7	2.6	6	m-0.4435	7	2.5	7	
LD1			0.4938	8	0.1315	8	1.9	7	1.5744	8	1.8	8	
RFT			0.3095	9	0.2956	9	1.0	9	2.5734	9	1.0	9	
HLD			0.2532	10	0.5151	10	0.1	12	3.7539	10	0.2	10	
1T36			0.0315	11	0.7638	11	0.0	18	5.3988	11	0.0	18	
1L56			0.0160	12	0.8349	12	0.1	13	5.8525	12	0.1	15	
1L57			0.0047	13	0.8761	13	0.1	14	6.0889	13	0.1	16	
1T49			0.0044	14	0.8779	14	0.1	15	6.1020	14	0.1	17	
1T12			0.0007	15	0.8954	15	0.1	16	6.2339	15	0.2	11	* = on-line design
1LSG			0.0004	16	0.9136	17	0.2	10	6.3389	16	0.2	12	** = off-line design
1L23			0.0001	17	0.9125	16	0.1	17	6.3391	17	0.2	13	m = misclassified
1L59			0.0000	18	0.9225	18	0.2	11	6.3981	18	0.2	14	
CR	1		1		1				1				
FR	0		0		0				0.17				

(a) Light loading

1L23 : outage of one line between bus 2 and bus 3

1T12 : outage of one transformer between bus 1 and bus 2

1LSG : outage of one unit of LS station

BP1 : outage of bipole 1

LD1 : outage of load 1 (pole 1)

RFT : outage of Radisson filters

HFT : outage of Hendy filters

RLD : outage of Radisson load

HLD : outage of Hendy load

LST : outage of LS station

Table 7.4 (continued)

Outage	Nara [27]		Ejebe [28]		Discriminant*		MH-D*D*		Discriminant**		MH-D*D**		Voltage deviations
	PIV_1	Rank	PIV_2	Rank	PIV_3	Rank	PIV_4	Rank	PIV_3	Rank	PIV_4	Rank	outside limits (p.u.)
LST	DIVERGENCE												
BP2	3.1637	1	7.1796	1	-1.5376	1	40.6	1	-9.1304	1	22.5	1	0.0091, 0.0679, 0.0081
HFT	2.2877	2	5.1007	2	-1.2422	2	35.7	2	-6.4769	3	14.7	3	-.0040, -.0491, -.0031
BP1	1.3025	3	3.8938	3	-0.2049	5	24.4	4	-6.9625	2	20.6	2	0.0181, 0.0248
LD3	1.1466	4	2.4274	5	-0.5931	3	14.6	5	-2.8165	5	6.9	5	0.0246
3L56	1.0232	5	2.6276	4	-0.3787	4	34.7	3	-0.9925	6	4.0	7	-.0219
RLD	0.4479	6	1.9724	6	-0.0907	6	10.9	6	-3.1245	4	9.6	4	0.0041, 0.0090
LD1			1.1039	7	0.2761	7	5.3	7	m-0.6996	7	4.8	6	
RFT			0.3261	8	0.5195	9	0.9	10	2.5469	8	1.0	8	
HLD			0.2331	9	0.5079	8	1.2	8	3.6787	9	0.5	9	
1L57			0.1347	10	0.6064	10	1.2	9	4.1992	10	0.3	10	
1T36			0.0452	11	0.6354	12	0.1	16	5.3132	11	0.0	17	
1T49			0.0209	12	0.6295	11	0.2	14	5.7533	12	0.1	14	
1L56			0.0158	13	0.7339	13	0.1	17	5.8592	14	0.1	15	
1T12			0.0076	14	0.7686	14	0.2	15	5.8526	13	0.1	16	
1L23			0.0006	15	0.7976	15	0.3	13	6.2140	15	0.2	11	
1LSG			0.0005	16	0.8092	17	0.4	11	6.2601	16	0.2	12	
1L59			0.0001	17	0.8034	16	0.4	12	6.3849	17	0.2	13	
CR	1		1		1				1				
FR	0		0		0				0.17				

(b) Middle loading

Table 7.4 (continued)

Outage	Nara [27]		Ejebe [28]		Discriminant*		MH-D*D*		Discriminant**		MH-D*D**		Voltage deviations
	PIV_1	Rank	PIV_2	Rank	PIV_3	Rank	PIV_4	Rank	PIV_3	Rank	PIV_4	Rank	outside limits (p.u.)
LST	DIVERGENCE												
BP2	3.9421	1	9.7469	1	-1.5741	1	32.3	1	-11.6639	1	30.6	1	0.0158, 0.0821, 0.0146
1L57	3.2057	2	5.0752	4	-0.3218	6	11.7	5	-6.5549	4	18.5	3	-0.0667
HFT	2.3945	3	5.2729	2	-1.0904	3	19.5	3	-6.5756	3	14.9	4	-.0043, -.0498, -.0034
BP1	1.7523	4	5.1250	3	-0.6542	4	12.4	4	-8.8095	2	27.2	2	0.0248, 0.0323, 0.0033
LD3	1.6168	5	3.4046	5	-0.6404	5	11.2	6	-4.4814	6	9.8	6	0.0337
3L56	1.0705	6	2.7769	7	-1.4175	2	31.0	2	-1.1312	8	4.1	8	-0.0226
RLD	0.8694	7	2.8554	6	-0.2579	7	6.4	7	-4.9450	5	14.2	5	0.0109, 0.0165
LD1	0.1681	8	1.4694	8	-0.0769	8	2.9	8	-1.7248	7	6.6	7	0.0033
HLD	0.1431	9	0.7627	9	m 0.4276	9	1.1	9	-1.0386	9	2.0	9	0.0030
RFT			0.3358	10	0.5160	10	0.4	10	2.5273	10	1.0	10	
1T36			0.0452	11	0.6632	11	0.1	13	5.2474	11	0.0	16	
1T49			0.0336	12	0.7459	13	0.0	17	5.5558	12	0.1	12	
1L56			0.0166	13	0.7326	12	0.0	16	5.8262	14	0.1	13	
1T12			0.0143	14	0.8381	14	0.1	14	5.6554	13	0.0	17	
1LSG			0.0015	15	0.8973	16	0.2	11	6.1549	16	0.1	14	
1L23			0.0011	16	0.8885	15	0.1	15	6.1418	15	0.1	15	
1L59			0.0001	17	0.9062	17	0.2	12	6.3652	17	0.2	11	
CR	1		1		0.88				1				
FR	0		0		0				0				

(c) Heavy loading

they are: on-line mode and off-line mode. In the on-line mode the cases used for the design are those only corresponding to the post-outage contingency conditions for the current system state. On the other hand, the off-line design mode means the generation of many post-outage conditions related to different system states.

