

**Investigation of novel methodologies using reactive power reserves for online voltage
stability margin monitoring and control**

by

Bruno Henrique Silveira Leonardi

A dissertation submitted to the graduate faculty
in partial fulfillment of the requirements for the degree of

DOCTOR OF PHILOSOPHY

Major: Electrical Engineering

Program of Study Committee:
Venkataramana Ajjarapu, Major Professor
Chen-Ching Liu
Manimaran Govindarasu
Sigurdur Olafsson
Umesh Vaidya

Iowa State University
Ames, Iowa
2011

Copyright © Bruno Henrique Silveira Leonardi, 2011. All rights reserved.

DEDICATION

This thesis is dedicated to my parents, Jamil and Vitoria Leonardi, and my fiancée Andrea Bekic, for their unconditional love and support.

TABLE OF CONTENTS

LIST OF FIGURES	vii
LIST OF TABLES	xi
ABSTRACT	xiv
CHAPTER 1. OVERVIEW OF THE PROBLEM AND RESEARCH CONTRIBUTIONS	1
1.1 Introduction.....	1
1.2 Research contributions.....	4
1.2.1 A novel online VSM monitoring tool	4
1.2.2 Sensitivity based real-time control to enhance RPRs and VSM.....	5
CHAPTER 2. LITERATURE REVIEW	7
2.1 Online voltage stability margin monitoring techniques.....	7
2.2 Multilinear regression models and machine learning techniques	9
2.3 Multiclass classification techniques.....	11
2.4 Online RPR and VSM control techniques	12
CHAPTER 3. ON THE RELATIONSHIP BETWEEN REACTIVE POWER RESERVES AND VOLTAGE STABILITY MARGIN	14
3.1 Introduction.....	14
3.2 Aspects of voltage stability	16
3.2.1 Classification of voltage instability	16
3.3 Definition of voltage stability margin.....	17
3.4 Voltage instability mechanisms	19
3.5 Reactive power reserve definitions	21
3.6 Relationship between RPRs and VSM	23
3.7 Effect of capacitor/reactor banks	27

3.8	Investigation of RPR definitions on performance	30
3.8.1	Effect of polynomial order.....	30
CHAPTER 4. DEVELOPMENT OF THE ONLINE VSM MONITORING		
	METHOD	37
4.1	Database generation and information flow	37
4.2	PSS/E® automation using python	38
4.3	A two-Stage VSM estimation methodology	42
CHAPTER 5. THE USE OF MULTI-LINEAR REGRESSION MODELS FOR		
	ONLINE VSM ESTIMATION	45
5.1	Mathematical formulation of MLRMs.....	45
5.2	Uncertainty in system stress direction	47
5.3	Power system modeling and simulation	49
5.4	Using MLRMs to relate RPRs and system VSM.....	51
5.5	MLRM developmental procedure.....	53
5.6	MLRM validation procedure	55
5.6.1	Homoskedasticity.....	56
5.6.2	Normality of the residuals.....	58
5.6.3	Hypothesis test	58
5.6.4	Multicollinearity	60
5.7	Results on sample systems.....	60
5.7.1	Description of test systems	61
5.7.2	The design of MLRMs.....	64
5.8	Validation of MLRMs.....	67
5.8.1	Homoskedasticity verification	67
5.8.2	Verification of normality	70

5.8.3	Test of overall regression aptness	72
5.9	Conclusions	73
CHAPTER 6. MACHINE LEARNING TECHNIQUES FOR MULTICLASS		
	CLASSIFICATION	75
6.1	Data mining overview	75
6.2	Investigated algorithms	76
6.2.1	Decision trees	76
6.2.2	Artificial Neural Networks (ANN)	81
6.2.3	Instance based learning and K-Nearest-Neighbor (KNN)	85
6.2.4	Support vector machines	89
6.3	Multiclass classification problem	91
CHAPTER 7. SIMULATION RESULTS FOR THE MLRM-IDTOOL		
7.1	Results on the IEEE-30 bus test system	97
7.1.1	Decision tree based classifiers	97
7.1.2	Artificial Neural Network based classifiers	102
7.1.3	KNN based classifiers	108
7.1.4	SVM based classifier	113
7.1.5	Stacked classifier	115
7.2	Results on the 22k bus system	117
7.2.1	Development of the MLRM-IDtool	117
7.3	Conclusions	135
CHAPTER 8. APPLICATION OF REACTIVE POWER RESERVE		
SENSITIVITIES FOR REAL TIME VOLTAGE STABILITY		
	MARGIN CONTROL	136
8.1	Introduction	136

8.2	Reactive Power Reserve Sensitivities	139
8.3	On the selection of reactive power reserves	144
8.4	Proposed control methodology	145
8.5	Simulation results.....	150
8.5.1	IEEE 30 bus test system.....	150
8.5.2	IEEE 118 bus system	159
8.6	Conclusions.....	165
CHAPTER 9. FINAL CONCLUSIONS.....		167
9.1	Discussions on the importance of this work	167
9.2	Future research.....	168
PUBLICATIONS.....		171
ACKNOWLEDGEMENTS		172
BIBLIOGRAPHY		174

LIST OF FIGURES

Figure 2.1. Different multiclass classification methods.....	11
Figure 3.1. Identification of the type of phenomena studied (shaded boxes)	16
Figure 3.2. Effect of different contingencies on system voltage stability margin	18
Figure 3.3. Capability curve of synchronous generator	21
Figure 3.4. PV diagram	22
Figure 3.5. The IEEE30-bus test system.....	24
Figure 3.6. RPR and VSM relationship for generator 5	25
Figure 3.7. RPR and VSM relationship for generator 11	26
Figure 3.8. Local portion of the 1648 bus test system	28
Figure 3.9. Effect of capacitor reactor banks on rotating RPRs	28
Figure 3.10. Capacitor bank reactive reserve variations with VSM.....	29
Figure 3.11. Comparison between linear model and higher order model.....	31
Figure 3.12. Comparison between linear and higher order model.....	32
Figure 3.13. Comparison between linear and nonlinear model	32
Figure 3.14. Comparison between linear and nonlinear model	33
Figure 3.15. Comparison between linear model and quadratic model	34
Figure 3.16. Comparison between linear model and quadratic model	35
Figure 3.17. Comparison between linear and quadratic model	36
Figure 4.1. Overall methodology description	37
Figure 4.2. Database automation process	39
Figure 4.3. Python - PSS/E® interface developed for software automation	40

Figure 4.4. Overall implementation of the proposed online VSM estimation tool.....	43
Figure 5.1. Variations in VSM due to different load increase directions and network topology	48
Figure 5.2. RPR and VSM relationship for generator 11	51
Figure 5.3. Flowchart representing MLRM development and validation stages.....	54
Figure 5.4. Homoskedastic residual distribution	56
Figure 5.5. Heteroskedasticity residual distribution	57
Figure 5.6. North American interconnections.....	62
Figure 5.7. Residual plot to verify presence of heteroskedasticity or statistical inconsistency on the residuals – IEEE30/MLRM-1	68
Figure 5.8. Residual plot to verify presence of heteroskedasticity or statistical inconsistency on the residuals – EI/MLRM-2	69
Figure 5.9. Residual histogram and best normal <i>pdf</i> of MLRM-1.....	70
Figure 5.10. Residual histogram and best normal <i>pdf</i> of MLRM-2.....	71
Figure 6.1. Partial representation of decision tree brach	77
Figure 6.2. Attribute line active power flow after discretization	81
Figure 6.3. Artificial neural network structure	82
Figure 6.4. Mathematical representation of an ANN.....	83
Figure 6.5. Comparison of learning approaches	86
Figure 6.6. The K-nearest neighbor approach	88
Figure 6.7. Support vector machine concept	90
Figure 6.8. Different multiclass classification methods.....	93
Figure 6.9. Hierarchical structure of the MLRM-IDtool	94

Figure 7.1. Attribute analysis under unforeseen LIDs for DTs	119
Figure 7.2. DT created using P_{flow} and V_{mag} as attributes	121
Figure 7.3. DT attribute analysis under unforeseen contingencies and LIDs	122
Figure 7.4. Representation of different amounts of noise added to the actual data	124
Figure 7.5. DT performance on noisy data for unforeseen LIDs	125
Figure 7.6. DT created using P_{flow} and V_{mag} as attributes and trained with 15% noise	126
Figure 7.7. DT performance on noisy data for unforeseen contingencies and LIDs	127
Figure 7.8. KNN attribute analysis under unforeseen LIDs	129
Figure 7.9. KNN attribute analysis under unforeseen contingencies and LIDs	130
Figure 7.10. KNN performance on noisy data for unforeseen LIDs	132
Figure 7.11. KNN performance on noisy data for unforeseen contingencies and LIDs	133
Figure 8.1. Capability curve of a synchronous generator: amount of RPRs depends on operating condition (A, B or C) and machine limits	140
Figure 8.2. PV curves depicting the linear control strategy proposed.	141
Figure 8.3. Flowchart describing the proposed approach.	148
Figure 8.4. Sensitivities of RPR of generator 11 (Q_{R11}) with respect to various control parameters	152
Figure 8.5. Effect of generation active power reduction on RPRs	153
Figure 8.6. Effect of switched shunts on RPRs	155
Figure 8.7. Voltage profile enhancement with switched shunts	156
Figure 8.8. VSM enhancement with shunt switch	156

Figure 8.9. Effect of load shedding on RPRs.....	157
Figure 8.10. Voltage profile enhancement with load shedding	158
Figure 8.11. VSM enhancement with load shedding.....	159
Figure 8.12. Effect of diverse set of control actions on RPRs – same control costs	161
Figure 8.13. Voltage profile enhancement with a diverse set of control actions – same control costs	162
Figure 8.14. VSM enhancement at with a diverse set of control actions – same control costs	162
Figure 8.15. Effect of diverse set of control actions on RPRs – different control costs.....	164
Figure 8.16. Voltage profile enhancement with a diverse set of control actions – different control costs	164
Figure 8.17. VSM enhancement at with a diverse set of control actions – different control costs	165

LIST OF TABLES

Table 4.1: System variables sampled during VSA	41
Table 5.1. Description of the IEEE 30 bus test system.....	61
Table 5.2. Description of the United States eastern interconnection	62
Table 5.3. Reduced case of the United States eastern interconnection.....	63
Table 5.4. Multilinear regression models and respective confidence bounds for IEEE 30 bus test system.....	65
Table 5.5 Multilinear regression models and respective confidence bounds for reduced case of the Eastern Interconnection.....	66
Table 6.1. Discretization of a numeric attribute	80
Table 6.2. Distances commonly used in the KNN approach	87
Table 7.1. Confusion matrix of DT.....	98
Table 7.2. Confusion matrix of DT 1.....	98
Table 7.3. Confusion matrix of DT-2	99
Table 7.4. Confusion matrix of DT-3	99
Table 7.5. Final classification precision of the DT based hierarchical classifier	99
Table 7.6. Confusion matrix of OVA DT-based classifier	100
Table 7.7. Confusion matrix of OVO DT-based classifier	101
Table 7.8. Confusion matrix of DT-based AdaBoost classifier.....	101
Table 7.9. Code words used to transform nominal class attributes into numeric	102
Table 7.10. Confusion matrix of single ANN classifier	103
Table 7.11. Confusion matrix of ANN-1	104

Table 7.12. Confusion matrix of ANN-2	104
Table 7.13. Confusion matrix of ANN-3	105
Table 7.14. Final classification precision of the ANN based hierarchical classifier	105
Table 7.15. Confusion matrix of OVA ANN-based classifier	106
Table 7.16. Confusion matrix of OVO ANN-based classifier	107
Table 7.17. Confusion matrix of ANN-based AdaBoost classifier	108
Table 7.18. Confusion matrix of single KNN classifier	108
Table 7.19. Confusion matrix of KNN-1 classifier	109
Table 7.20. Confusion matrix of KNN-2 classifier	110
Table 7.21. Confusion matrix for the KNN-3 classifier	110
Table 7.22. Final classification precision of the KNN based hierarchical classifier	111
Table 7.23. Confusion matrix of OVA KNN-based classifier	111
Table 7.24. Confusion matrix of OVO KNN-based classifier	112
Table 7.25. Confusion matrix of KNN-based AdaBoost classifier	112
Table 7.26. Confusion matrix of OVA SVM-based classifier	114
Table 7.27. Confusion matrix of OVO SVM-based classifier	114
Table 7.28. Confusion matrix of SVM-based AdaBoost classifier	115
Table 7.29. Confusion matrix of DT-1	116
Table 7.30. Confusion matrix of DT-2	116
Table 7.31. Confusion matrix of ANN-3	116
Table 7.32. Final classification precision of the hierarchical stacked classifier	117
Table 7.33. Classification precision of DT developed using P_{flow} and V_{mag}	121

Table 7.34. Classification precision of DT trained using P_{flow} and V_{mag} and 15% noise	127
Table 7.35. Classification precision of DT developed using P_{flow} and V_{mag}	130
Table 7.36. Classification precision of KNN trained using P_{flow} and V_{mag} and 15% noise	134
Table 8.1. Sensitivity of RPR with respect to different control actions	150
Table 8.2. Control amounts using load shed as control action	154
Table 8.3. IEEE 30 – Control Amounts Using Load shed as control action	158
Table 8.4. IEEE 118 – Amount of control considering same weight for ΔP_G , ΔB_{sh} , ΔP_L and ΔQ_L	160
Table 8.5. IEEE 118 – Amount of control considering weight 1 for ΔB_{sh} /weight 50 for ΔP_L and ΔQ_L	163

ABSTRACT

As the amount of uncertainty in online power system operations grows, new methodologies need to be devised in order to timely monitor and control the power grid. In this work, novel techniques for online voltage stability margin monitoring and control have been developed with a focus on reactive power reserves.

The maintenance of adequate reactive power reserves (RPRs) is a critical step in avoiding a voltage collapse. A thorough investigation of the relationship between different definitions of reactive power reserves and how they are related to voltage stability margin (VSM) is performed.

Multi-linear regression models are used to relate RPRs and VSM. Several operating conditions and a significantly large number of different network topologies, including NERC category B, C and D outages are considered as well. A classification tool is then developed in order to identify which regression model needs to be used based on system conditions and network topology. The approach is tested in the IEEE 30 bus test system and in a reduced case of the eastern power system interconnection of the United States. Results have shown that the approach can monitor voltage stability margin in real time based on the amount of system wide reactive power reserves.

In case degenerative system conditions are identified, control actions need to be put in place to increase the amount of RPRs and system VSM. A novel control method is proposed here in order to identify the location and amount of control necessary to recover RPRs, VSM and to remove existing voltage violations. The approach is based on the identification of a critical set of generators that, if exhausted, will directly contribute to a voltage collapse.

Potential control actions are investigated in order to recover those critical reactive power reserves, namely: active power re-dispatch, capacitor switching, active and reactive power load shedding. The effectiveness of each control variables on RPRs is calculated using reactive power reserve sensitivities, a concept introduced in this work. Once these sensitivities are calculated, the problem of recovering RPRs and VSM is formulated as convex quadratic optimization problem with a reduced dimension.

Results on the IEEE 30 bus test system and the IEEE 118 bus test system are used to illustrate the efficacy of the approach.

CHAPTER 1. OVERVIEW OF THE PROBLEM AND RESEARCH CONTRIBUTIONS

1.1 Introduction

The current lack of investment of transmission expansion associated with the aging of the current network infrastructure creates significant challenges for the stable operation of the power grid in the United States. Growing environmental concerns, combined with the lack of clear economical incentives for private investors have delayed and slowed the expansion of the transmission network.

On the other hand, load continues to grow in the system despite the physical constraints imposed by the transmission grid. If investments are not properly executed on time, the US power grid will slowly be pushed to its operational limits. In such an uncertain and critical scenario, voltage instability becomes a serious threat and large-scale blackouts are prone to occur.

In the past 15 years, the North American power grid has undergone two major reactive power related instability events. A description of the blackout which occurred in the western electric coordinating council (WECC) on august 10, 1996 can be found in Taylor, C. W. (1998). According to the author, the lack of proper reactive power support and loss of voltage control have caused a classic case of voltage collapse.

Another example of a recent large-scale event caused in part by improper reactive power support is the North American blackout in August 2003, NERC (2004). A detailed investigation of the event pointed out that not only a single factor caused the widespread blackout, but rather a combination of failures. The report has concluded that inappropriate reactive power support

played a major role in the events. An analysis performed by Anderson, P. and Geckil, I. (2003) has estimated the economical impact associated with the blackout to be in the range of 4.5 to 8.5 billion dollars. These costs include lost income to workers and investors, extra costs to government agencies (e.g., due to overtime and emergency service costs), the affected utilities, and lost or spoiled commodities, not accounting for the direct impact on people's lives.

In order to improve reactive power management and prevent voltage collapse events as the ones presented above, the North American Electric Reliability Corporation (NERC) has issued several reliability standards related to real time RPR monitoring and voltage control VAR-001-1 (2006), VAR-001-2 (2010), TOP-006-1 (2006) and TOP-006-2 (2008). Real time RPR monitoring has also been identified as one of the recommended actions in order to reduce the likelihood of future system blackouts as described in NERC recommendations (2004). Although the aforementioned standards and remedial actions may enhance operator's awareness regarding the amount of reactive power reserve available, it may not provide quantitative information regarding how far the system is from a voltage collapse.

However, the Federal Energy Regulatory Commission (FERC) has questioned the efficacy of the aforementioned standards by saying that system operators cannot gain *situational awareness* by simply viewing massive amounts of raw data. According to FERC (2006), "...while the requirements identify the data to be gathered, they fail to describe the tools necessary to turn that data into critical reliability parameters." Therefore, there is a need to develop alternative tools that can process the massive amount of data gathered into the supervisory control and data acquisition (SCADA) system. These tools have the objective to identify situations where voltage instability is a real threat.

In addition to that, the development and integration of new and improved operation techniques and methods is one of the goals of the smart grid initiative (SGI). In short, the SGI has the objective to modernize and maintain a reliable, efficient and secure operation of the United States power grid, according to the *Energy independence and Security Act of 2007*, EISA (2007), USDOE (2007).

Based on the aforementioned needs and objectives, a methodology to transform real time monitored RPRs and other SCADA measurements into VSM information is devised in this research. A weighted summation of system wide RPRs is used to estimate how much VSM is still available. The approach uses multi-linear regression models (MLRMs) to relate RPRs and VSM for a large set of contingencies and different load increase scenarios. Statistical analysis of the MLRMs is further performed in order to verify basic statistical properties for these models.

In case a few MLRMs are necessary, a decision tool needs to be developed in order to facilitate operator's choice while selecting the appropriate MLRM. Such tool can use other raw system measurements, such as line flow and bus voltage magnitudes. Several multiclass classification methods have been studied and analyzed. It is important to mention that the MLRM-IDtool will not only facilitate the operators to choose the right MLRM t, but it will also reduce the likelihood of human error involved in the selection process.

Several meta-learner techniques have been tested in order to identify the most suitable approach for this problem. A detailed comparison of the complexity and accuracy of each technique to the problem at hand is performed to determine the most appropriate technique. The results have shown that for certain cases, a single decision tree can successfully classify all MLRMs accurately. More complex classification techniques like a stacked hierarchical classification method are used when the MLRM identification is harder.

The final online VSM tool, which is composed of the MLRM-IDtool and the MLRMs, is then tested on the IEEE 30 bus and on a reduced case of the eastern interconnection with promising results.

Additionally, a control methodology based on RPR sensitivities to control actions is proposed to improve RPRs and VSM, while maintaining system voltages within normal operational limits. These linear RPR sensitivities with respect to control actions are calculated in order to identify the most effective controls. Active power generation, shunt capacitors/reactors and load shedding are investigated as potential candidates for control. A few critical RPRs are exhausted at the point of collapse and only the control actions that are most effective in reestablishing those RPRs are included in the control search.

A convex quadratic optimization problem is formulated to identify the minimal amount of control necessary to recover the critical RPRs and VSM to pre-specified offline limits. Tests on the control methodology on the IEEE30 and IEEE 118 help to illustrate the efficacy of the methodology.

1.2 Research contributions

1.2.1 A novel online VSM monitoring tool

- The online VSM monitoring approach can handle all different NERC types (B, C and D) of contingencies, from N-1 to N-k contingencies. The number of contingencies is not limited and can be as large as required by system operators.
- A reduced number of MLRMs is necessary to cover a wide variety of operating conditions and network topologies.

- Confidence intervals help the models account for uncertainty in load behavior, a situation commonly experienced in practical applications due to imprecision of load forecasting tools.
- The methodology converts raw data (RPRs) into meaningful information about proximity to voltage collapse. Margin estimation is provided in MW so that system operators can easily interpret and take appropriate actions, thereby allowing VSM estimation to be done in the online operating environment.
- The methodology can be applied to large, real-sized networks as demonstrated in the results section, enabling operators to perform wide area VSM monitoring.

1.2.2 Sensitivity based real-time control to enhance RPRs and VSM

- A fast and efficient method to identify control actions for RPR and VSM enhancement in real time.
- The methodology introduced the concept of RPR sensitivity to control action taking into account the capability curve of synchronous machines. These sensitivities can help operators identify the most effective control actions for each individual RPR.
- By identifying the most effective control actions, it is possible to significantly reduce the dimensionality of the optimization problem. Such dimensionality reduction is useful for real time operations since it reduces the number of control variables in the optimization problem, thereby allowing the approach to be implemented in an online fashion. Moreover, the use of linear sensitivities enables the search for optimal control to be formulated as a convex quadratic optimization problem.

- The solution of the optimization problem determines the minimal amount of control to be applied to the system in order to recover safe levels of RPRs and VSM. Simulation results on the IEEE 30 bus and IEEE 118 bus test systems are used to demonstrate the efficacy of the method.

CHAPTER 2. LITERATURE REVIEW

2.1 Online voltage stability margin monitoring techniques

Although there is large number of publications in the literature that relates the impacts of RPR management and improvements of VSM, very little research has been done in the area of online VSM monitoring through RPRs.

It is well known that the depletion of reactive power reserves (RPRs) directly impact voltage stability margin, Taylor, C. (1994), Van Cutsen, T. and Vournas, C. (1998) and Ajjarapu, V. (2006). NERC has issued several standards related to online monitoring of RPRs available in control areas across the United States power grid. The argument used by NERC is that by monitoring RPRs in real time, operator's awareness regarding proximity to voltage collapse may increase.

However, the Federal Energy regulatory adverts in FERC (2006) that system operators cannot gain situational awareness by simply viewing massive amount of real time data. It also indicates that new data applications are needed in order to convert raw data into meaningful information to operators. Therefore, there is a clear need for tools that can process the massive amount of data that is currently being streamed into SCADA/EMS systems into sensitive information to system operators. Aiming to fulfill the aforementioned needs, this research has developed a technique to convert online monitored RPRs into real time VSM estimation to system operators.

Previous studies have shown that proper RPR management directly improves system wide VSM. In Dong, F. et al. (2005), an optimal strategy attempts to maximize the amount of VSM while maintaining adequate voltages and thermal limits in the transmission lines. The problem is

formulated as a security constrained optimal power flow with Bender's decomposition accounting for different topologies of the grid. Results have demonstrated that an increase in overall RPR leads to an increase in VSM.

El-Keib, A. A. and Ma, X. (1995) have attempted to relate inputs from SCADA system to the amount of voltage stability margin. The authors used an Artificial Neural Network (ANN) for a specific loading condition and network topology, demonstrating that accurate estimations of VSM could be obtained. The downside of the approach is that different operating conditions and network topologies can emerge, causing the number of needed ANNs to grow out of proportion.

Jeyasurya, B. (2000) has attempted to use an ANN to estimate voltage stability margin in real time. The training data is composed of generator terminal voltage, real and reactive power output of the generators, reactive power reserve of the generators and active/reactive load demands. Principal component analysis is used here to reduce the dimensionality of the problem before the ANN design. Although good accuracy has been obtained, the author only attempted simulations on three different network topologies. A different ANN is used on each network topology, thereby indicating that the practical implementation of the approach may become cumbersome in case hundreds of topologies are considered.

Jimenez, C.A. and Castro, C. A. (2005) and have also used artificial neural networks (ANN) to estimate VSM directly. System variables such as voltage magnitudes, active/reactive power generation, real and reactive system load and active and reactive power in transmission lines are used as inputs to the ANN. The approach is tested on small systems with satisfactory performance.

Bao, L. et al. (2003) were the first to explore the RPRs as potential indicators of VSM. In this initial work, the authors mention that a linear relationship between RPRs and VSM seems to

exist. An online voltage stability monitoring system using the concept of equivalent reactive power reserve is proposed. The methodology is tested in a real-sized system for a particular operating condition and a selected list of N-1 contingencies. Results have shown that reactive reserves can be valuable indicators of voltage stability margin.

However, there is a need to account for uncertainty in loading conditions and changes in system network topology during online VSM estimation. System loading conditions can vary significantly during the day or seasons (spring, summer, fall and winter). The effect of load increase direction must be investigated and taken into account for it directly affects the amount of VSM. The methodologies proposed up to date have not demonstrated their performance when different load increase conditions are considered. Moreover, the effect of other NERC category B (N-1), C (N-2) and D (N-k) contingencies must be accounted for since contingencies more severe than N-1 can occur at any time in the system.

2.2 Multilinear regression models and machine learning techniques

Multilinear regression models (MLRM) are among one of the most widely used and well established statistical tools for inference making and prediction, Kleinbaum, D. et al. (1998) and Kutner, M. et al. (2004). Successful applications can be found in fields ranging from semiconductor manufacturing to economics Lin, Z. C. and Wu, W. J. (1999) and Studenmund, A. H. (2006). Similar to other statistical models, it has the ability to model uncertainty by adding confidence intervals to the point estimates. This feature can be of great value in order to handle load growth and network topology uncertainties.

Therefore, the design of a robust online VSM estimation tool is still a challenging topic. In this study, we have overcome some of the current limitations present in the previously described

techniques. The proposed approach incorporates uncertainties related to load increase direction (LID) and network topological changes and a wide range of operating conditions are considered.

A comprehensive voltage stability assessment (VSA) is performed to capture the behavior of RPRs, VSM and other SCADA data. Several LIDs and network topologies are used to account for system network configuration changes and uncertainty in load growth. Once a large amount of data is gathered from simulations, MLRMs are trained offline using the database and are further used to estimate VSM in online operations.

As more than one MLRM may be necessary, a decision tool is designed in order to identify the most appropriate MLRM for given system conditions. This identification tool is named multi-linear regression model identification tool (MLRM-IDtool). Several meta learning techniques have been investigated to form the multiclass classifier. The investigated algorithms are decision trees (DTs) Quinlan, J. R. (1993) and Breiman, L. et al. (1984); artificial neural networks (ANNs) Bishop, C. M., (1995); k-nearest neighbors (KNN) Aha, D. et al. (1991) and Aha, D. (1992); and support vector machines (SVM) Cortes, C. and Vapnik, V. (1995), Burges, C. J. (1998).

Although four different techniques have been investigated, ANN and SVM based classifiers have shown to require longer training times compared to DTs and KNNs. On the other hand, ANNs have shown to produce good results with high classification accuracies. Overall, the accuracy of the four investigated techniques varies significantly depending on how they are used to perform multiclass classification. Therefore, preference will be given to the techniques that are simpler to implement, easier to visualize and obviously can maintain the specified performance requirements.

2.3 Multiclass classification techniques

Multiclass classification problems are usually developed in two general ways: by *extending* binary classification algorithms or by *decomposing* the problem into several binary classification problems, Duda, R. O. et al. (2000), Aly, M. (2005).

The advantage of ANNs, KNNs and DTs is that they can be used in both extension and decomposition approaches. This characteristic gives them great flexibility compared to other techniques as they can be fit into different methods for performance enhancement. On the other hand, SVMs can only be used for multiclass classification if the multiclass problem is decomposed into several binary problems as it is an inherently binary classifier.

Among the most popular decomposition techniques, there approaches have demonstrated accurate performance. These approaches are the *one versus one* method, Hastie, T. and Tibshirani, R. (1998) and Friedman, J. (1996), the *one versus all* method, Rifkin, R. and Klautau. A. (2004) and the *hierarchical classification* method, Kumar et al. (2002) and Chen et al. (2004).

Figure 2.1 shows how these three different techniques can be differentiated.

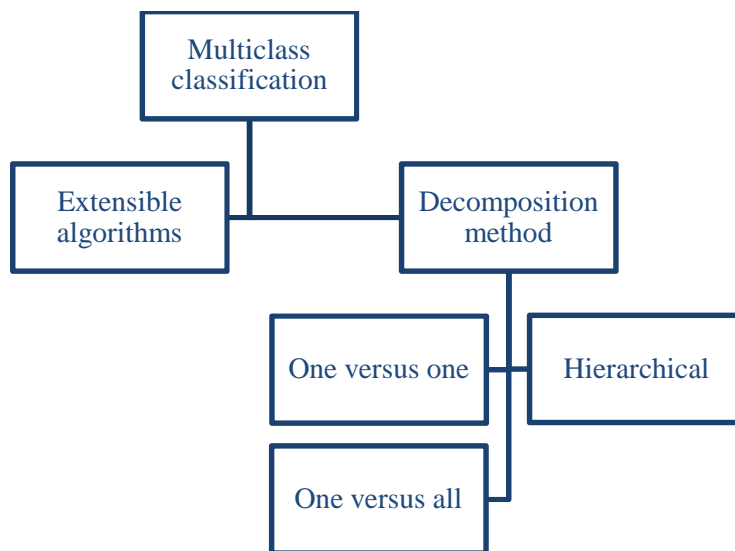


Figure 2.1. Different multiclass classification methods

The major objective is to develop a classifier that makes use of any of the above mentioned methods and achieves very high classification accuracy. It is important to mention that the exhaustive investigation of multiple meta learning techniques for multiclass classification is not the objective of this work.

2.4 Online RPR and VSM control techniques

In the second part of this research, a method is devised in order to identify control actions that can enhance RPRs and VSM in real time. Several studies have demonstrated that the amount of VSM directly increases with an increase of RPRs and different preventive/corrective control approaches have been proposed.

Vaahedi, E. et al. (2001) have proposed a planning VAR method considering possible contingencies in a planning horizon. Margin requirements are incorporated in the approach which is formulated as a nonlinear optimization problem and solved using Bender's decomposition method.

In Dong, F. et al. (2003), a dual objective optimization approach to maximize the amount of RPRs and reduce system losses is proposed. Simulation results have shown that the amount of voltage stability margin increased with an increase of reactive power reserves. The approach used a nonlinear optimization framework based on optimal power flow and Bender's decomposition to determine the best current operating condition.

The concept of reactive reserve based contingency constrained optimal power flow (RCCOPF) is introduced in Song, H. et al. (2003). An optimal power flow framework is used to identify the minimal amount of RPRs necessary in order to improve the amount of voltage stability margin for various contingencies and operating conditions. Implementation of the

approach shows that the amount of voltage stability margin is improved and that the found system state (power flow solution) corresponds to the minimum effective RPR.

The aforementioned approaches are based on variations of nonlinear optimal power formulation, thus being adequate for day ahead planning and/or offline applications. They can also be used to determine adequate levels of RPRs based on the study of different scenarios and contingencies. However, if uncertainties involved with real time operations reduce the RPRs beyond safe limits, control actions should be quickly identified and deployed to avoid further voltage profile degeneration and, in the worst case, a voltage collapse.

In this study, a methodology is proposed to address the problem of real time voltage stability through the enhancement of critical RPRs and system VSM. The method is expected to be used in emergency situations when low amounts of RPRs, VSM or voltage violations are observed.

Sensitivities of control actions of critical RPRs will be used to determine the optimal amount and location of control. The control search is then formulated as a quadratic convex optimization problem, which can be solved quickly.

Now that the review of the state of the art methods in the area is presented, the relationship between RPRs and VSM will be studied next.

CHAPTER 3. ON THE RELATIONSHIP BETWEEN REACTIVE POWER RESERVES AND VOLTAGE STABILITY MARGIN

3.1 Introduction

It is well known that voltage instability events usually take place in conditions where there is a shortage of reactive power supply and consequent loss of voltage control. Nonetheless, there are not many references studying the relationship between RPRs and VSM in the literature.

Voltage collapse usually occurs in heavily loaded systems that do not have satisfactory local reactive reserves and consequently cannot maintain a secure voltage profile across the system. Heavily loaded systems not only have high active and reactive power demand, but also have high reactive power losses in the transmission lines. This high demand of reactive power requires generating units to push reactive power production to their limits, hence causing them to eventually hit their capability curve limit and lose the ability to control terminal voltage.

Several voltage stability related incidents have occurred in the US and worldwide in the past decades. Improper reactive power management combined with lack of situational awareness of the local and regional grid played an important role in recent 2003 North America power blackout, as described in NERC (2004).

While the effect of RPRs on system stability is widely acknowledged, few studies have been conducted to investigate how RPR levels could be used to indicate the amount of VSM. In Dong et al. (2005), a methodology for maximizing reactive power reserves in critical areas is proposed. A nonlinear optimal power flow strategy is formulated considering various system scenarios. Bender's cut decomposition method is used to model the different system conditions. The

problem is later solved using the interior point method. Results have shown that by increasing system wide reactive power reserves, the amount of voltage stability margin also increases.

Bao, L. et al (2003) initially proposed a voltage stability margin tool that uses RPRs to estimate VSM. The authors used the concept of equivalent reactive power reserve and margin estimations are obtained by modeling these equivalent reactive power reserve. Static (capacitor banks) and dynamic (synchronous generators) RPRs are selected and an optimization procedure defines how much each machine contributes to the equivalent RPR.

Attempting to further investigate and better understand the relationship between RPRs and VSM, a detailed investigation of different types of RPR and system VSM is performed in Leonardi, B. and Ajjarapu, V. (2008). The major objective of this investigation is to determine the nature of the relationship between RPR and VSM.

Four different definitions of RPRs have been tested and practical aspects related to practical implementation have been explored. A thorough investigation of how different RPR definitions are related to VSM is reported in this chapter. Results demonstrated that although some RPR definitions have a better relationship with VSM, aspects related to practical implementation may prevent their use for online RPR and VSM monitoring.

Initially, a description of the aspects involved in a voltage collapse incident and the various voltage stability classifications are introduced. A summary of the results obtained from the study is given in detail along this section.

3.2 Aspects of voltage stability

3.2.1 Classification of voltage instability

Researchers in the area of voltage stability know that adequate reactive power support has vital importance on voltage regulation and control. A power system can be considered relatively safe (as far as voltage collapse is concerned) if a reasonable gross amount of RPRs are available in order to maintain voltages during emergency conditions in the system. These emergency conditions may arise from several situations, with the most common being equipment failure, system faults and high load conditions.

Voltage instability phenomenon can be broadly divided into two forms: *short term and long term voltage instability* Kundur, P. et al. (2004). In this research, we have addressed problems related to long term voltage instability events for both small and large disturbances. Figure 3.1 shows the CIGRE/IEEE international classification standard of all voltage instability phenomena.

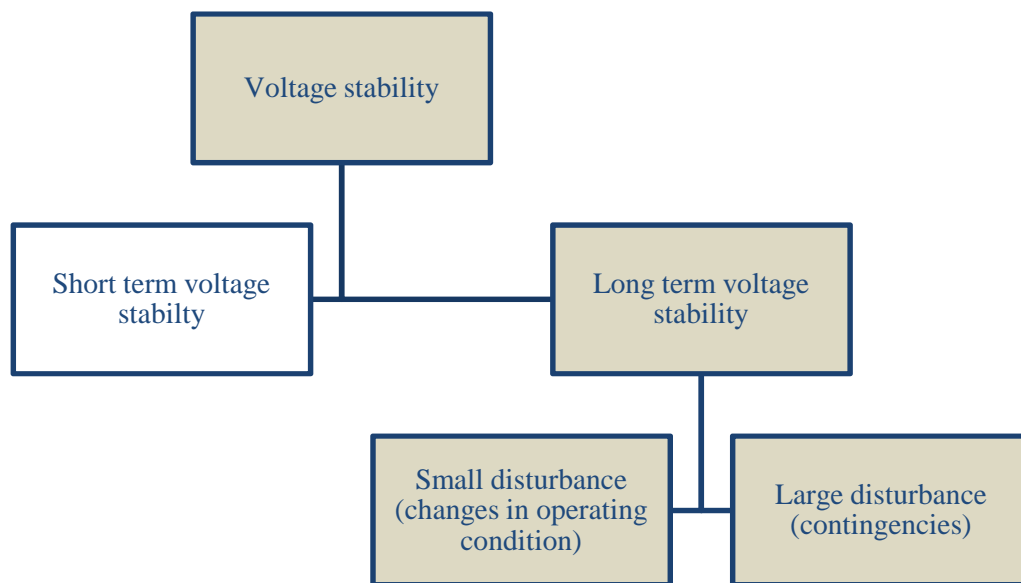


Figure 3.1. Identification of the type of phenomena studied (shaded boxes)

Long term voltage instability problems can be further divided into two sub areas named *structural failures* or *progressive load increase*. *Structural failures* address outages of any equipment in the power grid. System outages can be planned or unplanned, i.e., they depend on maintenance schedule and weather conditions, among other factors. A non-exhaustive list of system equipment would include generators, transmission lines, transformers, capacitor/reactor banks and static VAR compensators (SVCs).

If the outages are caused due to maintenance, it is called a planned outage. On the other hand, if the outage is caused by system faults or equipment failure during regular operations, it is called an unplanned outage. Removal of system equipment may prevent certain areas of the system from receiving an adequate amount of reactive power support, which can in turn lower the voltage profile across the area. Similarly, outage of vital equipment used to provide voltage support (e.g., generators, capacitor/reactor banks, SVCs, etc) would further exacerbate voltage problems.

Progressive load increase is usually used to see until what point the system can sustain load demand. It is by increasing system load so that generating units respond to it until the system cannot find a converged solution and a voltage collapse is observed. This is a typical stress method in order to analyze long term voltage stability.

In order to determine the effects caused by equipment outage on system voltages and VSM, we first need to introduce and define those concepts. Definition of VSM and the mechanisms that may lead the system into voltage instability are described next.

3.3 Definition of voltage stability margin

In order to explain how long term voltage instability occurs, we first need to introduce the concept of voltage stability margin (VSM). Voltage stability margin is defined as the difference

between the maximum possible load in the system and the current load of the system (in MW) as shown in Figure 3.2. This graph is widely known as the PV curve, where P represents the total amount of active power increased in the respective system (or area) and V represents the voltage at one load bus in the system.

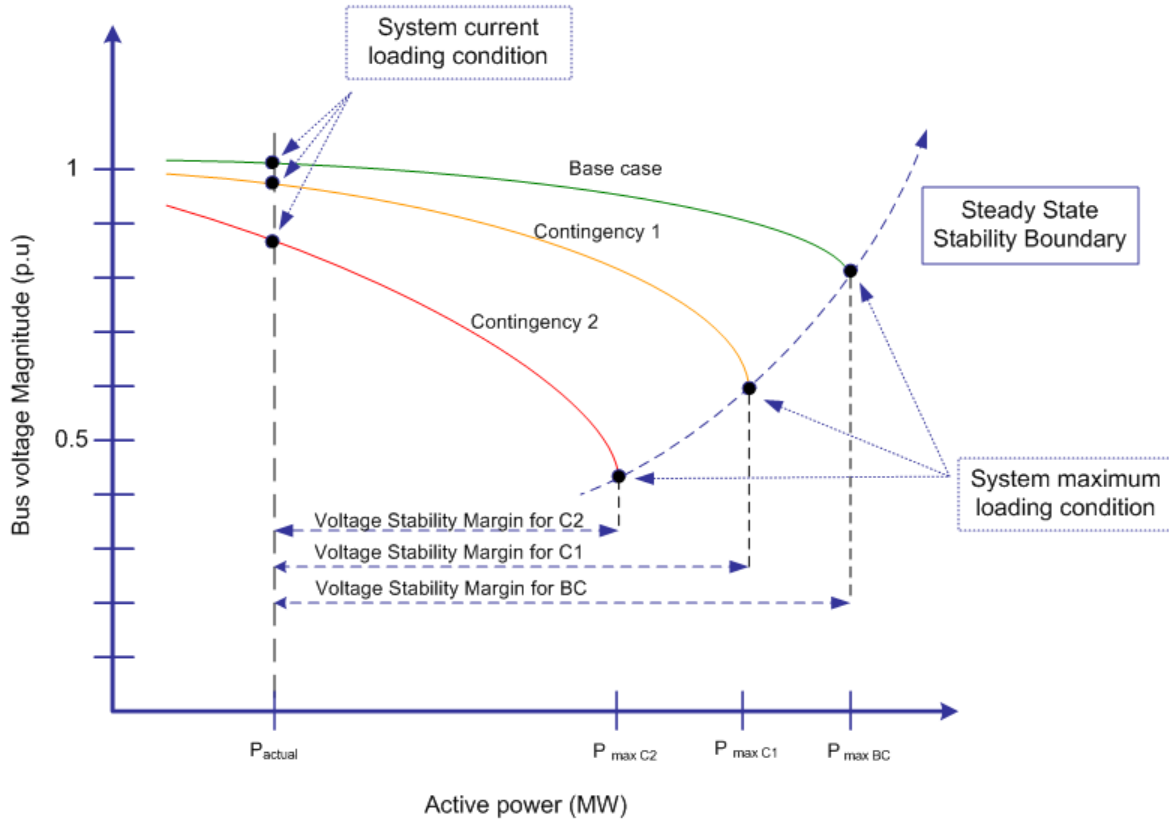


Figure 3.2. Effect of different contingencies on system voltage stability margin

Figure 3.2 shows how different contingencies affect the amount of VSM of a system. A typical example of voltage collapse due to topology change is given next. Assume the system is currently loaded at P_{max-C1} MW (on the x axis) and the current system topology is the base case (represented by the green curve). In case contingency 2 occurs, there will be no intersection between the loading level P_{max-C1} and the curve representing contingency 2. This means that a

post contingency stable equilibrium point would not be found and the system would face a voltage collapse due to unstable operation.

Another way of long term voltage instability can occur when the amount of load grows progressively until the system reaches a saddle node bifurcation point Ajjarapu, V. (1992), Ajjarapu, V. (2006). For instance, let us assume that the actual load in the system is represented by P_{actual} in Figure 3.2. Consider also that the PV curve representing the current topology is the base case. A voltage collapse happens when the system load grows in such a way that the operating point moves along the curve until it reaches the system maximum loading condition point (nose point). At that point, the voltage at the bus (y axis) shows a steep drop even for small load increments (x axis), characterizing thus a voltage collapse.

3.4 Voltage instability mechanisms

Although different time frames and disturbance severity are involved, most events have a quite similar mechanism: *the inability of the power system to meet the demand for reactive power*. This inability can be caused by two main reasons as discussed in Schlueter, R.A. (1998).

Loss of voltage control: “The *loss of voltage control* voltage instability is caused by exhaustion of reactive supply with resultant loss of voltage control on a particular set of generators, synchronous condensers, or SVCs. The loss of voltage control not only cuts off the reactive supply to a sub region requiring reactive power, but also increases reactive network losses that prevent adequate reactive supply from reaching sub-regions in need of reactive power.”

Clogging voltage instability: “*Clogging voltage instability* usually occurs due to high I^2X series reactive losses, tap changers reaching tap limits, switchable shunt capacitors reaching susceptance limits, and shunt capacitive reactive withdrawal due to decreasing voltage. These

network reactive losses that result from the above possibilities can completely choke off the reactive flow to a sub region needing reactive supply without any exhaustion of reactive reserves and loss of voltage control on generators, synchronous condensers, or SVCs.”

Since both phenomena may occur in the system, any methodology devised for online VSM estimation must be able to capture system behavior for both cases. According to Kundur, P. (1994), two main aspects must be analyzed during a voltage stability study.

Proximity to voltage instability: “The calculation of how far a system is from a voltage measures how many MW away from a collapse the system is. The distance to instability is usually measured in MW and represents the total system load. However, other system measurements such as total active power flow across certain interfaces, or reactive power reserves in the studied area can also be used as indicators.”

Characteristics of voltage instability: “This analysis has the objective to investigate the mechanisms that contribute to a voltage collapse and cause instability in the system. Steady state simulations of a large number of scenarios are usually helpful to determine critical system components and areas. This analysis is also capable of determining voltage-weak areas, i.e., areas in the system where voltage support is poor and are thus potential candidates to voltage instability.”