In the case of on-line design, it was found that two voltage features are satisfied and the prior probability of the insecure (critical) contingencies is equal to 0.6 . On the other hand, a two voltage features and a 0.8 prior probability of critical cases are considered for off-line design case.

It can be seen from Table 7.4a that CR equals 1; FR equals 0 for the on-line design and equals 0.17 for the off-line design mode due to one false alarm. Also from Table 7.4b, it can be observed that CR equals 1 for all algorithms and FR equals 0.17 for the off-line design. Moreover, from Table 7.4c it is clear that CR equals 0.88 for the on-line design case. It must be noted that the CR value can be maximized by increasing of the prior probability of the critical contingencies to a value greater than 0.5 .

Insecure cases can be easily recognized from Table 7.4 as those with values less than zero (using PIV_3). Those marked with m are misclassified. However, they can be simply minimized by checking if $V^{(i)}$ is greater than $1.03\ VD^{(i)}$ or less than $0.97\ VD^{(i)}$ for any monitored bus i . If so the case is noted as voltage insecure state, otherwise, it is a voltage secure state.

7.5 Conclusions

Pattern recognition based voltage contingency selection algorithms are developed. The proposed algorithms are examined on a sample system and on a real system. Results obtained proved the effectiveness of these algorithms.

A comparison with other selection methods is presented. The developed algorithms are efficient for single and multiple outages. They have slightly a lower speed

of computation (for on-line design), which could be considered less a problem with the fast advanced computer technology.

The developed algorithms, off-line design approach, is efficient and reliable compared with Nara and Ejebe methods. However, it requires an excessive computational burden.

The application of the proposed algorithms on a large scale real system is being investigated.

CHAPTER 8

GENERAL CONCLUSIONS AND RECOMMENDATIONS

8.1 General Conclusions

The following contributions have been achieved throughout this dissertation:

1- A new pattern-recognition based prediction system has been successfully designed. This design adapts to new system conditions. Using this system, a power system operator can anticipate abnormal conditions. The design scheme has been applied to the Manitoba Hydro Northern system in order to predict self-excitation and dynamic overvoltages.

2- A fast, efficient corrective action algorithm has been developed. This algorithm provides the operator with the suggested corrective action in case of anticipated trouble, i.e. insecure condition. It has been applied to improve the system security against self-excitation and dynamic overvoltages as well.

3- The design of a multi-class prediction system using a multi-class pattern-recognition approach has been applied to a three-class power-system load-rejection overvoltages problem. The application is new and it has many advantages. One of the major advantages is the reduction in telemetering communications which simplifies the information system structure. On the other hand, it posses some limitations:

   a- The estimation of the prior probability is not an easy task in this design scheme.

   b- The prediction time (time delay from receiving information to taking decision) could be longer compared to the two-class approach.

c- The degradation in the system performance due to telemetering failure.

4- An efficient, relatively fast voltage contingency selection algorithm based on pattern-recognition discriminant hyperplane has been developed. This algorithm proved to be effective in case of single as well as multiple outages study. On the other hand, the on-line design scheme has slightly lower speed of computation while the off-line has an excessive computation burden. The algorithm has been tested on a sample system and a real power system.

8.2 Recommendations for Future Extensions

Future research at this stage could be extended to include:

1- Application of the prediction system on a large scale power system e.g. Manitoba Hydro Northern system as a whole, for the detection of load-rejection self-excitation as well as dynamic overvoltages.

2- Extension of the pattern-recognition based voltage contingency analysis on a large scale system, e.g. IEEE 25-bus system. An investigation for its feasibility study is also recommended.

3- Application of the pattern recognition based prediction scheme to other power system operating problems.

APPENDIX (A)

ESTIMATION OF STATISTICAL RELATIONS

(1) Estimate of Means

Given the N by M matrix X, consisting of N patterns each of M variables. Then, computation of the means of the M variables uses the following formula :

$$\bar{X}_j = \sum_{i=1}^{N} \frac{X_{ij}}{N}, \; j = 1, 2, ..., M \qquad \text{(A.1)}$$

where:

X_{ij} is the pattern i on variable j ,

N is the number of patterns per variable.

(2) Estimate of Standard Deviations

Standard deviations are computed using the following formula:

$$s_j = \left(\sum_{i=1}^{N} \frac{x_{ij}^{\,2}}{(N-1)}\right)^{1/2}, \; j = 1, 2, ..., M \qquad \text{(A.2)}$$

where: $x_{ij} = X_{ij} - \bar{X}_j$

(3) Estimate of Correlation Coefficients

Computation of the simple correlation coefficients of the M variables is done as follows:

$$r_{ij} = \sum_{k=1}^{N} \frac{x_{ki} x_{kj}}{(N-1) s_i s_j}, \qquad i \neq j \qquad \text{(A.3)}$$

where: $r_{ij} = 1$ if $i = j$, $r_{ij} = r_{ji}$ and $i, j = 1, 2, ..., M$.

(4) Estimate of Covariance Matrix

The covariance matrix can be calculated as :

$$A = (a_{ij}), \qquad i, j = 1, 2, ..., M \tag{A.4}$$

where: $a_{ij} = r_{ij} s_i s_j$, $a_{ij} = a_{ji}$

(5) Determination of Required Features

The determination of the required number of features using the feature extraction technique is described herein. The objective of the feature extraction is to derive a linear transformation that will emphasize the differences among patterns belonging to different classes. In other words, it is required to define new coordinate axes in the direction of high information content useful for classification purposes.