Consistent and extensive work has already been done in studying the characteristics of voltage stability, especially in the field of voltage control areas identification as presented in Zhong, J et al. (2004) and Morison, K. et al. (2008). Therefore, this research will focus on the development of a real-time technique to estimate the distance from a voltage collapse. The method will make use of system wide RPRs in order to estimate the amount of VSM. A definition and investigation of four different types of RPRs is given next.

3.5 Reactive power reserve definitions

Before analyzing the relationship between RPRs and VSM, it is necessary to determine how RPRs are defined. In order to address this question, a study has been conducted in considering different definitions of RPRs, Leonardi, B. and Ajjarapu, V. (2008). The objective of the study is to investigate how different definitions of RPRs are related to VSM and also if any of these definitions can be used for online VSM monitoring. The four investigated definitions are introduced in Figure 3.3 and Figure 3.4 below.

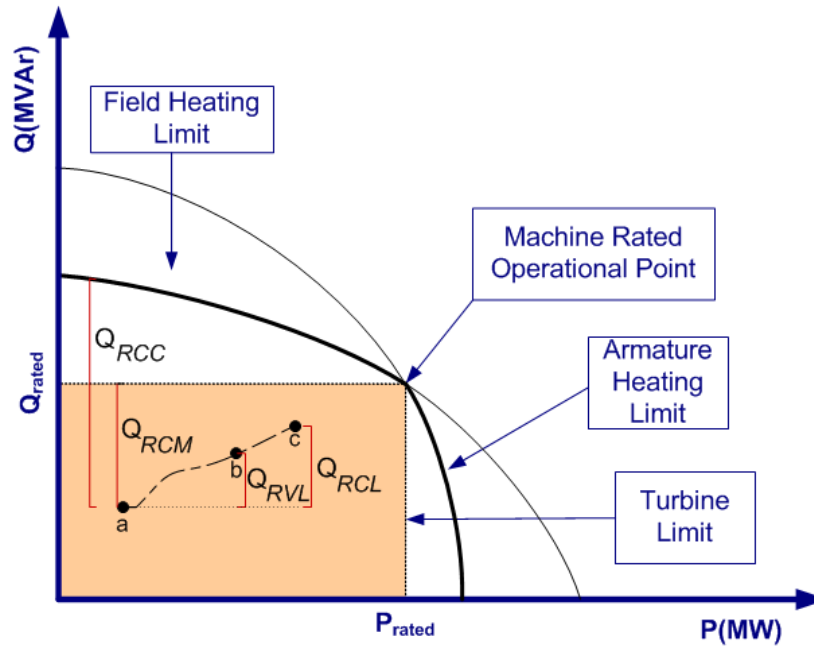


Figure 3.3. Capability curve of synchronous generator

The figures represent the capability curve of a synchronous generator and the PV curve frequently used in voltage stability studies, respectively. Using these two figures, four definitions of RPR can be identified.

Q_{RCC} represents the RPR definition with respect to the capability curve, Q_{RCM} represents the RPR definition with respect to a constant maximum value for reactive power dispatch, Q_{RVL} represents the RPR definition with respect to a minimum voltage level and Q_{RCL} represents the RPR definition with respect to the collapse point.

The points a , b and c in Figure 3.3 and Figure 3.4 are the same and represent the current operating condition, minimum voltage violation limit (V_l in Figure 3.4) and the voltage at the point of collapse (V_c in Figure 3.4). The voltage violation limit (V_l) represents the point where bus voltage magnitudes reach the minimum voltage limit. Although this value may vary among transmission operators, typical values lie in the range of 0.90 – 0.95 p.u.

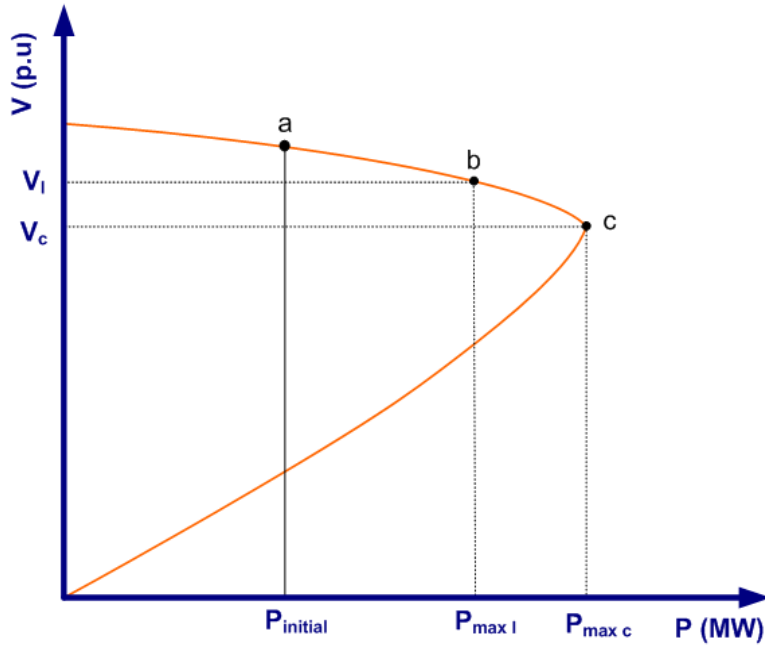


Figure 3.4. PV diagram

$$\begin{aligned}
Q_{RCC} &= Q_{Cap.Curve} - Q_a \\
Q_{RCM} &= Q_{Const.Max.} - Q_a \\
Q_{RVL} &= Q_b - Q_a \\
Q_{RCL} &= Q_c - Q_a
\end{aligned} \tag{3.1}$$

$Q_{Cap.Curve}$, $Q_{Const.Max.}$, Q_a , Q_b and Q_c are the reactive power limits due to the capability curve, reactive power limit at a constant maximum (commonly used in power flow simulations), current reactive power dispatched by the machine, reactive power dispatched at the point of minimum acceptable voltage and the reactive power dispatched at the point of voltage collapse, respectively.

3.6 Relationship between RPRs and VSM

As mentioned earlier, the relationship between RPRs and VSM was not very clear until recent studies have investigated it. Preliminary results indicated that a linear relationship between reactive power reserves and system voltage stability margin is found to occur as shown in Bao, L. (2003).

After studying the interaction of both variables in more detail, it has been observed that the relationship between RPRs and VSM can be linear or quadratic, depending on several system characteristics, such as: proximity between generator and load center, size of generating units and presence of nearby generators, Leonardi, B. and Ajjarap, V. (2008). In order to investigate the nature of this relationship, the IEEE30 bus test system shown in Figure 3.5 is used.

The system is stressed by increasing load until the point of collapse under different network topologies. All RPRs except the slack bus are monitored, as well as system VSM. Load is increased in the entire system to represent stressed operating conditions. System load is increased

in different directions in order to account for uncertainty in load behavior. The base case load is randomly perturbed and load is increased based on their proportion to the total initial load.

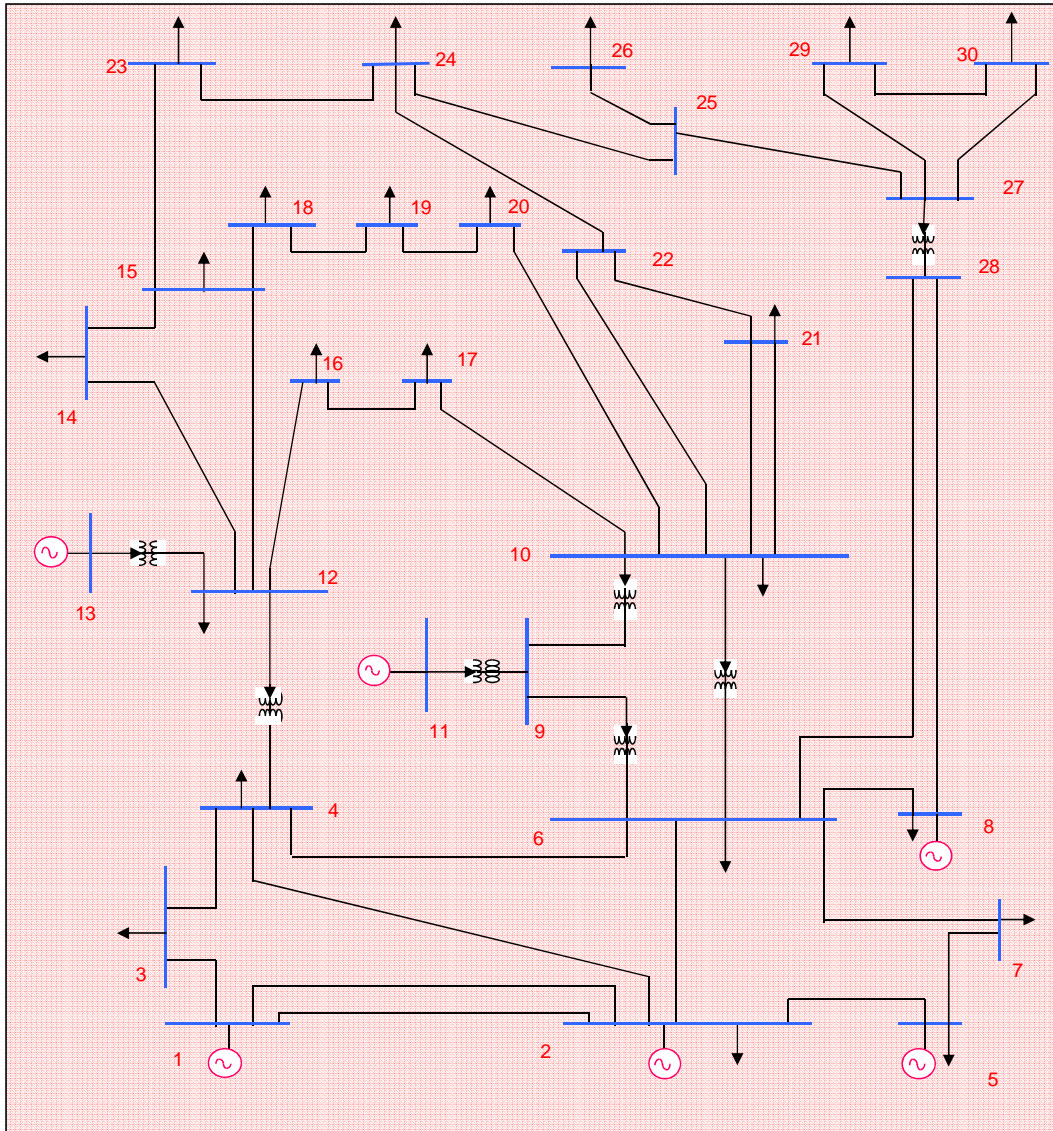


Figure 3.5. The IEEE30-bus test system

Several contingencies including all N-1 and various N-k contingencies with generator and transformer outages are considered. Generators are dispatched following a certain pattern which can be obtained from market clearing auctions or a merit order file.

Figure 3.6 and Figure 3.7 show the relationship between the RPRs and VSM for generators 5 and 11, respectively. The y -axis contains the RPR, whereas the x -axis contains the amount of VSM. From the pictures below, the RPR of generator at bus 5 shows a linear relationship with system VSM, whereas the RPR of generator at bus 11 has a more quadratic relationship. It is important to remember that the VSM is obtained from the PV curves, which have been already introduced in Figure 3.2 and Figure 3.4.

During normal load conditions, the system is operating at $P_{initial}$, represented as point a in Figure 3.4. As load is increased in the area, the current operating point shifts to the right on the curve until it reaches point c , with a total system load of $P_{max\ c}$. VSM is defined as the difference between the maximum load that the system can withstand and the current load, i.e., $VSM = P_{initial} - P_{max\ c}$.

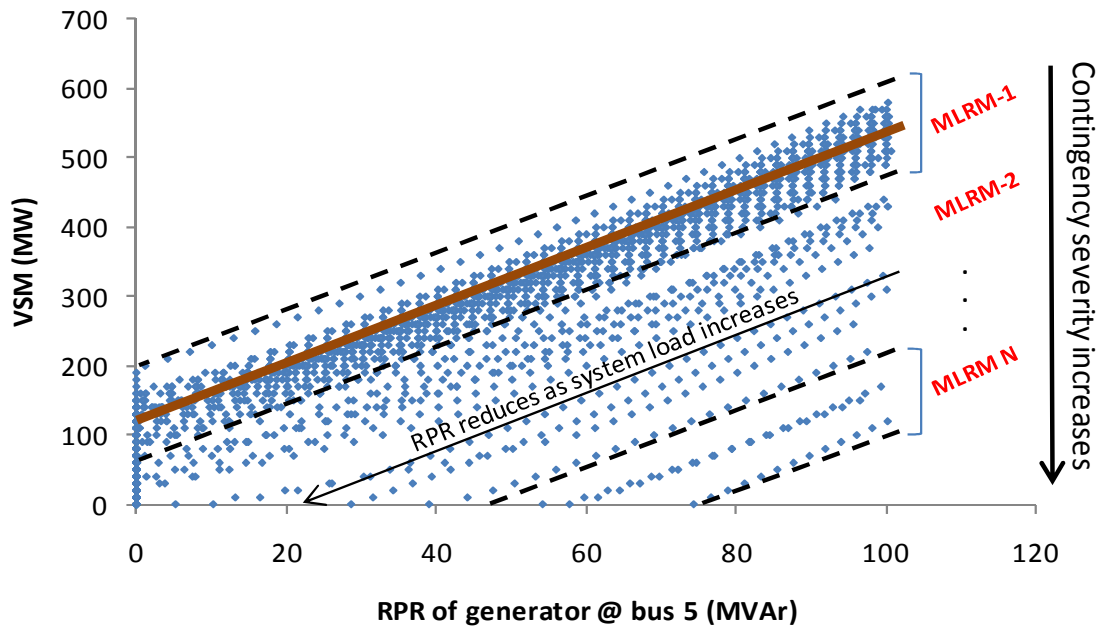


Figure 3.6. RPR and VSM relationship for generator 5

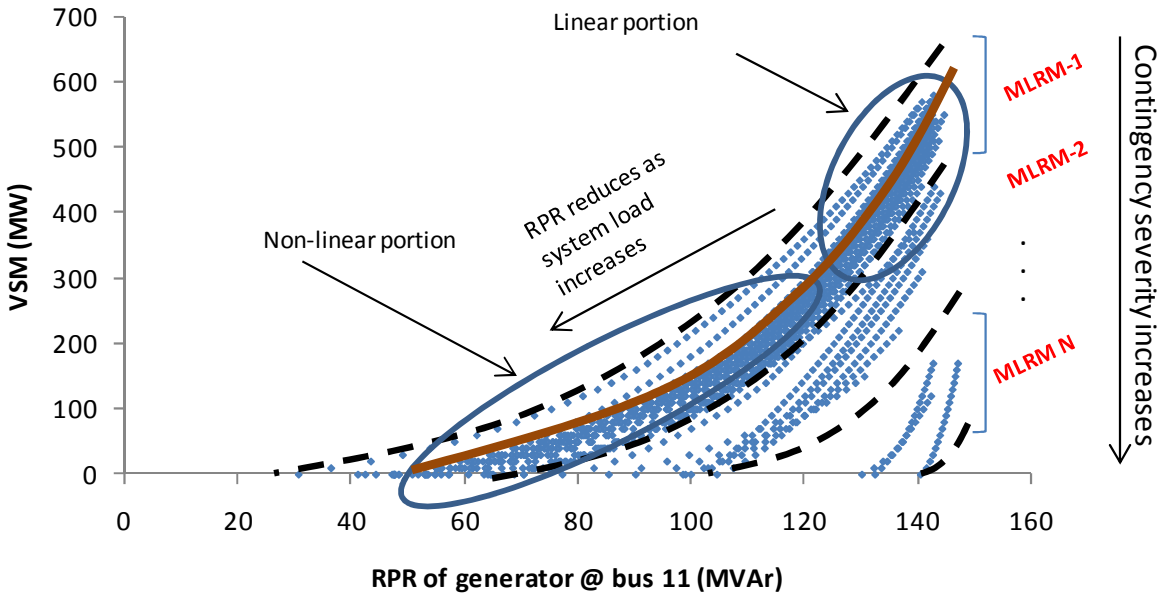


Figure 3.7. RPR and VSM relationship for generator 11

In order to understand how RPRs vary with VSM, the location of point *a* shown in Figure 3.4 must be translated to Figure 3.6 and Figure 3.7. When the system is loaded at point *a* in Figure 3.4, the RPRs of generator 5 and 11 in Figure 3.6 and Figure 3.7 will be around 100MVar and 140MVar, respectively. As system load is increased to reach the point of instability, the operating point moves along the dotted lines represented in Figure 3.6 and Figure 3.7 from upper right corner to the bottom left corner.

For some contingencies, the RPRs of generators 5 and 11 will be completely exhausted before the voltage collapse happens (cases where the dotted lines intercept the y – axis in Figure 3.6). However, for some other contingencies, there will still be some RPR left when the system reaches the point of voltage collapse (cases where the dotted lines intercept the x – axis of Figure 3.6 and Figure 3.7).

Cases like the ones shown in Figure 3.7 have been observed in the literature and are deemed to be caused by a *choke off* of reactive support due to excessive transmission IX^2 losses. Excessive

losses in transmission lines do not allow reactive power to be delivered to areas in need of voltage support.

Despite the causes of voltage instability, both Figure 3.6 and Figure 3.7 indicate that linear, as well as quadratic relationship between RPR and VSM may occur. The characteristic of the relationship will determine the order of the model used to relate both variables.

3.7 Effect of capacitor/reactor banks

Capacitor and reactor banks must have their effects accounted for as they are not used as regressors in the multi-linear regression models. The objective of not including those is neither to neglect nor to diminish the importance of capacitor/reactor banks as additional reactive power support sources. Nonetheless, there are two main reasons that contributed to the decision of not using capacitor/reactor banks as regressors in the MLRMs.

The first reason is that capacitor/reactor banks directly affect the amount of RPR available at dynamic VAR sources. In fact, there are reported cases where capacitor banks are switched in order to preserve dynamic reactive power reserves such as generators and synchronous condensers as indicated by Sandberg, L. et al. (1994) and Nirenberg, S. A. et al. (1992).

In order to visualize such influence, a portion of a 1648 bus test system is used and represented in Figure 3.8. The simulation demonstrates the effect of the capacitor bank located at bus 28 on rotating reactive power reserves nearby. In order to do so, simulations considering both the cap bank *on* and *off* have been performed. The cap bank is a 4x50MVar (4 blocks of 50 MVar each). Load is then increased in the area to simulate a system stress under a fixed network topology. The RPR of generator located at bus 26 is monitored during the load increase process. Other generators are affected to a different extent, based on their electrical distances from the cap bank and will not be shown here for simplicity.

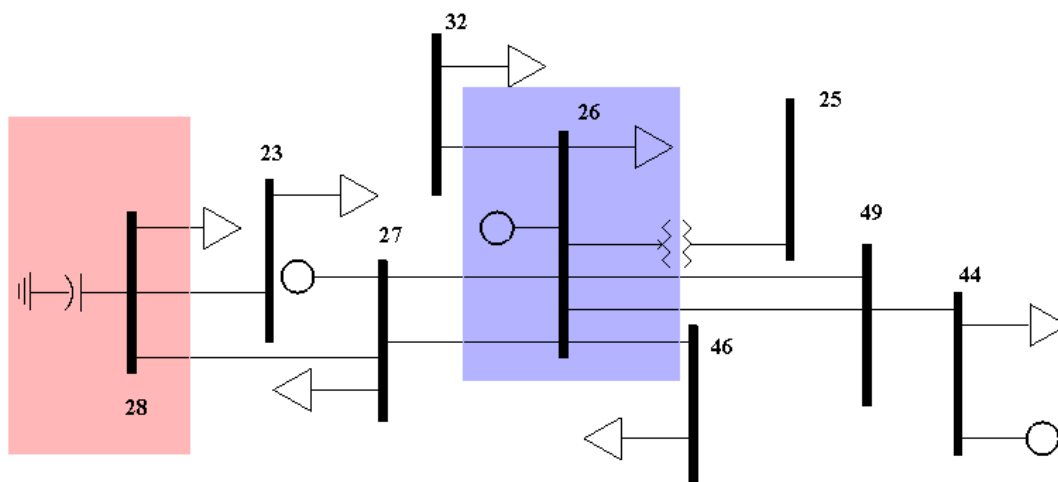


Figure 3.8. Local portion of the 1648 bus test system

Figure 3.9 shows that the RPR of generator 26 is shifted to the right when the cap bank is turned on. This effect is normally expected and more RPR is available at generator 26 when the cap bank is on than when it is off.

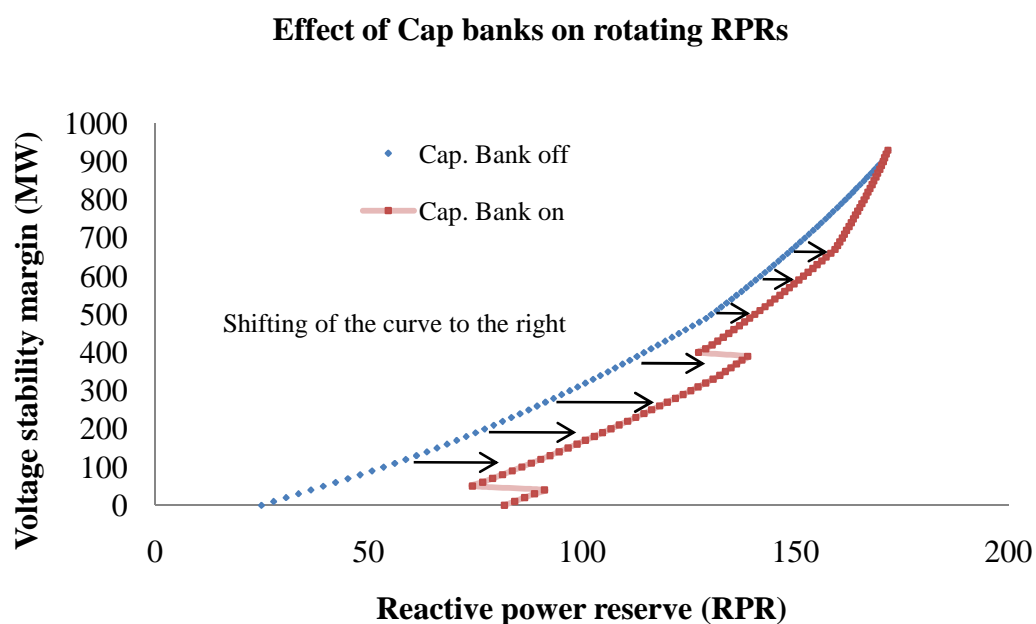


Figure 3.9. Effect of capacitor reactor banks on rotating RPRs

Therefore, the effect of the capacitor bank is captured by the RPRs of generator so that they do not need to be used as additional regressors in the derivation of MLRMs. Another reason why capacitor banks should not be included as additional RPRs comes from the inherent discrete characteristic of these devices. Capacitor bank RPRs do not have a smooth variation with system VSM.

Figure 3.10 shows how the capacitor bank RPR varies as system load increases for a given network topology. Such non-continuous behavior (discrete steps) reduces the quality of the fit between RPR and VSM and produces large residuals (errors). This non-continuous characteristic also distorts some of the statistical properties of the model as it concentrates the data points on the values represented by the step positions of the bank. Such distortion on statistical properties will cause a significant departure from the normality assumption of the residuals distribution Kutner, M. et al. (2004) and Lin, Z. C. and Wu, W. J. (1999).

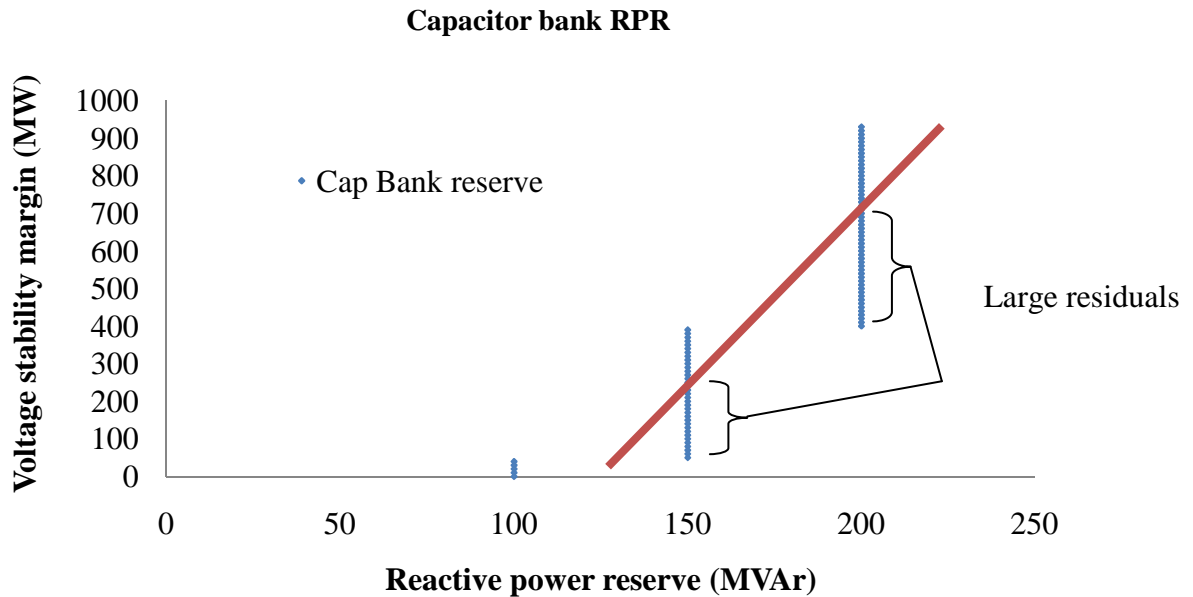


Figure 3.10. Capacitor bank reactive reserve variations with VSM

Based on the two aforementioned characteristics, it has been decided that capacitor and reactor banks will not be used as regressors in the MLRMs. Nonetheless, their effect on system performance is still captured as they relieve rotating reactive power sources and consequent impact on their RPRs.

3.8 Investigation of RPR definitions on performance

In order to investigate the effect of higher order terms on the precision of VSM estimation, a nonlinear regression model is used to relate each RPR in the system to the amount of VSM as shown in (3.2). A third degree polynomial is initially used in order to account for the nonlinearity of the relationship.

$$VSM \approx \alpha + \beta Q_R + \gamma Q_R^2 + \mu Q_R^3 + \dots \quad (3.2)$$

In case several RPRs are available, equation (3.2) can be extended to (3.3) in order to fit a polynomial relating VSM and all RPRs in the system. In (3.3), N_{MS} represents the number of RPRs, s represents the scenario or contingency and i represent the machine selected.

$$VSM^s \approx \alpha + \sum_{i=1}^{N_{MS}} \beta_i \cdot Q_R^{i,s} + \sum_{i=1}^{N_{MS}} \gamma_i \cdot (Q_R^{i,s})^2 + \sum_{i=1}^{N_{MS}} \mu_i \cdot (Q_R^{i,s})^3 \quad (3.3)$$

After a voltage stability assessment, the amounts of VSM and RPRs (Q_R) can be obtained for a few scenarios. By using a simple polynomial fit algorithm, the coefficients α , β , γ and μ can be identified and a preliminary assessment of the effect of higher order polynomial models can be inferred for some test systems.

3.8.1 Effect of polynomial order

All four definitions of RPR are investigated in order to verify how they vary with the amount of VSM. It has been noticed that the last two definitions of RPR introduced in (3.1) presented a more accurate relationship with VSM compared to the first two. In another words, their

relationship with system VSM had a better fit and produced more precise estimations when compared to the first two definitions as reported in Leonardi, B. (2008).

However, those definitions cannot be practically implemented since Q_b and Q_c are not known in online operations. Therefore, the RPR further used in these studies was the one with respect to the constant maximum value Q_{RCM} . The definition Q_{RCC} can also be used once the capability curves of generators are made available.

3.8.1.1 System A: 23Buses

- Q_{RCM} - Constant maximum reactive power definition

A total of fifteen contingencies are applied to this system, while the RPR and VSM have been sampled from the PV curve calculated for each contingency. Polynomials of different order have been fit in order to analyze the effect of the model order.

Figure 3.11 shows that the by increasing the order of the model, the accuracy of estimations also increases.

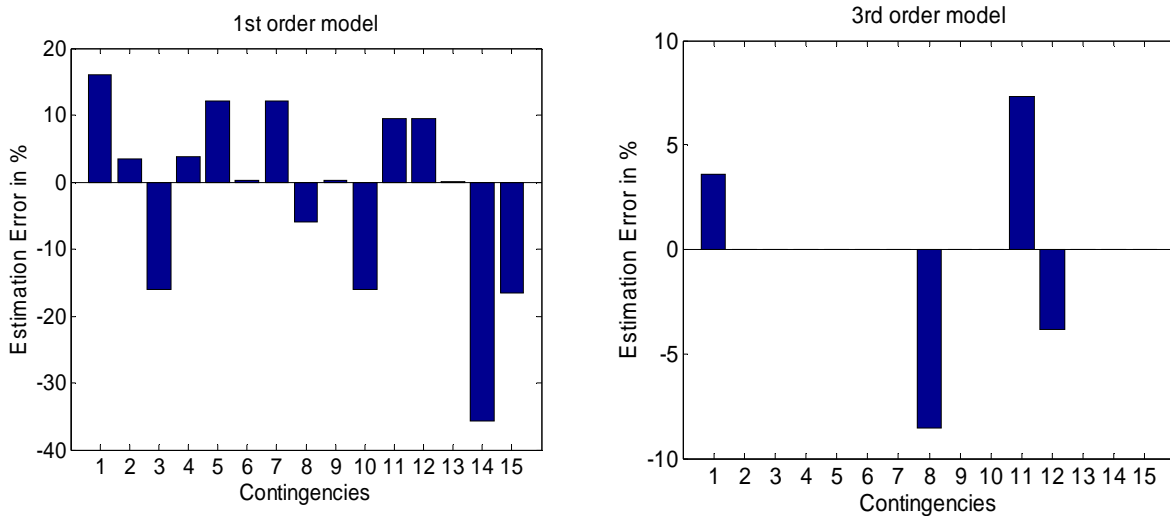


Figure 3.11. Comparison between linear model and higher order model

- Q_{RCC} - *Capability curve definition*

The second definition of RPR tested considers the capability curve of the machine. Similarly to the first definition of RPR, the higher order model provides a better fit than the simple linear model, indicating that capturing the nonlinearity of the relationship may be helpful as far as accuracy is concerned.

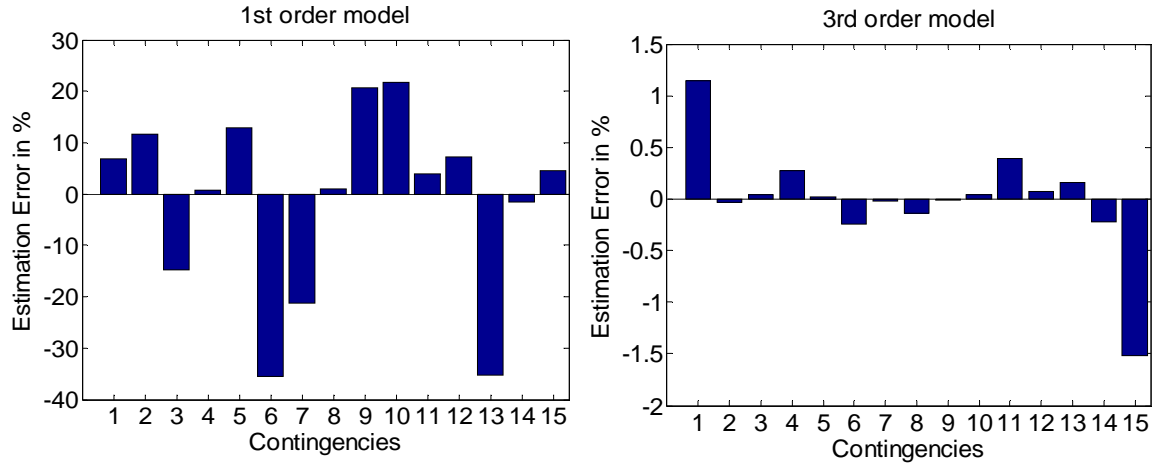


Figure 3.12. Comparison between linear and higher order model

- Q_{RVL} - *Minimum voltage limit definition*

The use of this definition of RPR to estimate VSM is presented next in Figure 3.13.

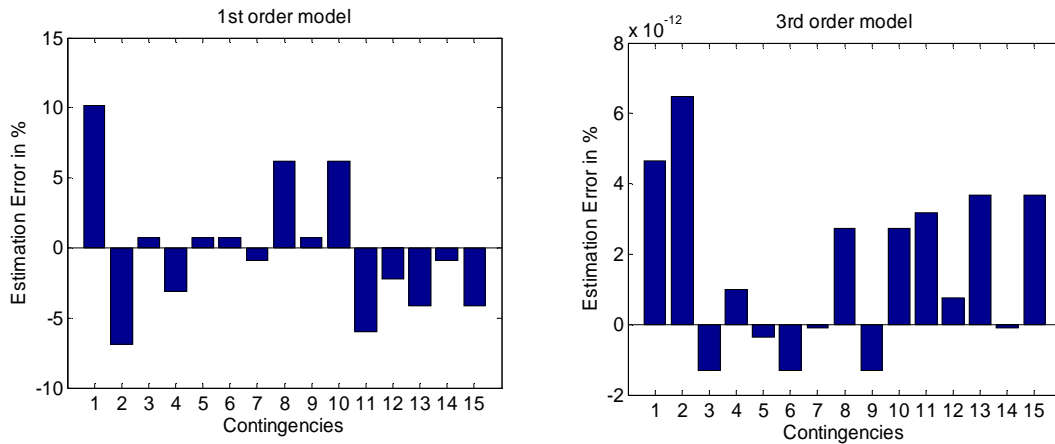


Figure 3.13. Comparison between linear and nonlinear model

It can be noticed that the higher order model presents a much more accurate estimation of VSM when compared to the lower order model. However, since this definition of RPR uses information about the point of minimum voltage in real time, it cannot be implemented in practice.

- Q_{RCL} - Voltage collapse limit definition

Similarly to the previous case, the RPR definition at the point of collapse shows a more precise estimation of VSM when a higher order model is used as shown in Figure 3.14. However, since the amount of reactive power dispatched at the point of collapse cannot be identified in real time, this definition also cannot be implemented in practice.

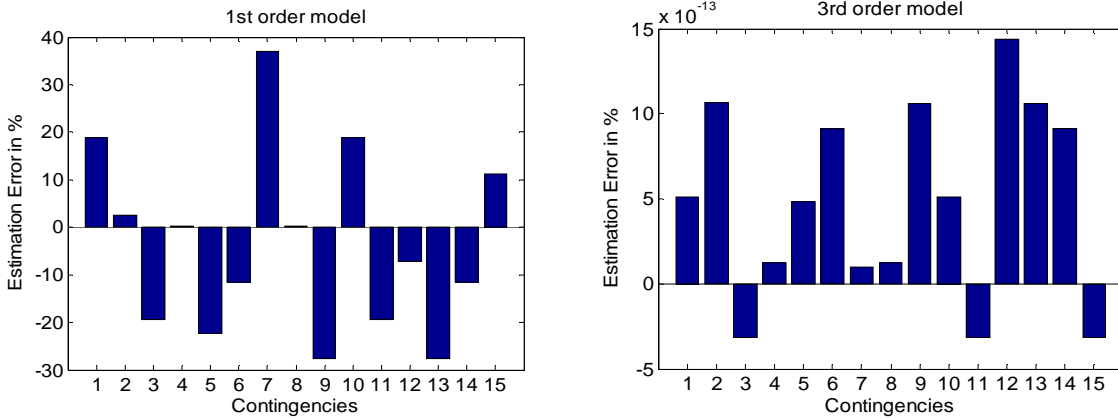


Figure 3.14. Comparison between linear and nonlinear model

Q_{RCL} and Q_{RVL} definitions are also known as *effective* RPR definitions, for they represent the effective amount of RPR available until system conditions are violated (voltage and stability conditions, respectively). On the other hand, Q_{RCM} and Q_{RCC} are also known as *technical* RPR definitions, for they are technically existent even in some cases even after the system has faced a voltage collapse.

3.8.1.2 System B: 1648 Buses

The same methodology has been tested in a larger system in order to verify if the improvements in performance would be seen on networks of realistic size. All four definitions except the capability curve are tested due to unavailability of data regarding the curves. The reason is that the curves were not available for this case.

This system has total of 49 generators and 23 have been selected to be used in the prediction model. Generators with best fit are selected. One hundred and fifty critical contingencies are used to generate the data. Linear and quadratic polynomials are considered in this case and results are presented in Figure 3.15, Figure 3.16 and Figure 3.17.

- Q_{RCM} - Constant maximum reactive power

It can be observed that the quadratic polynomial has a better fit with smaller estimation errors when compared to the linear model. This result reinforces the fact that a higher order model can indeed improve VSM estimation performance.

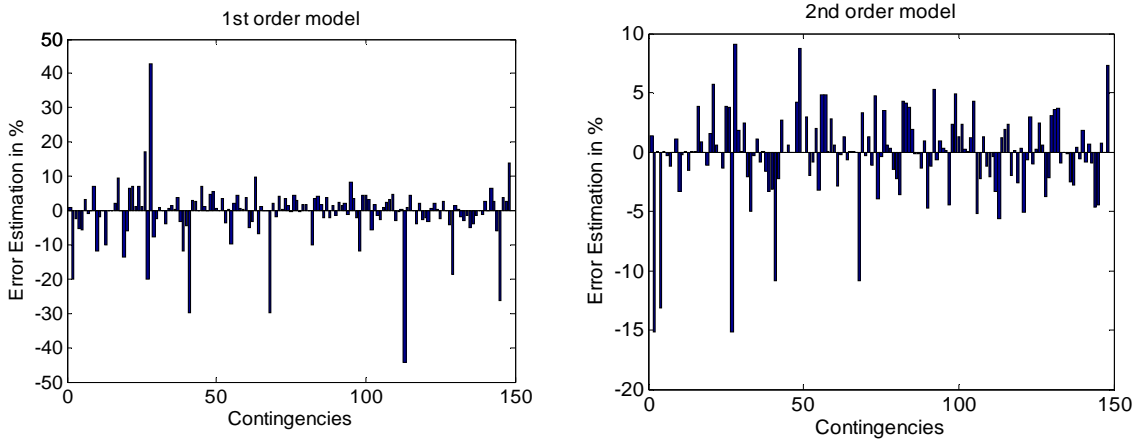


Figure 3.15. Comparison between linear model and quadratic model

- Q_{RVL} - Minimum voltage limit definition

Similarly to what has been done for the smaller system, the RPR definition with respect to minimum voltage values is also investigated here. The figure below not only shows that this definition of RPR enhances the precision of VSM estimation, but also shows that the quadratic model has better performance than the linear one.

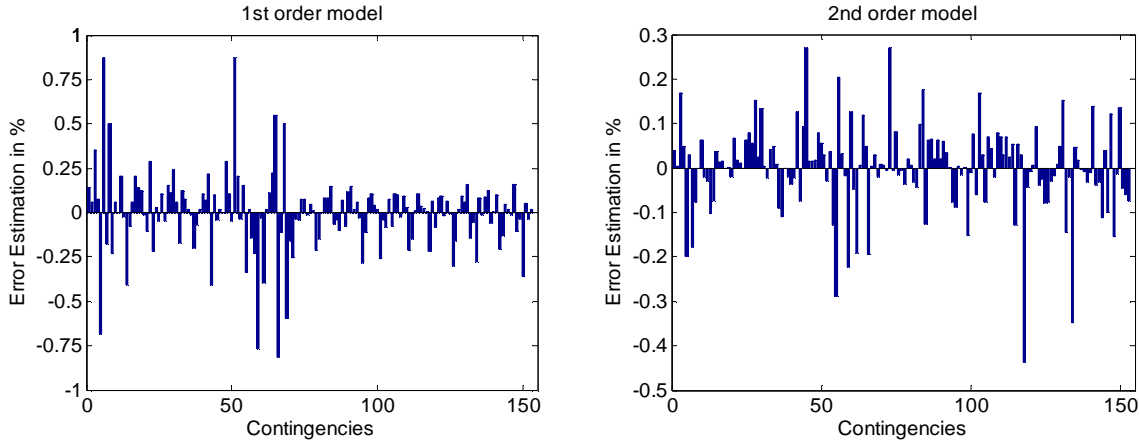


Figure 3.16. Comparison between linear model and quadratic model

This is an important observation as it indicates that higher order models tend to improve the accuracy in estimations.

- Q_{RCL} - Voltage collapse limit definition

Similarly to the previous definition, the precision of the quadratic model is higher than the precision of the linear model. However, this definition of RPR uses the amount of reactive power produced at the point of collapse. Since system operators do not know those values beforehand, it is not possible to use this definition in practical application for online VSM.

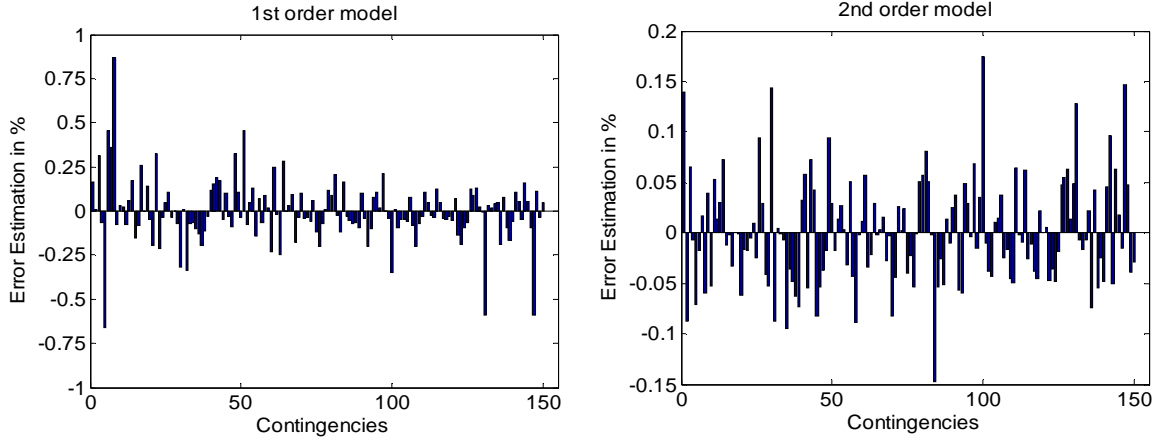


Figure 3.17. Comparison between linear and quadratic model

Overall, we can conclude that the minimum voltage and the collapse point definitions showed a more precise estimation of VSM when compared to the maximum constant reactive power and the capability curve definitions. However, the online implementation of these last two definitions (Q_{RCL} and Q_{RVL}) is not practical since system operators do not know the value of the maximum dispatchable reactive power at those points, thus not being able to calculate the RPRs.

CHAPTER 4. DEVELOPMENT OF THE ONLINE VSM MONITORING METHOD

4.1 Database generation and information flow

Before developing the MLRMs and the MLRM-IDtool, a thorough VSA needs to be performed in order to gather data that will be used for model development. The VSA takes into account various practical aspects involved with the system while capturing a massive amount of data.

Initially, essential information regarding system scenario, critical network topologies and load forecast is given to system planners as shown in the figure below.

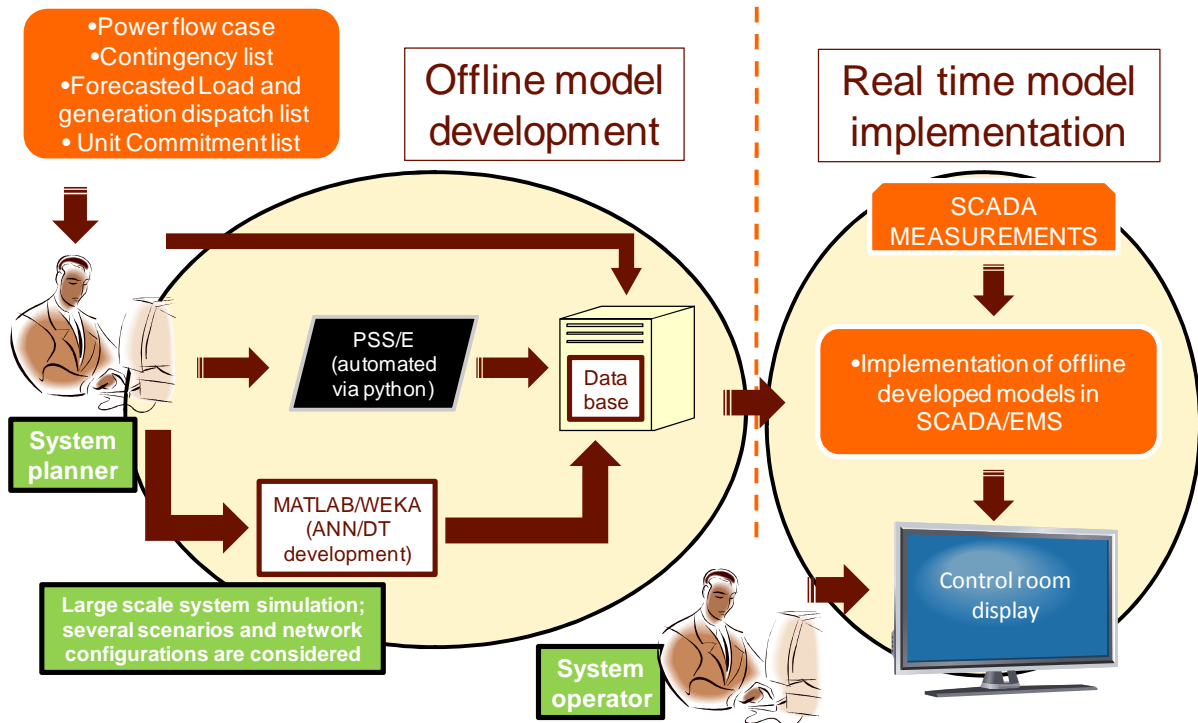


Figure 4.1. Overall methodology description

Next, all the system simulations are performed using PSS/E® and the variables of interest are stored in a database for future MLRM and MLRM-IDtool development. Figure 4.1 describes all

the steps involved in the database development process, from its creation until the final step of online visualization in SCADA/EMS system.

System scenarios are created using information related to future loading conditions and a contingency list that includes outages with higher probability of occurrence. A list of the generating units used in the generation increase pattern is also provided, as well as the loads that are increased. Once the operator has all this information in hand, extensive system simulations are performed in order capture the behavior of the system for the specified scenarios and contingencies. In order to accelerate the data base development, PSS/E® has been automated used python language. More details regarding the automation process will be provided in the next section.

Once the database is generated, the design of MLRMs and the MLRM-IDtool can finally begin. After both tools are properly developed and tested, they are used in real time operations via SCADA system as shown in Figure 4.1 above.

4.2 PSS/E® automation using python

In order to enable the methodology to be applied to larger networks, commercial grade software needed to be used for system simulation. In this research, PSS/E® has been automated via python for large system simulation.

Python is an open source, flexible, object-oriented programming language. It has been recently gaining popularity due to its simple syntax and capability of building high quality applications. PSS/E® has a python application program interface (API) that enables python scripts to control and utilize several activities/functions in PSS/E. These APIs enable any script written in python to access internal PSS/E functions and variables, execute and output simulation data into files.

The automation process not only speeds up the simulation of several scenarios but also enables the methodology to model various system components in greater detail. Figure 4.2 shows how the automation process generates the database and controls the flow of information.

Initially, all simulation parameter are provided by system planners to a python interface which controls the operation of PSS/E. Iterative scenario and contingency processing generates all the data which is stored in the database. All variables stored are listed in the picture and described in Table 4.1.

Once the database has been created, WEKA[®] and Matlab[®] are used to develop the MLRMs and the MLRM-IDtool.

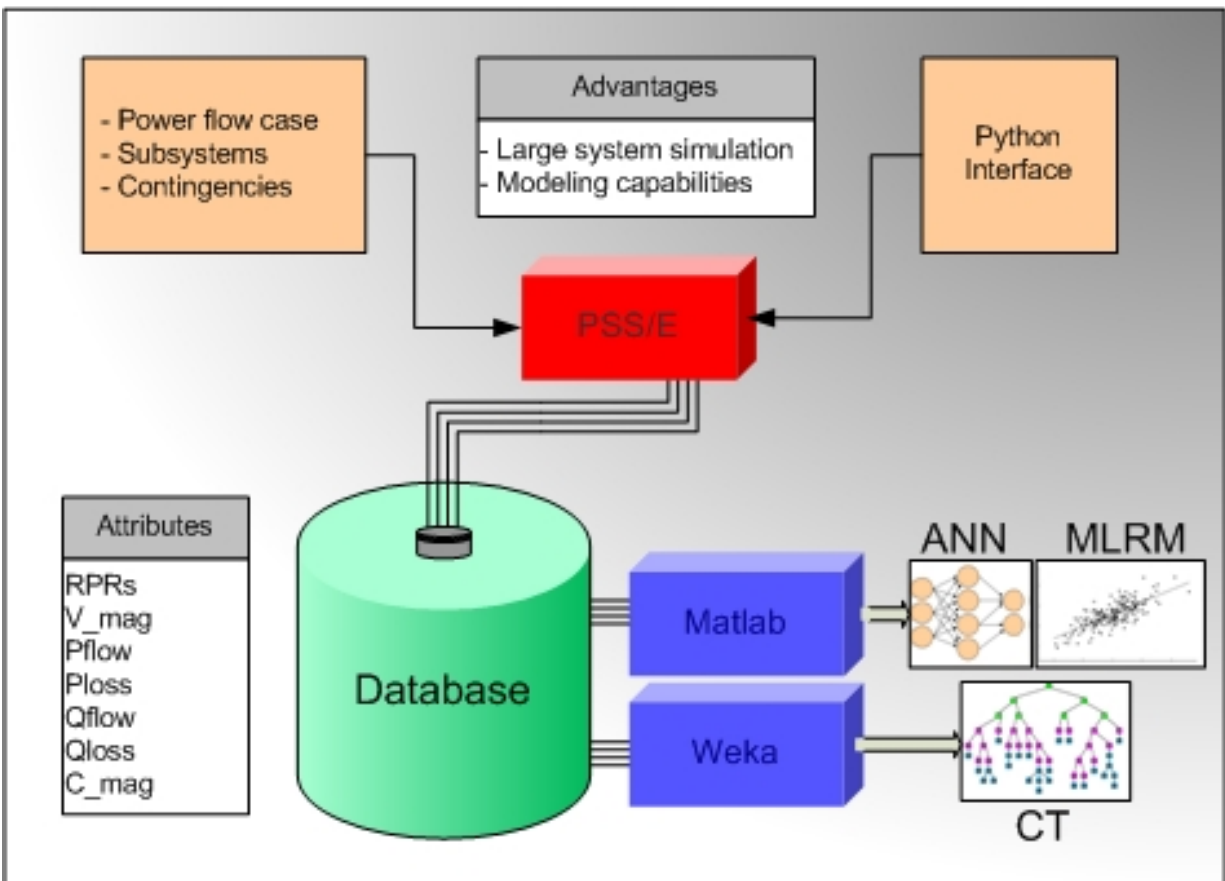


Figure 4.2. Database automation process

Figure 4.3 shows the interface developed in python and used to automate PSS/E® for efficient scenario processing. The objective to develop a more user friendly interface comes from the possibility of practical implementation of the approach in a utility scale system.

Although the developed application interface is quite simple, it avoids users from having to make tedious modifications into the python script. Only a few files are needed in order to begin the simulation. A snapshot of the developed interface is shown in Figure 4.3.

The first blank field is responsible for loading the power flow file selected by the user and defining the directory path to store output files. A simple click on the open button will enable the user to select the power flow case that he wants to run the application on.

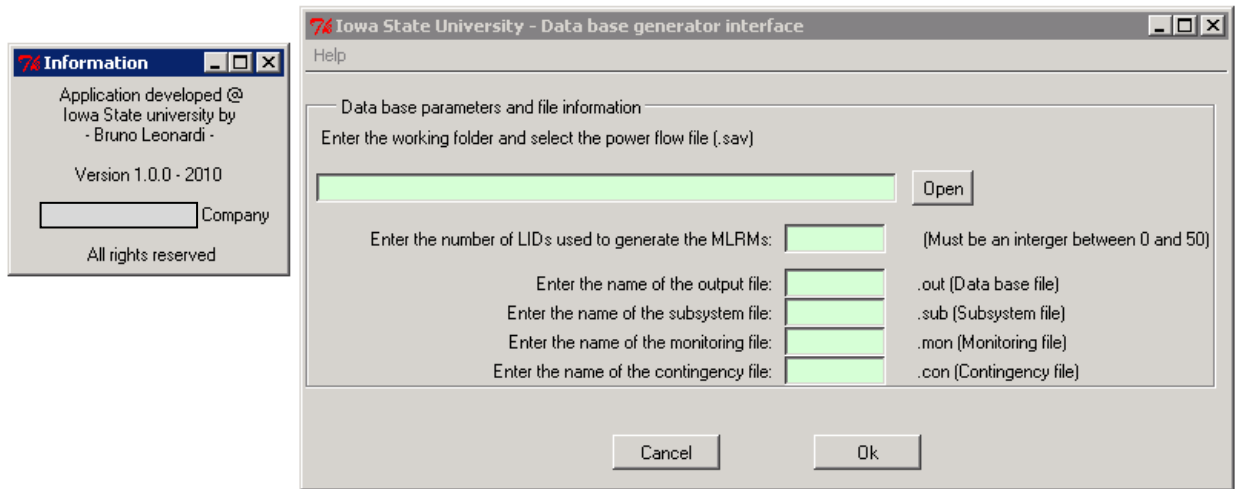


Figure 4.3. Python - PSS/E® interface developed for software automation

After selecting the power flow case, the number of load increase directions is provided in the second field. Any number between 1 and 50 can be provided in this field. In fact, more LIDs can be used depending on the user's need.

The third field represents the name of the output files and is determined by the user. During simulation, several variables are stored in the output files under their respective name. A complete list of sampled variables is given in Table 4.1. Given the large number of simulations,

several files are necessary in order to store all data when accounting for several LIDs and contingencies.

The next three fields are similar to various PSS/E® activities such as PV curve and contingency analysis engines. The three files given to the fourth, fifth and sixth fields represent the studied subsystems (.sub), the monitored elements (.mon) and the contingency list (.con) used during the VSA.

Table 4.1: System variables sampled during VSA

<i>Variable name</i>	<i>Description</i>
Vmag	Voltage magnitudes at selected buses (p.u.)
Pflow	Active power flow on selected lines (MW)
Ploss	Active power loss on selected lines (MW)
Qflow	Reactive power flow on selected lines (MVar)
Qloss	Reactive power loss on selected lines (MVar)
Cmag	Current magnitude on the lines (p.u.)
RPR	Reactive power reserves at selected generators (MVar)

Once all files are made available to the application, a click on the “OK” button will make PSS/E® accept these files and run the application. A complete VSA considering all contingencies in the “.con” file is performed for each different LID. Once the simulation is completed, all selected variables are stored in output files for each LID and all contingencies. The process is repeated until all LIDs are considered. As of now, there is no commercial grade software available in the market that performs the tasks herein described. Therefore, the automation of PSS/E represented not only an important step in the methodology but also a critical one.

Once the simulation is finished, data preprocessing is necessary to prepare the output files so that they can be used in Matlab and WEKA. Visual basic macros are used for efficient data preprocessing. This is done in order to clean the output files from headers and other simulation information used to monitor the flow of the application.

After all data is gathered, preprocessed and stored in the data base, different programs are used in order to generate the MLRMs and MLTM-IDtool. In the proposed method, Matlab is used to design the MLRMs and also the ANNs used in the MLRM-IDtool. WEKA is used to develop the DTs, KNNs and SVMs used in the MLRM-IDtool. The WEKA program is an open source program that contains several machine learning algorithms for data mining. The software has been developed and maintained by the University of Waikato in New Zealand, Hall, M. et al. (2009).

4.3 A two-Stage VSM estimation methodology

The methodology proposed in this research project is composed of two main stages formed by the MLRMs and the MLRM-IDtool. The MLRM-ID-tool uses system variables presented in Table 4.1 to identify what MLRM should be used in VSM estimation. Once the MLRM-IDtool properly identifies which model to use, the selected MLRM uses RPRs monitored in real time to provide an estimation of VSM. After developed and trained offline, both MLRMs and the MLRM-IDtool are used in the online environment for online VSM estimation.

Figure 4.4 shows how the complete scheme is to be implemented. Monitored data from SCADA/EMS system is passed onto the MLRM-IDtool, which will select the appropriate MLRM. RPRs are then passed onto the selected MLRM and an estimation of system VSM is made. The estimated VSM value can then be shown in a display in real time operations.

In this example, it is assumed that four MLRMs are necessary to account for all contingencies and load increase directions¹. Based on the current operating condition and network topological structure, the MLRM-IDtool will select the most appropriate MLRM to estimate VSM in real time. The corresponding MLRM selected represents the closest association between the trained MLRM-IDtool and the current operating conditions. Once the combined architecture is in place, only real time SCADA/EMS data available in the control center is necessary in order to estimate the amount of VSM.

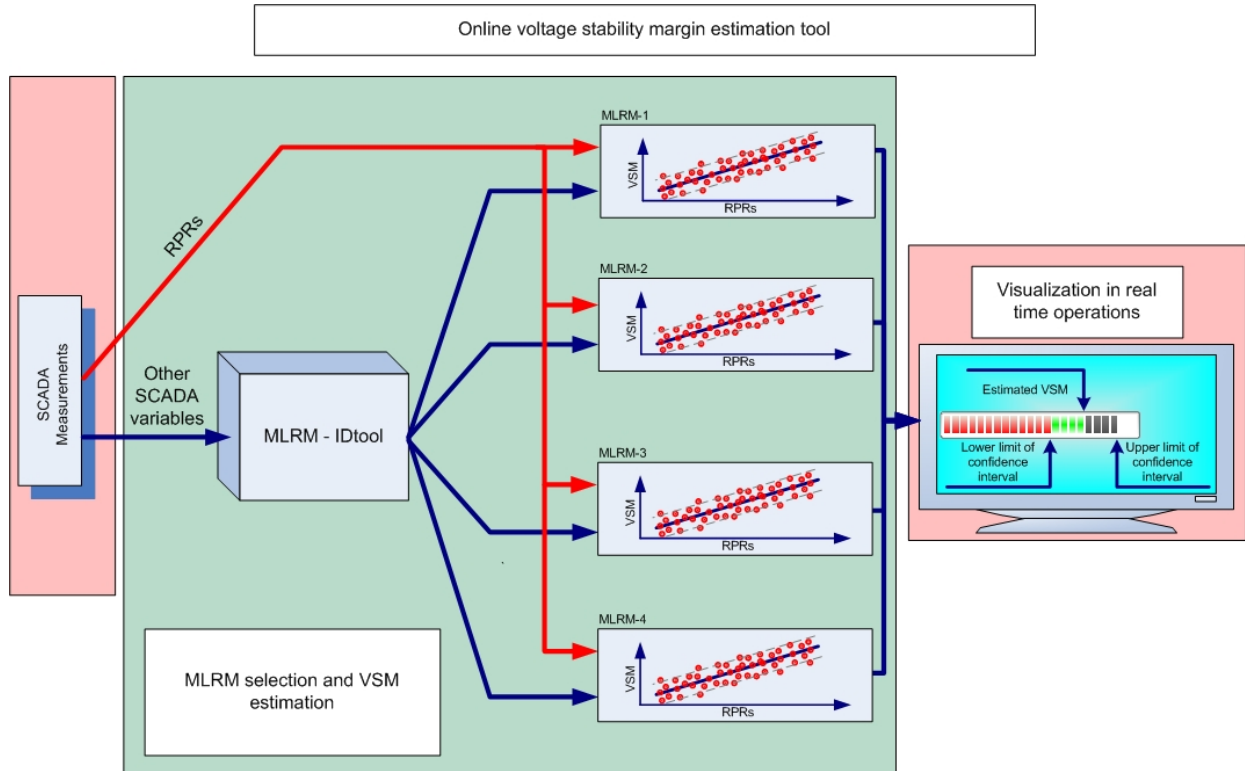


Figure 4.4. Overall implementation of the proposed online VSM estimation tool

The MLRMs are composed of several coefficients that weight the participation of each RPR to estimate VSM. A sum of all contributing RPRs will determine the existent amount of system VSM. Practical results have demonstrated that that a few MLRMs are required in order to

¹ The number of MLRMs varies according to design requirements, making the number of MLRMs vary from case to case. In this case, four MLRMs are used to exemplify how the approach is implemented.

achieve more accurate results. More importantly, it has been observed that the number of MLRMs does not increase with the size of the system or the number of contingencies considered, Leonardi, B. and Ajjarapu, V. (2011).

Figure 4.4 indicates that more than two MLRMs will often exist since a wide range of contingencies and operating conditions are considered. In the presence of more than one MLRM, the task of identifying the correct model becomes critical. The identification of the proper MLRM when more than two models are available is a typical case of multiclass classification.

Binary classification problems only have two classes to be distinguished (e.g., classify system condition as secure/insecure, stable/unstable, etc). Multiclass classification problems are characterized for having more than two classes to be distinguished (e.g., classify job candidates according to their education: elementary school, high school, college, graduate school, classify quality of a product based on some aspects: very good, good, regular, poor, very poor).

An introduction to the problem of multiclass classification will be given next, as well as an introduction on the theory of multilinear regression models.

CHAPTER 5. THE USE OF MULTI-LINEAR REGRESSION MODELS FOR ONLINE VSM ESTIMATION

5.1 Mathematical formulation of MLRMs

The general formulation of a MLRM for a given observation i is given in (5.1), with all variables described in the sequence. The variable y represents the dependent variable (VSM), the x variables represent system RPRs, the variables α , γ and ω represent the coefficients for each RPR in the model and ε_i represents the error term. The index i accounts for the number of samples available, whereas the indexes j , l and k account for the number of RPRs available.

$$y_i = \alpha_0 + \alpha_j x_{ij} + \gamma_{kl} x_{ik} x_{il} + \omega_j x_{ij}^2 + \varepsilon_i, \quad (5.1)$$

$$\text{for } \{i = 1, \dots, n; j = 1, \dots, p; l = 1, \dots, p; k = 1, \dots, p; \text{ with } k \neq l\}$$

Although quadratic and crossed terms are present (x_{ij}^2 and $x_{ik}x_{il}$), the model is still linear on the coefficients α , γ and ω and hence can be solved by the method of least square or robust least square, Kutner, M. et al. (2004).

In case several observations or samples are available to the model, equation (5.1) can be represented in the vector-matrix form as shown in (5.2), where the *coefficient vector* $\boldsymbol{\beta}$ is given by $\boldsymbol{\beta} = [\alpha_0, \alpha_p, \gamma_1, \gamma_{p(p-1)/2}, \omega_1, \omega_p]^T$.

$$\mathbf{y} = \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\varepsilon} \quad (5.2)$$

Adapting the formulation given in (5.2) to the problem at hand, vector \mathbf{y} will represent VSM measurements obtained from offline system simulation; \mathbf{X} will be a matrix containing monitored RPRs and $\boldsymbol{\varepsilon}$ represents the residual or errors.

The first column of matrix \mathbf{X} is formed by a unitary vector (it contains 1's from the first until the last vector position) to account for the linear interception coefficient α_0 . All the remaining columns of \mathbf{X} represent a RPR, a product of RPRs, or a squared RPR as described in (5.1). Each row of matrix \mathbf{X} and row of vector \mathbf{y} represents a sample of the RPRs and system VSM, respectively.

The samples of RPRs and VSM are taken at different points along the PV curve, enabling the MLRM to be used at different loading levels along the LID. The *coefficient vector* $\boldsymbol{\beta}$ is found by minimizing the sum of the square of the residual as follows.

$$\text{Min}_{\boldsymbol{\beta}} \|\boldsymbol{\varepsilon}\|^2 = \text{Min}_{\boldsymbol{\beta}} \frac{1}{2} \|\mathbf{y} - \mathbf{X}\boldsymbol{\beta}\|^2 \quad (5.3)$$

The solution of problem (5.3) is defined as the least square solution, Kutner, M. et al. (2004). The best linear unbiased estimation (BLUE) for the vector of coefficients $\boldsymbol{\beta}$ is given by equation (5.4).

$$\hat{\boldsymbol{\beta}} = (\mathbf{X}^T \mathbf{X})^{-1} (\mathbf{X}^T \mathbf{y}) \quad (5.4)$$

Once the vector of coefficients $\boldsymbol{\beta}$ is found, the MLRM regression model can be used online to estimate VSM. An estimation of the VSM vector ($\hat{\mathbf{y}}$) is obtained by multiplying the vector of coefficients $\boldsymbol{\beta}$ by the matrix of monitored regressors \mathbf{X} as follows.

$$\hat{\mathbf{y}} = \mathbf{X}\hat{\boldsymbol{\beta}} \quad (5.5)$$

The difference between the estimated VSM values ($\hat{\mathbf{y}}$) and the actual VSM values (\mathbf{y}) is defined as *residuals* or *errors* ($\boldsymbol{\varepsilon}$).

$$\boldsymbol{\varepsilon} = \mathbf{y} - \hat{\mathbf{y}} \quad (5.6)$$

Confidence intervals for the estimated VSM ($\hat{\mathbf{y}}$) can be obtained by modeling the residual probability density function (*pdf*). Once obtained, the confidence interval (c.i.) can then be used to handle uncertainty of VSM estimation in the following manner.

$$\hat{\mathbf{y}} = \mathbf{X}\hat{\boldsymbol{\beta}} \pm c.i. \quad (5.7)$$

Equation (5.7) represents how the MLRM are to be used in the online environment to estimate VSM. The online monitored RPR vector (\mathbf{X}) is multiplied by the regression coefficient vector ($\boldsymbol{\beta}$) to get an estimation of VSM. The confidence interval is obtained by modeling the residual *pdf* and is later added as bounds to the estimated VSM.

5.2 Uncertainty in system stress direction

In order to explain how uncertainty in load increase direction and changes in network topology affect the relationship between RPRs and VSM, different contingencies and load increase directions (LIDs) have been used to stress the system.

Figure 5.1 shows how these two different types of stress affect the amount of VSM. In long term voltage stability studies, a voltage collapse may occur in two different ways as previously mentioned. A voltage collapse situation can occur by increase load constantly followed by RPR exhaustion and consequent loss of voltage control, or through equipment outages, which makes the stability boundary to move closer to the current operating point, Kundur, P. et al. (2004).

Uncertainty in LID has been incorporated through the consideration of each load as a random variable following a normal *pdf*, with mean equals to the base case load and standard deviation equals to 15% of base case load. After randomly perturbing the base case, different LIDs are

obtained by systematically increasing the loads proportionally to their initially perturbed amounts.

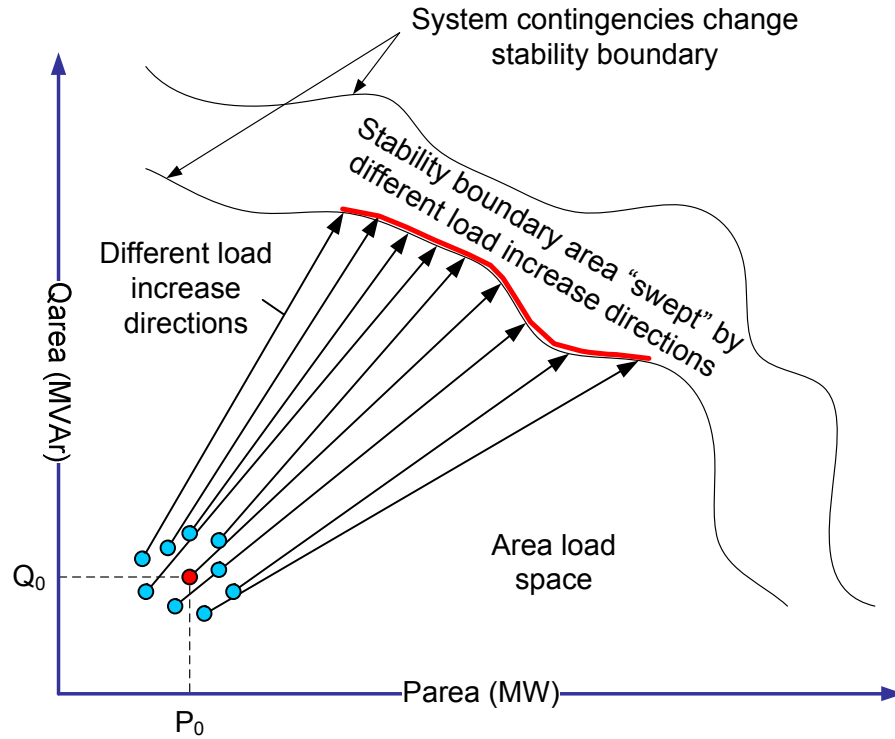


Figure 5.1. Variations in VSM due to different load increase directions and network topology

The idea of increasing load in different directions has the objective to account for uncertainties in load forecasting tools, which ranges from $\pm 5\%$ to $\pm 10\%$ for day ahead forecasts, Makarov, Y. V (2010). In case more precise statistical information about load increase direction is available, it can certainly be used to provide a more realistic stress direction to the system.

An extensive contingency list is used in order to cover various possible network topologies that may be encountered in daily system operations. This includes all NERC category B (N-1), C (N-2) and D (N-k) outages. Category B represents the situations where one system component (out of N system components) is offline, category C represents the situations where two system components (out of N system components) are offline and category D represents the situations

where k system components (out of a total of N system components) are offline, NERC-TPL-001-0.1 (2008).

5.3 Power system modeling and simulation

In order to investigate the phenomenon of long term voltage stability, a static modeling of the power system is considered in this study. The circuit equations of the power system are described in (5.8) and its solution is commonly known as the power flow solution, which determines the flows of power and values of network variables.

$$\begin{aligned} P_k &= V_k \sum_{j=1}^{N_{PQ} + N_{PV}} V_j \left(G_{kj} \cos(\theta_k - \theta_j) + B_{kj} \sin(\theta_k - \theta_j) \right) \\ Q_k &= V_k \sum_{j=1}^{N_{PQ}} V_j \left(G_{kj} \sin(\theta_k - \theta_j) - B_{kj} \cos(\theta_k - \theta_j) \right) \end{aligned} \quad (5.8)$$

Where:

N_{PV}	-	Number of generating buses
N_{PQ}	-	Number of load buses
P_k	-	Active power on bus k
Q_k	-	Reactive power at bus k
V_k	-	Voltage magnitude at bus k
V_j	-	Voltage magnitude at bus j
G_{kj}	-	Conductance of KJ element on Ybus
B_{kj}	-	Susceptance of kj element on Ybus
θ_k	-	Angle of bus voltage phasor of bus k
θ_j	-	Angle of bus voltage phasor of bus j

A reformulation of the power flow equations described above can be done in order to express the system in the classical nonlinear algebraic system of equations form shown in (5.9), where x represents the unknown bus voltage magnitudes and phase angles and λ represents the load increase parameter.

$$f(x, \lambda) = 0 \quad (5.9)$$

In order to collect measurements of RPRs and VSM under several different LIDs and network topologies, a systematic voltage stability assessment is performed. As mentioned before, PSS/E[®] is automated in order to reduce the processing time of a large number of scenarios and enhance modeling capabilities of system components. The effect of discrete slow dynamics devices such as ULTC, switchable capacitors/reactors banks is also taken into account during the simulations.

The base case load vectors $(\mathbf{P}_G, \mathbf{P}_L, \mathbf{Q}_L)$ contain all generators and individual loads that will be increased during VSA. Each generator and load i is varied according to (5.10).

$$\begin{aligned} P_{G_i} &= P_{G_i}^0 (1 + \lambda K_{PG_i}) \\ P_{L_i} &= P_{L_i}^0 (1 + \lambda K_{PL_i}) \\ Q_{L_i} &= Q_{L_i}^0 (1 + \lambda K_{QL_i}) \end{aligned} \quad (5.10)$$

For each different contingency, system load and generation are increased until the system reaches the voltage collapse point. In the above equation, $P_{G_i}^0, P_{L_i}^0$ and $Q_{L_i}^0$ represent the base case active power generation, active power load and reactive power load at bus i . The variables P_{G_i}, P_{L_i} and Q_{L_i} represent the total active power generation, active power load and reactive power load at bus i , respectively. The variables $K_{PG_i}, K_{PL_i}, K_{QL_i}$ represent the proportions or

rates at which active power generation; active power load and reactive power load are increased, respectively. Generators are dispatched according to a merit order file or based on market clearing auctions.

After the VSA is performed for all contingencies and LIDs, the database used for MLRM and MLRM-IDtool development is ready and the design of the models can finally begin.

5.4 Using MLRMs to relate RPRs and system VSM

After observing the relationship between RPRs and system VSM for different LID and network topologies, it has been observed that MLRMs could be successfully used to capture the relationship between those two variables. Figure 5.2 is shown here once again to evidence that the relationship between RPRs and system VSM can be linear for some operating region, but that as system conditions change, that relationship can become quadratic.

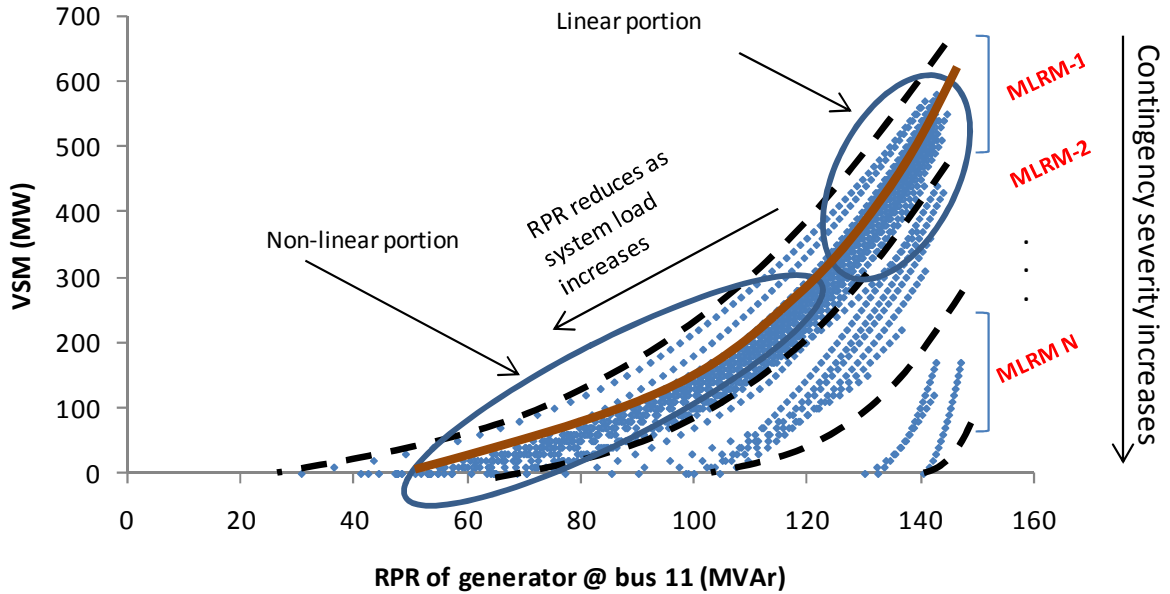


Figure 5.2. RPR and VSM relationship for generator 11

The relationship between the RPR of generator 11 and system VSM in Figure 5.2 is indeed close to linear for certain range of operating points, as pointed out in Bao, L. et al. (2003).

However, for loading conditions closer the point of collapse, the relationship between RPR and VSM can become very nonlinear, evidencing that a quadratic model would do a better job at predicting VSM than a linear one.

After performing a detailed investigation on how RPRs are related with system VSM, it was noticed that some generators have a linear relationship (see Figure 3.6) whereas others have a quadratic one (see Figure 5.2). Simulation results will show that the inclusion of quadratic terms in the MLRM not only improves the accuracy but also enhances some statistical properties of the model. Therefore, a more appropriate MLRM can be obtained if those quadratic terms are included.

Another important observation from Figure 5.2 is that the data spread along the y-axis is significantly wide in case several contingencies are considered. Therefore, due to the need to account for several credible contingencies, more than one MLRM is usually necessary as a single MLRM cannot accurately capture the relationship between RPR and VSM for all considered topologies. By considering more than one MLRM, the error involved with VSM estimation can be reduced as the confidence intervals for each model are made smaller.

Although the effect of changes in LID is not shown in Figure 5.2, practical observations have demonstrated that the effect is similar a network topological change. It is important to mention that the data spread caused by different LIDs is more localized, rather than widely spread as network topology changes.

Contingencies with similar amounts of VSM are grouped in the same MLRM range as shown in Figure 5.2. The data collected from those contingencies will be used in the development of

that specific MLRM. In order to clarify system behavior according to RPR variation, the point of maximum RPR in Figure 5.2 (around 150 MVar) represents the initial loading condition of the system, whereas the point of minimum VSM (0 MW) represents the point where system loses stability and voltage collapse occurs.

5.5 MLRM developmental procedure

The procedure to develop and verify the MLRMs is described in detail by the flowchart in Figure 5.3. Initially, one MLRM is used to account for all the LID and contingencies in the entire VSM range and the $\pm 2\sigma$ confidence interval of the MLRM is calculated. Due to the fact that the data is widely spread, the $\pm 2\sigma$ confidence interval for the residuals will tend to be large. If the residual confidence interval does not meet the design requirements, the VSM interval is divided in two and the process of MLRM development is repeated iteratively until the desired specification are met.

In this study, a $\pm 2\sigma$ confidence interval smaller than $\pm 10\%$ of the upper limit for each VSM range has been considered. For instance, a MLRM covering scenarios with VSM in the range 1000-600MW has to be further divided into two other MLRMs in case its $\pm 2\sigma$ confidence interval exceeds $\pm 100\text{MW}$. The procedure is repeated until all the designed MLRMs meet the specified confidence interval accuracy.

Another reason to select this approach to determine the amount of MLRMs is that MLRMs that account for more critical contingencies will have narrower confidence intervals, since the confidence interval is dependent upon the upper limit of the range. As more critical contingencies have lower upper values for their ranges, the MLRMs that account for those contingencies will have narrower confidence intervals.

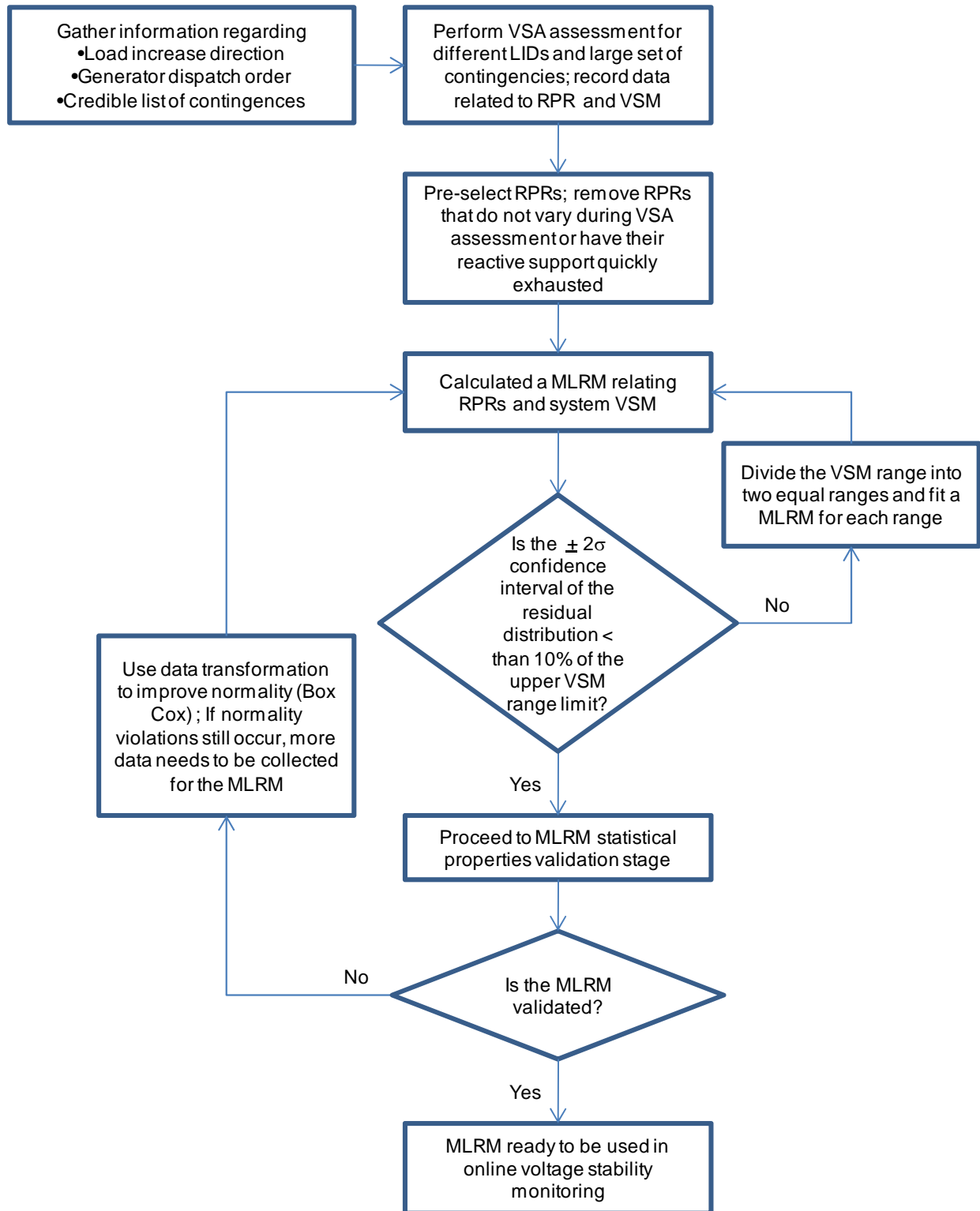


Figure 5.3. Flowchart representing MLRM development and validation stages

Another important observation is that the width of confidence intervals directly affects the amount of required MLRMs. If narrower confidence intervals are specified (e.g. $\pm 5\%$), a larger number of MLRMs would be necessary to cover all the studied cases. However, the VSM estimations of the models would be more accurate as the added confidence interval is smaller.

On the other hand, wider confidence intervals (e.g. $\pm 15\%$) would require less MLRMs to cover all contingencies and LIDs. Nonetheless, the models would have lower accuracy due to the bigger size of the confidence interval. For practical applications, some of the non-critical contingencies (usually N-1) can be eliminated so that the amount of needed MLRMs is reduced.

In case the MLRM validation process fails due to violation of some statistical assumptions, variable transformations (Box Cox) can be employed to correct for the violations, Box, G. E. P. and Cox, D. R. (1964). If the variable transformation is not able to remove the statistical violations encountered, more simulation can be performed and added to the training set so that it captures the behavior of the system under all scenarios.

5.6 MLRM validation procedure

Once the MLRMs are developed, a few statistical properties must be validated beforehand so that the models can be used in practice. Model adequacy verification investigates whether basic statistical properties about the MLRMs are valid and is usually performed through residual analysis. Among the main properties involved with MLRMs, *homoskedasticity* and *normality distribution of the residuals* are two important properties that must be attained to a minimum level, Kleinbaum, D. et al. (1998) and Kutner, M. (2004).

5.6.1 Homoskedasticity

Homoskedasticity is the assumption that observations of the error term (ε_i) are drawn from a probability density function of constant variance, Studenmund, A. H. (2006). In other words, the residual vector ($\mathbf{\varepsilon}$) has constant variance throughout the entire range of the regressed variable (\hat{y}).

Figure 5.4 and Figure 5.5 show the residual distribution in the presence of homoskedasticity and heteroskedasticity, respectively.

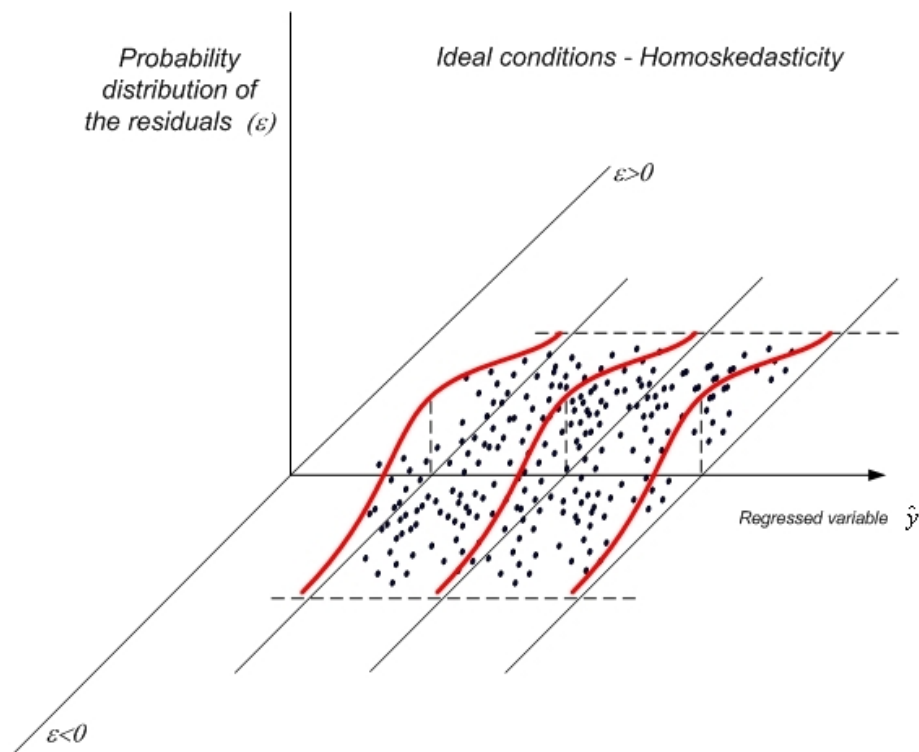


Figure 5.4. Homoskedastic residual distribution

An equivalent interpretation for homoskedasticity is when the dependent variable (y) has constant variance for any combination of the independent variables (x_i 's) as shown in equation (5.11), Kleinbaum, D. et al. (1998). Violation of the homoskedasticity assumption is defined as *heteroskedasticity*.

$$\text{Var}(\varepsilon)_{|X_1, X_2, \dots, X_n} = \sigma^2 \quad (5.11)$$

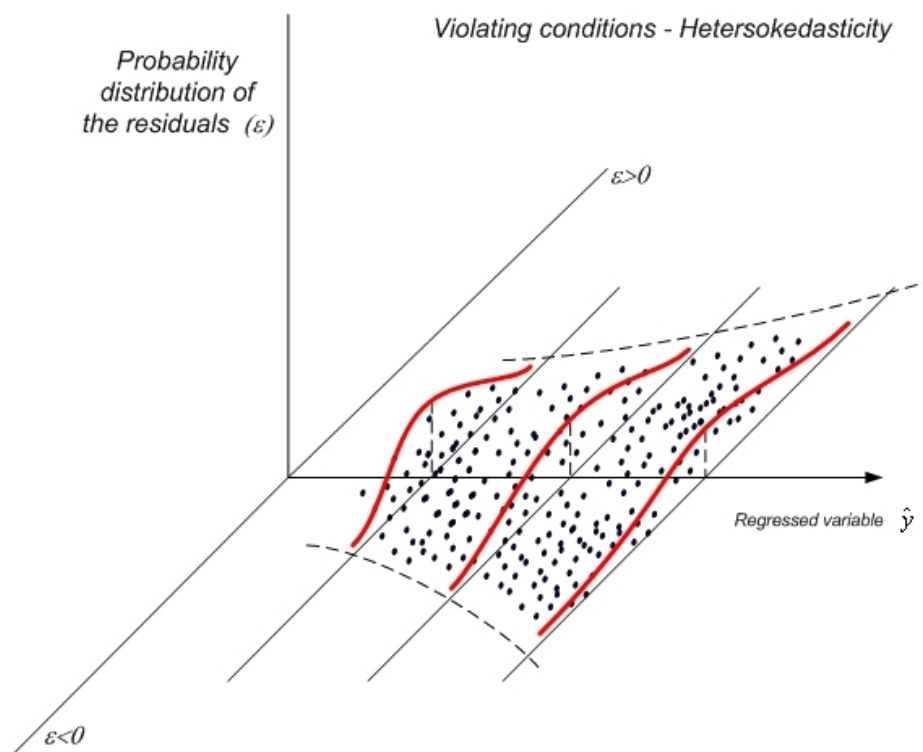


Figure 5.5. Heteroskedasticity residual distribution

Although the assumption of homoskedasticity may seem very restrictive, correction for *heteroskedasticity* only needs to be applied when the data shows significant departure from *homoskedasticity* Kleinbaum, D. et al. (1998) and Kutner, M. (2004). In case corrections are needed, Box Cox power transformation is the traditional remedial action used to correct for heteroskedasticity, Studenmund, A. H. (2006).