The algorithm used here is called the principal components, Karhunen-Loeve transformation (KLT) in pattern recognition terminology [41]. The procedure is described as follows:

1- Determine the M by M covariance matrix among classes.

2- Find the M eigen-values and the corresponding M eigen-vectors of the covariance matrix.

3- Extract the f largest eigen-values such that their sum is almost equal to the sum of the M eigen-values.

4- The transformation matrix can be obtained from the f eigen-vectors corresponding to the f eigen-values in a descending order.

5- The M-dimensional pattern vector can be transformed to the f vector using the transformation matrix, i.e. the required feature vector.

APPENDIX (B)

STEPWISE DISCRIMINANT ANALYSIS

(1) Feature Selection

The selection of features using this algorithm [48,49] is based on the ratio $R(X)$ of the within generalized dispersion $|W(X)|$ to the total generalized dispersion $|T(X)|$ as:

$$R(X) = |W(X)|/|T(X)| \tag{B.1}$$

Large values of $R(X)$ indicate poor separation between classes, while small values indicate good separation.

where:

$$W(X) = (w_{rs}/r, s = 1, 2, ..., f) \tag{B.2}$$

and

$$w_{rs} = \sum_{i=1}^{c} \sum_{j=1}^{N_i} (x_{ijr} - \bar{x}_{ir})(x_{ijs} - \bar{x}_{is}) \tag{B.3}$$

$$T(X) = (t_{rs}/r, s = 1, 2, ..., f) \tag{B.4}$$

and

$$t_{rs} = \sum_{i=1}^{c} \sum_{j=1}^{N_i} (x_{ijr} - \bar{x}_{r})(x_{ijs} - \bar{x}_{s}) \tag{B.5}$$

The multiplicative increment MI_k in Fisher's ratio $R(X)$ resulting from the addition of a variable y_j to the set X can be formulated as follows:

$$MI_j = R(X, y_j)/R(X) = \frac{|W(X, y_j)|\,|T(X)|}{|T(X, y_j)|\,|W(X)|}$$

$$= \alpha_{jj} / \beta_{jj} \tag{B.6}$$

where

$$\alpha_{jj} = |W(X, y_j)| \; / \; |W(X)|$$

$$\beta_{jj} = |T(X, y_j)| \; / \; |T(X)|$$

while for the case of deletion of feature x_i from the set of X, it can be given as:

$$R(X, -x_i) \; / \; R(X) = \frac{|W(X, -x_i)| \, |T(X)|}{|T(X, -x_i)| \, |W(X)|}$$

$$= \beta_{ii} \; / \; \alpha_{ii} = MI_i^{-1} \tag{B.7}$$

Using the F-statistic as a measure to guide the selection or deletion of variable x_k, it can be written as:

$$F = \frac{N - c - f}{c - 1} \; (1 - MI_k) \; / \; MI_k \tag{B.8}$$

Therefore, using Eq. (B.8) the selection measure of a variable x_j can be written in the form:

$$SM_j = \frac{N - c - f}{c - 1} \; (1 - MI_j) \; / \; MI_j$$

$$= (a/b) \; [1/MI_j - 1] \tag{B.9}$$

where $a = N - c - f$ and $b = c - 1$. Similarly, the removing measure of a variable x_i can be obtained as:

$$RM_i = \frac{N - c - f + 1}{c - 1} \; (1 - MI_i^{-1}) \; / \; MI_i^{-1}$$

$$= ((a + 1)/b) \; [MI_i - 1] \tag{B.10}$$

The within class tolerance t_j for a variable y_j not included in the set X is given by:

$$t_j = \alpha_{jj} \; / \; w_{jj} \tag{B.11}$$

(2) Stepwise Discriminant Design

When the stepping is complete, or when the number of features selected is equal to the one specified, then the discriminant design starts. The design of the stepwise discriminant function can be easily formulated [48,49]. Let us assume that:

$M_i = [m_{i1}, m_{i2}, ..., m_{if}]^T$ is the mean vector for class i,

$M = [m_1, m_2, ..., m_f]^T$ is the overall mean vector, and

$X = [x_1, x_2, ..., x_f]^T$ is the feature vector.

The class i discriminant function is:

$$d_i(X) = (N - c)\, M_i^T \; W(X)^{-1} X - 0.5(N - c)\, M_i^T \; W(X)^{-1} M_i + \ln p_i \tag{B.12}$$

therefore, the class weighting vector W_i and the threshold weight w_0 are given by:

$$W_i^T = (N - c)\, M_i^T \; W(X)^{-1} \tag{B.13}$$

$$w_{i0} = -0.5(N - c)\, M_i^T \; W(X)^{-1} M_i + \ln p_i \tag{B.14}$$

APPENDIX (C)

PATTERN-RECOGNITION LEARNING ALGORITHMS

In this appendix three learning algorithms are introduced:

1- Bayes learning theory;

2- Linear discriminant method; and

3- K-nearest neighbor method.

(1) Bayes Learning Theory

In this method the conditional probability density function for each class must be known in order to construct the likelihood ratio

$$l(X) = f(X/i)/f(X/j) \tag{C.1}$$

where i, j are any two classes. Assuming a normal distribution for the primary variables, the density function can be estimated as:

$$f(X/i) = (2\pi)^{-f/2} |\Sigma_i|^{-1/2} \exp[-0.5 (X - M_i)^T \Sigma_i^{-1} (X - M_i)] \tag{C.2}$$

$$i = 1, 2, ..., c$$

The main objective of Bayes theory is to minimize the average or expected value of the loss function $L(d_i, j)$ (i.e. the loss incurred for taking decision d_i for pattern of class j). The Bayes rule could be written as:

$$d(X) = d_i \quad \text{if} \quad \sum_{j=1}^{c} L(d_i, j)\, p(j/X) < \sum_{r=1}^{c} L(d_m, r)\, p(r/X) \tag{C.3}$$

for all $m \neq i$, $i = 1, 2, ..., c$.