5.6.2 Normality of the residuals

Along with homoskedasticity, the distribution of the residuals is also expected to be approximately *normally distributed*, with zero mean and constant variance σ^2 .

$$\varepsilon \sim N(0, \sigma^2) \quad (5.12)$$

Since normally distributed random variables contain around 95% of its observations within the ± 2 standard deviation range ($\pm 2\sigma$), a confidence interval for the residuals can then be established.

The assumption that the residuals are normally distributed is not *strictly* necessary for the least square fitting of the regression model to hold. Moreover, the hypothesis tests used to validate the model are robust in the sense that only extreme departures from normality would cause spurious results. This assumption is based on theoretical and experimental results as pointed out in Kleinbaum, D. et al. (1998). However, since the confidence intervals for estimations of VSM are based on the assumption that the residuals are normally distributed, significant departure from normality would render poorly accurate confidence intervals for the MLRMs.

5.6.3 Hypothesis test

After checking for homoskedasticity and normality conditions, a hypothesis test is carried out in order to assess the significance of overall regression, Kutner, M. et al. (2004). The objective of the test is to verify if all considered independent variables (x_i) are meaningful in explaining the dependent variable (y). The test is formulated as shown in (5.13) – (5.19).

- H_0 - *Null hypothesis*: All independent variables (x_i) considered *do not* explain a significant amount of the variation in (y):

$$\alpha_j = \gamma_k = \omega_j = 0 \text{ for } \forall j \in \{1, p\}, \forall k \in \{1, p(p-1)/2\} \quad (5.13)$$

- H_a - *Alternative hypothesis*: The model is *well explained* by the independent variables (x_i):

$$\alpha_j \neq \gamma_k \neq \omega_j \neq 0 \text{ for } \forall j \in \{1, p\}, \forall k \in \{1, p(p-1)/2\} \quad (5.14)$$

The F statistics used in the test is defined in (5.15), where and MSR (regression mean square) and MSE (error mean square) are defined as shown in equations (5.16) and (5.17), respectively.

$$F = \frac{MSR}{MSE} \quad (5.15)$$

$$MSR = \frac{\sum_{i=1}^n (\hat{y}_i - \bar{y})^2}{k} \quad (5.16)$$

$$MSE = \frac{\sum_{i=1}^n (y_i - \hat{y}_i)^2}{n - k - 1} \quad (5.17)$$

The formal hypothesis test can be written formally as shown in equation (5.18).

$$F = \frac{MSR}{MSE} > F_{k, n-k-1, 1-\alpha}, \text{ a reject } H_0 \quad (5.18)$$

After calculating the F statistics, the result is compared with the critical point of the F distribution $F_{k,n-k-1,1-\alpha}$, where: α represents the preselected significance level, k is the number of regressors and n is the total number of samples used. If the calculated value for the F statistics exceeds the critical point for the F distribution, then the null hypothesis (H_0) is rejected.

5.6.4 Multicollinearity

If two or more independent variables show strong linear dependence, the matrix $(\mathbf{X}^T \mathbf{X})$ can become ill conditioned and computational problems may arise while calculating the regression coefficients □. Therefore, in case two or more RPRs are highly correlated, only one RPR is included in the model. This action helps us to avoid the occurrence of multicollinearity and enhance computational aspects of model calculation. A practical situation where multicollinearity occurs is when two identical machines are placed in the same bus (or nearby buses), thereby behaving in the exact same way as far as reactive power production and reactive power reserve depletion is concerned.

Once the validation process ends and all the MLRMs have been created, they are made available to online VSM estimation.

5.7 Results on sample systems

The proposed methodology has been tested on two different systems in order to verify its applicability to real sized systems. The IEEE30 bus test system and a larger case representing the eastern interconnection of the United States are used to test the methodology. It is important to mention that this large scale system represents an actual system, which is composed of more than 22 thousand buses. Results obtained from the simulations will further strengthen the practical capabilities of the method.

5.7.1 Description of test systems

The IEEE 30 bus test system was first used to implement the methodology presented on the flowchart in Figure 3.5. This system is composed of 30 buses, 6 generators, 22 loads, 35 lines and 6 transformers. A list of 50 contingencies including NERC category B (N-1), C (N-2) and D (N-K) contingencies is used to account for different network topologies, including all (N-1) contingencies. Fifteen random LIDs are used in order to account for uncertainty in load increase direction, totaling 750 scenarios (contingency + LID). Once the list of contingencies and LIDs are defined, the VSA begins in order to generate the database used to design the MLRMs. A summary of the IEEE 30 bus test system is given in Table 5.1.

Table 5.1. Description of the IEEE 30 bus test system

BUSES	PLANTS	MACHINES	LOADS	BRANCHES	TRANSFORMERS	DC LINES	FACTS DEVICES
30	6	6	22	35	6	0	0

The next test system is a real representation of the eastern interconnection of the United States. According to NERC, the North American power grid can be divided into three main regions: the western interconnection (WI), the eastern interconnection (EI) and the largest part of the state of Texas (ERCOT). Figure 5.6 shows a pictorial image of the national power grid and its three major administrative regions.

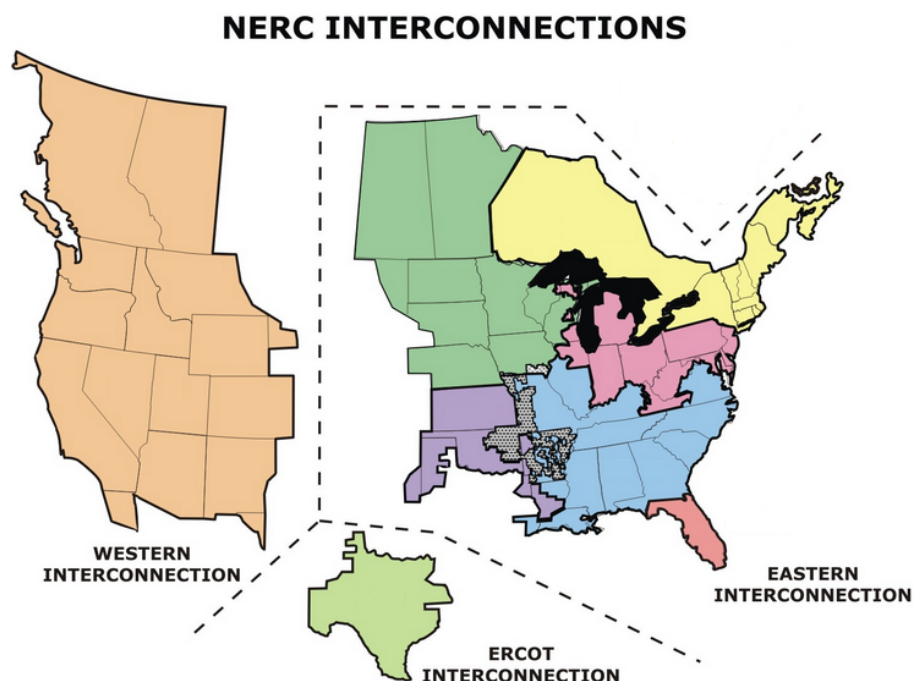


Figure 5.6. North American interconnections

The EI contains a total of 136 different areas, which can represent different generating companies, transmission companies and load serving entities. Some of these areas are geographically large and need to be further divided into zones. A total of 474 zones are represented in the EI, with a total forecasted load demand of approximately 650 GW for summer 2010, NERC TSD. (2010). Table 5.2 contains a detailed description of all system components present the Eastern interconnection.

Table 5.2. Description of the United States eastern interconnection

BUSES	PLANTS	MACHINES	LOADS	BRANCHES	TRANSFORMERS	DC LINES	FACTS DEVICES
48282	5862	5312	28613	62033	18675	23	0

Given the continental characteristics of the EI, the original power flow case used to represent the entire interconnection can be reduced around an area of interest. Moreover, considering the fact that reactive power cannot be transferred long distances without being consumed, the reduced system should be able to capture major local reactive power flows and voltage stability behavior. This reduction in system size is important because it simplifies the VSA and also reduces the computational burden of running hundreds of scenarios. A summary of the reduced EI case is shown in Table 5.3.

Table 5.3. Reduced case of the United States eastern interconnection

BUSES	PLANTS	MACHINES	LOADS	BRANCHES	TRANSFORMERS	DC LINES	FACTS DEVICES
21388	4648	4815	14766	44622	7709	23	0

A total of ten LIDs are considered to account for uncertainty in load increase direction. The contingency list comprises 190 critical outages is used to represent network topological changes. NERC category B, C and D were included in the list. All (N-1) contingencies in the 161kV network and above have been considered. A total of 1900 scenarios (contingency + LID) are thus used in this case. After reducing the size of the system, one particular area is selected to conduct the study. The studied area contains 815 buses, 46 machines, 570 loads, 1005 transmission lines and 149 transformers

The objective to use systems of different sizes is to verify if the number of required MLRMs would significantly increase with an increment in system size. It will be show later that the number of MLRMs does not increase with the size of the studied area, thus creating the expectancy that the technique can be successfully implemented on even larger networks.

5.7.2 The design of MLRMs

The MLRMs designed for the IEEE30 bus test system are presented in Table 5.4. Five MLRMs are found to be necessary in order to cover the entire set of selected contingencies and LIDs, according to the specified accuracy. Since a small number of RPRs is available in this test system, all RPRs but the slack bus are selected as regressors, totaling 5 RPRs. If all 5 linear terms of the RPRs are considered along with their crossed and squared terms, each MLRM will have a total of 20 regressors for this case.

The VSA identified that the most critical contingency had a VSM of 124 MW, whereas the less harmful one had a VSM of 620MW. As expected, the majority of the contingencies do not reduce the VSM of the system significantly, a fact that can be noticed by the number of test data available in each model.

The independent test samples have been obtained from different LIDs than the ones utilized on MLRM design. Each LID is randomly created by assuming each load is varied using a normal *pdf* with mean equal to the base case load and a standard deviation of 15% of the base case load. After randomly perturbed, the loads are increased proportionally to their initial value until the voltage collapse point is reached.

Each MLRM is able to estimate VSM with high accuracy for various load levels (any point along the PV curve) and for different LIDs than the ones used during the design phase. This is an important result since uncertainty in LID is an inevitably characteristic in any system, despite of how accurate load forecasting tools may be.

Contingencies belonging to the range presented in column two are used to design each one of the models. Column three presents the confidence intervals calculated for each model using the

training set. Column four contains the size of the dataset used for testing and column five contains the accuracy of the MLRMs on the test sets.

It can be observed from Table 5.4 that the overall accuracy of the models on their respective test sets is significantly high, thereby indicating that most models can accurately predict VSM and correctly incorporate the uncertainty through confidence intervals. In case the confidence intervals are found to significantly below 95%, more data can be generated and included in the training set in order to improve performance. A more detailed description of this procedure can be found in Leonardi, B. and Ajjarapu, V. (2011). Another option to improve the accuracy of the confidence interval is to move the cases with poor performance from the test set to the training set, thereby allowing the model to recognize those cases with greater precision.

Table 5.4. Multilinear regression models and respective confidence bounds for IEEE 30 bus test system.

<i>Model</i>	<i>VSM range (MW)</i>	<i>Confidence interval in MW</i>	<i>Independent test set</i>	
			<i>Number of estimated VSMs</i>	<i>Estimated VSMs within confidence interval (%)</i>
MLRM-1	620/496	± 38.8	1976	94.23
MLRM-2	495/372	± 36.8	258	97.7
MLRM-3	371/310	± 26.9	217	99.54
MLRM-4	309/248	± 18.7	59	100.00
MLRM-5	247/124	± 19.1	73	98.7

The reduced case of the eastern interconnection is used next to test the proposed methodology. The area under study is around 27 times larger than the IEEE 30 bus system and

the MLRM derived for this case are presented in Table 5.5. The meaning of each column is the same as that for Table 5.4, thus being omitted here for the sake of simplicity.

Despite the increment in size of the studied area, only four MLRMs are necessary to handle all LIDs and 190 contingencies considered. Similarly to what has been observed for the IEEE30 bus system, the accuracy of all four designed MLRMs is reasonably high, varying between 94-96%.

Table 5.5 Multilinear regression models and respective confidence bounds for reduced case of the Eastern Interconnection

<i>Model</i>	<i>VSM range (MW)</i>	<i>Confidence interval in MW</i>	<i>Independent test set</i>	
			<i>Number of estimated VSMs</i>	<i>Estimated VSMs within confidence interval (%)</i>
MLRM1	5425/4341	± 389	7567	95.85
MLRM2	4340/3257	± 332	799	94.11
MLRM3	3256/2173	± 265	119	94.96
MLRM4	2172/ 1089	± 159	70	94.29

The same design standard used in the IEEE30 bus test system is adopted here. According to the proposed procedure, the number of MLRM is determined by how many times the data (VSM range) needs to be split. Similarly to the previous case, the threshold considered for this system also tries to maintain the 2 sigma confidence interval smaller than 10% of the upper VSM.

After all MLRMs have been properly designed, validation steps are taken in order to enable these models to be used in practice.

5.8 Validation of MLRMs

In order to verify basic statistical properties of the models are held, MLRM-1 is selected from the IEEE30 test system, whereas MLRM-2 is selected from the reduced EI. These models will initially be tested for homoskedasticity and adequacy of quadratic fit. In the sequence, verification of normally for the residuals and the aptness of overall regression through hypothesis test will be performed.

5.8.1 Homoskedasticity verification

In order to check if heteroskedasticity is present, Figure 5.7 and Figure 5.8 contain plots of the residuals (ϵ) versus the regressed variable (\hat{y}) for MLRM-1 and MLRM-2, respectively. This plot is a helpful indicator of the level of heteroskedasticity that may be present in the regression model, Kleinbaum, D. et al. (1998).

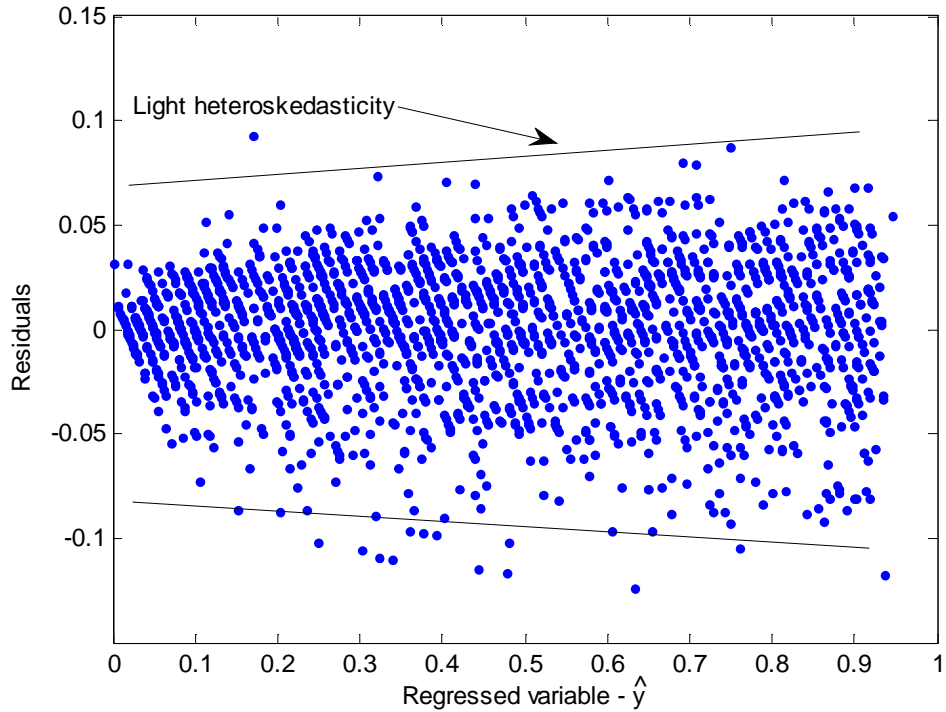


Figure 5.7. Residual plot to verify presence of heteroskedasticity or statistical inconsistency on the residuals – IEEE30/MLRM-1

The funneling of the residuals along the range of \hat{y} values in Figure 5.7 indicates that light heteroskedasticity is likely to occur in MLRM-1. Although a tenuous level of heteroskedasticity is present, remedial actions are only necessary when the data shows significant departure from homoskedasticity, Kleinbaum, D. et al. (1998).

On the other hand, Figure 5.8 shows that the residual distribution of EI/MLRM-2 is practically homoskedastic. In case the presence of heteroskedasticity is significant, *Box Cox* power transformations of the dependent variable (y) can be used to correct for heteroskedasticity, Kutner, M. et al. (2004) and Box, G. E. P. and Cox, D. R. (1964).

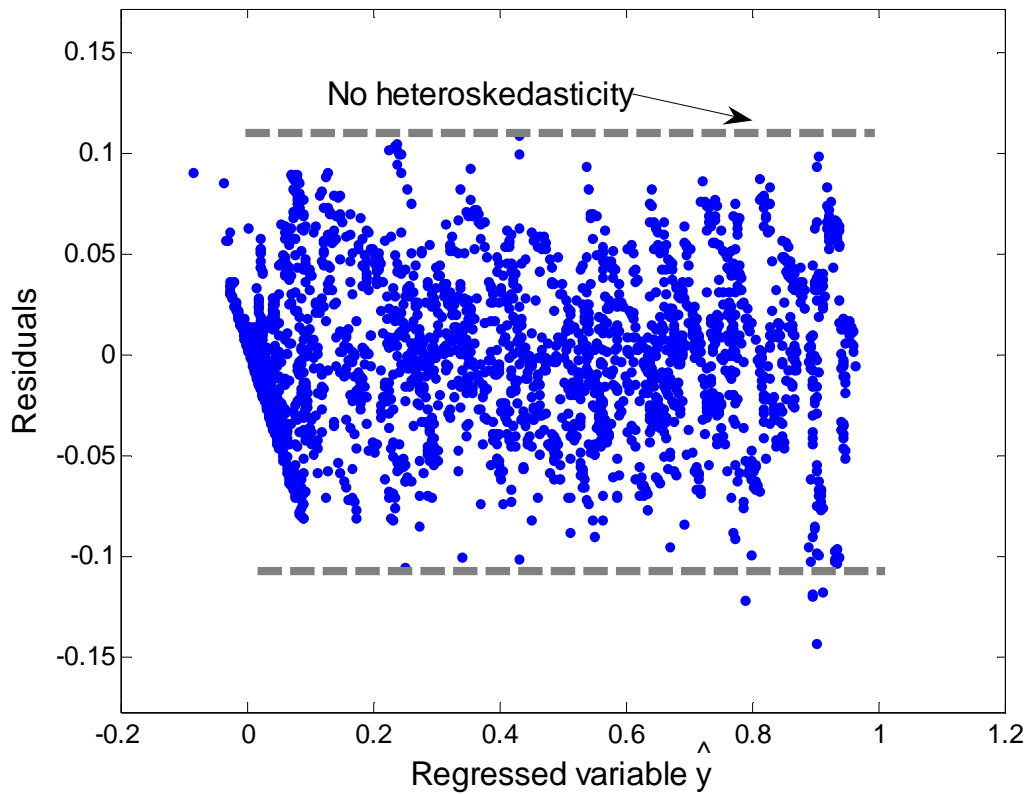


Figure 5.8. Residual plot to verify presence of heteroskedasticity or statistical inconsistency on the residuals – EI/MLRM-2

Another important aspect of these plots is regarding the proportionality of data distributed above and below zero residual. This observation indicates that skewness is not significantly present in both cases, as required by the normal *pdf*. The absence of patterns (linear or curvilinear) in the residual plot also reinforces the fact that the degree of regressors considered in the MLRMs is appropriate. The presence of patterns in the residuals is an indication that either the order of the MLRM (linear or quadratic) or the selected set of regressors is inadequate.

5.8.2 Verification of normality

After checking for the presence of heteroskedasticity, the histogram of the residuals is plotted along with the best normal fit for both MLRM-1 and MLRM-2 in Figure 5.9 and Figure 5.10, respectively.

Figure 5.9 shows that the residual distribution follows a bell shaped curve which closely resembles a normal *pdf*.

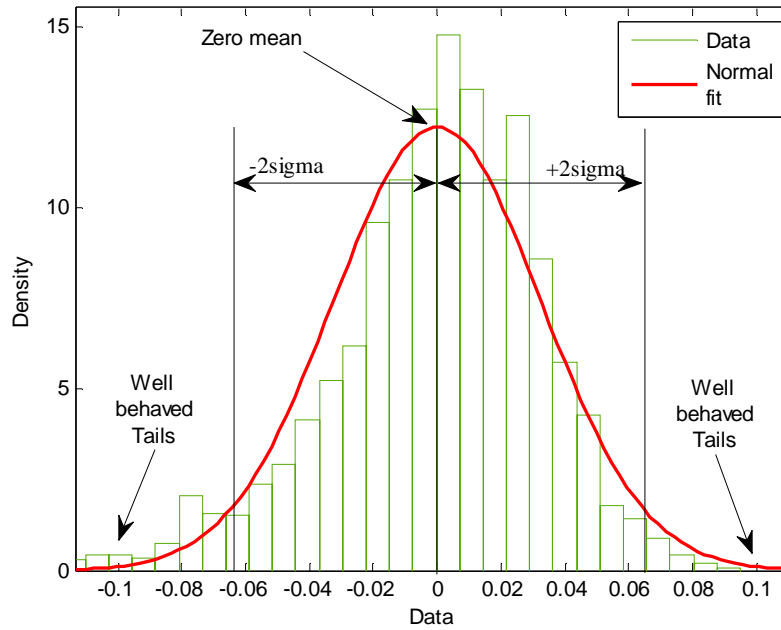


Figure 5.9. Residual histogram and best normal *pdf* of MLRM-1

The standard deviation (σ) for the best normal distribution fit is 0.0313 and the mean (μ) is $2.9 \cdot 10^{-16}$, practically zero as required by equation (5.12). After transforming the normalized σ to MWs, the $\pm 2\sigma$ confidence interval found for MLRM-1 was ± 38.8 MW, as previously indicated in Table 5.4.

Figure 5.10 shows that the histogram of the residuals for MLRM-2 also follows a bell shaped curve closely resembling a normal distribution. The best normal fit in this case is plotted along

with the $\pm 2\sigma$ confidence interval. The tails of the normal distribution fit are well behaved, further supporting the condition that only a few observations will fall out of the $\pm 2\sigma$ range shown in Figure 5.9 and Figure 5.10.

The calculated standard deviation of the residuals for MLRM-2 is 0.0385 with a mean equal to -6.8×10^{-14} , close to zero as required by equation (5.12). This standard deviation will represent a $\pm 2\sigma$ confidence interval of ± 332 MW , as previously indicated in Table 5.5.

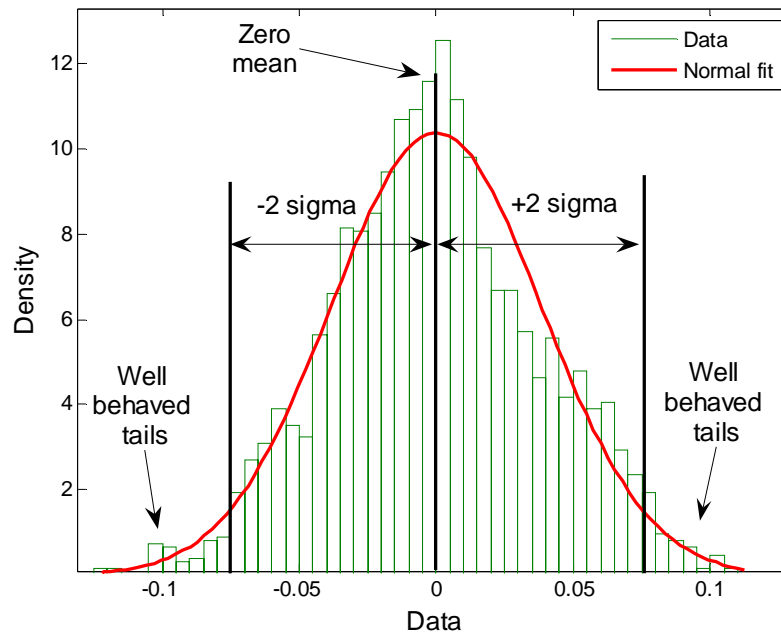


Figure 5.10. Residual histogram and best normal *pdf* of MLRM-2

According to Kleinbaum, D. et al. (1998), only extreme departures from normality would cause spurious results of the models. Moreover, the hypothesis tests used to validate the model are robust in the sense that only extreme deviations from normality would cause spurious results. This assumption is based on theoretical and experimental results as pointed out in Kleinbaum, D. et al. (1998).

Therefore, it can be assumed that the tested MLRMs comply with the minimal requirements of homoskedasticity and normality.

5.8.3 Test of overall regression aptness

After testing the residuals of the MLRMs for the presence of heteroskedasticity, a final hypothesis of overall regression is performed to validate the model. Since only 5 machines are available to design MLRM-1, the total number of regressors is 20, representing 5 linear, 10 crossed and 5 quadratic terms as defined in (5.1). The *F statistic* obtained from the ANOVA table for MLRM-1 is equal to 6740.3.

After calculating the critical value of the F distribution for $k=20$ (number of regressors), $n=2000$ (number of samples used to generate the model) for a significance level $\alpha=5\%$ ($F_{k,n-k-1,1-\alpha}$), the hypothesis test is performed as shown in equation (5.19).

$$F_{statistic} = \frac{MSR}{MSE} = 6740.3 > F_{20,1979,0.95} = 1.575 \quad (5.19)$$

Therefore, since the F statistic is larger than the critical value of the F distribution as shown in equation (5.19) above, the null hypothesis is rejected and the coefficients are considered to explain the model satisfactorily.

Once the model for the IEEE30 bus case has been validated, we turn our attention to the second model. Since this system is significantly larger than the IEEE30 bus test system, more RPRs are available for MLRM development and thus only the most effective reserves can be selected. Following the steps proposed in the flowchart on Figure 5.3, nine machines have been selected out of a total of 46. The pre-selection of RPRs removes those reserves that do not significantly vary during the VSA, or quickly exhaust their reactive support capability, hence not

being able to provide useful information about VSM as system load increases. Another criterion for RPR selection is to select those machines which have a better individual relationship with VSM. The quality of the relationship can be measured with simple statistical parameters, such as the coefficient of multiple determination (R^2) and size of residual variance.

A total of 54 regressors compose MLRM-2, representing the 9 linear terms of each RPR, 36 crossed and 9 squared RPRs terms, as described in (5.1). The hypothesis test carried out with $k=90$ (number of regressors), $n=2000$ (number of samples used to design the model) for a significance level of $\alpha=5\%$. The F statistic of MLRM-3 obtained from the ANOVA table is equal to 2763.1. The hypothesis test is formally performed as indicated in equation (5.20).

$$F_{statistic} = \frac{MSR}{MSE} = 2763.1 > F_{54,1946,0.95} = 1.343. \quad (5.20)$$

Therefore, it can be concluded from (5.18), (5.19) and (5.20) that the hypothesis test rejects the null hypothesis (H_0) and the alternative hypothesis (H_a) is accepted. These results reinforce the assumption that the dependent variable (y) is effectively explained by the regressors (x_i) in both MLRM-1 and MLRM-2.

After validation steps have been taken and basic statistical properties have been met, the MLRMs are ready to be used in online VSM estimation.

5.9 Conclusions

The results have demonstrated how multilinear regression models can be used to estimate VSM in real time based on the amount of RPR. A small number of MLRMs is necessary even for practical real-sized systems. The addition of confidence interval to the estimated VSM value can help operators to account for uncertainty involved with changes in LID and network topology. However, since more than one MLRM is available, a tool needs to be developed in

order to help system operators to identify the adequate model based on current system conditions.

CHAPTER 6. MACHINE LEARNING TECHNIQUES FOR MULTICLASS CLASSIFICATION

This chapter introduces the basis of data mining and machine learning techniques investigated in this research. Initially, a brief explanation of the mathematical aspects of each technique is given. Simulations results provided in the next chapter are used to demonstrate the effectiveness of each technique when applied to the problem of MLRM identification.

6.1 Data mining overview

Data mining is usually defined as the process of discovering and extracting patterns and knowledge from data. The word *data* can have broad meaning and will be defined as any measurement, variable or information available from the power system. Several successful applications of data mining have been documented in a wide range of fields ranging from profiling practices (for market strategies) to scientific discovery in various areas of knowledge, Frank, I. and Witten, E. (1999).

The process of mining data begins once a database containing samples that represent a process is made available for mining. It basically consists in learning and extracting patterns from data in order to be able make non-trivial predictions on unforeseen instances. This process of knowledge extraction is commonly accomplished by machine learning techniques.

Machine learning algorithms represent the practical applications of rules and mathematical methods used to obtain knowledge from a database. In this study, four machine learning techniques have been investigated in order to solve the practical problem of MLRM identification. Despite the fact that each approach has its own advantages and disadvantages, a detailed analysis of the results will help us to analyze and identify which technique is more

appropriate for the problem at hand. The investigated techniques are artificial neural ANNs, DTs, SVMs and KNNs. An introduction to each one of the algorithms is presented next.

6.2 Investigated algorithms

6.2.1 Decision trees

Decision tree is a powerful machine learning technique commonly employed in classification and regression tasks. It has been applied to different areas of power systems as described by Diao, R. et al. (2009), Goubin, M. (1996). Van Cutsem, T. et al. (1993) have used decision trees to perform a voltage security assessment of the grid to identify potential voltage violations. Morison, K. et al. (2008) used decision trees in order to identify voltage control areas based on current system operating conditions.

Figure 6.1 shows a pictorial representation of a decision tree branch applied to the problem at hand. Two different types of attributes are used in this case: line active power flow (Line_i) and bus voltage magnitude (Bus_i). The first node of the tree is usually known as the *root node* and nodes originated from the root node are referred as *child nodes* or *splitting nodes*. The nodes in the extremities of the tree are usually known as *leaf nodes* or *terminal nodes* and contain the class attribute, which in this case represents a MLRM. In the example below, the root node is not shown since the branch is only a partial representation of the tree.

As large power systems are composed of thousands of variables, one of the objectives of this research is to identify the most effective variables and use them to develop the MLRM-IDtool.

Although several variables are made available to the decision tree algorithm, only a few of them are selected to be a part of the tree. The selection process is done based on the amount of

information gain that each attribute provides. Attributes with higher information gain are selected to be a part of the tree.

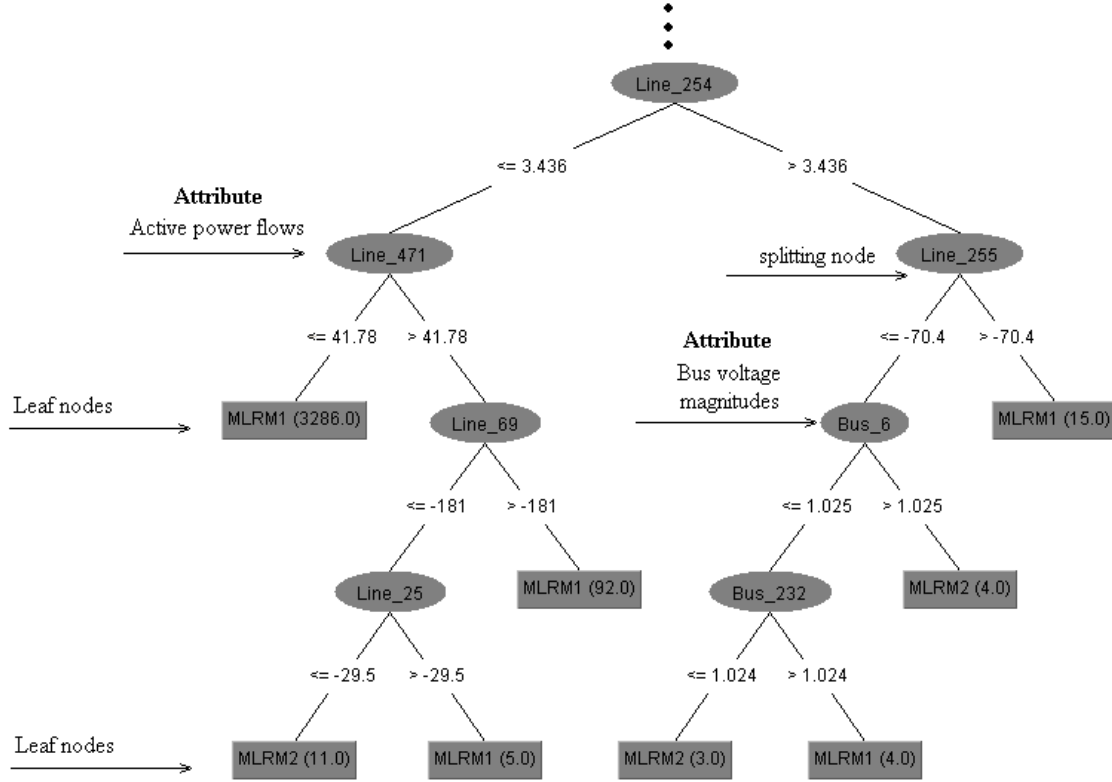


Figure 6.1. Partial representation of decision tree branch

Once the DT is constructed, the classification of an instance begins at the root, which will further activate one of its branches based on the splitting decision. In the sequel, the instance to be classified is passed onto a child node. The process follows until the instance reaches one of the leaf nodes and is properly classified.

In order to reduce the size and complexity of the tree, a mechanism known as pruning is commonly employed. The major objectives to prune a tree are to reduce the complexity and create a tree that can offer better generalization. Moreover, smaller trees are easier to handle and to analyze. In addition to that, smaller trees usually have better predictive accuracy since they do

not *overfit* the training data set, thereby showing enhanced performance on unforeseen instances, Frank, I. and Witten, E. (1999).

6.2.1.1 *DT algorithms*

Various algorithms have been proposed to induct decision trees up to date, Morgan, J.N. and Sonquist, J.A. (1963), Breiman, L. et al. (1984) and Quinlan, J.R. (1986). However, two of algorithms have received great attention in the machine learning society and practitioners for having robust performance and demonstrated scalability: CART and C4.5.

The classification and regression trees (CART) algorithm proposed by Breiman, L. et al. (1984) is a decision tree building technique proposed in the 80's. It can produce either a classification or a regression tree, depending on whether the class attribute is nominal or numeric, respectively. The tree is formed by a collection of rules based on values of certain variables in the modeling data set and it takes both nominal and numeric attributes as inputs during its construction.

The C4.5 algorithm developed by Ross Quinlan and is an extension of Quinlan's previous ID3 algorithm, whose development dates back to the late 70's. The *divide and conquer* approach represents the core of the ID3 algorithm. Although the ID3 algorithm works well, it can only consider nominal attributes during the development of the tree, hence not being able to incorporate most SCADA/EMS numeric attributes available in this project. Therefore, the evolved version of the ID3 algorithm (the C4.5) is used since it can handle numeric attributes.

However, differently from CART, the C4.5 is only capable of creating decision trees for classification purposes, also using both numerical and nominal attributes as inputs.

Although both CART and C4.5 have been successfully employed in power system applications showing similar performances, the C4.5 algorithm has been selected to be used in this project. Two main reasons led us to make a decision towards C4.5.

Firstly, the problem at hand (identification of what MLRM to use) is a classical classification problem, where the class attribute (output of the tree) is nominal and represents a MLRM. Secondly, the algorithm is already implemented and has been made available at no cost in WEKA. The WEKA program is an open source machine learning software developed and maintained by the University of Waikato, New Zealand, Hall, M. et al. (2009).

Another reason that supported our decision to select the C4.5 algorithm has been its successful employment in the Powertech Labs[®] *Voltage Control Area* (VCA) software, hence demonstrating the necessary scalability and robustness required by large scale power system applications, Morison, K. et al. (2008).

6.2.1.2 Discretization of numeric attributes

In order to make use of numeric attributes in the tree building process, the C4.5 algorithm uses a method called attribute discretization to transform numeric attributes into nominal ones. The idea is simple: to separate the numerical attributes into in binary intervals in order to attain the highest purity level, i.e., to represent one single class on each interval.

Several methods to discretize attributes are available in the literature. Nonetheless, in the discretization process, decision tree algorithms usually use the same entropy-based method to identify which attribute to use, Frank, I. and Witten, E. (1999).

A practical discretization of an attribute is presented next in order to illustrate the methodology. The attribute line active power flow presented in Figure 6.1 will be discretized using the concept of information value described next.

Frequently, the dividing thresholds (represented from (A) to (H) in this case) are placed halfway between the values that delimit the boundaries of the attribute. However, enhancement in classification performance might be gained by adopting a more sophisticated method.

Table 6.1. Discretization of a numeric attribute

Line active power flow (MW)	64	65	68	69	70	71	72	75	80	81	83	85
Class	M1	M2	M1	M1	M1	M2	M1 M2	M1 M1	M2	M1	M1	M2
	(A)	(B)			(C)	(D)	(E)	(F)	(G)		(H)	

Considering the example shown in Figure 6.1, there are only eleven possible positions for the breakpoint (or eight if the breakpoint is not allowed to separate items of the same class (for instance, M1-M1 or M2-M2)). The information gain calculated for each breakpoint is shown in (6.1).

$$\begin{aligned} \text{(A) } \mathbf{info}([1, 0], [8, 5]) &= -(\mathbf{1/14}) \cdot ([1/1 \cdot \log(1/1)] + [0/1 \cdot \log(0/1)]) + L \\ &\quad - (\mathbf{13/14}) \cdot ([8/13 \cdot \log(8/13)] + [5/13 \cdot \log(5/13)]) = 0.893bits \end{aligned}$$

$$\text{(B) } \mathbf{info}([1, 1], [8, 4]) = 0.930bits$$

$$\text{(C) } \mathbf{info}([4, 1], [5, 4]) = 0.895bits$$

$$\text{(D) } \mathbf{info}([4, 2], [5, 3]) = 0.939bits \quad (6.1)$$

$$\text{(E) } \mathbf{info}([5, 3], [4, 2]) = 0.939bits$$

$$\text{(F) } \mathbf{info}([7, 3], [2, 2]) = 0.915bits$$

$$\text{(G) } \mathbf{info}([7, 4], [2, 1]) = 0.939bits$$

$$\text{(H) } \mathbf{info}([9, 4], [0, 1]) = 0.827bits$$

The breakpoints with lowest information are usually located at the extremities of the interval. In this example, the break point located at 84MW ($\mathbf{info}([9, 4], [0, 1]) = 0.827bits$) has the lowest

information value (hence the highest information gain) and should be the one used to break the attribute into two ranges.

Ideally, each range should be pure after separation. However, in case both classes are present in a range, the majority class will be used to determine the classification of further instances. In the previous example, the breakpoint of the attribute line active power flow is selected as shown in Figure 6.2.

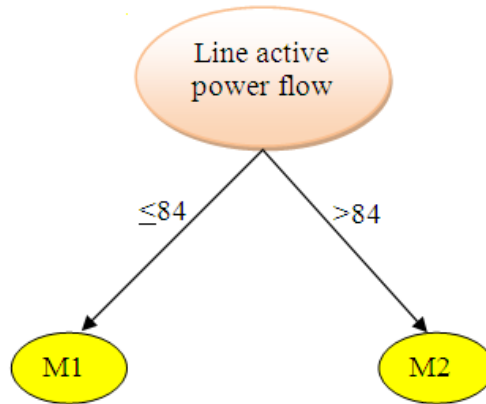


Figure 6.2. Attribute line active power flow after discretization

This process is repeated until all numeric attributes have been discretized and the tree completely developed. Once the tree is validated, it can then be used in real time operations for online MLRM identification.

6.2.2 Artificial Neural Networks (ANN)

ANNs have emerged in the late 80's as a practical technology with successful applications in many fields. Although several different topologies of ANNs have been proposed in the literature, the multilayer perceptron network and radial basis function network have demonstrated great success in pattern recognition and classification problems Bishop, C. M., (1995). In this work, we have only considered multilayer perceptron ANN for its widely known capabilities and

demonstrated performance on classification problems. Moreover, computationally efficient methods to train this type of ANN such as the back propagation algorithm are available and widely employed in practical designs.

A typical multilayered feed forward ANN is shown in Figure 6.3. Throughout this dissertation, the ANNs utilized will be formed of three layers unless stated otherwise. The layers that compose an ANN are: the input layer, the hidden layer (also called intermediate layer) and the output layer. Except for the input and output layers, an ANN can have as many hidden layers as necessary. Each layer has a certain number of neurons, which are interconnected through the links as shown in the picture below.

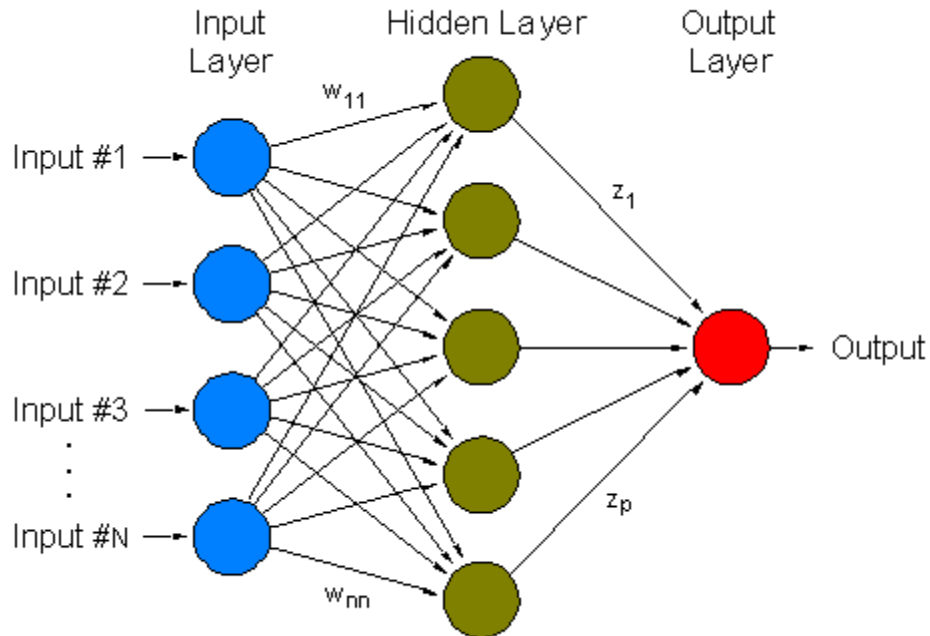


Figure 6.3. Artificial neural network structure

Data is initially supplied to the input layer so it can propagate through the network. After the inputs are provided, sequential multiplication by the weigh factors and addition to the local

neuron biases occurs as the impulse signal moves along the intermediary layers. In the end, all signals are added up to produce the output value. A mathematical representation of the ANN utilized in this work is give in Figure 6.4.

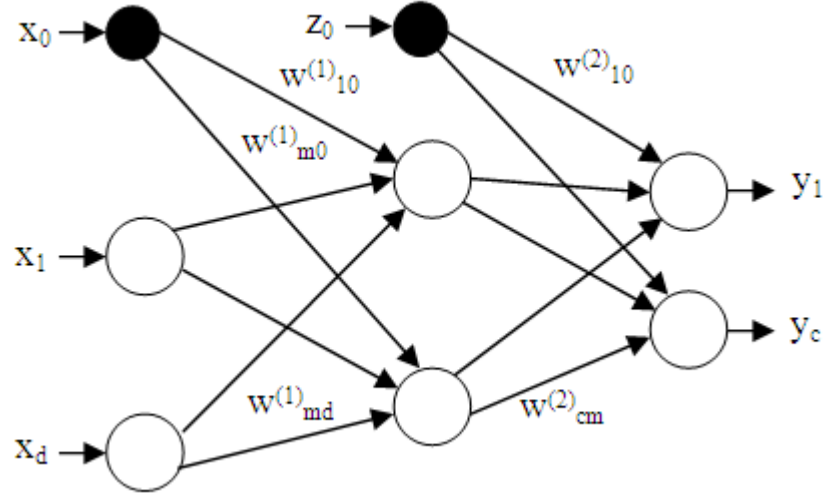


Figure 6.4. Mathematical representation of an ANN

Data is presented at the input layer in the form of a vector $x = [x_1 \ x_2 \ L \ x_d]$. Each dimension is then multiplied by weighting factors w_{ji} and summed up with a bias factor w_{j0} as shown in equation (6.2).

$$a_j = \sum_{i=1}^d w_{ji}^{(1)} x_i + w_{j0}^{(1)} \quad (6.2)$$

The quantities a_j are known as *activations* and are obtained by the product of the input values by the weighting factor, added with the bias factor. Once the activation value is obtained, it will be transformed using a nonlinear activation function $h(\cdot)$. Common examples of activation functions include the sigmoid and hyperbolic tangent functions, Bishop, C. M. (1995). The activation value a_j is transformed into the value z_j as shown in (6.5).

$$z_j = h(a_j) \quad (6.3)$$

In this work, a hyperbolic tangent function is used as the activation function in the hidden layer and a linear function is used in the output layer. The hyperbolic tangent function is described in equation (6.4).

$$h(a_j) = \frac{e^{2a_j} - 1}{e^{2a_j} + 1} \quad (6.4)$$

After vector the hidden layer outputs z_k are generated, they are multiplied by the weighing factors linking the hidden layer to the output layer as shown in Figure 6.4. The output activation terms c_k are then formed as shown in equation (6.5).

$$c_k = \sum_{j=1}^m w_{kj}^{(2)} z_j + w_{j0}^{(2)} \quad (6.5)$$

Finally, the output terms y_k are generated by transforming the activation values c using another hyperbolic tangent sigmoid function $\tilde{\sigma}$ as described in equation (6.6).

$$y_k = \sigma(c_k) \quad (6.6)$$

The final formulation representing the multilayered perceptron ANN model is described in equation (6.7).

$$y_k(x, w) = \sigma \left(\sum_{j=1}^m w_{kj}^{(2)} h \left(\sum_{i=1}^d w_{ji}^{(1)} x_i + w_{j0}^{(1)} \right) + w_{j0}^{(2)} \right) \quad (6.7)$$

Once the ANN model is defined, the network is trained using the back propagation algorithm where the deviations from the target outputs are used to adjust the weights $w_{kj}^{(1)}$ and $w_{kj}^{(2)}$ in the network. The back propagation algorithm is repeated until the mean square error falls below a certain threshold or the maximum number of training epochs is realized.