For the two-class problem i.e. $c = 2$, considering the loss function equals one in case of taking a wrong decision and zero in case of right decisions, then the decision rule could be written as:

$$d(X) = d_1 \quad \text{if} \quad f(X/1)\, p_1 > f(X/2)\, p_2 \quad ,$$

$$d(X) = d_2 \text{ if } f(X/2)\, p_2 > f(X/1)\, p_1 \qquad \text{(C.4)}$$

where:

$$p(i/X) = f(X/i)\, p_i / f(X) \qquad \text{(C.5)}$$

The Bayes recognition rule is optimal from the point of view of the probability of error, but its practical performance depends on the degree of validity of assuming normal distribution for the primary variables.

(2) Linear Discriminant Method

The linear discriminant method derived here [50] is based on a non-parametric estimation of the class density function $f(X/i)$. In order to take into account the contribution of each design pattern in the estimation of the density function, the following form is the most suitable for this situation, that is:

$$f(X/i) = \frac{1}{N_i\, \sigma^f\, (2\pi)^{f/2}} \sum_{j=1}^{N_i} \exp\left[-\frac{(X - X_{ij})^T (X - X_{ij})}{2\sigma^2}\right] \qquad \text{(C.6)}$$

$$i = 1, 2, \ldots, c$$

where:

f is the number of features,

N_i is the number of patterns in class i,

σ is a smoothing factor.

As explained in [50] the class density function could be written as:

$$f(X/i) = \frac{1}{\sigma^f\, (2\pi)^{f/2}} \left[\exp\left(-\frac{X^T X}{2\sigma^2}\right)\right] G^i(X) \qquad \text{(C.7)}$$

$$i = 1, 2, \ldots, c$$

where:

$$G^i(X) = g^i{}_{0\ldots0} + g^i{}_{10..0}\, x_1$$

$$+ g^{i}{}_{010..0}\, x_2 + ... + g^{i}{}_{z_1 z_2 \cdots z_f}\, x_1^{z_1} x_2^{z_2} \cdots x_f^{z_f}. \qquad (C.8)$$

$$g^{i}{}_{z_1 z_2 \cdots z_f} = [\frac{1}{(z_1!\, z_2! \cdots z_f!)\, \sigma^{2h}\, N_i}] \sum_{j=1}^{N_i} x_{ij1}^{z_1} x_{ij2}^{z_2} .... x_{ijf}^{z_f}\, B_{ij} \qquad (C.9)$$

where:

$h = z_1 + z_2 + ... + z_f$

$B_{ij} = \exp(-\frac{X_{ij}{}^T X_{ij}}{2\sigma^2})$

$X^T = [x_1, x_2, ..., x_f]$

Taking only first order terms and neglecting higher ones, then Eq. (C.8) can be reformulated as:

$$G^i(X) = g_0{}^i + g_1{}^i\, x_1 + g_2{}^i\, x_2 + ... + g_f{}^i\, x_f \qquad (C.10)$$

where

$$g_0{}^i = \frac{1}{N_i} \sum_{j=1}^{N_i} B_{ij} \quad ,$$

$$g_k{}^i = \frac{1}{\sigma^2 N_i} \sum_{j=1}^{N_i} x_{ijk}\, B_{ij} \qquad (C.11)$$

Once the density functions are estimated on the basis of the design patterns, Bayes rule could be derived as:

$$d(X) = d_i \quad \text{if} \quad p_i\, f(X/i) > p_j\, f(X/j) \quad \text{for } all \quad j \neq i \qquad (C.12)$$

substituting with $f(X/i), f(X/j)$ from Eqns. (C.7) and (C.10) we can write the discriminant rule as:

$$d(X) = d_i \quad \text{if} \quad p_i\, G^i(X) > p_j\, G^j(X) \quad \text{for } all \quad j \neq i \qquad (C.13)$$

$$i = 1, 2, ..., c .$$

(3) K-Nearest Neighbor Method

In this algorithm the density function is estimated as follows:

$$f(X/i) = \frac{k_i - 1}{N_i V} \tag{C.14}$$

where V is the volume of the hypersphere that contains the set of all k-th nearest neighbor patterns to the given pattern X, N_i is the number of patterns from class i, k_i is the number of nearest patterns from class i i.e.

$$K = \sum_{i=1}^{c} k_i \tag{C.15}$$

Then using Bayes rule given by eq. (C.12) and substituting with $f(X/i)$ from eq. (C.14) we can obtain the following K-NN rule:

$$d(X) = d_i \quad \text{if} \quad k_i > k_j \quad \text{for} all \; j \neq i \tag{C.16}$$

$$i = 1, 2, ..., c$$

where $p_i = \frac{N_i}{N}$ and $N = \sum_{i=1}^{c} N_i$ is the number of patterns on the design set.

Therefore, the K nearest neighbor rule is very simple and it is just a comparison of k_i with k_j (i.e. comparison between the number of nearest neighbor patterns from each class). Its disadvantage is the need to store all patterns and to compare the distance measure between each with the unknown one.

APPENDIX (D)

PERFORMANCE EVALUATION METHODS

The following methods are discussed:

(1) The Resubstitution Estimate

In this method, patterns used for design are also used for the performance evaluation of the pattern prediction system. This method could be considered an optimistic estimate since the patterns used for design are the same used for the evaluation. However, when a large design set is available, this error estimate is probably as good an estimator as any other one.

(2) The Hold-out Estimate

The most obvious alternative to the resubstitution scheme is to *partition* the pattern set into two mutually exclusive subsets and to use one subset for *designing* the predictor and the other one to *test* it. This approach has the disadvantage of making a poor use of the available patterns since a predictor designed on the entire pattern set will, on the average, perform better than the one designed on only a portion of the whole set.

This approach is relying on the available patterns. These patterns should be divided into the design patterns and the test patterns i.e. they should be statistically independent or at least different. Also it must be emphasized that if most of the patterns are used for designing the predictor, there will be a little confidence in the testing stage and vice versa.