6.2.3 Instance based learning and K-Nearest-Neighbor (KNN)

The nearest neighbor method is one of the most simple and yet widely used machine learning techniques. More popularity has been gained after the work developed by Aha, D et al. (1991) and Aha, D. (1992), where it has been shown that the removal of noisy attributes and the weighting of selected attributes could make the KNN approach perform better than other popular machine learning techniques.

The technique explores the concept of memorization, one of the simplest and yet most powerful forms of learning. Similarly to what happens in the human brain, machines can also use the concept of memory to classify objects. Basically, once a set of instances (samples) has been memorized, the classification of a new instance can be achieved by matching the closest instance/s in the database.

This type of learning is named *instance based learning* and *K nearest neighbors* (KNN) stands as one of the most popular instance based learning techniques. The most notable difference between this type of techniques and other classification approaches is the instant at which learning takes place.

Contrary to decision trees and other classification methods, where a model (tree, ANN or SVs) is obtained (learned) from the training set as soon as the set is made available, instance based learners only learn when they are required to. In other words, learning only occurs when a new instance is presented for classification.

Due to this characteristic, instance based learners like the K nearest neighbors are also known as *lazy methods*. While other approaches produce a generalized model based on the training data that is made available, instance based learners defer the real work as long as possible. Figure 6.5 pictorially explains the different approaches carried by both techniques.

For instance, in the decision tree algorithm, a tree is developed based on information (attributes and instances) contained in the training database. After designed, the model (decision tree) is used to classify new instances that are presented to it. Therefore, there is no learning at the time which the instance is presented for classification.

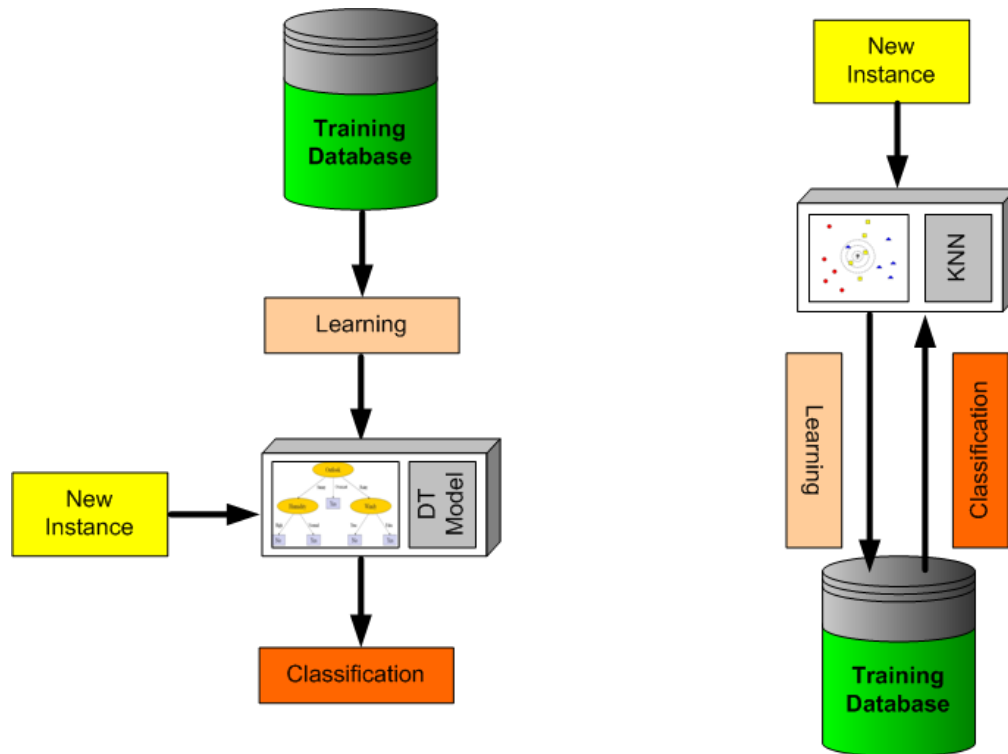


Figure 6.5. Comparison of learning approaches

On the other hand, a K nearest neighbor classifier only classifies a new instance is when it is presented for classification. After the new instance is made available, the KNN method compares it with the K closest instances on its database. The closest neighbors are identified by measuring the distance between the new instance and all of its neighbors. After the K nearest neighbors have been identified, a simple count of their classes can determine the most frequent class, which will in turn be assigned to the new instance.

Conceptually, the KNN approach is quite simple: it tries to classify the new instance based on its nearest neighbors. In order to identify the nearest neighbors, the concept of distance must be properly understood. Consider two vectors $x = [x_1, x_2, \dots, x_n]$ and $y = [y_1, y_2, \dots, y_n]$ belonging to a vector space of dimension n (for instance, \mathbb{R}^n). Several definitions of distance commonly employed by the KNN method are presented in Table 6.2.

Table 6.2. Distances commonly used in the KNN approach

Distance	Formula
<i>Euclidean</i>	$D(\mathbf{x}, \mathbf{y}) = \sqrt{\sum_{i=1}^n (x_i - y_i)^2}$
<i>Manhattan</i>	$D(\mathbf{x}, \mathbf{y}) = \sum_{i=1}^n x_i - y_i $
<i>Chebyshev</i>	$D(\mathbf{x}, \mathbf{y}) = \max_i x_i - y_i $

Although various distances are available, the standard Euclidean distance is most frequently used in the literature and is thus employed in this project. An analysis of the impact of different distances on KNN performance can be made in order to find the best one for the current application.

A pictorial description of how the KNN approach works in practice is given in Figure 6.6. Attempting to classify the new instance (represented by ?), the KNN algorithm calculates the distance of the K nearest neighbors.

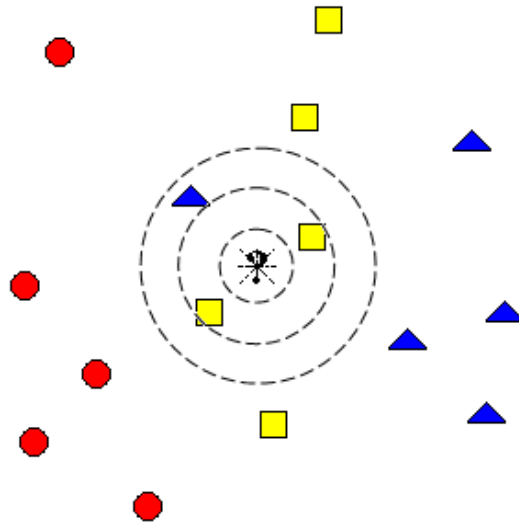


Figure 6.6. The K-nearest neighbor approach

The number of neighbors (K) can vary from case to case and must be selected based on a tradeoff between performance and computational time. For the application herein developed, K is set to 3 after good results have been obtained.

Assuming that $K=3$ for the classification case shown in Figure 6.6, there is a total of two squares and one triangle neighbors around the new instance (?). Therefore, since the majority of neighbors are “squares”, the new instance is classified as a “square”.

6.2.3.1 Processing speed and storage requirements

The KNN method tends to be slow in very large data bases since the distances between the new instance and all instances in the data base need to be calculated. However, in case only a single instance is presented for classification, all distances can be calculated quickly even on regular desktop computers. Since the application developed here only requires the identification of a single operating condition at the time, classification speed does not adversely affect the performance of the approach.

Another common drawback of the KNN approach mentioned in the literature is related to large amounts of storage requirements. Although storage capacity might have been a problem at the time when the technique was developed (and even in the early 90's), current computational advancements have lifted this limitation. For instance, the problem addressed in this project contains thousands of attributes and tens of thousands of instances. Nonetheless, no special storage has been necessary and all simulations are performed on a desktop computer with a Pentium® 2.6 GHz processor, 1 GB of RAM and an 80GB hard drive, a relatively modest configuration compared to currently available technology.

6.2.4 Support vector machines

Support vector machine is a powerful supervised classification technique initially proposed by Cortes, C. and Vapnik, V. (1995). The idea is based on the identification of the best hyper plane used which separates two different classes. Identification of the best hyperplane involves the maximization of the margin ($\frac{2}{\|w\|}$) between the two support hyper planes. Figure 6.7 shows the support vectors that are used to create the support hyper planes.

A quadratic optimization problem is formulated to identify the plane coefficients (denoted by w). The addition of slack variables ξ_i is needed when the datasets are not completely separable, i.e., when the instances cannot be completely separated by the hyper plane. The objective function will thus be formed by the margin term plus a penalty function in the form $C \sum_{i=1}^n \xi_i$, which can measure “how bad” the misclassifications are. Penalty functions in this form have the advantage of maintaining the optimization problem quadratic and convex. Therefore, the final

optimization used to determine the coefficient vector \mathbf{w} problem can be formulated as shown in equation (6.8).

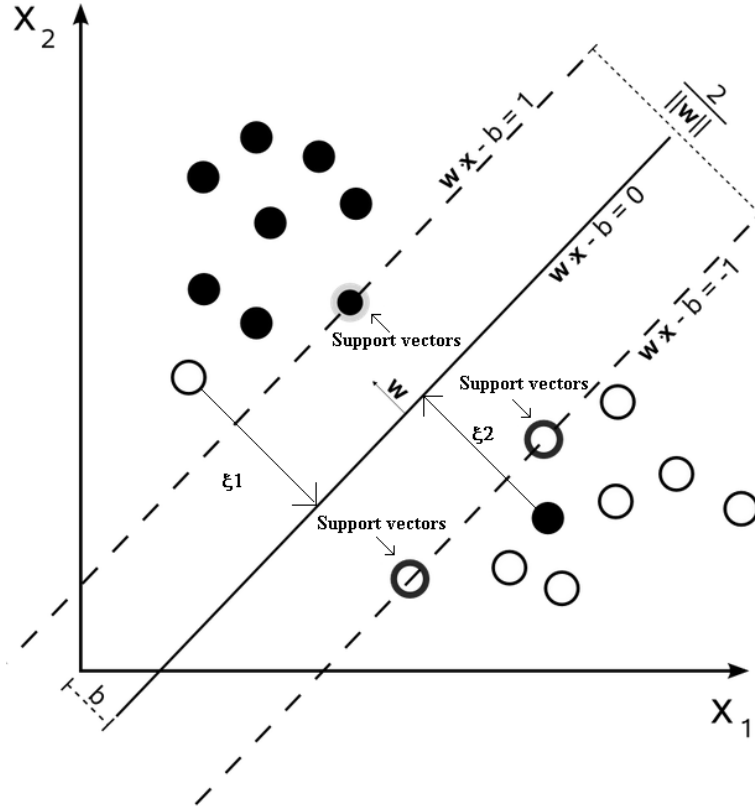


Figure 6.7. Support vector machine concept

$$\begin{aligned}
 & \text{Min } \frac{1}{2} \|\mathbf{w}\|^2 + C \sum_{i=1}^n \xi_i \\
 & \text{s.t.} \\
 & y_i (\mathbf{w}^T \mathbf{x}_i - b) - 1 + \xi_i \geq 0, \quad \forall i \\
 & \xi_i \geq 0, \quad \forall i
 \end{aligned} \tag{6.8}$$

The corresponding Lagrangian function is represented in equation (6.3).

$$L(\mathbf{w}, b, \xi, \lambda, \mu) = \frac{1}{2} \|\mathbf{w}\|^2 + C \sum_{i=1}^n \xi_i - \sum_{i=1}^n \lambda_i (y_i (\mathbf{w}^T \mathbf{x}_i - b) - 1 + \xi_i \geq 0) - \sum_{i=1}^n \mu_i \xi_i \tag{6.9}$$

The constant C represents the cost of each slack variable, \mathbf{x}_i is an input vector with \mathbf{y}_i being the class it belongs to (± 1); \mathbf{w} is the weight factor vector, b is the plane coefficient, λ_i and μ_i are the Lagrangian coefficients for the inequality constraints. Solution of the quadratic optimization problem (6.8) will identify the coefficients for the hyperplane that maximizes margin. A more detailed description of SVMs and their applications to classification problems can be found at Bishop, C. M. (2006).

Once the investigated machine learning techniques have been introduced, we shall focus our attention on how we can use them to perform multiclass classification. A detailed description of the most common multiclass classification methods is given next.

6.3 Multiclass classification problem

Multiclass classification problems arise frequently in various situations of daily life. For instance: assume that a professor wants to rank students grades (from A to E) based on their exam scores, homework scores and class participation. Once he determines his rules, the students are classified into 5 different groups based on their exam scores, homework scores and class participation. The set of rules imposed by the professor are equivalent to the classification tool used to classify the students.

Differently from binary classification, where only two classes are present (in the above example, five classes are available: A, B, C, D and E), multiclass problems involve several classes and are usually harder to solve than binary classification problems.

The complexity of the classification techniques employed will depend on the degree of difficulty to distinguish among the classes. Initially, the simplest approaches are used to address the classification problem. In case poor classification performance is obtained, more complex approaches are utilized in order to improve classification precision.

Several approaches have been proposed in the literature to address these kinds of problems, with the most popular being the extension of binary classification algorithms to *multiclass classification*, the *binary decomposition* of the multiclass classification problem into several two class problems and *hierarchical classification* methods.

The *multiclass classification* approach uses one classifier to classify all the classes at the same time. This approach can only be used if the classification technique can be modified to handle multiple classes in its formulation (e.g., DTs, ANNs and KNN). Some powerful binary classification techniques (e.g., support vectors machines) are not commonly employed since their extensions to multiclass problems are still at a developing stage, Fürnkranz, J. (2007) and Bishop, C. M. (2006). The main advantage of this approach relies on the fact that a single classifier is needed to classify all classes present in the problem. However, it is common to observe a lower classification precision when compared to the other two classification approaches for problems where the classes are difficult to separate.

The so-called *binarization methods* divide a multiclass classification problem into several binary ones. The *one versus all* (OVA) approach trains a classifier using one of the classes against all the other classes at the time. For instance, if n classes are present, the OVA approach will require a total of n classifiers to perform the classification. The fact that the OVA method has shown superior performance compared to the single multiclass technique has made the approach one of the most popular for multiclass classification, Fürnkranz, J. (2007).

The *one versus one* (OVO) approach trains a classifier to distinguish a class from every other class. If n classes are present, a total of $n*(n-1)/2$ classifiers are needed. This approach usually performs better than the multiclass and the OVA approaches mentioned above. A voting scheme using the output of each binary classifier combines all outputs of the classifiers into a final

classification result. The major drawback of the OVO strategy is the fact that the number of classifiers grows quadratically with the number of classes, hence making the approach cumbersome in case many classes are present. A pictorial representation of the three aforementioned classification methods is given in Figure 6.8.



Figure 6.8. Different multiclass classification methods

While the approaches mentioned above are more common in the literature, a hierarchical classification method has shown very good performance on multiclass classification problems Ananda, R. et al. (1995). The main advantage of such hierarchical classification approach compared to the OVA and OVO approaches is that it may reduce the total number of classifiers needed for classification.

The hierarchical classification scheme investigated in this work is schematically represented in Figure 6.9. The figure shows that classes C1 and C5 are classified with high precision by the first classifier, whereas class C2 is accurately classified by the second classifier and classes C3 and C4 are passed to the third classification stage. Therefore, only three classifiers are necessary in order to distinguish among all classes with high precision. The modification of the hierarchical classification methods relies on the fact that more than one class can be successfully identified at

each classification stage, thereby reducing the number of classifiers needed. Such reduction is only possible because the algorithms used at each stage are capable of performing multiclass classification.

Each block in Figure 6.9 contains a classifier which can be a DT, an ANN, a KNN or a combination of them. The combination of different classification methods in order to improve accuracy is also known as stacking, Frank, I. and Witten, E. (1999). Only MLRMs with classification precision lower than the threshold are passed on to the next classification stages. A precision and a recall threshold of 90% have been considered in this work.

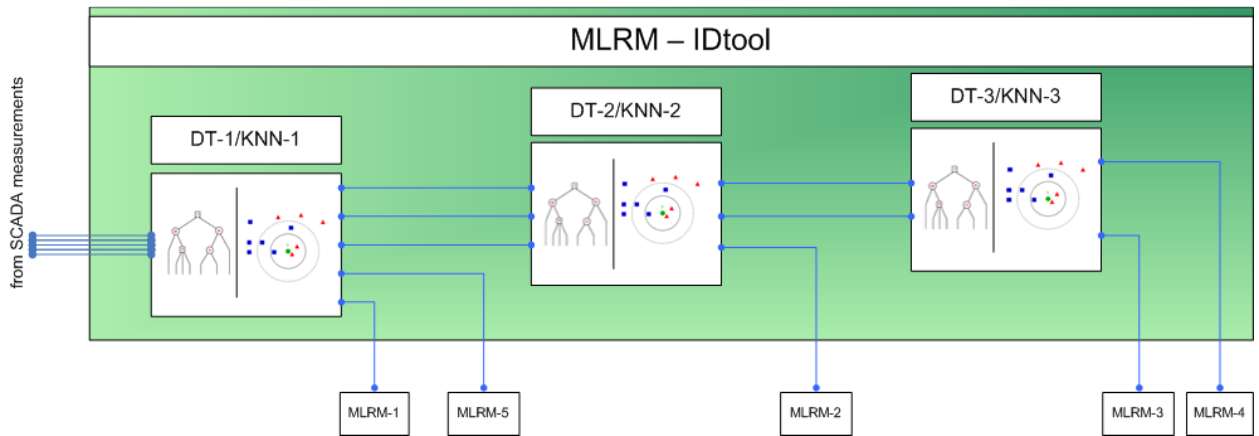


Figure 6.9. Hierarchical structure of the MLRM-IDtool

The variables passed to the classifiers represent several power system attributes (SCADA measurements shown in Table 4.1 along with class attributes (MLRMs). The MLRMs are then classified and the models with high classification precision (and low recall) are removed from the process, leaving only models with low classification precision to be classified in the next stages.

Two major aspects have been considered while selecting techniques to be used in the hierarchical classifier: complexity of the method and overall classification performance of the

tool. Ideally, the simplest machine learning method would be used and high classification performance would be achieved. However, that is not the case most of the time and complex classification schemes may be necessary.

We have observed that DTs, ANNs and KNNs are simpler for practical implementations and need less training time compared to SVMs. It has been observed that they are able to provide reasonable classification performance within a small amount of training time. Moreover, since they are able to perform multiclass classification, they contribute to a reduction in the number of classifiers needed inside the MLRM-IDtool. The SVM method cannot be considered under this topology as it cannot perform multiclass classification at each stage, thereby requiring a large number of classifiers and reducing classification performance.

After talking with utility partner, we have noticed a higher interest in the use of decision trees as the classifier. The reason is that it can provide easy to understand and meaningful visual information to system operators. Therefore, an attempt will be made in order to assemble DTs in several different classification schemes in order to improve classification performance.

A description of how the method can be applied to the MLRM identification problem is given next. Let us assume that current system operating condition is made available to the MLRM-IDtool. The information is passed onto the hierarchical classification system for MLRM identification and classification is done sequentially on each classifier. If one of the classes is properly classified with precision/recall higher than 90%, that class (MLRM) is then excluded from the next classification step and the process continues with the remaining imprecisely classified MLRMs.

However, it is important to remember that the order of classification in Figure 6.9 is not necessarily the same for each studied system and variations on the number of classification

stages may occur. The structure presented above is relative to the implementation of the methodology on the IEEE30 bus test system, Leonardi, B. and Ajjarapu, V. (2010).

Although the scheme presented above represents a potential architecture for the MLRM-IDtool, the final complexity of the tool will depend on the difficulty to classify the MLRMs. Therefore, practical implementation on large networks may have higher or lower complexity than the method presented in Figure 6.9.

All multiclass classification methods are then compared to a standard boosting algorithm called AdaBoost, Freund, Y. and Schapire, R. E (1997). Results will show that although all techniques work well at the problem, some of them may enhance classification precision with a smaller number of classifiers. Simulation results of the hierarchical method have been reported in the next chapter and in Leonardi, B and Ajjarapu, V. (2011).

CHAPTER 7. SIMULATION RESULTS FOR THE MLRM-IDTOOL

This chapter contains tests of the MLRM-IDtool on the IEEE30 bus test system and on the reduced case of the eastern interconnection of the United States. All four machine learning techniques described above are tested and investigated. A detailed description of the results is given next.

7.1 Results on the IEEE-30 bus test system

This test system is formed by considering a total of 50 contingencies including NERC category B, C and D along with 15 LIDs for training and 5 LIDs for testing, totaling 750 training scenarios and 250 testing scenarios, respectively. Different multiclass classification processes are investigated as proposed. The results and performance of each machine learning techniques is presented next.

7.1.1 Decision tree based classifiers

7.1.1.1 *Single multiclass DT*

As previously mentioned, every machine learning technique will be investigated under different topologies for multiclass classification. The simplest multiclass classification method considering a DT is by extending the algorithm to handle multiple classes. Table 7.1 shows the classification accuracy of a single DT.

It can be observed that MLRM-1, MLRM-4 and MLRM-5 have good classification precision, although MLRM-4 has a high recall rate. MLRM-2 and MLRM-3 have low classification precision and indicated that a more complex classification system might be necessary.

Table 7.1. Confusion matrix of DT

		Outputs					Individual precision (in %)	Overall precision (in %)
		MLRM-1	MLRM-2	MLRM-3	MLRM-4	MLRM-5		
Target	MLRM-1	1363	32	2	0	0	94.6	84.2
	MLRM-2	78	1456	24	0	0	71.9	
	MLRM-3	0	536	895	73	0	65.4	
	MLRM-4	0	0	448	1108	0	93.8	
	MLRM-5	0	0	0	0	1543	100.0	

7.1.1.2 DT-based hierarchical classifier

In order to improve classification precision of poorly classified models, a DT-based hierarchical classifier is proposed. The threshold considered to remove a MLRM from the classification process is a precision and recall above 90%. According to this criterion, three DTs are found to be necessary in order to achieve the desired classification accuracy. DT-1 is used in the first stage and its confusion matrix is shown in Table 7.2 below.

Table 7.2. Confusion matrix of DT 1

		Outputs					Individual precision (in %)	Overall precision (in %)
		MLRM-1	MLRM-2	MLRM-3	MLRM-4	MLRM-5		
Target	MLRM-1	1363	32	2	0	0	94.6	84.2
	MLRM-2	78	1456	24	0	0	71.9	
	MLRM-3	0	536	895	73	0	65.4	
	MLRM-4	0	0	448	1108	0	93.8	
	MLRM-5	0	0	0	0	1543	100.0	

MLRM-1 and MLRM-5 have precision and recall higher than 90% and are thus removed from the classification process. In the second stage, DT-2 removes MLRM-2 as it meets the requirements and DT-3 finally distinguishes between MLRM-3 and MLRM-4 as shown in Table 7.3 and Table 7.4, respectively.

The final classification precision of each MLRM will depend not only of their classification precision at the stage where they are classified, but also their rejection precision in the previous stages.

Table 7.3. Confusion matrix of DT-2

		Outputs			Individual precision (in %)	Overall precision (in %)
		MLRM-2	MLRM-3	MLRM-4		
Target	MLRM-2	473	21	0	93.3	84.4
	MLRM-3	34	503	36	74.1	
	MLRM-4	0	155	358	90.9	

Table 7.4. Confusion matrix of DT-3

		Output		Individual precision (in %)	Overall precision (in %)
		MLRM-3	MLRM-4		
Target	MLRM-3	308	14	82.8	85.9
	MLRM-4	64	169	92.3	

The final classification accuracy for the hierarchical DT-based classifier is given in

Table 7.5 below. It can be seen that the classification precision of some MLRMs have increased significantly, whereas others have either maintained the same classification accuracy or slightly reduced.

Table 7.5. Final classification precision of the DT based hierarchical classifier

Final classification precision (in %)	
MLRM-1	94.6
MLRM-2	92.6
MLRM-3	80.5
MLRM-4	89.8
MLRM-5	100.0

Overall, the gains in classification precision are meaningful and the hierarchical method is successful in increasing classification precision.

7.1.1.3 OVA DT based classifier

In order to assess other classification methods that could provide enhanced classification precision, a one versus all DT based classifier is investigated. A total of five DTs are used in order to train each class against all the other classes. The classification precision of each one of the MLRMs, as well as the overall classification precision, is given in Table 7.6.

Table 7.6. Confusion matrix of OVA DT-based classifier

		Outputs					Individual Precision (in %)	Overall precision (in %)
		MLRM-1	MLRM-2	MLRM-3	MLRM-4	MLRM-5		
Target	MLRM-1	1270	60	67	0	0	92.4	77.6
	MLRM-2	105	1281	172	0	0	66.8	
	MLRM-3	0	578	790	136	0	50.1	
	MLRM-4	0	0	548	985	23	87.9	
	MLRM-5	0	0	0	0	1543	98.5	

As can be noticed from the table above, the classification accuracy of individual MLRMs has not improved with the increased number of classifiers as compared with the single and

hierarchical DT classifiers. Therefore, we can conclude that the OVA approach is not useful to enhance classification precision when DTs are used as the core classification techniques.

7.1.1.4 OVO DT based classifier

Another possible approach is to implement the OVO classification approach using DTs as classifiers. The approach requires a total of ten DTs since five MLRMs are to be classified. The results of the OVO DT based classifier are shown in Table 7.7.

Table 7.7. Confusion matrix of OVO DT-based classifier

		Outputs					Individual precision (in %)	Overall precision (in %)
		MLRM-1	MLRM-2	MLRM-3	MLRM-4	MLRM-5		
Target	MLRM-1	1306	91	0	0	0	75.7	79.3
	MLRM-2	216	1334	8	0	0	66.6	
	MLRM-3	59	577	802	66	0	66.6	
	MLRM-4	144	0	405	1007	0	93.8	
	MLRM-5	0	0	0	0	1543	100.0	

The classification accuracy of some MLRMs has improved from the case whereas others have reduced. Overall, the classification precision has been reduced from the case where only a single DT is used for classification. Therefore, it can be concluded that neither the OVO nor the OVA DT based classifiers have shown improvements in classification precision.

7.1.1.5 AdaBoost using DT as classifier

In order to boost performance on poorly classified instances, the AdaBoost algorithm is investigate in this study. The method was proposed by Freund, Y. and Schapire, R. E (1997) and has the objective of improving classification in cases where the classifier shows poor classification precision. The results for the AdaBoost classifier are presented in Table 7.8.

Table 7.8. Confusion matrix of DT-based AdaBoost classifier

		Output					Individual precision (in %)	Overall precision (in %)
		MLRM-1	MLRM-2	MLRM-3	MLRM-4	MLRM-5		
Target	MLRM-1	1363	32	2	0	0	94.6	84.2
	MLRM-2	78	1456	24	0	0	71.9	
	MLRM-3	0	536	895	73	0	65.4	
	MLRM-4	0	0	448	1108	0	93.8	
	MLRM-5	0	0	0	0	1543	100.0	

The classifier does not show a very significant improvement classification precision when compared to the single DT-classifier. Under these circumstances, we have concluded that the hierarchical classifier has shown the best classification precision among all the investigated techniques and hence should be used for practical applications.

7.1.2 Artificial Neural Network based classifiers

7.1.2.1 Single multiclass ANN

The simplest way to apply ANNs to the problem at hand is by generalizing the binary classification algorithm to the multiclass case. In this study, the considered ANNs structure considered in this work has 5 output layers, 30 hidden layers and 230 attributes in the input layer. In order to perform online MLRM identification, all SCADA variables of interest are presented to the inputs of the ANN and the selected MLRM is produced in the output.

Since the MLRMs represent nominal classes, they must be represented numerically so that the ANN can process them. Therefore, the output labels need to be numerically coded so that the ANN can distinguish among them during the training stage. Table 7.9 shows how each MLRM is represented in the output of the ANN.

Table 7.9. Code words used to transform nominal class attributes into numeric

Model	Code word
MLRM-1	1 0 0 0
MLRM-2	0 1 0 0
MLRM-3	0 0 1 0
MLRM-4	0 0 0 1
MLRM-5	0 0 0 0

Basically, the presence of a “1” in one of the output neurons and “0” in all the other outputs will indicate which MLRM is identified. For instance, MLRM-1 is identified when the ANN produces a “1” on its first output and “0” in all other outputs. After modeling the MLRM accordingly, the ANN training process considering a simple multiclass ANN can begin.

Table 7.10 shows the classification performance when a single multiclass ANN is used to differentiate among the five MLRMs present in this case.

Table 7.10. Confusion matrix of single ANN classifier

		Outputs					Individual precision (in %)	Overall precision (in %)
		MLRM-1	MLRM-2	MLRM-3	MLRM-4	MLRM-5		
Target	MLRM-1	1325	72	0	0	0	94.4	81.6
	MLRM-2	78	1313	167	0	0	78.4	
	MLRM-3	0	289	975	236	4	63.5	
	MLRM-4	0	0	394	1013	149	81.1	
	MLRM-5	0	0	0	0	1543	91.0	

Although the overall performance of the network is high (81.6%), some MLRMs (specifically MLRM-3) have low classification precision and therefore would be poorly identified most of the time.

7.1.2.2 ANN-based hierarchical classifier

In order to enhance the classification accuracy for some of the MLRMs, more powerful classification schemes have been used as mentioned in the previous chapter. The hierarchical classification method used here considered a sequence of ANNs in order to improve classification accuracy.

The thresholds considered to remove a MLRM from the classification process are precision and recall higher than 90%. Therefore, if a MLRM has a precision and recall higher than 90% it is assumed to be classified with high accuracy and is removed from the process. Table 7.11 shows that MLRM-1 and MLRM-5 have high classification precision and are removed from the next step.

Table 7.11. Confusion matrix of ANN-1

		Outputs					Individual precision (in %)	Overall precision (in %)
		MLRM-1	MLRM-2	MLRM-3	MLRM-4	MLRM-5		
Target	MLRM-1	1325	72	0	0	0	94.4	81.6
	MLRM-2	78	1313	167	0	0	78.4	
	MLRM-3	0	289	975	236	4	63.5	
	MLRM-4	0	0	394	1013	149	81.1	
	MLRM-5	0	0	0	0	1543	91.0	

It is important to notice that despite MLRM-4 has a good classification precision, it still has a high recall rate (around 61.9%). Therefore, it is left in the process to be classified in the next stage.

Table 7.12 shows that MLRM-2 has good classification precision and is thus removed from the process by second classifier.

Table 7.12. Confusion matrix of ANN-2

		Outputs			Individual precision (in %)	Overall precision (in %)
		MLRM-2	MLRM-3	MLRM-4		
Target	MLRM-2	445	49	0	95.7	83.3
	MLRM-3	20	553	0	69.4	
	MLRM-4	0	195	318	100	

Table 7.13 shows the precision of the ANN-3 when classifying MLRM-4 and MLRM-5. It can be observed that the classification precision and recall rates increased significantly from the previous classification step.

Table 7.13. Confusion matrix of ANN-3

		Output		Individual precision (in %)	Overall precision (in %)
		MLRM-3	MLRM-4		
Target	MLRM-3	305	17	97.1	95.3
	MLRM-4	9	224	92.9	

The final classification precision of the method takes into account all steps involved in the methodology and is shown in Table 7.14.

Table 7.14. Final classification precision of the ANN based hierarchical classifier

Final classification precision (in %)	
MLRM-1	94.4
MLRM-2	94.2
MLRM-3	91.3
MLRM-4	87.4
MLRM-5	91.0

By comparing the results from Table 7.10 and Table 7.14, it can be noticed that there is a significant enhancement in classification precision in case the hierarchical approach is employed.

However, it is important to remember that three ANNs are used in the later approach in comparison to the former case, thus increasing the overall complexity and computational effort.

7.1.2.3 OVA ANN based classifier

In order to investigate how a more complex classifier would perform in this test case, an OVA ANN-based classifier is designed and investigated. The method makes use of five ANNs which are trained to differentiate one class from all the others at the time. The results obtained by the OVA ANN-based classifier are shown in Table 7.15.

Table 7.15. Confusion matrix of OVA ANN-based classifier

		Output					Individual precision(in %)	Overall Precisionn (in %)
		MLRM-1	MLRM-2	MLRM-3	MLRM-4	MLRM-5		
Target	MLRM-1	1303	86	8	0	0	97.2	82.3
	MLRM-2	37	1311	210	0.00	0	74.2	
	MLRM-3	0	357	1117	30	0	64.2	
	MLRM-4	0	13	405	949	189	96.9	
	MLRM-5	0	0	0	0	1543	89.1	

It can be noticed that even though the performance of the classifier is good, it cannot overcome the hierarchical ANN classifier. Moreover, this approach uses five ANNs whereas the hierarchical method uses only three to achieve higher performance.

Another down side of this method is that each one of the five classifier are trained using an unbalanced training set. This happens because even if all five classes are balanced (thus each representing roughly 20% of the data), the classifier is trained using one class versus the other four, thereby changing the training ratio to 20/80. This may bias each individual classifier and reduce its generalization ability.

7.1.2.4 OVO ANN based classifier

Another possible use of ANN classifiers is the one versus one approach. In this method, one classifier is trained to distinguish between every two classes. Since there are five classes, a total of ten classifiers will be needed. Simulation results of the method are presented in Table 7.16.

Table 7.16. Confusion matrix of OVO ANN-based classifier

		Output					Individual precision (in %)	Overall precision (in %)
		MLRM-1	MLRM-2	MLRM-3	MLRM-4	MLRM-5		
Target	MLRM-1	1355	42	0	0	0	91.9	82.1
	MLRM-2	120	1295	143	0	0	81.0	
	MLRM-3	0	261	1224	19	0	63.1	
	MLRM-4	0	0	574	793	189	97.7	
	MLRM-5	0	0	0	0	1543	89.1	

Comparing the results obtained by the OVO and the OVA approaches, it can be noticed that there is not much gain in performance despite the increase in the number of ANNs used by the later approach.

It is also important to mention that in case several classes (MLRMs) are considered, the OVO method may require a larger number of classifiers to be used. Practical implementation aspects must then be taken into account while deciding which method to use. Preference will be given to the methods that achieve highest performance at lowest complexity levels.

7.1.2.5 Adaboost using ANN as classifier

In order to compare the hierarchical classifier with existing standard boosting techniques, the AdaBoost algorithm has been tested using ANNs as classifiers. It is expected that by using a meta-learner technique, the classification precision of MLRM will increase. The algorithm uses several classifiers which are adjusted in favor of those instances misclassified by previous classifiers. The results of using ANN – based AdaBoost method is shown in Table 7.17.

As observed in the table above, the overall classification precision and the individual classification precision are similar to OVA and OVO methods previously mentioned. However, the classification precision of the AdaBoost algorithm cannot surpass the precision of the hierarchical ANN classifier.

Table 7.17. Confusion matrix of ANN-based AdaBoost classifier

		Outputs					Individual Precision (in %)	Overall precision (in %)
		MLRM-1	MLRM-2	MLRM-3	MLRM-4	MLRM-5		
Target	MLRM-1	1376	21	0	0	0	87.7	80.0
	MLRM-2	193	1142	223	0	0	67.2	
	MLRM-3	0	536	859	109	0	57.8	
	MLRM-4	0	0	405	1125	26	91.2	
	MLRM-5	0	0	0	0	1543	98.3	

7.1.3 KNN based classifiers

7.1.3.1 Single multiclass KNN classifier

Similarly to the ANN and DT methods proposed above, we begin our investigation by considering a multiclass KNN classifier that utilizes the three closest neighbors to classify all classes simultaneously. The confusion matrix containing the classification precision of the single KNN classifier is shown in Table 7.18.

Table 7.18. Confusion matrix of single KNN classifier

		Outputs					Individual precision (in %)	Overall precision (in %)
		MLRM-1	MLRM-2	MLRM-3	MLRM-4	MLRM-5		
Target	MLRM-1	1320	77	0	0	0	92.9	85.8
	MLRM-2	101	1325	132	0	0	76.4	
	MLRM-3	0	332	1147	25	0	74.6	
	MLRM-4	0	0	259	1155	142	97.9	
	MLRM-5	0	0	0	0	1543	91.6	

After comparing the results for the single KNN classifier with the single ANN and DT classifier, we notice that the KNN method outperforms both of its competitors in overall classification precision. Interestingly, this is one of the simplest methods and yet most powerful methods for multiclass classification. However, it is important to remember that since there is no model derived from the database, all samples must be store and classification occurs only when a new sample is submitted for classification. Due to this characteristic, the KNN classifier is also known as *lazy method*.

Overall, all the individual classification precisions are around 75% and above, thereby indicating that the KNN method can be successfully employed in MLRM identification.

7.1.3.2 KNN based hierarchical classifier

In order to improve classification precision, a KNN based hierarchical is proposed in the sequel. The classifier has a similar structure as the one presented before for the ANN and DT based ones.

The confusion matrix for the KNN-1 is given in Table 7.19. Although MLRM-3 has a high classification precision, its recall rate surpasses the 90% value specified in this project and therefore it is retained for a later classification stage.

Table 7.19. Confusion matrix of KNN-1 classifier

		Outputs					Individual precision (in %)	Overall precision (in %)
		MLRM-1	MLRM-2	MLRM-3	MLRM-4	MLRM-5		
Target	MLRM-1	1320	77	0	0	0	92.9	85.8
	MLRM-2	101	1325	132	0	0	76.4	
	MLRM-3	0	332	1147	25	0	74.6	
	MLRM-4	0	0	259	1155	142	97.9	
	MLRM-5	0	0	0	0	1543	91.6	

However, the precision of MLRM-2 in the second classifier cannot reach the minimum precision and recall of 90% and the approach should only have two classifiers. Since the MLRM-3 and MLRM-4 are not well distinguished (look at the high recall rate in KNN-2 classifier), their classification precision can be enhanced if a third classifier is added. Therefore, KNN-3 is created for it improves the classification performance of MLRM-3 and MLRM-4, although the design requirements would not suggest its development. The confusion matrix of KNN-2 is shown in Table 7.20.

Table 7.20. Confusion matrix of KNN-2 classifier

		Outputs			Individual precision (in %)	Overall precision (in %)
		MLRM-2	MLRM-3	MLRM-4		
Target	MLRM-2	455	39	0	88.0	75.1
	MLRM-3	62	388	123	65.0	
	MLRM-4	0	170	343	73.6	

Since the recall rate of MLRM-3 and MLRM-4 is high, the inclusion of a third classification stage may enhance individual classification precision of those models. Therefore, a third stage is developed and its confusion matrix is show in Table 7.21.

Table 7.21. Confusion matrix for the KNN-3 classifier

		Output		Individual precision (in %)	Overall precision (in %)
		MLRM-3	MLRM-4		
Target	MLRM-3	318	4	83.5	87.9
	MLRM-4	63	170	97.7	

Once all classification stages have been identified, the final classification precision for each MLRM can be calculated. The final classification precision of the hierarchical KNN based classifier is shown in Table 7.22.

Table 7.22. Final classification precision of the KNN based hierarchical classifier

Final classification precision (in %)	
MLRM-1	92.9
MLRM-2	86.5
MLRM-3	79.0
MLRM-4	92.5
MLRM-5	91.6

By comparing the classification precision of the hierarchical KNN classifier with the single KNN classifier, it can be noticed that the classification precision of most MLRMs went up, therefore indicating a better classification performance.

7.1.3.3 OVA KNN based classifier

In order to investigate wide range potential classifiers, an OVA KNN based multiclass classifier is developed as well. The confusion matrix of the OVA KNN is shown in Table 7.23.

Table 7.23. Confusion matrix of OVA KNN-based classifier

		Outputs					Individual Precision (in %)	Overall precision (in %)
		MLRM-1	MLRM-2	MLRM-3	MLRM-4	MLRM-5		
Target	MLRM-1	1320	77	0	0	0	92.8	85.8
	MLRM-2	101	1325	132	0	0	76.4	
	MLRM-3	0	332	1147	25	0	74.5	
	MLRM-4	0	0	259	1155	142	97.8	
	MLRM-5	0	0	0	0	1543	91.5	

Despite the fact that five classifiers are used in this case, the performance of the OVA KNN based classifier is very close to the case where a single KNN classifier is used. Therefore, there is practically no gain in precision by using a larger and more complex classification scheme in this case.

7.1.3.4 OVO KNN based classifier

Another potential topology to be tested is the OVO KNN based classifier. The confusion matrix for this classifier is shown in Table 7.24.

Table 7.24. Confusion matrix of OVO KNN-based classifier

		Outputs					Individual precision (in %)	Overall precision (in %)
		MLRM-1	MLRM-2	MLRM-3	MLRM-4	MLRM-5		
Target	MLRM-1	1313	84	0	0	0	93.0	85.4
	MLRM-2	98	1317	143	0	0	76.1	
	MLRM-3	0	328	1139	37	0	72.7	
	MLRM-4	0	0	283	1145	128	96.8	
	MLRM-5	0	0	0	0	1543	92.3	

It can be noticed that the OVO KNN based classifier does provide a slight improvement in performance compared to the single and OVA classifiers previously shown. However, a total of ten KNN classifiers are used in this approach, thereby increasing the complexity for practical applications.

7.1.3.5 AdaBoost KNN based classifier

The last and final methodology tested using the KNN approach is the AdaBoost algorithm with KNN as its main classification technique. The confusion matrix for the AdaBoost KNN based classifier is shown in Table 7.25.

Table 7.25. Confusion matrix of KNN-based AdaBoost classifier

		Outputs					Individual precision (in %)	Overall precision (in %)
		MLRM-1	MLRM-2	MLRM-3	MLRM-4	MLRM-5		
Target	MLRM-1	1376	21	0	0	0	87.7	80.0
	MLRM-2	193	1142	223	0	0	67.2	
	MLRM-3	0	536	859	109	0	57.8	
	MLRM-4	0	0	405	1125	26	91.2	
	MLRM-5	0	0	0	0	1543	98.3	

By comparing the confusion matrix of the KNN based AdaBoost classifier with the performance of a single KNN, it can be noticed that it does not achieve superior classification performance. In fact, the classification precision is reduced for most of the MLRMs. Therefore, this is a strong indication that this type of classifier may not be suited for classification in this case.

7.1.4 SVM based classifier

In order to investigate the performance of SVM, a binary classifier is used for multiclass classification in this study. The reason why a binary SVM is used is due to the fact that multiclass SVM is still an open subject and an efficient method is yet to be developed, Bishop, C. (2006). Therefore, the three possible methods that can be employed in this case are the OVA SVM based, OVO SVM based and the AdaBoost SVM based classifiers. A detailed analysis of these three methods is given next.

7.1.4.1 OVA SVM based classifier

The first implemented approach using SVM is the OVA method. Similarly to the DT, ANN and KNN based methods, this technique employs five SVM which will try to distinguish each MLRM from all the rest. The confusion matrix of the OVA SVM based classifier is shown in Table 7.26.