For small to moderate sample sizes, very significant discrepancies between the resubstitution and hold-out estimates may be observed. The later one being an order of magnitude larger than the former. As it turns out the hold-out method has a definite tendency to over-estimate the actual error rate.

To summarize, the hold-out estimate is making poor use of the available patterns and it gives pessimistic error estimates.

(3) The Leave-One-Out Estimate

The so-called *leave−one−out or deleted* method is formed in the following manner:

1- remove one pattern, (X) from the pattern set S_d .

2- design the predictor using remaining patterns S_{d-1} and estimate its performance using removed pattern (X).

3- return the removed pattern (X) to the pattern set S_{d-1} .

4- repeat the above procedure for all patterns in the pattern set.

Clearly, with this method, all patterns are used in each design, and also all of them are used in the tests, though each design and test sets may be regarded as independent. This estimate could be considered as unbiased, that is because the design and test sets distributions are essentially identical. Also, another advantage of this method is the efficient making use of the available patterns. On the other hand, there are two disadvantages: the first is that the error bias reduction is achieved at the expense of an increase in the variance of the estimator, and the second is the excessive computation involved as N design sessions are required.

(4) The Rotation Estimate

This error estimate method is a compromise between the hold-out and leave-one-out methods. The procedure of this method can be summarized as follows:

1- partition the N patterns of the pattern set into N/L disjoint subsets (L is an integer and a divisor of N).

2- remove $j-th$ subset from the pattern set.

3- design the predictor with the remaining subsets and then assess its performance using the $j-th$ subset.

4- return the subset j to the pattern set.

5- repeat the above operations for all subsets ($j = 1, 2, ..., N/L$).

6- calculate the rotation estimate as the average frequency of misrecognitions over the N/L test sessions.

In this estimate, if $L = 1$, it reduces to the leave-one-out method, where as when $L = N/2$ it reduces to the hold-out method. The rotation estimate reduces both the bias inherent to the hold-out method and the computational burden associated with the leave-one-out method.

APPENDIX (E)

NORTHERN SYSTEM BLOCK DIAGRAMS AND DATA

I: System Block Diagrams

The dynamics of the power system used for this study is explained in Fig. 1. From this diagram [6] it can be seen that the system model consists mainly of: machine model; excitation system model; governor-turbine system model; and transmission system model. The system model and block diagrams [52] are given as shown in Fig.2.

II: System Data

Machine data

$X_d = 0.922$, $X_q = 0.535$, $X_d' = 0.251$, $X_q' = 0.535$,

$X_d'' = X_q'' = 0.194$, $X_l = 0.193$, $T_{do}' = 4.1$ sec,

$T_{do}'' = 0.019$ sec, $T_{qo}'' = 0.048$ sec, $H = 4.1$ sec

Exciter data

$KA = 289.0$, $TA = 7.04$ sec, $TB = 1.43$ sec,

$TC = 0.02$ sec, $TD = 0.012$ sec, $TE = 0.0$,

$KE = 1.0$, V_A max = 5.0, V_A min =-3.5

Governor data

$1/fR = 0.88$, $C_1 = 4.8$, $C_2 = 0.1$

$C_3 = 0.0$, $C_4 = 0.02$, $C_5 = 0.04$,

$T_C = 3.7$ sec, $T_3 = 0.44$ sec, $T_4 = 0.22$ sec,

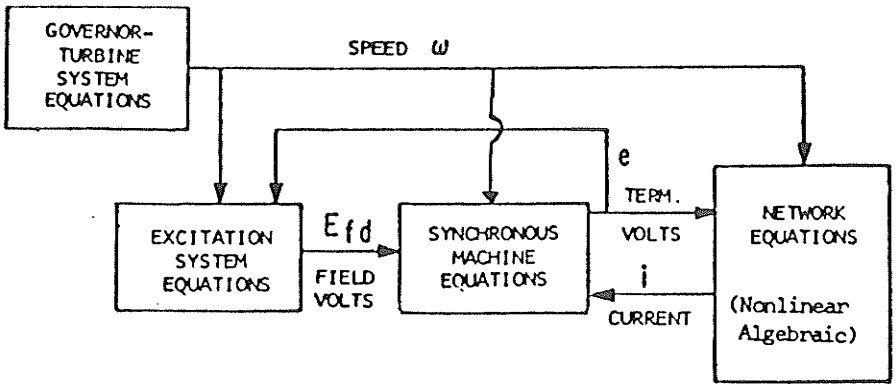
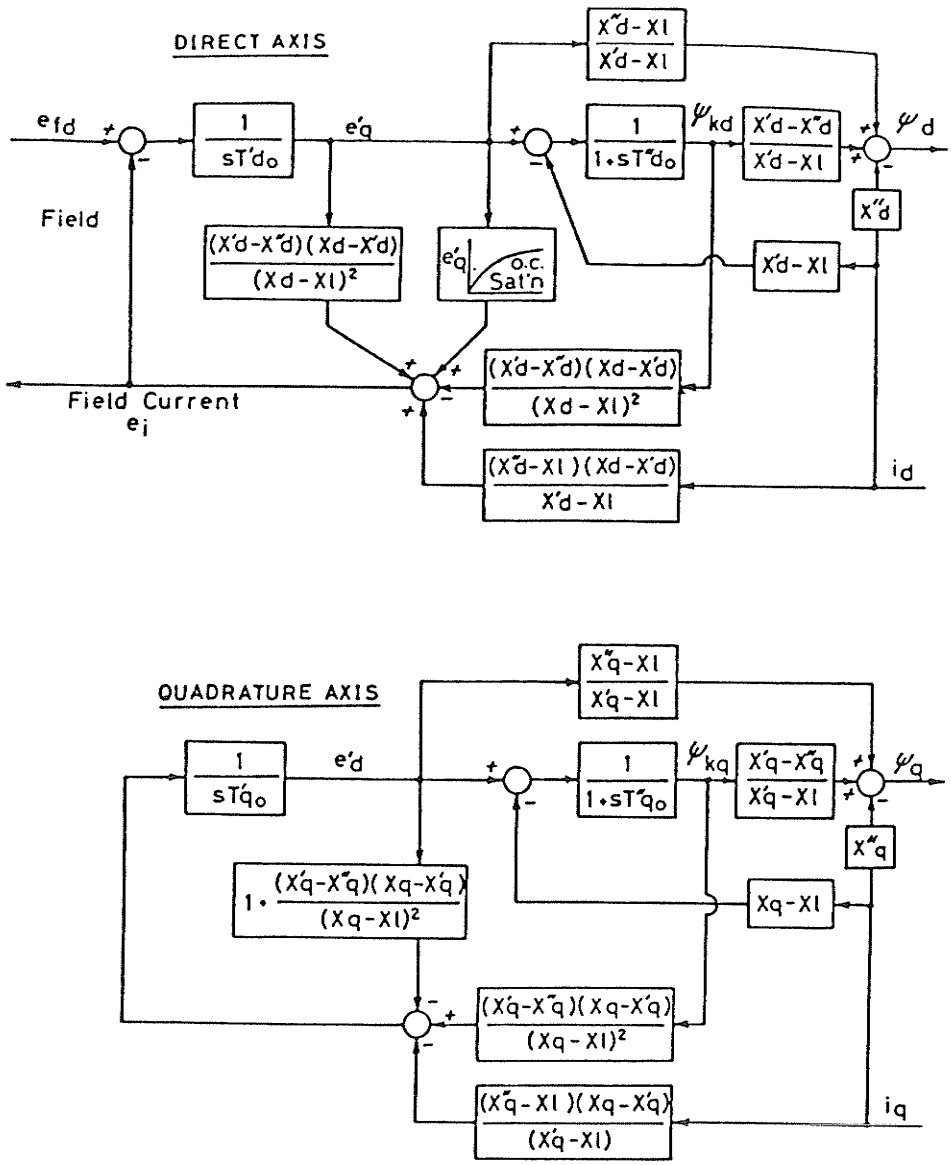
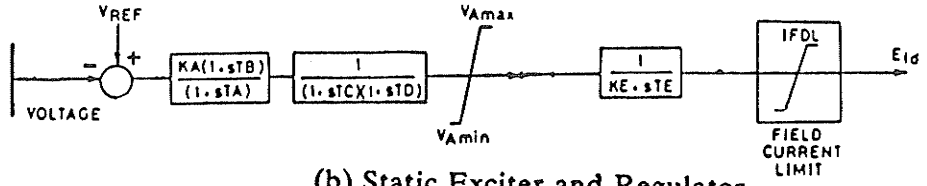