As noticed below, the classification precision of the OVA SVM based method is not significantly superior to the other techniques investigated before. This is an indication that this type of classifier is not able to distinguish among the classes very well, hence not being efficient for practical applications.

Table 7.26. Confusion matrix of OVA SVM-based classifier

		Outputs					Individual precision (in %)	Overall precision (in %)
		MLRM-1	MLRM-2	MLRM-3	MLRM-4	MLRM-5		
Target	MLRM-1	1354	43	0	0	0	62.2	77.4
	MLRM-2	150	1284	124	0	0	85.5	
	MLRM-3	616	175	713	0	0	60.2	
	MLRM-4	57	0	348	962	189	61.8	
	MLRM-5	0	0	0	0	1543	89.1	

Moreover, the time involved in the design of this classifier has been significant superior than the other classifiers, which may represent another difficulty for application to real-sized/large networks.

7.1.4.2 OVO SVM based classifier

The next methodology tested is the OVO SVM based classifier. A total of ten SVM are used in this approach and the confusion matrix is shown in Table 7.27. Once again, the classification precision of the method is not significantly superior compared to any of the previous methods.

Table 7.27. Confusion matrix of OVO SVM-based classifier

		Outputs					Individual precision (in %)	Overall precision (in %)
		MLRM-1	MLRM-2	MLRM-3	MLRM-4	MLRM-5		
Target	MLRM-1	1358	39	0	0	0	52.4	72.0
	MLRM-2	300	1192	66	0	0	84.8	
	MLRM-3	755	175	394	170	10	57.5	
	MLRM-4	180	0	225	962	189	85.0	
	MLRM-5	0	0	0	0	1543	88.6	

However, the amount of time involved with training this classifier has been significantly higher than any of the previous methods, which may represent a drawback for practical applications.

7.1.4.3 AdaBoost SVM based classifier

Differently than the previously investigated methods, the AdaBoost classifier has shown a very good classification performance while trying to classify which MLRM to use. The confusion matrix of this classifier is shown in Table 7.28.

Table 7.28. Confusion matrix of SVM-based AdaBoost classifier

		Outputs					Individual precision (in %)	Overall precision (in %)
		MLRM-1	MLRM-2	MLRM-3	MLRM-4	MLRM-5		
Target	MLRM-1	1369	28	0	0	0	94.4	87.9
	MLRM-2	81	1405	72	0	0	86.2	
	MLRM-3	0	197	1303	4	0	75.9	
	MLRM-4	0	0	340	1027	189	99.6	
	MLRM-5	0	0	0	0	1543	89.0	

The classification precision of all models is significantly high using the AdaBoost approach (above 75%), although other classification techniques have shown better results. Overall, the use of SVM for classification has not shown drastically enhanced performance compared to other methods. On the other hand, the time involved in training SVM methods is significantly higher than the other three techniques.

Therefore, applications of the technique to large power system networks may be difficult due to the large amount of data available.

7.1.5 Stacked classifier

Another powerful way of enhancing the classification precision of the hierarchical classification method is by combining different classifiers into the process. For instance, for the three classifiers used in the hierarchical process in the previous study, the first classifier could be a DT, with the second one being a KNN and the third one being an ANN.

In order to enhance performance, we have combined DTs and ANNs together as a stacked classifier to improve classification precision. The classifiers used at each stage are described in Table 7.29, Table 7.30 and Table 7.31, respectively.

Table 7.29. Confusion matrix of DT-1

		Outputs					Individual precision (in %)	Overall precision (in %)
		MLRM-1	MLRM-2	MLRM-3	MLRM-4	MLRM-5		
Target	MLRM-1	1363	32	2	0	0	94.6	84.2
	MLRM-2	78	1456	24	0	0	71.9	
	MLRM-3	0	536	895	73	0	65.4	
	MLRM-4	0	0	448	1108	0	93.8	
	MLRM-5	0	0	0	0	1543	100.0	

Table 7.30. Confusion matrix of DT-2

		Outputs			Individual precision (in %)	Overall precision (in %)
		MLRM-2	MLRM-3	MLRM-4		
Target	MLRM-2	473	21	0	93.3	84.4
	MLRM-3	34	503	36	74.1	
	MLRM-4	0	155	358	90.9	

Table 7.31. Confusion matrix of ANN-3

		Output		Individual precision (in %)	Overall precision (in %)
		MLRM-3	MLRM-4		
Target	MLRM-3	305	17	97.1	95.3
	MLRM-4	9	224	92.9	

The final classification accuracy of the hierarchical stacked classifier is further enhanced from previous schemes and the results are given in Table 7.32.

Table 7.32. Final classification precision of the hierarchical stacked classifier

Final classification precision (in %)	
MLRM-1	94.6
MLRM-2	92.6
MLRM-3	94.4
MLRM-4	90.4
MLRM-5	100.0

Therefore, it can be observed that by using stacked classifiers in the hierarchical classification method, the final classification precision of all MLRM is higher than 90%. This is a very good result since the designer can enhance classification performance at no extra cost, as all individual classifiers have already been designed.

7.2 Results on the 22k bus system

7.2.1 Development of the MLRM-IDtool

Although four techniques have been investigated previously, only two of them will be used in this practical system for their simplicity and possibility of practical implementation. These methods are easy to understand and provide good visual information, specially in the DT case. Therefore, the two methods investigated for this system are the DTs and KNNs due to their suitability to practical implementation and good classification performance,

Before the development of the MLRM-IDtool, the most descriptive attributes (variables) among the ones described in Table 4.1 need to be identified. The reduction in the number of variables is important in this case of thousands of attributes are available. If an excessive number of attributes are used during the design of the MLRM-IDtool, adverse effects such as curse of dimensionality and increased complexity may occur.

In order to handle the excessive number of variables, the performance of each attribute will be analyzed separately and only attributes that demonstrate superior classification performance will be selected.

Two different datasets have been considered to test the MLRM-IDtool. The first dataset contains the same contingencies used during the training phase and independent LIDs are used to generate unforeseen operating conditions. However, in order to investigate the performance of the tool for unforeseen contingencies and scenarios, a different contingency set is also used in this case. It will be shown later that the KNN based classifier has better performance in case unforeseen contingencies occur.

7.2.1.1 DT-based classifier

A DT-based MLRM-IDtool is developed using the algorithm C4.5 previously described. As mentioned before, two different scenarios are considered in this case. Initially, only unforeseen load increase directions are considered similarly to the cases used to test the MLRMs. Next, ten unforeseen N-k contingencies and five independent LIDs are used to generate fifty completely unforeseen scenarios.

Figure 7.1 shows a summary of the classification precision of each variable when used separately. The elements labeled from top to bottom in the legend are the same from right to left on the bar groups for each one of the variables.

The x – axis contains the variables used in the DT development. A DT is constructed considering each one of the variables types separately. There are a total of 603 variables of each type indicated as P_{flow} , P_{loss} , Q_{flow} , Q_{loss} and C_{mag} , which represent all 161kV and above transmission lines and transformers in the studied area. A total of 276 bus voltage magnitudes are represented by the V_{mag} variable type with rated voltages of 161kV and above.

The classification precision of the developed DTs can be seen in Figure 7.1. Although all variables have shown satisfactory performance, P_{flow} and V_{mag} have shown slightly superior results. When only P_{flow} variables are used, the classification precision has an overall test value of around 97%, and the MLRM with lowest classification precision is M1 with 95%. The DT created using V_{mag} as attributes has an overall classification precision of 96% and the MLRM with M1 having the lowest classification precision of 92.5%.

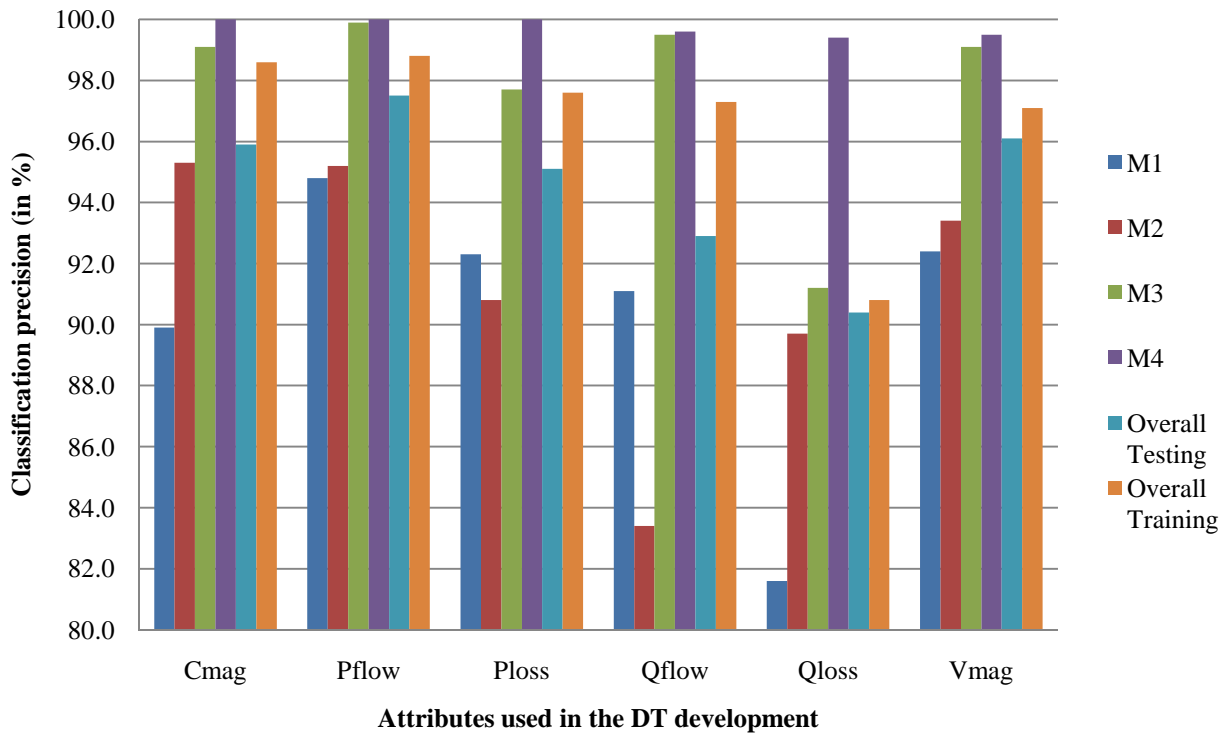


Figure 7.1. Attribute analysis under unforeseen LIDs for DTs

Another important point to be noticed is the fact that M3 and M4 have shown higher classification precision than M1 and M2. This is a very desirable characteristic since those models contain the most critical contingencies, i.e., the contingencies with lowest VSM.

Another advantage of using P_{flow} and V_{mag} is that they are easier to be obtained and monitored as they are readily available in SCADA/EMS. Other variables like P_{loss} , Q_{loss} and C_{mag} can only be obtained after processing SCADA data since they are not directly measured most of the time. Therefore, both P_{flow} and V_{mag} variables are used to build the final DT-based MLRM-IDtool.

In order to further reduce the amount of data used by the tree, only major 161-345kV buses and transmission lines are considered in the process. This simplification helped to reduce the total amount of data necessary during model development. An excessive amount of available attributes at the DT development stage can cause a phenomenon known as curse of dimensionality, where several good predictor attributes compete against each other. This attribute competition is likely to adversely affected the DT development and reduce its classification abilities. The total number of attributes (variables) available for training the classifiers is thus reduced from 3291 to 122, significantly facilitating the development of the tool.

Figure 7.2 shows the decision tree built using only P_{flow} and V_{mag} on a few selected transmission lines and buses rated 161-345kV.

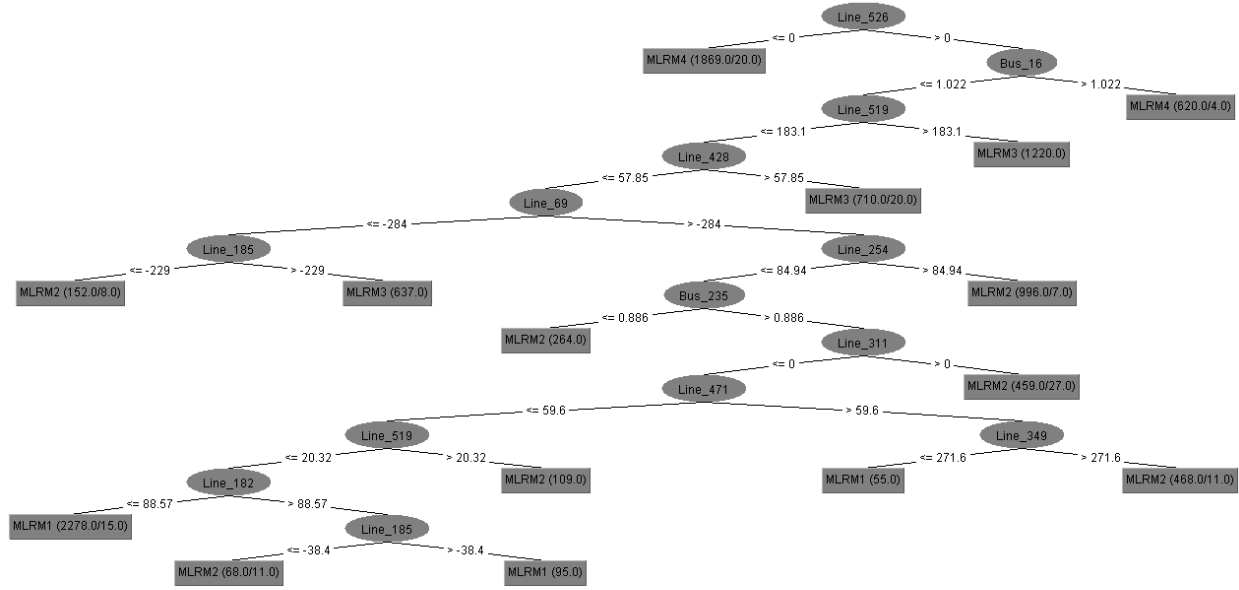


Figure 7.2. DT created using P_{flow} and V_{mag} as attributes

The confusion matrix for the tree including the classification precision for each MLRM is shown in Table 7.33. As can be observed from the table, individual classification precision for all models is above 90% as required by the design specifications.

Table 7.33. Classification precision of DT developed using P_{flow} and V_{mag}

		Outputs				Individual Precision (in %)	Overall Precision (in %)
		MLRM-1	MLRM-2	MLRM-3	MLRM-4		
Target	MLRM-1	2136	34	0	0	93.1	97.7
	MLRM-2	124	2007	2	0	98.3	
	MLRM-3	35	0	2048	0	99.9	
	MLRM-4	0	0	0	2188	100.0	

- Effect of unforeseen contingencies**

In order to evaluate the classification precision of the technique on unforeseen contingencies and LIDs, a set of 10 contingencies and 5 independent LIDs is created. A VSA is performed using the aforementioned contingencies and LIDs. Around 500 samples are taken from the 50

scenarios and used to generate the independent testing dataset. Similarly to what has been done previously, a DT is developed for each attribute type and the results are summarized in Figure 7.3.

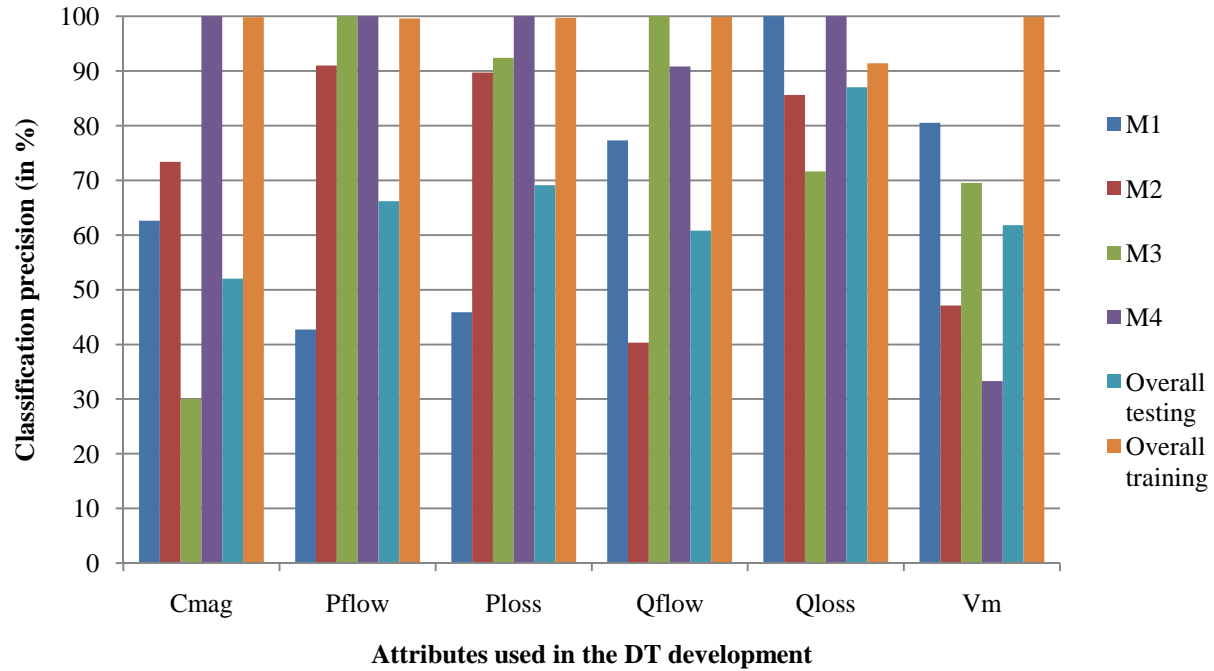


Figure 7.3. DT attribute analysis under unforeseen contingencies and LIDs

The results show that the classification precision of the DT can be significantly affected in case unforeseen contingencies are used during the DT testing stage. Overall classification precision has dropped to values ranging from 50-86%, indicating spurious classification performance compared to the case where only different LIDs are considered.

Attempting to solve this problem and improve classification precision for unforeseen contingencies, most techniques usually add the new unforeseen contingencies to the training set and retrain the DT on an extended set. This re-training on an extended dataset including data from unforeseen contingencies should be able to improve performance in those conditions.

Overall, the DT can be a powerful technique in the development of the MLRM-IDtool. Appropriate re-training can make the methodology flexible to handle unforeseen contingency and hence maintain high classification precision.

- **Effect of noise**

An undesired but rather common situation in SCADA/EMS control centers is the reception of data corrupted by noise. Various factors can cause noise to be added to the real variable measurements, with typical causes being defects or malfunctioning of field measurement devices weather, animals and vegetation.

Despite of the way how noise becomes present in the measurements, it is necessary to properly investigate and analyze the impact that it can cause on the performance of the MLRM-IDtool. Having that objective in mind, new training and testing data sets are developed to account for the influence of noise.

According to practical information, the amount of noise considered in SCADA/EMS measurements received in their control center is around 3%-5%. To account for that effect, the previously used training and testing data sets have been corrupted with white noise. Five different levels of noise added to the test set are considered in the study: 0%, 1%, 5%, 10% and 15%. These different amounts of noise represent a percentage of the actual variable value and are used to generate the Gaussian white noise.

Once the Gaussian curve representing noise is generated, a random value is sampled from it and added to the actual variable value. By doing so, system variables vary up and down around the actual measurement. A pictorial representation of the different amounts of noise as normal distributions is shown in Figure 7.4.

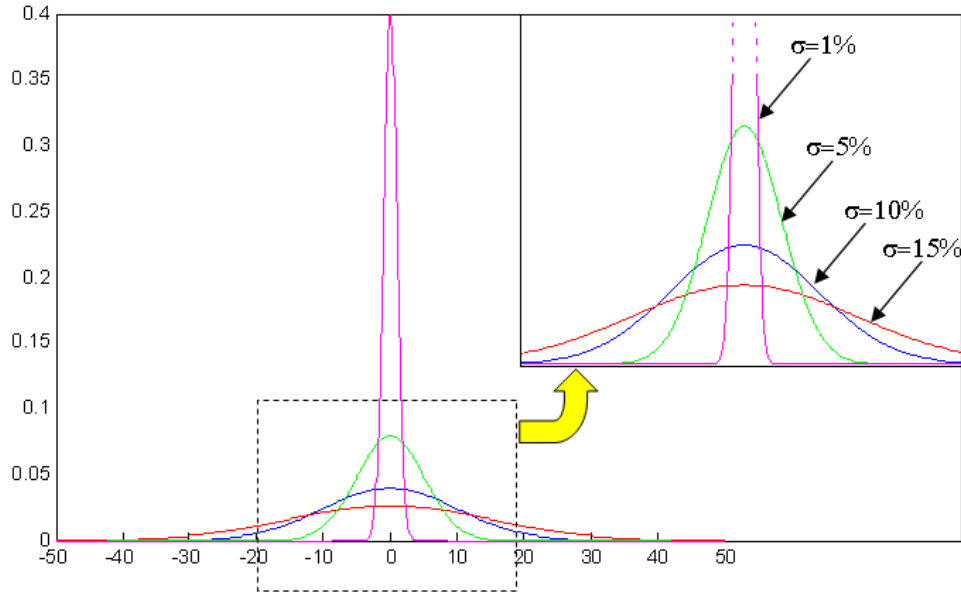


Figure 7.4. Representation of different amounts of noise added to the actual data

One of the most traditional approaches to make meta learners resilient to noise is to train them in the presence of noisy data, Bishop, C. M., (1995). In order to improve DT performance when noisy attributes are present, noise is added to the training set in a similar way it is added to the test set and results are shown in Figure 7.5. The legend shows different training conditions used to improve the performance of the DT; the x-axis shows the amount of noise included in the test set and the y axis shows the overall classification precision in the test set.

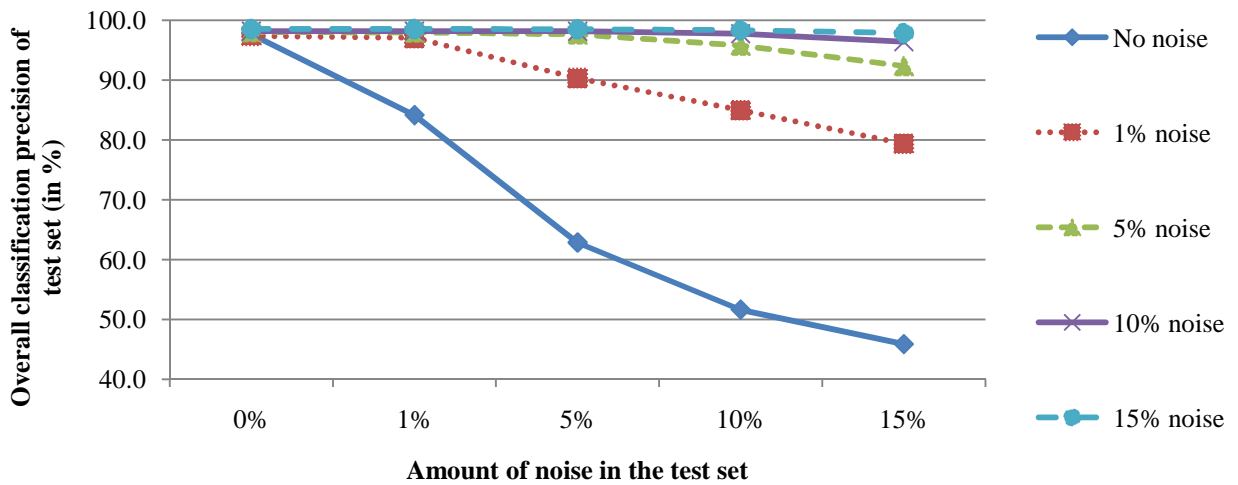


Figure 7.5. DT performance on noisy data for unforeseen LIDs

The first case represents the condition where the DT is trained without noise and tested on sets with different amount of noise. As observed in the picture, the DT performs well when there is no noise added to the test set and the precision decays continuously as more noise is added to it.

By training the DT with a training set considering 1% of noise, the precision is kept high for both when there is no noise and when there is 1% noise. If the test set contains more noise than what is used to train the DT, the performance still decays but at a slower rate compared to the case when no noise is added to the training set.

The results presented in Figure 7.5 indicate that the DT performs well when the amount of noise added to the measurements is smaller or equal to the amount of noise present measurements used in the training dataset. Another important characteristic is that even when the DT is trained with a significant amount of noise (15%), its performance does not decay for cases when no noise is present in the measurements. This is an important result because although the

SCADA/EMS can be corrupted with noise, there will be circumstances where the measurements are noise free or have a small amount of noise.

After incorporating 15% of noise in all measurements for P_{flow} and V_{mag} attributes, a new DT is developed and shown in Figure 7.6.

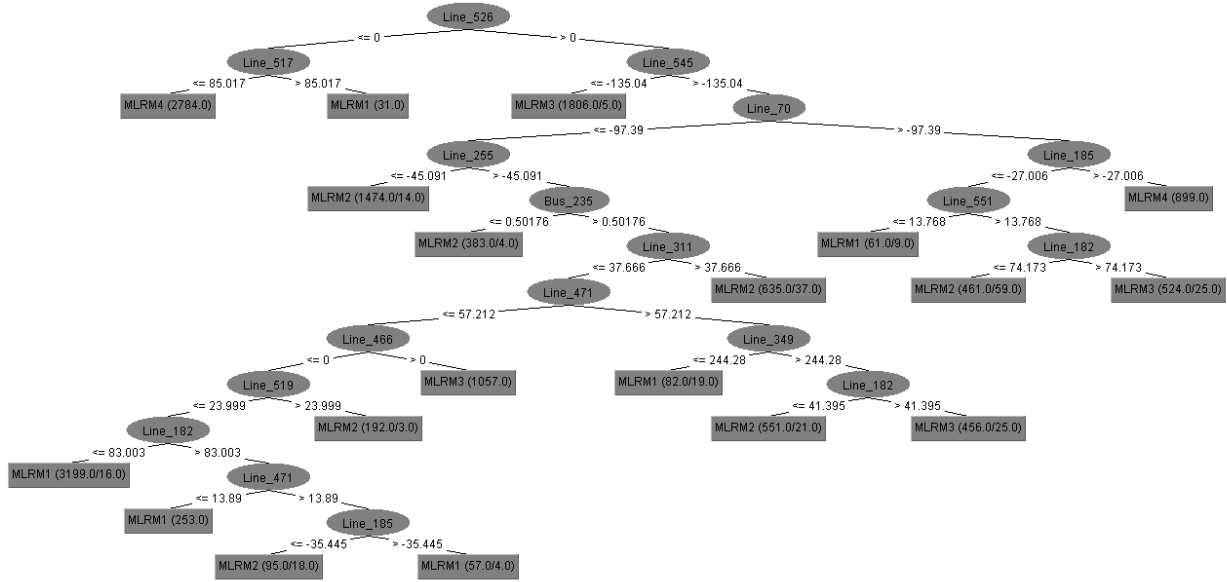


Figure 7.6. DT created using P_{flow} and V_{mag} as attributes and trained with 15% noise

In order to be able to maintain a good classification precision, the size of the tree had to be increased. However, the relative increment in size should not affect the implementation of the DT as it is still adequate for practical implementation. The confusion matrix for the final DT trained with noise is given on Table 7.34.

Table 7.34. Classification precision of DT trained using P_{flow} and V_{mag} and 15% noise

		Outputs				Individual Precision (in %)	Overall precision (in %)
		MLRM-1	MLRM-2	MLRM-3	MLRM-4		
Target	MLRM-1	2066	87	16	1	95.0	97.1
	MLRM-2	97	2016	20	0	95.5	
	MLRM-3	8	8	2067	0	98.1	
	MLRM-4	4	0	4	2180	100.0	

In order to evaluate the impact of noise in the presence of unforeseen contingencies and LIDs, a similar procedure is used to analyze the DTs. The results of the analysis are summarized in Figure 7.7. The figure below shows that the classification performance significantly decays for the case when unforeseen contingencies and LIDs are used. The only case when the overall precision surpassed 90% is when the DT is trained with 15% of noise.

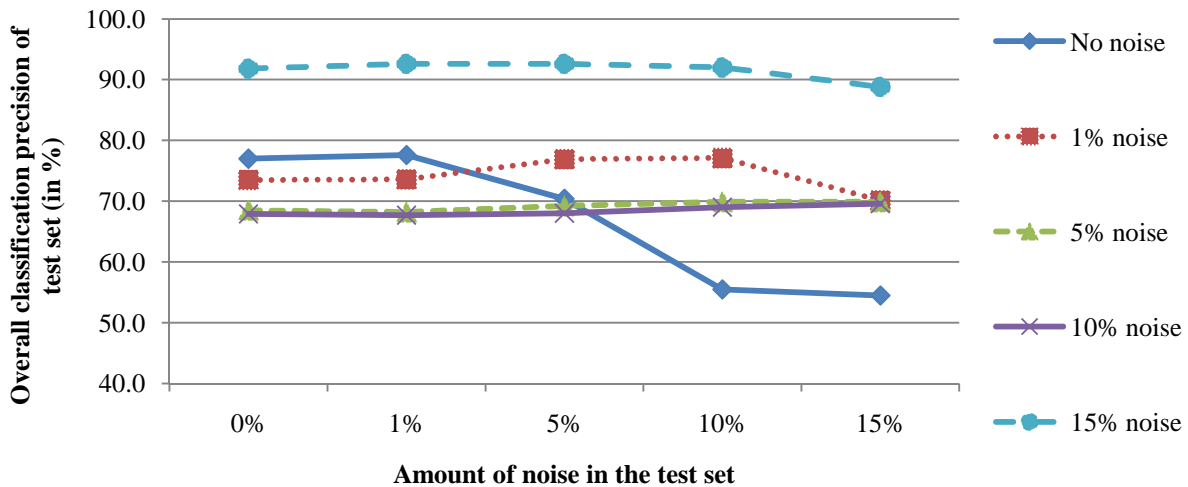


Figure 7.7. DT performance on noisy data for unforeseen contingencies and LIDs

These results were expected since the DT has shown a significant reduction in classification precision in case unforeseen contingencies are included in the test set as observed in Figure 7.3. Therefore, the DT is not able to generalize well in case unforeseen contingencies occur.

7.2.1.2 *KNN-based classifier*

After investigating the use of DTs as the core technique in the MLRM-IDtool, it has been noticed that the performance of the tool on unforeseen contingencies can be significantly affected. A significant improvement of the methodology can be achieved in case this limitation is overcome.

The need of frequent MLRM-IDtool re-training can be reduced in case the tool performs better on unforeseen contingencies. It will be shown later that by using a KNN classifier in the MLRM-IDtool, an improvement in classification precision can be obtained.

- **Tests on unforeseen LIDs**

Before we start testing the technique on unforeseen LIDs, an investigation of all attributes is made in order to identify the ones that give the best classification precision. Similarly to the DT case described in earlier sections, a KNN classifier is developed considering a single type of variable at a time.

The datasets used here are the same ones used in the DT development and testing. Therefore, there is a total of 603 attributes (variables) for the P_{flow} , P_{loss} , Q_{flow} , Q_{loss} and C_{mag} types, representing all 161kV and above transmission lines and transformers in the studied area. A total of 276 bus voltage magnitudes are represented in by V_{mag} attributes (variables) type with rated voltages of 161kV and above for the same MEC/ALTW area.

The tests of the KNN-based MLRM-IDtool are summarized in Figure 7.8. Overall, all the attributes have a very similar classification performance for both individual MLRMs and overall classification precision of the test set. Only the Q_{loss} attributes have shown inferior performance when compared to other attributes. In addition to that, the performance of KNN classifiers also slightly surpasses the ones using DT classifiers for the same testing set and shown in Figure 7.1.

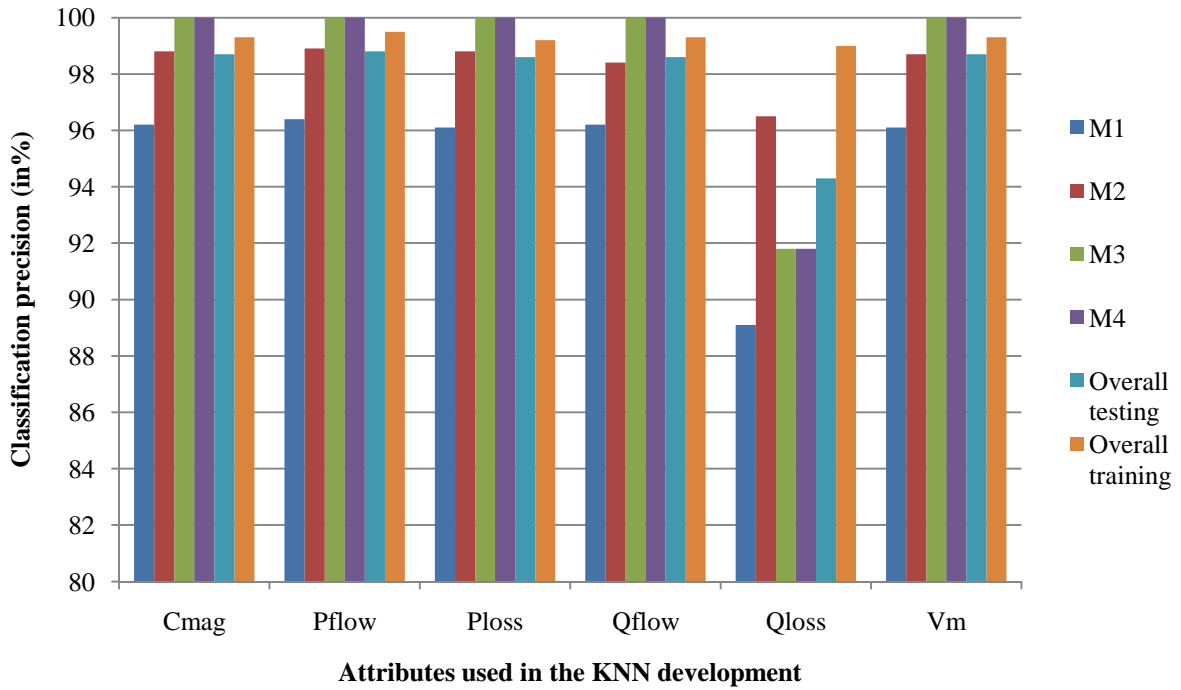


Figure 7.8. KNN attribute analysis under unforeseen LIDs

Once again, in order to further reduce the amount of data used by the KNN and consequently increase the classification speed, only major 345kV buses and transmission lines have been considered in the KNN development.

Similarly to the DT-based MLRM-IDtool, the attributes P_{flow} and V_{mag} are selected to reduce the number of attributes available during the KNN development and reduce the chances dimensionality issues. Since both attributes are directly available in SCADA/EMS, they can be directly used and thus eliminate the need of extra calculations. This simplification helps to reduce the total amount of attributes used in the KNN from 3291 to 122, hence reducing the computational time necessary to develop the KNN.

The confusion matrix for the KNN based classification tool using only P_{flow} and V_{mag} is shown in Table 7.35.

Table 7.35. Classification precision of DT developed using P_{flow} and V_{mag}

		Outputs				Individual	Overall
		MLRM-1	MLRM-2	MLRM-3	MLRM-4	Precision (in %)	Precision (in %)
Target	MLRM-1	2129	39	0	2	95.1	98.2
	MLRM-2	109	2024	0	0	98.1	
	MLRM-3	0	0	2083	0	100	
	MLRM-4	0	0	0	2188	99.9	

It can be noticed that the classification precision of each individual MLRM is above 95%, thus indicating the powerful capability of the KNN based MLRM-IDtool.

- **Effect of unforeseen contingencies**

In order to evaluate the classification precision of the KNN-based MLRM-IDtool on unforeseen contingencies and LIDs, the same contingency set used to evaluate the DT is used here. A KNN classifier is developed for each attribute type similarly to what has been done for the DT case and the results are summarized in Figure 7.9.

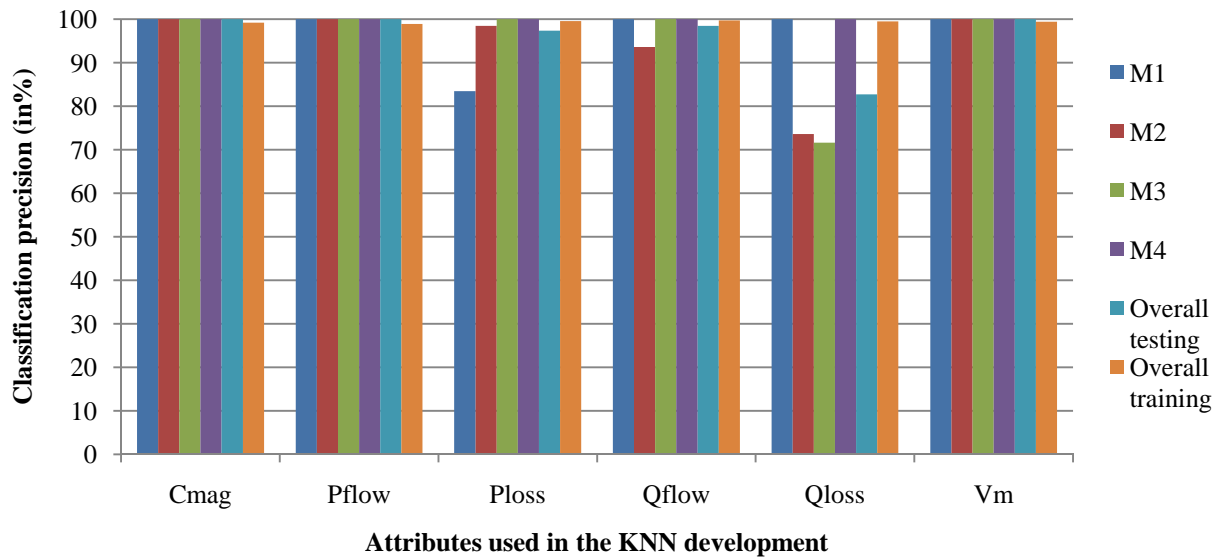


Figure 7.9. KNN attribute analysis under unforeseen contingencies and LIDs

Differently from what has been observed for the DTs, the classification precision of the KNN-based MLRM-IDtool is maintained high than the cases where only unforeseen LIDs are used in the test set. Such an outstanding performance represents an enormous improvement since it may significantly reduce the need for re-training the MLRM-IDtool.

By no means can we ensure that the classifier performs with high classification precision for every possible unforeseen network topology and that no retraining is required after the KNN-based MLRM-IDtool has been developed. However, we can say that if the MLRM-IDtool is developed with a KNN-based classifier rather than a DT-based one, the need for frequent re-training can be reduced.

In our opinion, this discrepancy in performance is caused by the inherent difference on how the algorithms work. The DT technique selects a reduced number of attributes of a large data set to construct the tree. As the tree is branched, the ramifications are constructed on a portion of the original database. Therefore, further expansions of the branches are done using local portions of the training dataset, instead of the complete dataset.

On the other hand, the KNN technique always uses all samples and attributes in the entire dataset. Every classified instance is based on the nearest neighbors of the entire data set, and not only on a partition of it. The downside of this would be time required to classify a new instance as all data is considered for classification. However, since only the current system operating condition has to be classified every few seconds (we just to identify the current system operating condition), the speed classification is on the order of tens of a second, fast enough for the time frame of interest.

By comparing the results obtained with the DT and the KNN techniques, it has been observed that the KNN-based MLRM-IDtool has shown superior performance than its DT-based

concurrent. Considering that no extra additional complexity is added, we see the KNN-based MLMR-IDtool as a good option to handle unforeseen contingencies with high classification precision.

- **Effect of noise**

Similarly to the investigation performed to find out the influence of noise in the DT-based MLRM-IDtool, tests have been performed with the KNN-based MLRM-IDtool. The objective is to evaluate its performance when there is noise present in the SCADA/EMS measurements.

The test data sets used here are the same ones used in to assess DT's performance and will not be described for the sake of simplicity. Noise has been added in the same manner and the reader should refer to the previous section for a more detailed explanation. The results obtained are presented in Figure 7.10 and Figure 7.11, respectively.

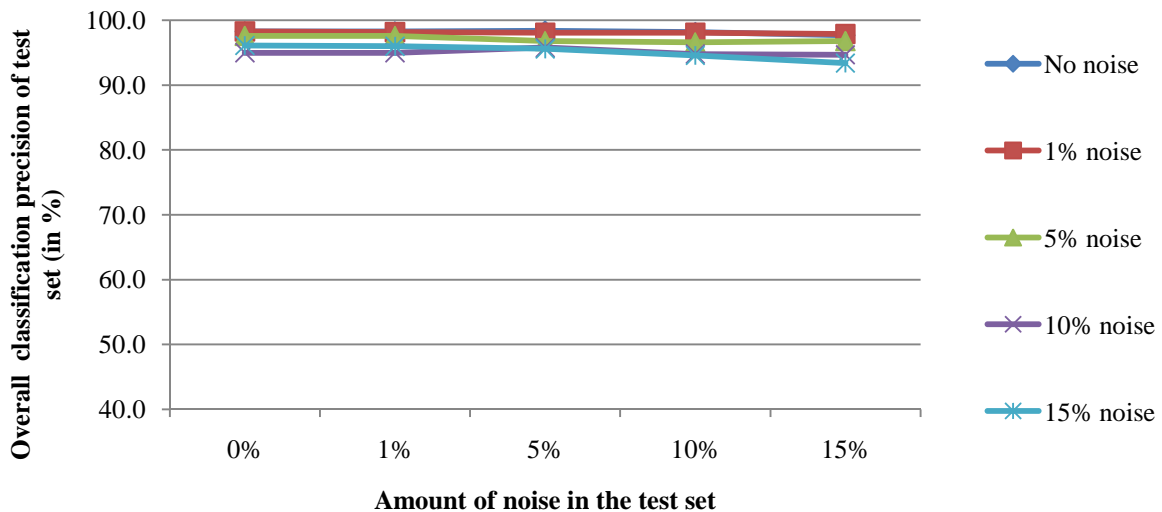


Figure 7.10. KNN performance on noisy data for unforeseen LIDs

Figure 7.10 shows the classification performance of the KNN-based MLRM-IDtool for different load increase directions. Each case shown in the legend represents a different amount of

noise added to the training set. It can be concluded from the graph that the technique has shown to be innately noise resilient. The addition of noise in the training sets neither improves nor decreases classification precision significantly, which is a completely different behavior compared to the DT-based classifier under the same circumstances.

The results shown in Figure 7.10 indicate a very good characteristic of the KNN-based MLRM-IDtool compared to the DT-based one. While DT-based classifier showed reduced classification precision when more noise is added to the measurements, the KNN-based classifier has shown to maintain classification precision unaffected throughout a wide range of noise.

Figure 7.11 shows the classification performance of the KNN-based MLRM-IDtool in the presence of noise and under different load increase directions and unforeseen contingencies.

Once again, the addition of noise in the training sets neither improves nor decreases classification precision significantly. It can be noticed that classification performance is maintained above 95% for all cases and does not significantly vary in the presence of noise.

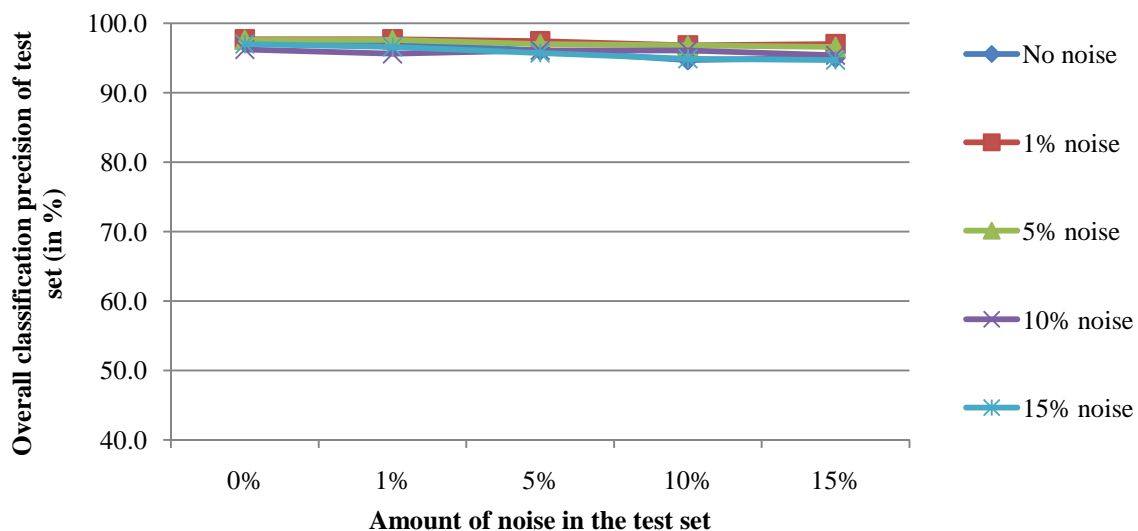


Figure 7.11. KNN performance on noisy data for unforeseen contingencies and LIDs

The classification performance has been maintained constant throughout a wide range of different noise values as shown in Figure 7.11. These results are significantly different than the ones observed for the DT-based classifier, where the classification precision significantly decays in the presence of noisy inputs and under unforeseen contingencies and LIDs as seen in Figure 7.7.