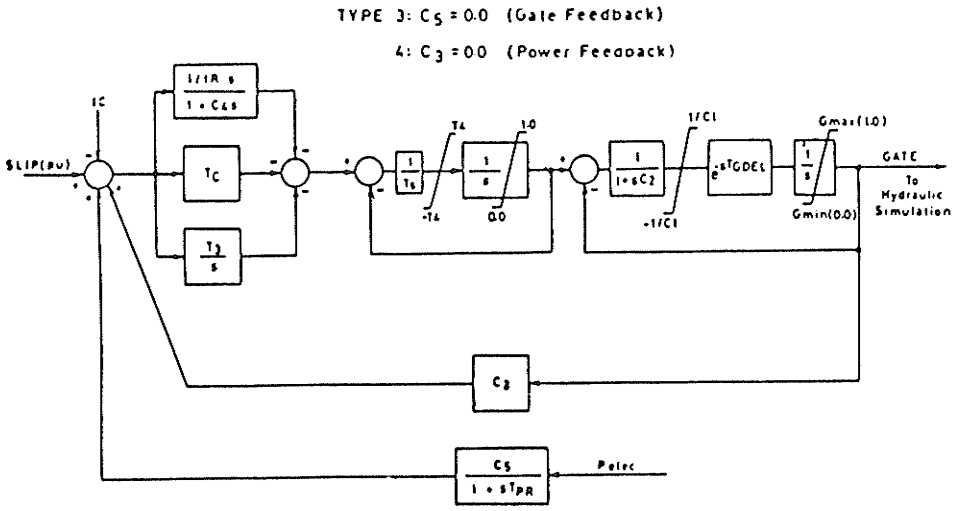
Fig. 1. Power system dynamics



(a) Salient Pole Generator



(b) Static Exciter and Regulator



(c) Hydro-Turbine and Governor

Fig. 2. Power system block diagrams

T_5 = 1.0 sec, T_s = 0.05 sec, T_{PR} = 0.05 sec,

T_{GDEL} = 0.0 sec

Line and Transformer data

R_{TL} = 0.00066, X_{TL} = 0.00593, R_{tr} = 0.00329,

X_{tr} = 0.0857, Line charging MVAR = 0.147

III: List of Variables

All quantities are in per unit unless noted.

ω_o = rated frequency

ω = actual frequency

X_l = stator leakage reactance

X_d'' = direct axis subtransient reactance

X_d' = direct axis transient reactance

X_d = direct axis synchronous reactance

X_q'' = quadrature axis subtransient reactance

X_q = quadrature axis synchronous reactance

ψ_d'' = direct axis subtransient flux linkages

ψ_q'' = quadrature axis subtransient flux linkages

e_q' = voltage proportional to field flux linkages

e_i = voltage proportional to field current

T_{do}' = open circuit field time constant

T_{do}'' = open circuit d-axis subtransient time constant

T_{qo}' = open circuit q-axis transient time constant

T_{qo}'' = open circuit q-axis subtransient time constant

E_{fd} = slip ring field voltage

I_{fd} = field current

E_x = exciter voltage

ψ_d, ψ_q = armature flux linkages, direct and quadrature axis components

i_d, i_q = armature current, direct and quadrature axis components

e_d, e_q = armature voltage, direct and quadrature axis components

H = inertia constant (sec)

P_g, Q_g = generator output powers

V_t, V_L = generator terminal and load voltages

$\Delta\omega$ = speed change

R_{TL} = transmission line series resistance

X_{TL} = transmission line series reactance

R_{tr} = transformer series resistance

X_{tr} = transformer series reactance

REFERENCES

(1) C. K. Pang, A. H. El-Abiad, et..al,"Application of Pattern Recognition to Steady State Security Evaluation in a Power System" IEEE Trans., Vol. SMC-3, No.6, pp. 622-631, Nov.1973 .