The confusion matrix for the KNN classifier trained with 15% of noise and tested on 15% of noise is shown Table 7.36. Similarly to the DT based classifier trained with noise, all MLRMs are classified with precision superior to 90%. Since this is the precision criteria considered in the design stage, the KNN classifier shown in the table is the one selected to be used in the MLRM-IDtool.

Table 7.36. Classification precision of KNN trained using P_{flow} and V_{mag} and 15% noise

		Outputs				Individual precision (in %)	Overall precision (in %)
		MLRM-1	MLRM-2	MLRM-3	MLRM-4		
Target	MLRM-1	1925	55	0	190	91.8	93.4
	MLRM-2	123	1885	99	26	96.1	
	MLRM-3	12	22	2045	4	95.4	
	MLRM-4	38	0	0	2150	90.7	

After analyzing the results obtained in this section, we believe that if a KNN-based classifier is used in the development of the MLRM-IDtool, a more robust classification tool can be obtained.

It has been noticed that the KNN based MLRM-IDtool is naturally noise resilient and is also able to identify the appropriate MLRM even when unforeseen network topologies and load increase directions occur.

Further research could focus on identifying the reasons why the KNN performs so well even for unforeseen contingencies. In addition to that, studies can be conducted in order to identify the cases where the KNN cannot provide accurate performance.

7.3 Conclusions

Several multiclass classification methods and machine learning techniques have been investigated in order to develop the MLRM-IDtool. It has been shown that the complexity of the tool will depend on how difficult it is to identify the right MLRM. A hierarchical stacked classifier combining DTs and ANNs has shown provide the best results for the IEEE30 bus test system. In case several variables are present, variable selection can significantly reduce the total number of variable used to identify the MLRMs. It has also been shown that the DT based MLRM-IDtool can be made noise resilient and achieve good accuracy if it is trained with noisy variables. The KNN-based MLRM-IDtool has shown to be inherently less susceptible to noise compared to the DT-based tool. This has been observed for test sets containing only unforeseen LIDs and for unforeseen LIDs and contingencies.

CHAPTER 8. APPLICATION OF REACTIVE POWER RESERVE SENSITIVITIES FOR REAL TIME VOLTAGE STABILITY MARGIN CONTROL

This study has the objective of determining the most effective control actions in order to reestablish critical RPRs and VSM during emergency conditions. Initially, the concept of reactive power reserve sensitivity with respect to control actions is introduced. In the sequence, a control approach based on a convex quadratic optimization problem is used to identify the minimal amount of control necessary to increase RPRs and VSM above offline pre-specified levels. The approach identifies the proper location and amount of control necessary to bring specific reactive power reserves to pre-specified levels. Simulation results have shown that by using reactive power reserve sensitivities, system operators can optimally determine a proper amount and location of control actions in order to restore critical RPRs and enhance VSM. Moreover, the optimization problem size can be made small by only selecting the most effective control actions, thereby facilitating real time implementation of the method.

8.1 Introduction

With the increased penetration of smart grid technologies and expansion of renewable generation portfolio, electric power systems are expected to operate under unprecedented levels of uncertainty, EISA (2007). Innovations in both transmission and distribution levels are certainly not only changing the way power systems behave, but also the way that the systems are operated.

Although the benefits of a more efficient system with a lower carbon footprint are innumerable, the challenges faced in order to implement and operate such a system are enormous, USDOE (2007) and USDOE (2008).

In order to maintain the grid efficient and reliable, operators will need to take quick and effective measures against degenerating system conditions. Therefore, methodologies for real time control will play a crucial role in maintaining safe operation. This research addresses the problem of real time voltage stability control in emergency conditions.

The influence of reactive power reserves in maintaining adequate voltage control and stability is widely known Taylor, C. (1994), Van Cutsem, T and Vournas, C. (1998), Ajjarapu, V. (2006). In the United States, NERC has issued standards that aim at monitoring reactive power reserves in real time, TOP-006-1 (2006), TOP-006-1 (2006). The standards also require transmission operators to maintain reactive resources to be used in case degenerative system conditions occur, VAR-001-1 (2006), VAR-001-1 (2010).

Monitoring RPRs is the first step to improving system reliability by observing whether RPR levels remain within pre-established limits Taylor, C. W. and Ramanathan, R. (1998), Bao, L.. et al. (2003), Leonardi, B. and Ajjarapu, V. (2011). However, in case RPRs start to drop below acceptable levels, system operators need to rapidly intervene to maintain RPRs within safe limits.

Several studies have demonstrated that the amount of VSM has a strong positive correlation with RPRs and different preventive/corrective control approaches have been proposed.

Vaahedi, E. et al. (2001) proposed a planning VAR method considering credible contingencies in a planning horizon. Margin requirements are incorporated in the approach

which is formulated as a nonlinear optimization problem with barrier functions and solved using Bender's decomposition method.

In Dong, F. et al. (2003), a dual objective optimization approach to maximize the amount of RPRs and reduce system losses is proposed. Simulation results have shown that the amount of VSM increased with an increase of RPRs. The approach used a nonlinear optimization framework based on optimal power flow and bender's decomposition to determine the best current operating condition.

The concept of reactive reserve based contingency constrained optimal power flow (RCCOPF) is introduced in Song, H. (2003). An optimal power flow framework is used to identify the minimal amount of RPRs necessary to improve the amount of VSM for various contingencies and operating conditions. Implementation of the approach shows that the amount of VSM is improved and that the found system state (power flow solution) corresponds to the minimum effective RPR.

The aforementioned approaches are based on variations of nonlinear optimal power formulation, hence being adequate for day ahead and offline applications. They can also be used to determine adequate levels of RPRs based on the study of different scenarios and contingencies. However, if uncertainties involved with real time operations reduce RPRs beyond pre-specified limits, control actions should be taken quickly in order to avoid further voltage profile degeneration and, in the worst case, a voltage collapse.

In this study, a methodology is proposed to address the problem of real time voltage stability through the enhancement of critical RPRs. The method is expected to be used in emergency situations, when low amounts of RPRs and VSM or voltage violations are observed. Sensitivities of control actions of these critical RPRs are used to determine the optimal amount and location

of control. Once the most effective control variables are identified, the solution of the quadratic convex optimization problem will determine the minimal amount of control necessary to recover RPRs and VSM.

8.2 Reactive Power Reserve Sensitivities

In order to develop a control approach that could be used to improve RPRs in real time operations, two major aspects need to be taken into consideration: computational burden and effectiveness of control. The approach needs to be based on a fast and reliable control algorithm, which is expected to converge in a reasonable time frame. Moreover, the control actions obtained must guarantee that critical RPRs will be restored to safe levels after a minimal amount of control is applied.

Before the concept of RPR sensitivities with respect to control actions is introduced, the definition of a generator RPR used in this work is presented in (8.1).

$$Q_{R_i} = Q_{\max_i}(P_g) - Q_{g_i}, \text{ where} \quad (8.1)$$

$$Q_{\max_i}(P_{gi}) = -V_{gi}^2 / X_{di} + \sqrt{V_{gi}^2 I_{fd\max}^2 / X_{di}^2 - P_{gi}^2}$$

In equation (8.1), Q_{Ri} is the amount of reactive power reserve in generator i , $Q_{\max i}(P_g)$ is the maximum reactive power limit given by the field heating limit in the capability curve and Q_{gi} is the current reactive power produced by generator i . A simplified capability curve of a synchronous generator is pictorially given in Figure 8.1. A typical PV diagram shows the amount of system VSM and is shown in Figure 8.2. Both figures will be used to introduce the qualitative effect of different controls on RPRs and VSM.

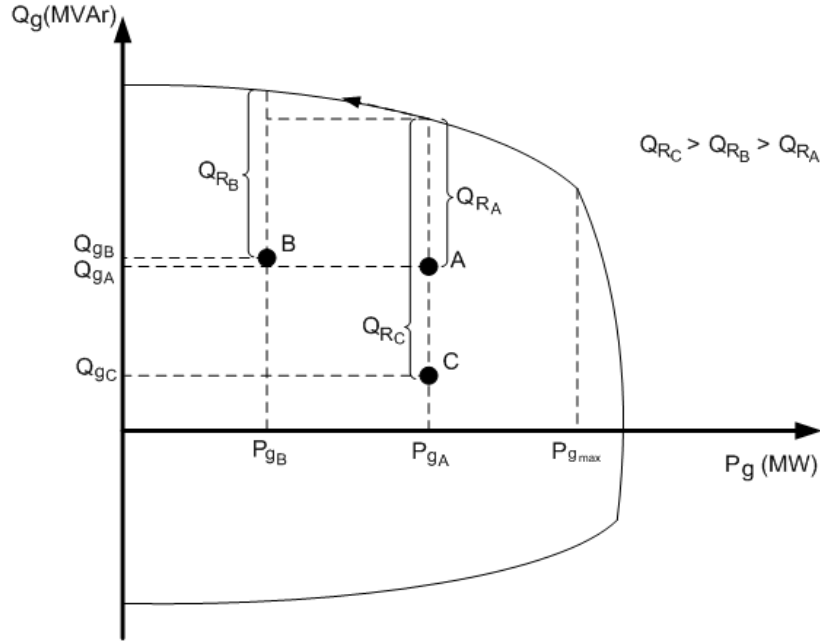


Figure 8.1. Capability curve of a synchronous generator: amount of RPRs depends on operating condition (A, B or C) and machine limits.

Initially, let us assume that the system is operating at point A both in Figure 8.1 and Figure 8.2. Consider that the amount of reserve that must be maintained by the generator at point A (Q_{RA}) is found to violate its minimum RPR requirements.

In order to bring the RPR to a safe level (usually determined via offline studies), control actions must be applied to the system. One option of control would be to reduce the active power generation from P_{gA} to P_{gB} , thus changing the amount of RPR from Q_{RA} to Q_{RB} . Another potential control action is to switch shunt capacitors/reactors or shed local load. In case shunt switching or load shedding are employed, the reactive power production of the machine changes from point A to point C as described in both in Figure 8.1 and Figure 8.2, with a final RPR of Q_{RC} .

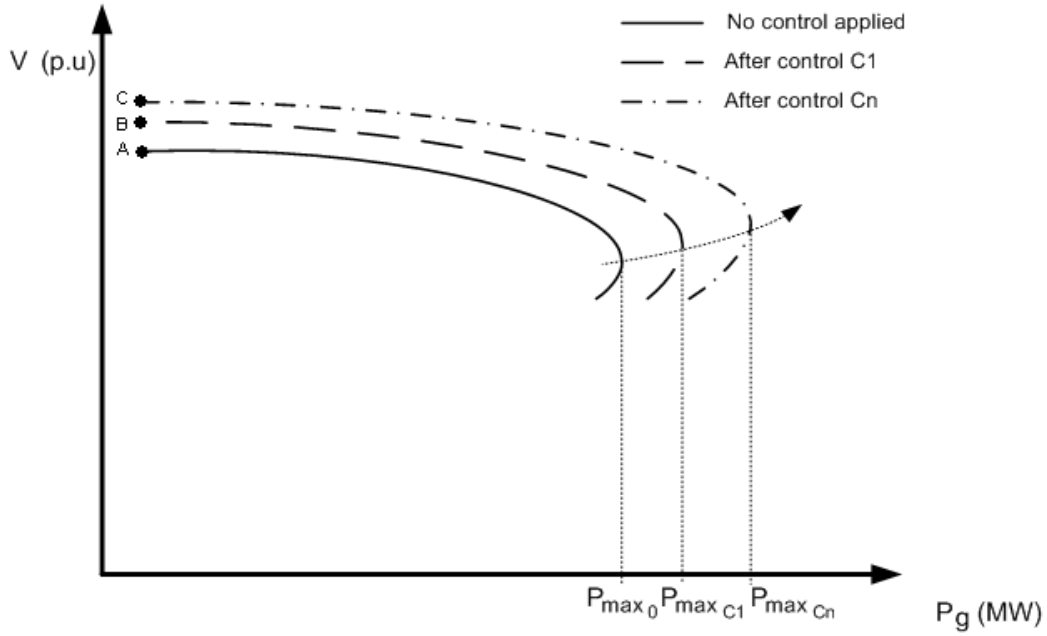


Figure 8.2. PV curves depicting the linear control strategy proposed.

In this study, an attempt is made to explore the concept that by increasing the amount of critical RPRs, the amount of system VSM will also increase. Therefore, once proper amount of control is applied on a critical set of generators, their RPRs should be brought back to levels specified in offline planning studies, as should the amount of system VSM.

Three potential control actions have been considered in this study: active power re-dispatch, shunt capacitor/reactor switching and load shedding. Other control actions can be investigated and the method proposed here can be naturally extended to accommodate other controls. Begovic, M. and Phadke, A. (1992) calculated the sensitivities of reactive power generation with respect to control actions. In this research, we extend their by defining RPR sensitivities as follows.

- *RPR sensitivity with respect to P_g*

The sensitivity of RPRs with respect to active power generation is introduced by equation (8.2), where P_{gi} corresponds to the current active power generation at unit i , Q_{Li} is the reactive power load at bus i , Q_{Ti} is the injected reactive power at node i ; V_i is the current terminal voltage, I_{fmax} is the maximum field current limit and X_d is the synchronous reactance of machine i . Bus voltage angles and magnitudes are given by θ_j and V_j , respectively.

$$\begin{aligned} \frac{\partial Q_{R_i}}{\partial P_{g_i}} &= \frac{\partial Q_{\max_i}(P_g)}{\partial P_{g_i}} - \frac{\partial Q_{g_i}}{\partial P_{g_i}} = \frac{\partial Q_{\max_i}(P_g)}{\partial P_{g_i}} - \left(\frac{\partial Q_{L_i}}{\partial P_{g_i}} + \frac{\partial Q_{T_i}}{\partial P_{g_i}} \right) = \dots \\ &- \frac{P_{g_i}}{\sqrt{(V_i I_{f\max} / X_d)^2 - P_{g_i}^2}} - \sum_{j=1}^n \left(\frac{\partial Q_{T_i}}{\partial \theta_j} \frac{\partial \theta_j}{\partial P_{g_i}} + \frac{\partial Q_{T_i}}{\partial V_j} \frac{\partial V_j}{\partial P_{g_i}} \right) \end{aligned} \quad (8.2)$$

- *RPR sensitivity with respect to B_{sh}*

Equation (8.3) represents the sensitivity of RPRs with respect to shunt compensation. All the variables used in equation (8.3) have been previously described, with exception of B_{shk} , which corresponds to the shunt capacitor bank at bus k .

$$\frac{\partial Q_{R_i}}{\partial B_{shk}} = \frac{\partial Q_{\max_i}(P_g)}{\partial B_{shk}} - \frac{\partial Q_{g_i}}{\partial B_{shk}} = \frac{\partial Q_{\max_i}(P_g)}{\partial B_{shk}} - \left(\frac{\partial Q_{L_i}}{\partial B_{shk}} + \frac{\partial Q_{T_i}}{\partial B_{shk}} \right) = - \sum_{j=1}^n \left(\frac{\partial Q_{T_i}}{\partial \theta_j} \frac{\partial \theta_j}{\partial B_{shk}} + \frac{\partial Q_{T_i}}{\partial V_j} \frac{\partial V_j}{\partial B_{shk}} \right) \quad (8.3)$$

- *RPR sensitivity with respect to P_L and Q_L*

Equations (8.4) and (8.5) represent RPR sensitivities with respect to active and reactive load shedding, respectively.

$$\frac{\partial Q_{R_i}}{\partial P_{L_k}} = \frac{\partial Q_{\max_i}(P_g)}{\partial P_{L_k}} - \frac{\partial Q_{g_i}}{\partial P_{L_k}} = \frac{\partial Q_{\max_i}(P_g)}{\partial P_{L_k}} - \left(\frac{\partial Q_{L_i}}{\partial P_{L_k}} + \frac{\partial Q_{T_i}}{\partial P_{L_k}} \right) = - \sum_{j=1}^n \left(\frac{\partial Q_{T_i}}{\partial \theta_j} \frac{\partial \theta_j}{\partial P_{L_k}} + \frac{\partial Q_{T_i}}{\partial V_j} \frac{\partial V_j}{\partial P_{L_k}} \right) \quad (8.4)$$

$$\frac{\partial Q_{R_i}}{\partial Q_{L_k}} = \frac{\partial Q_{\max_i}(P_g)}{\partial Q_{L_k}} - \frac{\partial Q_{g_i}}{\partial Q_{L_k}} = \frac{\partial Q_{\max_i}(P_g)}{\partial Q_{L_k}} - \left(\frac{\partial Q_{L_i}}{\partial Q_{L_k}} + \frac{\partial Q_{T_i}}{\partial Q_{L_k}} \right) = - \sum_{j=1}^n \left(\frac{\partial Q_{T_i}}{\partial \theta_j} \frac{\partial \theta_j}{\partial Q_{L_k}} + \frac{\partial Q_{T_i}}{\partial V_j} \frac{\partial V_j}{\partial Q_{L_k}} \right) \quad (8.5)$$

In equations (8.4)-(8.5), P_{Lk} (in 8.4) and Q_{Lk} (in 8.5) correspond to the active and reactive power load at bus k , respectively. The terms $\partial Q_{T_i} / \partial \theta_j$ and $\partial Q_{T_i} / \partial V_j$ in equations (8.2)-(8.5) can be obtained directly from the reactive power injection equation given in (8.6).

$$\begin{aligned}
P_{T_i} &= V_i \sum_{j=1}^n (G_{ij} \cos(\theta_{ij}) + B_{ij} \sin(\theta_{ij})) \\
Q_{T_i} &= V_i \sum_{j=1}^n (G_{ij} \sin(\theta_{ij}) - B_{ij} \cos(\theta_{ij}))
\end{aligned} \tag{8.6}$$

However, the terms $\partial \theta_j / \partial P_{g_i}$, $\partial \theta_j / \partial B_{sh_k}$, $\partial \theta_j / \partial P_{L_k}$ and $\partial \theta_j / \partial Q_{L_k}$ are obtained through the linearization of the power flow equations around the current operating point.

Let us assume that the nonlinear set of algebraic equations described in (8.6) can be represented in a compact form as (8.7), where \mathbf{x} represents the vector of system variables (θ , V) and p represents any control variable used (in this case, P_g , B_{sh} , P_L or Q_L).

$$f(\mathbf{x}, p) = 0 \tag{8.7}$$

Initially, let us assume that the system is operating at point A in Figure 8.1 and Figure 8.2. The linearization of (8.7) at the current operating condition yields the relationship shown in (8.8), where $f_x|_A$ represents the derivative of the power flow equations with respect to the system variables (Jacobian) and $f_p|_A$ is the derivative of the power flow equations with respect to the control variables.

$$f_x(\mathbf{x}, p)|_A \Delta \mathbf{x} + f_p(\mathbf{x}, p)|_A \Delta p = 0 \tag{8.8}$$

By rearranging the terms in (8.8), the relationship between system variables and the control parameters (vector $d\mathbf{x}/dp$) is given in (8.9).

$$\Delta \mathbf{x} = \left[-f_x(\mathbf{x}, p)^{-1} f_p(\mathbf{x}, p) \right] \Delta p = \frac{d\mathbf{x}}{dp} \Delta p \tag{8.9}$$

Calculation of $d\mathbf{x}/dp$ elements does not involve the inversion of the Jacobian matrix since efficient computational methods are available to find an analytic solution. These power flow sensitivities are further used to calculate the RPR sensitivities defined in equations (8.2) – (8.5).

8.3 On the selection of reactive power reserves

Previous studies have shown that RPRs are essential to maintain stability and voltage regulation across the system. However, only a reduced set of those RPRs will play a crucial role on system's voltage stability. Since reactive power supply is local in nature, RPRs are frequently grouped together on voltage control areas (VCA) across the system Morison, K. et al. (2008). In general, VCAs are composed by load buses, generating buses and a set of contingencies that affect that area. A generating bus is assigned to VCAs in case its participation factor on the most critical eigenvalue is different than zero. Generators usually have zero participation factors on the most critical eigenvalue unless they have reached their capability limits, Gao, B. et al. (1992).

Therefore, there are a reduced number of critical generators that, if exhausted, will cause voltage collapse. This group of critical RPRs has been called *basin reactive power reserves* Schlueter, R. A. (1998).

In this work, an attempt is made to enhance critical RPRs (the ones that form the basin) in order to improve the amount of voltage stability margin. According to Figure 5.2, the relationship between RPRs and VSM can be approximated by a linear function, Bao, L. et al. (2003), Leonardi, B. and Ajjarapu, V. (2011). By reformulating (5.1) in terms of RPRs and VSM, the relationship between these two variables can be expressed as shown in (8.10).

$$VSM \cong \sum_{i=1}^{N_{critical}^{Qr}} \alpha_i (Q_{Ri}) + \sum_{j,k=1}^{N_{critical}^{Qr}} \gamma_{jk} (Q_{Rj} Q_{Rk}) + \sum_{l=1}^{N_{critical}^{Qr}} \mu_l (Q_{Rl}^2) + \beta \quad (8.10)$$

For small perturbations, the higher order terms in (8.10) can be neglected and all individual coefficients of the linear terms can be assumed the same and equal to α . This simplification will cause incur in a larger error when estimating VSM. However, the error involved is relatively small if a single network topology is being considered for small perturbations. It will be shown

later that despite the approximation, the control method finds a converged solution in a relatively small number of iterations. Therefore, by making the aforementioned simplifications, amount of VSM and critical RPRs can be approximated as (8.11).

By taking the first order derivative of equation (8.11) with respect to Q_R^{Total} , the rates of VSM change with respect to total critical RPR changes is approximately constant.

$$VSM \approx \alpha \sum_{i=1}^{N_{critical}^{Qr}} Q_{R_i} + \beta = \alpha Q_R^{Total} + \beta \quad (8.11)$$

$$\frac{dVSM}{dQ_R^{Total}} \approx \alpha \quad (8.12)$$

Equation (8.12) indicates that a constant term α can be used to relate the changes in the amount of VSM with changes in the sum of all critical RPRs (Q_R^{Total}). This linear approximation is very important for it will enable us to estimate the VSM margin enhancements using changes in critical RPRs.

8.4 Proposed control methodology

So far, the concept of RPR sensitivity with respect to control actions has been introduced. The next step is to formulate the control problem that can make use of the most effective controls to enhance critical RPRs. The objective is to identify the minimal amount of control necessary to increase RPRs and VSM to adequate levels while maintaining a normal voltage profile.

It is important to ensure that voltage limits are not violated when control actions are applied. Moreover, in case localized low voltage violations occur, the control actions can also be used to remove voltage violations so that the system operates within the normal voltage band. In this study, the normal voltage band assumed is 0.92-1.05 p.u.

In order to minimize the amount of control, an optimal control approach is proposed in (8.12). The objective function is formed by the sum of the squares of each control action that can be applied to the system. The control actions are represented by ΔP_{gk} , ΔB_{shj} , ΔP_{Li} and ΔP_{Qm} , whereas their respective weights are given by w_k^{Pg} , w_j^{Bsh} , w_l^{Pl} and w_m^{Ql} . The indexes N_{pg} , N_{bsh} , N_{pl} and N_{ql} represent the total number of generators to be re-dispatched, total number of capacitor banks considered and total number of active and reactive loads considered to be shed, respectively.

The first inequality constraint ensures that the critical RPRs will be shifted back to their minimum reserve limits, whereas the second inequality ensures that a minimum amount of VSM is achieved.

RPR sensitivities with respect to changes in active power dispatch, shunt capacitor/reactor switch, active and reactive load shed are represented by $\partial Q_{R_i} / \partial P_{g_k}$, $\partial Q_{R_i} / \partial B_{sh_k}$, $\partial Q_{R_i} / \partial P_{L_i}$ and $\partial Q_{R_i} / \partial Q_{L_m}$, respectively.

Maintenance of adequate voltage limits is accounted for in the third inequality. The equality constraint ensures that load shedding is performed at constant power factor for any load u with phase angle given by θ_{Lu} . A total of tl buses are considered for load shedding.

Since the problem (8.13) has a quadratic positive definite objective function with linear inequality and equality constraints, it can be classified as a quadratic convex optimization problem. For this type of problems, a solution can be found (or shown not to exist) in a finite number of iterations. In other words, the optimization problem can be solved in a finite amount of time, Nocedal, J. and Wright, S. (2006).

For quadratic convex optimization problems, every local optimum solution can be shown to be the unique global optimal. This is a very important characteristic of these type problems as it

ensures that the search for the optimal amount of control will always find the global optimum solution, Nocedal, J. and Wright, S. (2006).

However, since this quadratic convex optimization problem needs to be solved iteratively, there is no guarantee of convergence for the successive iterations of the approach. In case the VSM, RPR or voltage magnitude requirements are significantly stringent, there may not be enough control resources in order to find a converged solution within the feasibility region of the problem. Therefore, it is important to make sure that enough control actions are considered in (8.13) in case significant enhancements of VSM, RPRs and voltage profile are required by the system.

$$\begin{aligned}
 & \text{Min} \sum_{k=1}^{N_{pg}} w_k^{P_g} (\Delta P_{g_k})^2 + \sum_{j=1}^{N_{bsh}} w_j^{B_{sh}} (\Delta B_{sh_j})^2 + \sum_{l=1}^{N_{pl}} w_l^{P_l} (\Delta P_{L_l})^2 + \sum_{m=1}^{N_{ql}} w_m^{Q_l} (\Delta Q_{L_m})^2 \\
 & \text{s.t.} \\
 & Q_{R_i}^{current} + \sum_{k=1}^{N_{pg}} \frac{dQ_{R_i}}{dP_{g_k}} \Delta P_{g_k} + \sum_{j=1}^{N_{bsh}} \frac{dQ_{R_i}}{dB_{sh_j}} \Delta B_{sh_j} + \sum_{l=1}^{N_{pl}} \frac{dQ_{R_i}}{dP_{L_l}} \Delta P_{L_l} + \sum_{m=1}^{N_{ql}} \frac{dQ_{R_i}}{dQ_{L_m}} \Delta Q_{L_m} \geq Q_{R_i}^{minimal} \\
 & VSM^{current} + \alpha \sum_{i=1}^{N_{Qcritical}} \left(\sum_{k=1}^{N_{pg}} \frac{dQ_{R_i}}{dP_{g_k}} \Delta P_{g_k} + \sum_{j=1}^{N_{bsh}} \frac{dQ_{R_i}}{dB_{sh_j}} \Delta B_{sh_j} + \sum_{l=1}^{N_{pl}} \frac{dQ_{R_i}}{dP_{L_l}} \Delta P_{L_l} + \sum_{m=1}^{N_{ql}} \frac{dQ_{R_i}}{dQ_{L_m}} \Delta Q_{L_m} \right) \geq VSM^{\min} \\
 & V_i^{\min} \leq \left(\sum_{k=1}^{N_{pg}} \frac{dV_i}{dP_{g_k}} \Delta P_{g_k} + \sum_{j=1}^{N_{bsh}} \frac{dV_i}{dB_{sh_j}} \Delta B_{sh_j} + \sum_{l=1}^{N_{pl}} \frac{dV_i}{dP_{L_l}} \Delta P_{L_l} + \sum_{m=1}^{N_{ql}} \frac{dV_i}{dQ_{L_m}} \Delta Q_{L_m} \right) + V_i^0 \leq V_i^{\max} \\
 & [\tan(\theta_{L_u})][\Delta P_{L_u}] = [\Delta Q_{L_u}], \quad \text{with } u = [1, L, tl] \\
 & \Delta P_{g_k}^{\min} \leq \Delta P_{g_k} \leq \Delta P_{g_k}^{\max} \\
 & \Delta B_{sh_j}^{\min} \leq \Delta B_{sh_j} \leq \Delta B_{sh_j}^{\max} \\
 & \Delta P_{L_l}^{\min} \leq \Delta P_{L_l} \leq \Delta P_{L_l}^{\max} \\
 & \Delta Q_{L_m}^{\min} \leq \Delta Q_{L_m} \leq \Delta Q_{L_m}^{\max}
 \end{aligned} \tag{8.13}$$

The last four inequalities ensure that limits are defined for each control variable so that it will remain within operational range, i.e., that the amount of control obtained is within physical and operational limits of each device.

A stepwise description of the method used identify the minimal amount of control is then given in Figure 8.3.

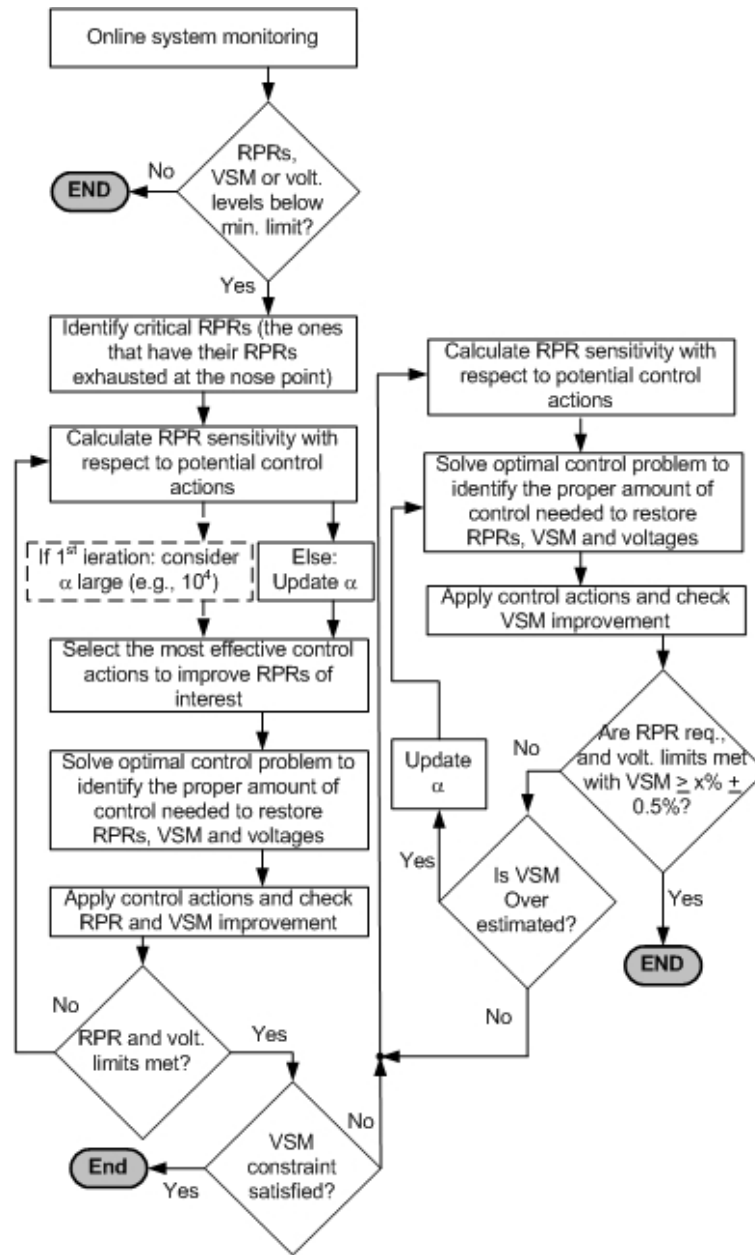


Figure 8.3. Flowchart describing the proposed approach.

Initially, system conditions are monitored in real time during operations. Once RPR/VSM limits are violated or unsafe operating conditions are identified, the methodology is applied to

identify critical control actions. Initially, the system is stressed until the collapse point and the RPR basin is determined. RPR sensitivities with respect to control actions are then calculated at the current operating point.

In the first iteration, the variation of the sum of critical RPRs with the amount of VSM (shown in eq. 8.11) is unknown. In order to keep the VSM margin constraint in (8.13) from binding, a very large value for α is assumed in the first iteration.

Only the most sensitive control actions on critical RPRs are included in the optimization problem (8.13). This makes the control search very effective and reduces the dimension of the optimization problem, valuable characteristics of real time problems.

Solution of the convex quadratic control problem (8.13) will determine the minimal amount of control necessary to recover critical RPRs. For this type of problem, it can be shown that every local optimum solution is also a global optimum solution, Nocedal, J. and Wright, S. (2006).

Once the control actions are identified by solving (8.13), they are applied to the system and some requirements are verified. In case all requirements are met, the approach is considered to have converged. Otherwise, parameter α is updated and the process continues until all requirements are met.

It is important to mention that since the control approach is based on a linearization of the system, it is likely that the amount of control found after the first solution of the optimization problem may not satisfy all RPR, VSM and voltage requirements imposed. Therefore, the approach is iteratively applied until all the constraints are satisfied. It will be shown later that a few iterations are usually needed in order to achieve all imposed requirements.

8.5 Simulation results

8.5.1 IEEE 30 bus test system

This system is composed of five generating units and a slack bus. The studied scenario is a high load condition case, where the system is operating with total load demand of 285.40MW and VSM of 37.08MW. At this load level, some of the RPRs are low and close to be exhausted as indicated by the *initial* bars in Figure 8.5.

After stressing the system to the collapse point, it has been noticed that Q_{R8} , Q_{R11} and Q_{R13} have their limits reached and lose voltage control capability. Therefore, these three units will form the reactive power reserve basin. The sensitivities RPRs with respect to control actions are calculated for the current operating point and shown in Table 8.1.

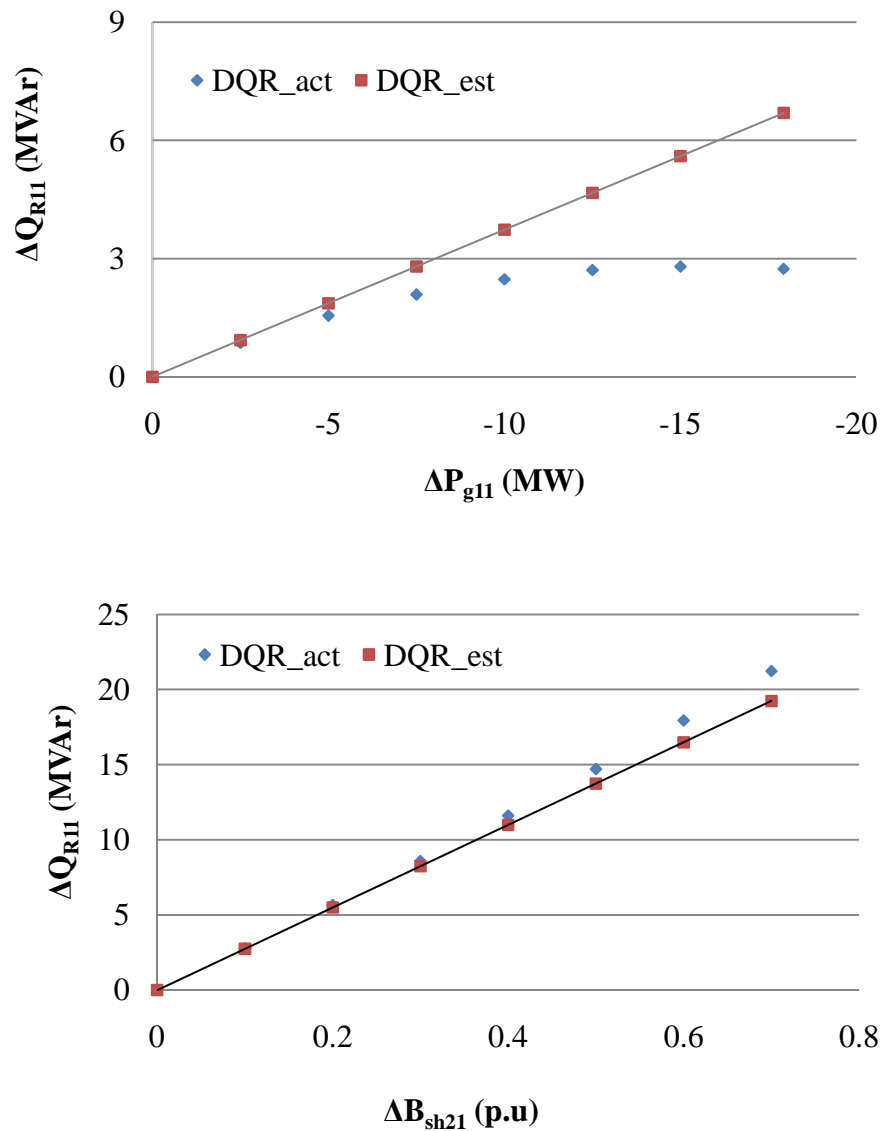
Table 8.1. Sensitivity of RPR with respect to different control actions

	Active power generation					Shunt capacitors switch					Load shedding									
	P_{G2}	P_{G5}	P_{G8}	P_{G11}	P_{G13}	B_{sh7}	B_{sh14}	B_{sh18}	B_{sh21}	B_{sh23}	P_{L16}	Q_{L16}	P_{L17}	Q_{L17}	P_{L21}	Q_{L21}	P_{L24}	Q_{L24}	P_{L26}	Q_{L26}
Q_{R2}	-0.10	0.20	0.15	0.23	0.25	0.12	0.11	0.11	0.12	0.12	-0.25	-0.12	-0.26	-0.13	-0.26	-0.14	-0.28	-0.15	-0.31	-0.18
Q_{R5}	0.00	0.07	0.01	0.05	0.04	0.48	0.04	0.05	0.06	0.05	-0.05	-0.05	-0.05	-0.06	-0.06	-0.06	-0.06	-0.07	-0.07	-0.08
Q_{R8}	0.00	0.02	-0.06	0.27	0.23	0.33	0.22	0.27	0.33	0.29	-0.28	-0.27	-0.31	-0.33	-0.34	-0.37	-0.38	-0.42	-0.52	-0.60
Q_{R11}	0.00	0.00	0.01	-0.37	0.02	0.03	0.11	0.18	0.26	0.16	-0.06	-0.18	-0.09	-0.26	-0.11	-0.29	-0.09	-0.23	-0.08	-0.18
Q_{R13}	0.00	0.00	0.01	0.02	-0.15	0.03	0.52	0.39	0.25	0.39	-0.11	-0.46	-0.07	-0.33	-0.07	-0.28	-0.12	-0.33	-0.10	-0.25

Various important observations can be made from the table above. For instance, reducing active power generation is likely to improve the RPR of the unit which has its active power production reduced. Moreover, reducing active power generation in unit five, its RPR is also reduced. This reduction is due to a higher increase in the amount of reactive power production at that unit, compared to the gain obtained on the capability curve shown in Figure 8.1. Although

this type of behavior has only been observed in one unit, it is a rather interesting observation since RPRs are usually expected to increase with a reduction on generator's active power production.

Figure 8.4 shows a comparison between the estimated RPR improvements and the actual improvement for each control variable. As can be seen from the picture, shunt compensation and reactive load shedding have shown a linear behavior for a wider range of control compared to active power reduction and active power load shed.



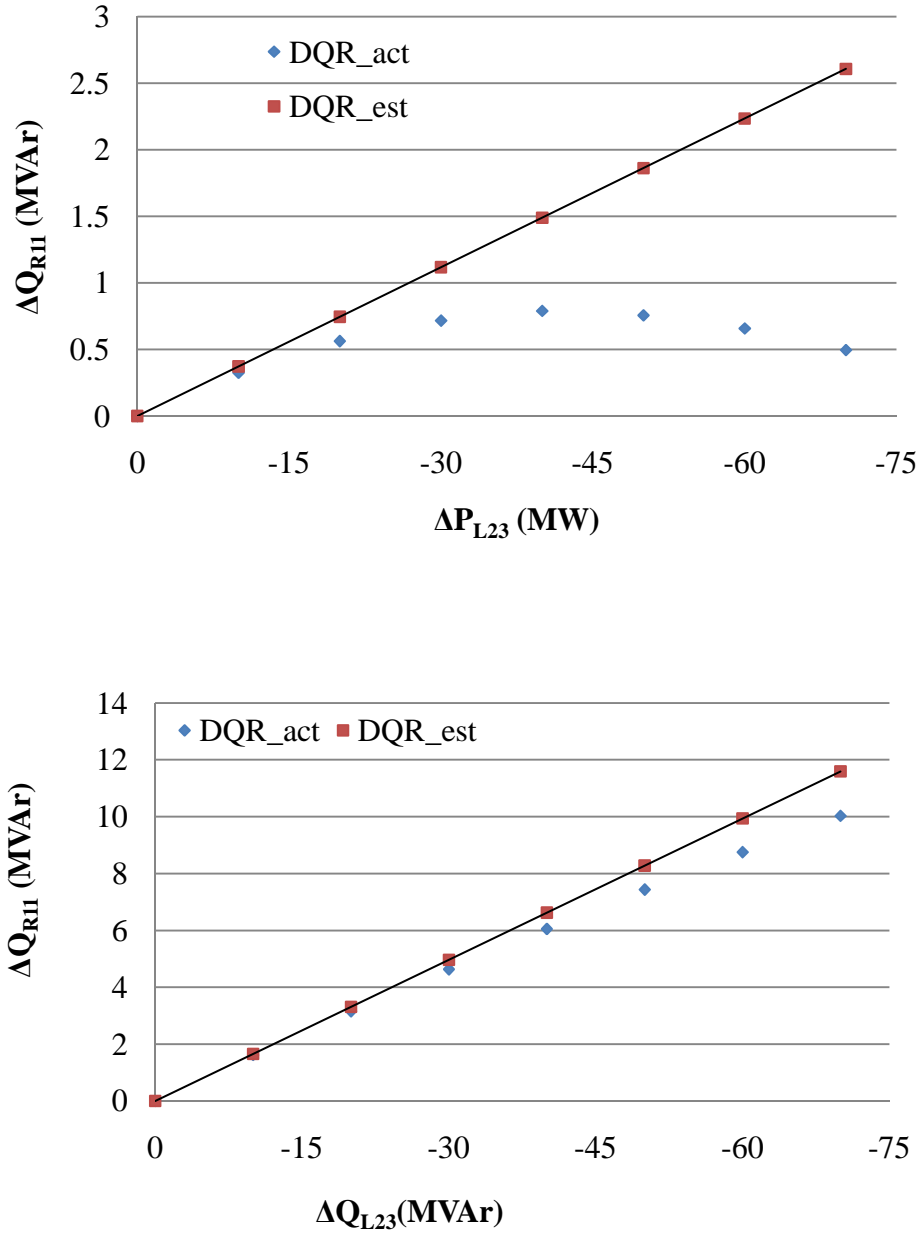


Figure 8.4. Sensitivities of RPR of generator 11 (Q_{R11}) with respect to various control parameters

These characteristics will directly affect the amount of iterations needed by the algorithm in while searching for the minimal amount of control. Inaccuracies due to these nonlinearities will require the optimization problem to be solved more than once. An investigation of how each control variable performs while enhancing RPRs and VSM is done next.

8.5.1.1 P_g as control action

As mentioned earlier, Q_{R8} , Q_{R11} and Q_{R13} are found to form the reactive power reserve basin in this system and are thus included in the optimization problem described in (8.13). RRP requirements for the three units are assumed to be 20, 10 and 10 MVar as shown by the horizontal bars above each unit in Figure 8.5. Three control actions are considered in this case: P_{g8} , P_{g11} and P_{g13} and the minimum active dispatch value of all three units is assumed to be 0MWs.

Due to the stringent RPR requirements, control variable limits and the small sensitivity of active power re-dispatch, no feasible solution could be found in this case. However, in order to analyze the effect of reducing active generation, all three units had their active power production set to zero and variations of their RPRs are shown in Figure 8.5. It can be noticed that the effect of reducing generation tends to cause an enhancement in RPR on only the generator being reduced, with a possible adverse effect on nearby units.