(2) C. K. Pang, et .. al,"Security Evaluation in Power Systems Using Pattern Recognition " IEEE Trans., Vol. PAS-93, pp. 969-76, May/June 1974 .

(3) A. Brameller and H. Rudnick," Transient Security Assessment Methods" , IEE Proceedings, Vol. 125, No. 2, pp. 135-140, Feb. 1978.

(4) F. P. de Mello, et .. al,"Analog Computer Studies of System Overvoltages Following Load Rejections " AIEE Transactions, Pt. III, PAS, pp. 42-49, April 1963.

(5) P. L. Dandeno, K. R. McClymont," Extra-High-Voltage System Overvoltages Following Load Rejection of Hydraulic Generation ", IEEE Trans. on PAS, Vol. 65, Pt. III, pp. 49-57, April 1963.

(6) F. P. de Mello, et .. al,"Load Rejection Overvoltages as Affected by Excitation System Control" IEEE Trans., Vol. PAS-94, No. 2, pp. 280-287, March/April 1975 .

(7) A. M. Gole, et..al,"Capacitive Load Induced Field Stress on Generators Connected into DC Systems" IEEE Trans. on Power Apparatus and Systems, Vol. PAS-104, No. 3, pp. 543-48, March 1985.

(8) E. C. Mc Clelland, et .. al,"Fast Voltage Prediction Using a Knowledge Based Approach " IEEE Trans., Vol. PAS-102, No.2, pp. 315-19 , Feb.1983 .

(9) O. Saito, et .. al,"Security Monitoring Systems Including Fast Transient Stability Studies "IEEE Trans., Vol. PAS-94, No.5, pp. 1789-1805, Sept./Oct. 1975 .

(10) M. Udo,"Design and Functional Characteristics of Hierarchical Power Control Systems " CIGRE, Report 32-77-69, Meeting 1977 in Dortmund.

(11) C. L. Gupta, et ..al,"Transient Security Assessment of Power System by Pattern Recognition - a Statistical Approach " , IEEE PES Summer Meeting, Mexico City, Mex.,July 17-22, 1977 .

(12) J. M. G. Sa Da Costa,"Application of Pattern Recognition to Transient Security Assessment in Power Systems "PH.D thesis, Univ. of Manchester, Institute of Science and Technology, 1982 .

(13) H. Hakimmashhadi, G. T. Heydt,"Fast Transient Security Assessment" IEEE Trans., Vol. PAS-102, No.12, pp.3816-24, Dec.1983 .

(14) J. M. G. Sa Da Costa, et..al "Pattern Recognition in Power System Security" Int. J. Electric Power and Energy Systems, pp. 31-36, Jan. 1984 .

(15) J. M. G. Sa Da Costa "Transient Security Assessment in Power Systems" Real Time Control of Large Scale Power Systems ,PATRAS, GREECE, pp. 254-61, July 9-12, 1984.

(16) K. Koizumi, O. Saito ,et..al,"Fast Transient Stability Study Using Pattern Recognition" 5-th PSCC (power system computation conference), September 1975.

(17) F. S. Prabhakara,"On Line Transient Stability and Security Evaluation Using Lyapunov and Pattern Recognition Methods" TR-EE 74-29, Purdue University, West Lafayette, August 1974.

(18) C. L. Gupta,"Transient Security Assessment of Power Systems by Pattern Recognition and Liapunov's Direct Method" Ph.D. Thesis, Purdue University, West Lafayette, August 1976.

(19) H. Hakimmashhadi,"Fast Transient Security Assessment" Ph.D. Thesis, Purdue University, West Lafayette, IN, August 1982.

(20) C. L. Gupta, et..al,"Transient Security Assessment of Power Systems by Pattern Recognition -A Pragmatic Approach" IFAC Symposium 1977 on Automatic Control and Protection of Electric Power Systems, Melbourne, Australia, pp. 359-63, Feb. 21-25, 1977.

(21) S. Mokhtari," Fast Transient Security Evaluation of Power Systems by Using Pattern Recognition Techniques",Ph.D. Thesis, University of Missouri-Columbia, Dec. 1983.

(22) N. S. Van Nielen,"New Energy Management System for The Dutch Power Pool," IEEE Trans. Power Systems, Vol. PWRS-2, No. 1, pp. 58-64, Feb. 1987.

(23) F. Mc Dyer, K. Herger,"The New National Control Center, IRELAND," IEEE Trans. Power Systems, Vol. PWRS-2, No. 1, pp. 85-91, Feb. 1987.

(24) R. L. Lugtu, et..al,"The Atlantic Electric System Control Center," IEEE Trans. PAS, Vol. 102, No. 11, pp. 3571-76, Nov. 1983.

(25) A. K. Subramanian, et..al,"Power System Security Functions of the Energy Control Center at the ORANGE and ROCKLAND Utilities "IEEE Trans., Vol. PAS-102, No.12, pp. 3825-34, Dec.1983 .

(26) R. L. Lugtu, et..al,"The OHIO Edison Energy Control Center," IEEE Trans. PAS, Vol. 102, No. 11, pp. 3577-81, Nov. 1983.

(27) K. Nara, R. R. Shoults, M. S. Chen, et..al, "On-line Contingency Selection Algorithm For Voltage Security Analysis," IEEE Trans. on PAS, Vol. 104, No. 4, pp. 847-856, April 1985.

(28) G. C. Ejebe, B. F. Wollenberg,"Automatic Contingency Selection," IEEE Trans. on PAS, Vol. 98, No. 1, pp. 97-109, Jan/Feb 1979.

(29) M. G. Lauby, T. A. Mikolinnas, N. D. Reppen,"Contingency Selection of Branch Outages Causing Voltage Problems," IEEE Trans. on PAS, Vol. 102, No. 12, pp. 3899-3904, December 1983.

(30) R. G. Wasley, M. Daneshdoost,"Identification and Ranking of Critical Contingencies in Dependent Variable Space," IEEE Trans. on PAS, Vol. 102, No. 4, pp. 881-892, April 1983.

(31) T. F. Halpin, R. Fischl, R. Fink,"Analysis of Automatic Contingency Selection Algorithms," IEEE Trans. on PAS, Vol. 103, No. 5, pp. 938-945, May 1984.

(32) G. D. Irisarri, A. M. Sasson,"An Automatic Contingency Selection Method for On-line Security Analysis," IEEE Trans. on PAS, Vol. 100, No. 4, pp. 1838-1844, April 1981.

(33) T. A. Mikolinnas, B. F. Wollenberg,"An Advanced Contingency Selection Algorithm," IEEE Trans. on PAS, Vol. 100, No. 2, pp. 608-617, Feb. 1981.

(34) I. Dabbaghchi, G. Irisarri,"AEP- Automatic Contingency Selector: Branch Outage Impacts on Load Bus Voltage Profile," IEEE Trans. on PAS , Summer Meeting, Vancouver, July 1985.

(35) G. Irisarri, A. M. Sasson, D. Levner,"Automatic Contingency Selection For On-line Security Analysis- Real-Time Tests," IEEE Trans. on PAS, Vol. 98, No. 5, pp. 1552-1559, Sept/Oct 1979.

(36) R. Fischl, T. F. Halpin, A. Guvenis,"The Application of Decision Theory to Contingency Selection," IEEE Trans. on Circuits and Systems, Vol. 29, No. 11, pp. 712-723 , Nov. 1982.

(37) S. Vemuri, R. E. Usher," On-line Automatic Contingency Selection Algorithms," IEEE Trans. on PAS, Vol. 102, No. 2, pp. 346-354, Feb. 1983.

(38) D. H. Foley,"Consideration of Sample and Feature Size" IEEE Trans., Vol. IT-18, No. 5,pp. 618-626, September 1972.

(39) G. F. Hughes,"Number of Pattern Classifier Design Samples Per Class" IEEE Trans., Vol. IT-18, No. 5, pp. 615-618, September 1972.

(40) S. Watanabe," Methodologies of Pattern Recognition " Academic Press, New York, 1969.

(41) P. A. Devijver, J. Kittler," Pattern Recognition: A Statistical Approach ", Prentic / Hall International, Inc., London, 1982.

(42) " Pattern Recognition and Scene Analysis " Course notes by E. Shwedyk and Pawlak, 1986, University of Manitoba.

(43) S. T. Bow," Pattern Recognition: Application to Large Data Set Problems ", Marcel Dekker, Inc., 1984.

(44) J. M. Mendel, K. S. Fu," Adaptive, Learning, and Pattern Recognition Systems: Theory and Applications " Academic Press, Inc. (LONDON) LTD., 1970.

(45) K. S. Fu," Pattern Recognition and Machine Learning " Plenum Press, New York, 1971.

(46) R. O. Duda and P. E. Hart," Pattern Classification and Scene Analysis",John Wiley and Sons, 1973.

(47) F. Denomme, et .. al,"Security Monitoring of the Hydro-Quebec Power System " CEA, Toronto, Ont., March 1984 .

(48) "BMDP Statistical Software Manual", University of California Berkeley, Berkeley, CA, 1985 Edition.

(49) R. I. Jeunrich," Stepwise Discriminant Analysis ", Statistical Methods for Digital Computers, Vol. III, John-willey and Sons, Toronto, 1977, pp. 76-95.

(50) D. F. Specht,"Generation of Polynomial Discriminant Functions for Pattern Recognition" IEEE Trans., Vol. EC-16, No. 3, pp. 308-319, June 1967.

(51) M. I. Olken, W. T. Woelfle," Load Rejection of Subcritical Steam Turbine_Generator Units ", IEEE Trans. PAS, Vol. PAS-85, Feb. 1973, pp. 336-340.

(52) "Self-Excitation Program" Manitoba Hydro User's Manual.

(53) W. Pyl, C. V. Thio," Nelson River Collector System Operating Restrictions and Protection for Self-Excitation and Second Harmonic Resonance ", 61D-01001, SPD Report file, Manitoba Hydro.

(54) L. A. Beattie, " Experience with Power System Security Analysis via Pattern Recognition Techniques ",paper A75 449-9 Presented at IEEE Summer Power Meeting, July 1975.

(55) S. Yamashiro," On-line Secure-Economy Preventive Control of Power Systems by Pattern Recognition ", Proceedings of PICA 1985, pp. 110-115.

(56) S. Yamashiro, T. Koike, A. H. El-Abiad," Fast Transient Security Assessment and Enhancement Using Pattern Recognition ", Proceedings of 8-th PSCC, August 1984.

(57) E. A. Mohamed, G. W. Swift," Security Assessment to Avoid Self-Excitation " IASTED conf., Bozeman, MT, Aug. 20-22, 1986.

(58) E. A. Mohamed, G. W. Swift," Power System Self-Excitation On-line Security Assessment" accepted paper (in press) to be published in the Int. J. of Electrical Power and Energy Systems.

(59) E. A. Mohamed, G. W. Swift,"Prediction of Load Rejection Overvoltages Using Pattern Recognition" a paper under preparation for IEEE Trans., Power Systems.

(60) T. E. Dy Liacco,"The Adaptive Reliability Control System," IEEE Trans. PAS, Vol. 86, No. 5, pp. 517-31, May 1967.

(61) EPRI report EL-2526 by on Power Technologies, Inc.," Transmission System Reliability Methods," July 1982.

(62) A. P. Meliopoulos, A. G. Bakirtizis, R. R. Kovacs, R. J. Beck," Bulk on Power System Reliability Assessment Experience With The RECS Program," proceedings PICA Conference 1985, pp. 38-45.