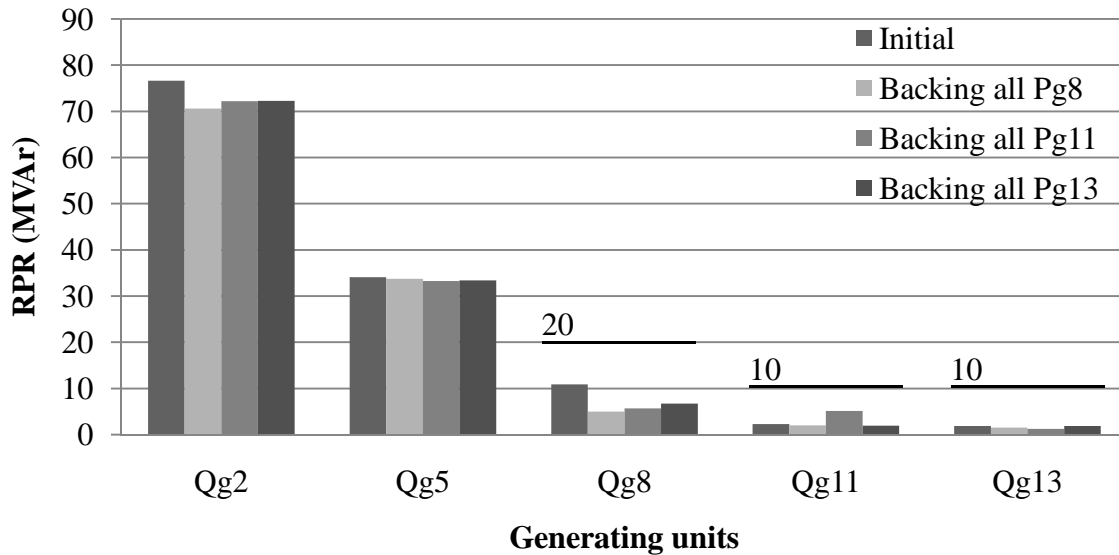


Figure 8.5. Effect of generation active power reduction on RPRs

Moreover, due to the nonlinearity of the capability curve and if large excursions of the control variable are imposed, reduction of active power generation may have an adverse effect on RPRs as the one observed in unit 8. Therefore, simply reducing active generation may not be enough to enhance several RPRs simultaneously.

8.5.1.2 B_{sh} as control action

According to the sensitivity values shown in Table 8.1, shunt capacitors should only cause positive increments on RPRs. Capacitor banks allocated at buses 7, 14, 18 21 and 23 are used to illustrate the approach. As indicated before, control actions need to be taken in order to bring Q_{R8} , Q_{R11} and Q_{R13} back to 20, 10 and 10MVar, respectively. The maximum amount of shunt control allowed is 0.5pu on each location.

Implementing the procedure described in the flowchart of Figure 8.3 and considering all the control weights equal to 1, three iterations are necessary in order to achieve all requirements imposed in the design. The amount of control found on each stage is summarized in Table 8.2.

Table 8.2. Control amounts using load shed as control action

	ΔB_{sh7}	ΔB_{sh14}	ΔB_{sh18}	ΔB_{sh21}	ΔB_{sh23}
<i>After C1(p.u)</i>	0.018	0.089	0.113	0.156	0.195
<i>After C2 (p.u)</i>	0.033	0.069	0.068	0.068	0.054
<i>After C3(p.u)</i>	0.141	-0.001	-0.009	0.168	-0.058
<i>Total (p.u)</i>	0.192	0.157	0.171	0.392	0.191

Figure 8.6 shows the effect of each round of control on the RPRs. Since control variables with a higher sensitivity on the critical RPRs are included in the optimization problem, there is a higher improvement of RPRs that belong to the basin than in other RPRs. It is important to notice also that all RPR requirements are met after the first round of control.

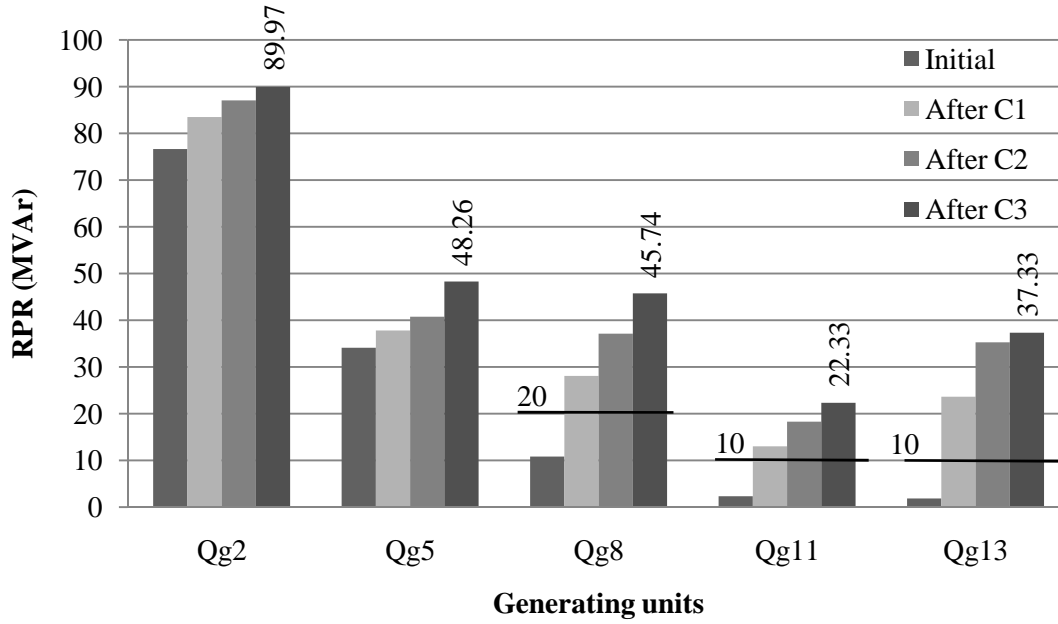


Figure 8.6. Effect of switched shunts on RPRs

Bus voltage magnitudes across the system are shown in Figure 8.7. As can be seen from the figure, low voltage magnitude at bus 26 is the binding constrain in the first round of control. It is important remember that the parameter α is always set to a large value in the first round of control to keep VSM constraint from binding. This is done because we cannot estimate the initial value of α until one round of control is applied. However, after the first round of control, an estimation of parameter α can be obtained by dividing the change in VSM obtained after C1 by the sum of the changes in all critical RPRs (8, 11 and 13 in this case), as shown in equation (8.12).

The second round of control is calculated since the amount of VSM is still under the limits specified in the approach ($30\% \pm 0.5\%$). In total, three rounds of control are necessary in order to meet all the specified requirements. Notice that the control includes reduction of shunt compensation in certain buses in order to avoid violation of voltage limits.

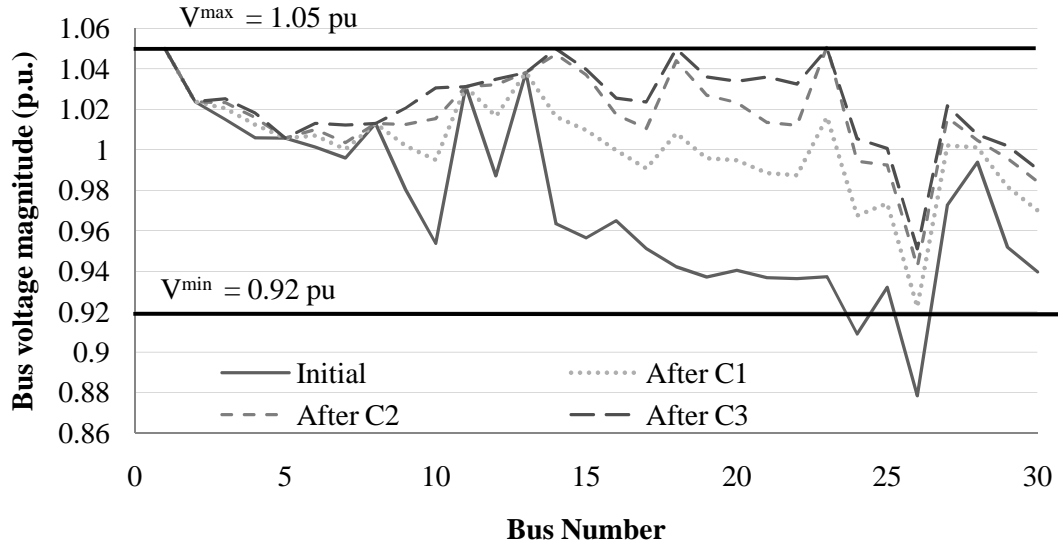


Figure 8.7. Voltage profile enhancement with switched shunts

After C3, the amount of margin is enhanced from 37.09 to 48.14MW, characterizing an improvement of 29.8% as shown in Figure 8.8.

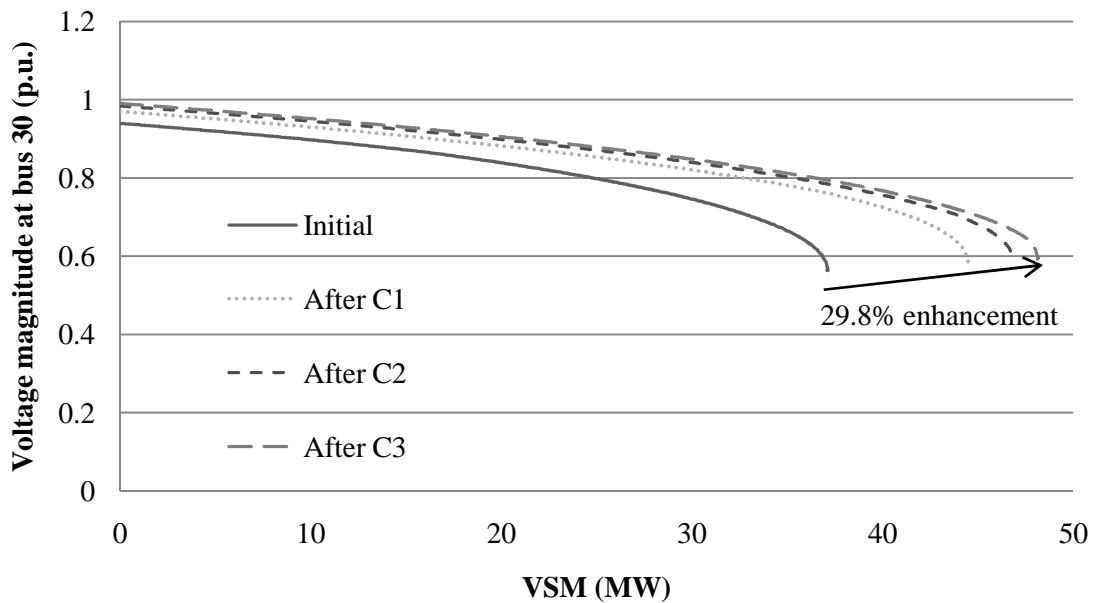


Figure 8.8. VSM enhancement with shunt switch

Therefore, differently than active power generation, switched shunts can be used alone in order to improve RPRs and VSM while maintaining an adequate voltage profile.

8.5.1.3 Load shedding as control action (P_l and Q_l)

The effectiveness of load shedding in enhancing RPR and VSM is investigated next. Although load shedding may be highly undesirable, it may sometimes be the last resort to improve RPRs and avoid a voltage collapse. Load relief can be seen as a demand side response and could be used in order to avoid further system degenerating condition. In the current deregulated energy market, demand side response schemes are already in place in order to reduce peak load and can also be considered for emergency control.

In this study, load shedding is considered on buses 16, 17, 21, 24 and 26, with load being shed at constant power factor. The RPRs improvements due to load shedding are presented in Figure 8.9.

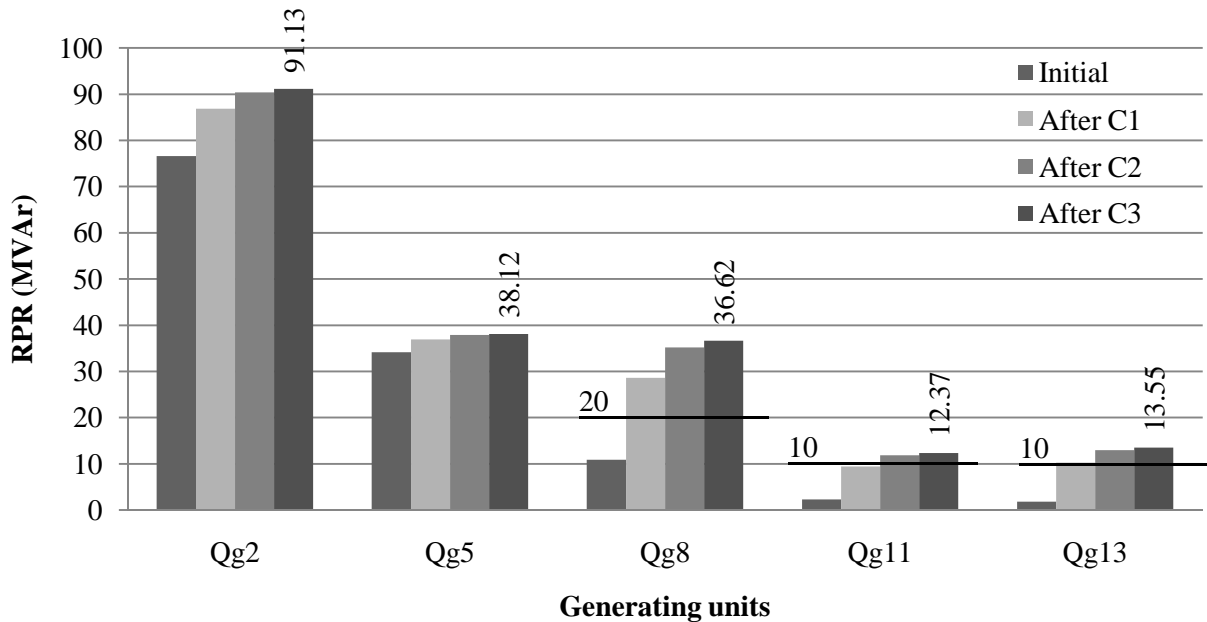


Figure 8.9. Effect of load shedding on RPRs

Three rounds of control are necessary to achieve all the requirements and are displayed in Table 8.3. Notice that only one round of control is necessary to increase critical RPRs to the pre specified levels shown in Figure 8.9. The subsequent control actions will only be used to enhance system VSM while maintaining an adequate voltage profile.

Table 8.3. IEEE 30 – Control Amounts Using Load shed as control action

	ΔP_{L16}	ΔQ_{L16}	ΔP_{L17}	ΔQ_{L17}	ΔP_{L21}	ΔQ_{L21}	ΔP_{L24}	ΔQ_{L24}	ΔP_{L26}	ΔQ_{L26}
<i>After C1(p.u)</i>	0.035	0.018	0.072	0.047	0.081	0.052	0.064	0.058	0.054	0.027
<i>After C2 (p.u)</i>	0.000	0.000	0.018	0.011	0.032	0.021	0.032	0.028	0.036	0.018
<i>After C3(p.u)</i>	0.000	0.000	0.000	0.000	0.009	0.006	0.009	0.008	0.009	0.005
Total (p.u)	0.035	0.018	0.090	0.058	0.122	0.078	0.105	0.094	0.099	0.050

Figure 8.10 shows system wide bus voltage magnitudes. After control C1 is applied, the system not only brings all RPRs to the required values, but is also removes existing voltage violations as shown in Figure 8.10.

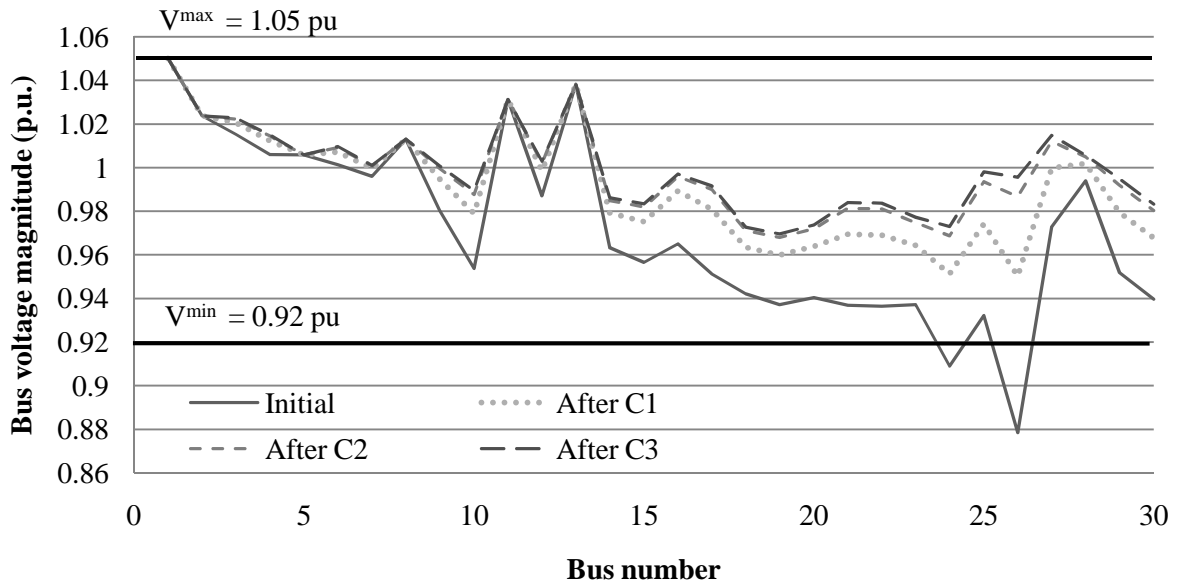


Figure 8.10. Voltage profile enhancement with load shedding

However, two rounds of control are not sufficient to bring the amount of VSM within the range of $30\% \pm 5\%$ and thus a third round of control is necessary. After C3 is applied, the amount of VSM is increased from 37.09MW to 48.07MW, representing an improvement of approximately 29.6% as shown in Figure 8.11. After three rounds of control the method converges with all requirements met and no more control actions are necessary.

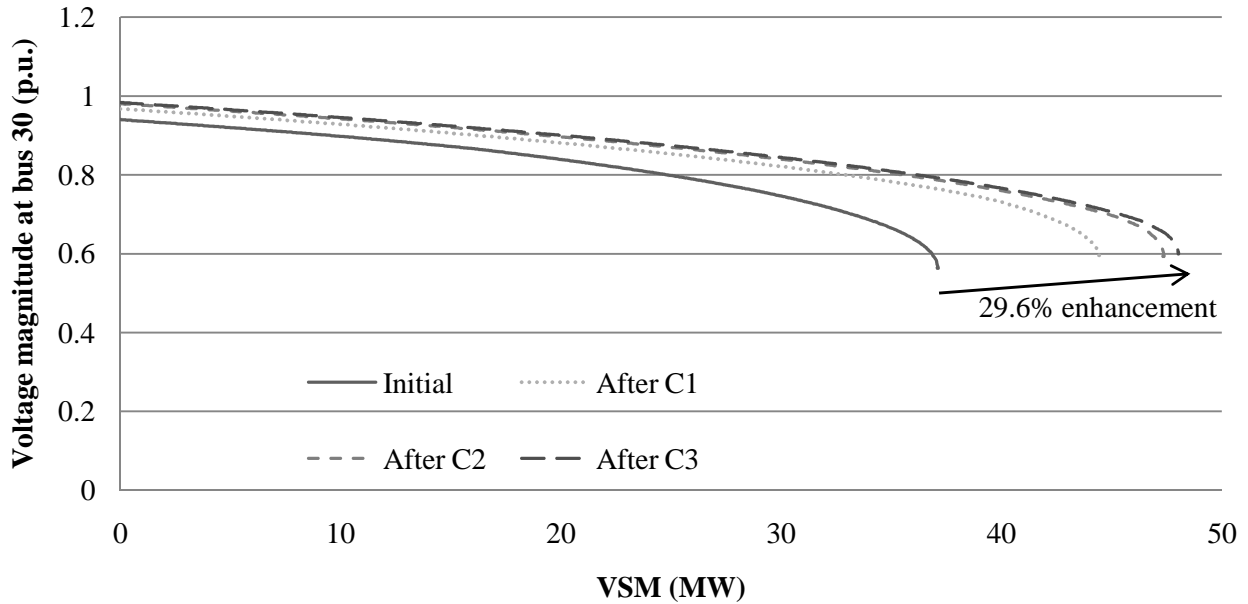


Figure 8.11. VSM enhancement with load shedding

Based on the aforementioned results, load shedding can also be used alone in order to improve RPRs and VSM. In the next section, the performance of all control variables will be assessed together in a larger test case.

8.5.2 IEEE 118 bus system

Once the efficacy of each control variable has been investigated, the approach is implemented considering all controls at once. The IEEE118 bus test system is used to

demonstrate the methodology. A total of 58 machines are available in the system and the current load level is 3867.00MW, with a VSM of 253.07MW. At this operating condition, some RPRs are at a very low level and a few voltage violations are found to occur on the system. In order to enhance RPRs, VSM and voltage profile, the methodology is applied to the system.

System load is increased in order to obtain the RPR basin. Seven generators (Q_{R4} , Q_{R6} , Q_{R8} , Q_{R10} , Q_{R15} , Q_{R18} and Q_{R19}) are found to exhaust their limits and hence form the RPR basin. Initially, the costs associated with each control are considered the same. The amount of control found after implementing the methodology is shown in Table 8.4. It can be seen that there is a significant generation reduction in unit 10 and load shedding on buses 7, 17, 20 and 33. After three iterations, all RPRs, voltage limits and VSM requirements have been met.

Table 8.4. IEEE 118 – Amount of control considering same weight for ΔP_G , ΔB_{sh} , ΔP_L and ΔQ_L

	ΔP_{G10}	ΔB_{sh1}	ΔB_{sh9}	ΔB_{sh14}	ΔB_{sh16}	ΔB_{sh17}	ΔB_{sh22}	ΔB_{sh30}	ΔP_{L7}	ΔQ_{L7}	ΔP_{L17}	ΔQ_{L17}	ΔP_{L20}	ΔQ_{L20}	ΔP_{L33}	ΔQ_{L33}
<i>After C1(p,u)</i>	-0.162	0.229	0.119	0.155	0.447	0.976	0.502	0.3735	-0.071	-0.007	-0.110	-0.030	-0.100	-0.100	-0.034	-0.013
<i>After C2 (p,u)</i>	-0.183	0.166	0.180	0.161	0.146	0.089	0.070	0.0829	-0.088	-0.009	-0.000	-0.000	-0.000	-0.000	-0.064	-0.025
<i>After C3(p,u)</i>	-0.162	0.151	0.164	0.146	0.132	0.081	0.064	0.0753	-0.028	-0.003	-0.000	-0.000	-0.000	-0.000	-0.057	-0.022
<i>Total (p,u)</i>	-0.324	0.381	0.283	0.302	0.580	1.057	0.566	0.4488	-0.099	-0.010	-0.110	-0.030	-0.100	-0.100	-0.091	-0.035

The amount of control after three rounds is displayed on Table 8.5, whereas RPR enhancements are shown in Figure 8.12 and system voltage profile is shown in Figure 8.13. All RPR requirements and system wide voltages requirements are brought back within the specified requirements after one round of control.

The next two rounds of control will be used to regain a minimal amount of VSM. Figure 8.14 shows how the VSM varies with the next two rounds of control. Notice that the amount of margin is increased from 253.07 to 332.43MW (31.4%), representing a small overestimation.

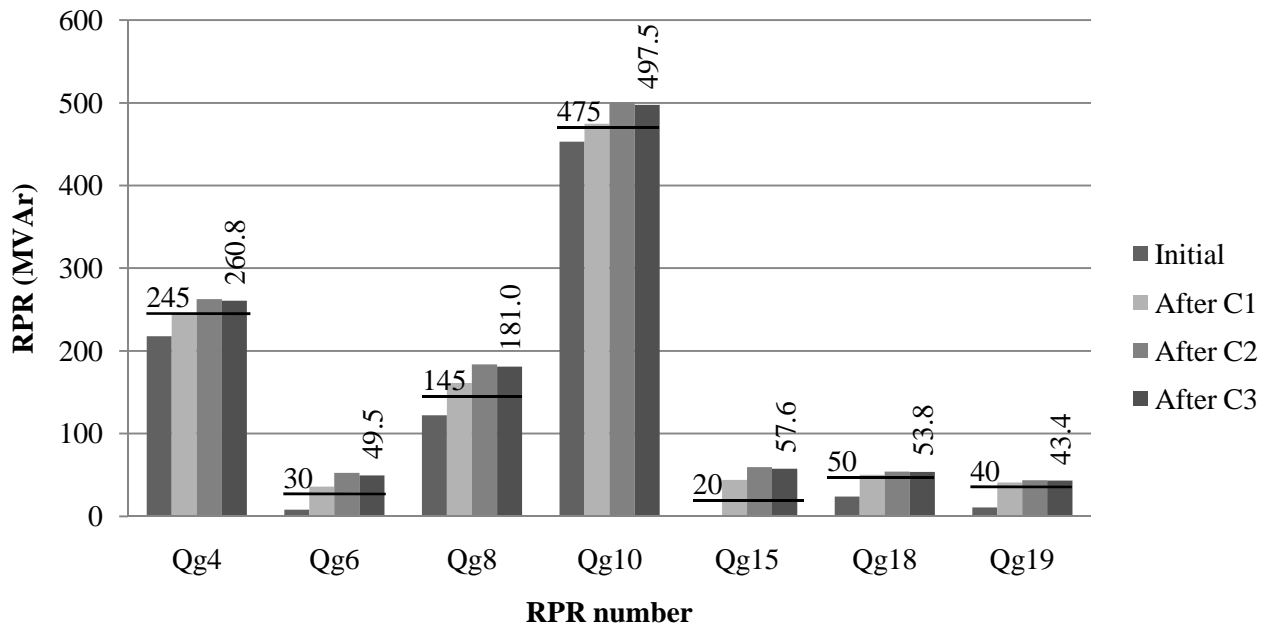


Figure 8.12. Effect of diverse set of control actions on RPRs – same control costs

Since this is more than 0.5% beyond the expected margin enhancement, another round of control needs to be applied in order to reduce some of the excess in margin correction. According to the flowchart above, in case of a small overestimation of margin, a correction of the parameter α should be enough to properly bring VSM to the specified limits.

The third round of control reduces the amount of load shed in the second round and brings the VSM within the $30\% \pm 0.5\%$. A total margin is of 330.2MW is found after C3, representing an enhancement of 30.4%. At this point, all requirements are met and the approach is said to converge.

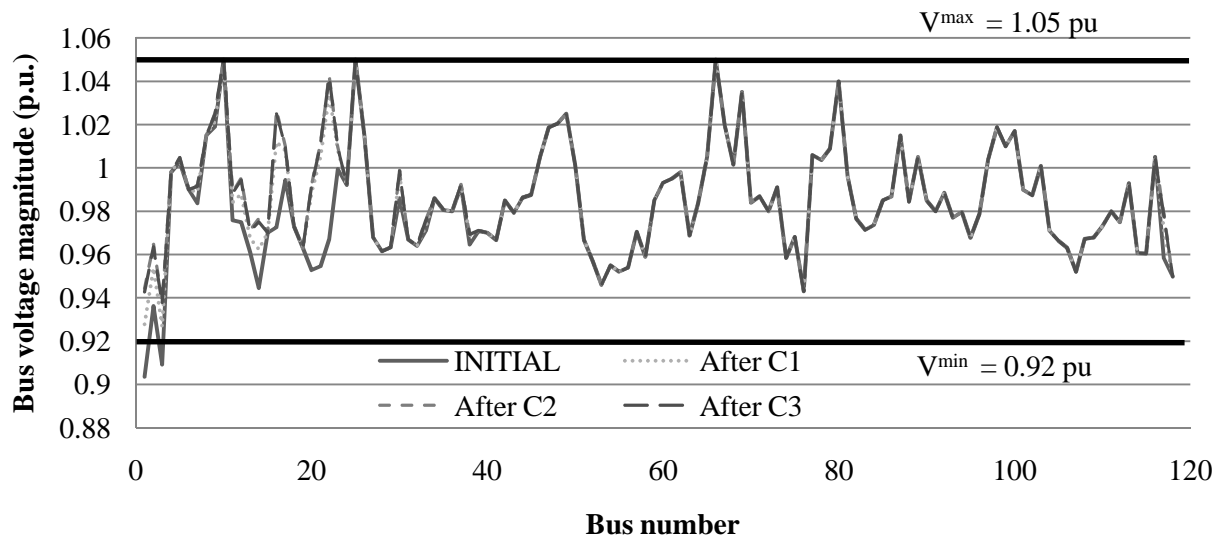


Figure 8.13. Voltage profile enhancement with a diverse set of control actions – same control costs

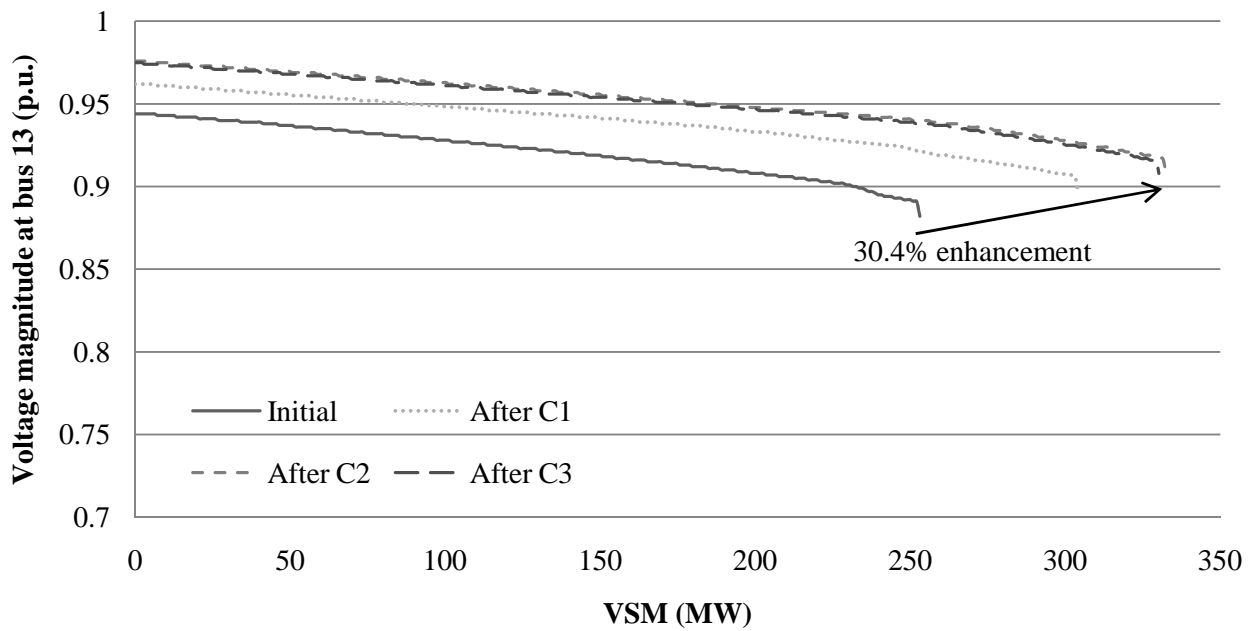


Figure 8.14. VSM enhancement at with a diverse set of control actions – same control costs

However, it is often desirable to avoid generation reduction and load shedding as this can result in economic losses and service interruptions. Another problem that may occur in deregulated energy markets is the absence of justification to tell energy producers that they need to reduce their generation output.

In order to reduce generation re-dispatch and load shedding, their weights in the optimization problem (8.13) are set to 50, whereas the weight of shunt compensation remains at 1. Seventeen possible control candidate variables are considered in (8.13) and only twelve control variables are selected by the method due to their higher sensitivities. The amount of control found at each round of control is shown at Table 8.5.

Table 8.5. IEEE 118 – Amount of control considering weight 1 for ΔB_{sh} /weight 50 for ΔP_L and ΔQ_L

	ΔP_{G10}	ΔB_{sh1}	ΔB_{sh9}	ΔB_{sh14}	ΔB_{sh16}	ΔB_{sh17}	ΔB_{sh22}	ΔB_{sh30}	ΔP_{L7}	ΔQ_{L7}	ΔP_{L17}	ΔQ_{L17}	ΔP_{L20}	ΔQ_{L20}	ΔP_{L33}	ΔQ_{L33}
<i>After C1(p.u)</i>	-0.012	0.223	0.387	0.152	0.457	1.016	0.722	0.389	-0.001	-0.0002	-0.009	-0.002	-0.019	-0.019	-0.000	-0.0003
<i>After C2 (p.u)</i>	-0.004	0.208	0.228	0.202	0.183	0.112	-0.053	0.103	-0.002	-0.0002	-0.001	-0.000	-0.001	-0.001	-0.001	-0.0006
<i>After C3(p.u)</i>	-0.003	0.168	0.184	0.163	0.148	0.090	-0.053	0.083	-0.001	-0.0002	-0.001	-0.000	-0.001	-0.001	-0.001	-0.0005
<i>Total (p.u)</i>	-0.016	0.391	0.571	0.316	0.605	1.107	0.669	0.472	-0.003	-0.0004	-0.011	-0.002	-0.021	-0.021	-0.002	-0.0008

The enhancements in RPRs, voltage profile and VSM can be seen Figure 8.15, Figure 8.16 and Figure 8.17, respectively. Similarly to the previous cases, all RPRs and voltage violations are removed after the first round of control.

However, due to error involved with the linear approximations used to calculate the sensitivities, voltage at bus 22 goes beyond the maximum limit after C1 is applied as shown in Figure 8.16. Therefore, the amount of control applied in the second round of control must remove the voltage violation and still be able to increase the amount of VSM.

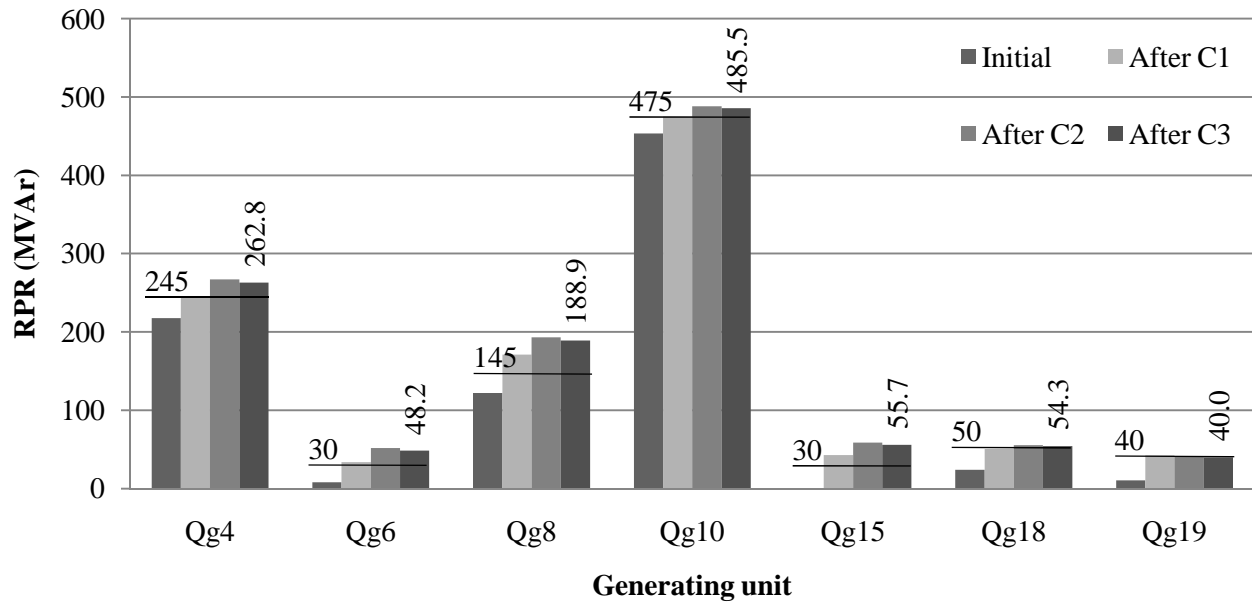


Figure 8.15. Effect of diverse set of control actions on RPRs – different control costs

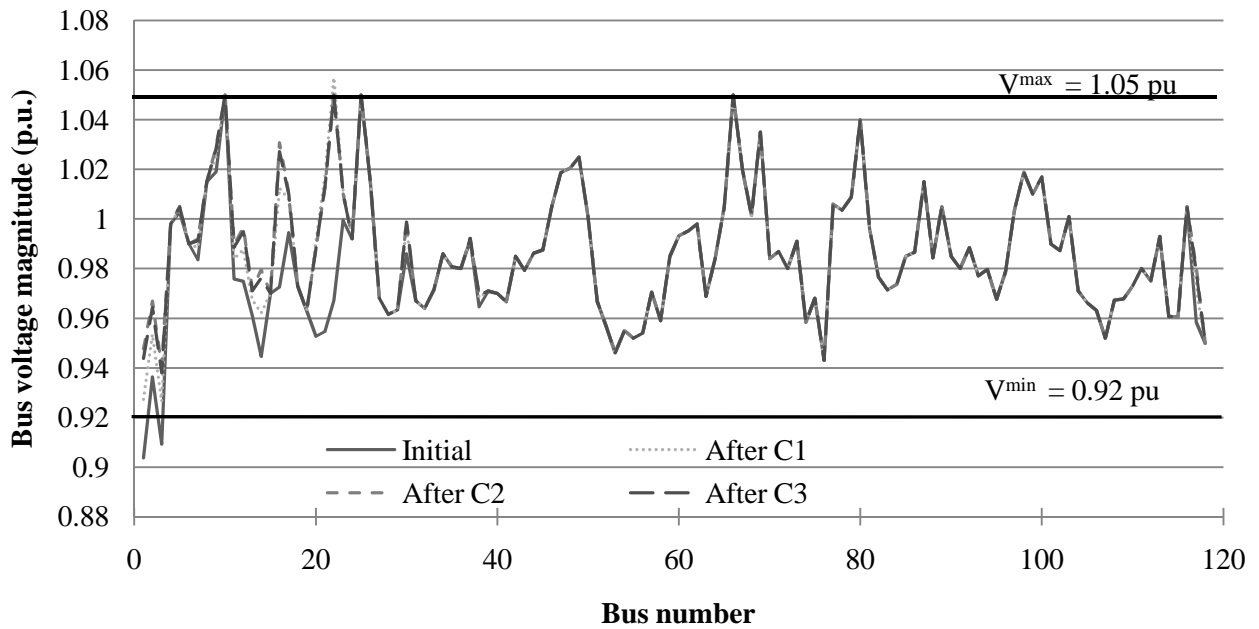


Figure 8.16. Voltage profile enhancement with a diverse set of control actions – different control costs

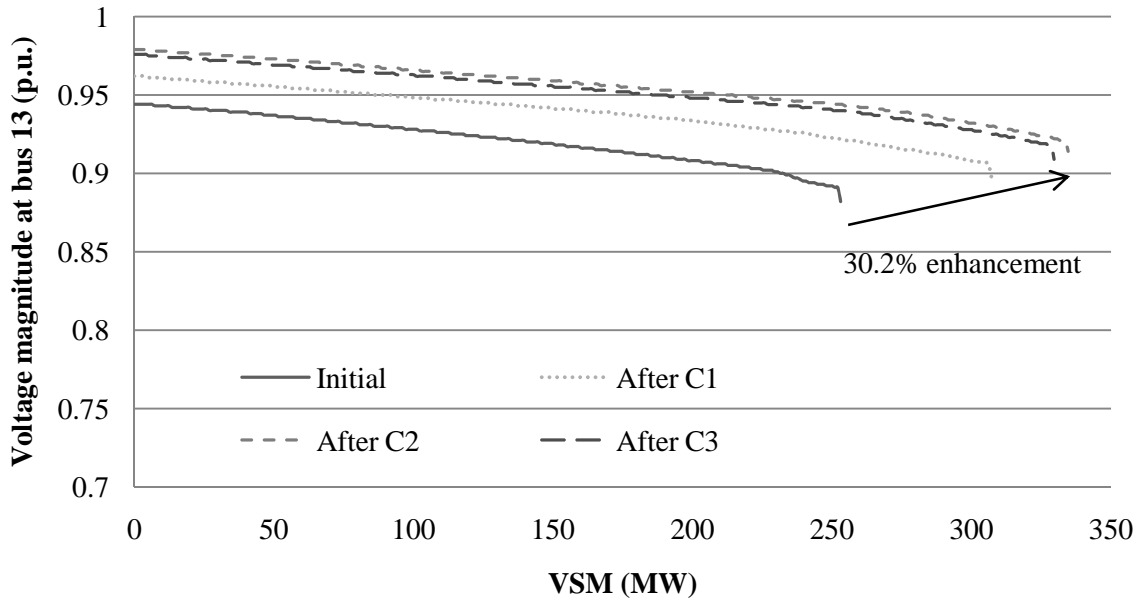


Figure 8.17. VSM enhancement at with a diverse set of control actions – different control costs

After applying C2 from Table 8.5, the voltage at bus 22 is brought back to the specified voltage limits by reducing the amount of shunt compensation at bus 22. Notice that the amount of VSM has increased from 253.07MW initially to 334.67MW, representing an enhancement of 32.2%.

Since the amount of VSM has been increased slightly beyond the desired margin of $30\% \pm 0.5\%$, the third round of control has the objective to adjust the control variables so that the amount of VSM falls within the predefined range. After C3, the total amount of margin is 329.56MW, representing an enhancement of 30.2% and thus finally reaching convergence.

8.6 Conclusions

In this work, a fast control method for online applications has been proposed. The methodology utilizes RPR sensitivities to find the most effective control variables at enhancing RPR, VSM and voltage profile. A quadratic convex optimization problem is then formulated to

identify the minimal amount of control necessary to attain the specified requirements. Results on the IEEE30 and IEEE118 bus test systems have shown that the methodology can successfully find the minimal amount of control in a few iterations. It has also been shown that by weighting the control variables differently, system operators can choose which controls variables to emphasize. Using less generation re-dispatch and load shedding may be preferable since these two actions may cause loss of revenues and service interruption, respectively.

CHAPTER 9. FINAL CONCLUSIONS

9.1 Discussions on the importance of this work

This research has investigated the possibility of using RPRs as an indicator to voltage stability margin. Results have shown that even though the relationship between RPR and VSM has been thought to be linear, some generators have a large quadratic component if a wide operating range is considered. The inclusion of those quadratic RPR terms in the MLRM not only contributed to an improvement in the accuracy, but also helped to enhance statistical properties of the MLRM.

The improvement in performance obtained when quadratic terms are used must also reduce the number of MLRMs necessary to cover all different scenarios. Even when a larger system is considered, only four MLRMs are needed to account for all 190 contingencies and 10 LIDs,. This shows that the methodology can be implemented in larger networks without an excessive number of MLRMs. Another important observation is that the number of MLRMs does not increase with the size of the system, thereby enabling the methodology to be used in large networks.

The incorporation of confidence intervals in the estimations of VSM helps system operators model the amount of uncertainty involved in online operations. Once the MLRM produces an estimation of VSM, the addition of a confidence interval will help system operators to determine the urgency and amount of remedial actions necessary.

Another advantage of the presented methodology is that it provides an absolute degree of stability regarding voltage stability. The estimation of VSM with confidence intervals in MW is

easier for system operators to comprehend. This could be of great usefulness in SCADA systems, where a massive amount of data is currently displayed and concise indexes are needed.

A novel approach to improve RPRs and VSM in real time is also proposed in this thesis. The use of RPR sensitivities helps to identify the location of the most effective control variables that affect each individual reserve. Only a selected set of RPRs is targeted due to their direct influence on system voltage collapse. Those critical reserves combined form the so-called RPR basin. An optimal control approach is formulated as a convex quadratic problem in order to identify the minimum amount of control necessary for these critical RPRs. The advantage of this type of formulation is that there are algorithms which can find a converged solution in a finite number of iterations. Moreover, every local optimal solution can be shown to be a global optimal solution for this type of problem, thereby guaranteeing that the minimal amount of control will be found at each stage.

The insertion of voltage limits and minimal VSM constraints help the method achieve a minimal amount of control while maintaining an adequate voltage profile. The reduced dimension of the optimization problem makes the approach suitable for real time emergency control. The formulation of the control search as a convex quadratic optimization problem has several desirable characteristics for the operations environment, including guaranteed convergence and convergence to the global optimum solution.

Overall, the methodology represents a potential approach to overcome challenges brought by the introduction of smart grid technologies and increased uncertainty.

9.2 Future research

Future research should focus on the practical implementation aspects of the methodology and its capability to generalize. A higher level of automation to preprocess data can also reduce the

total time involved in the MRLM and MLRM-IDtool development. Although good results been achieved, further enhancement of both the MLRMs and the MLRM-IDtool can contribute to more precise and accurate VSM estimations.

One of the points of improvement can focus on how to enable the models to perform well for unforeseen operating conditions and scenarios. An expansion of the training set to include a wider range of operating conditions can be a step towards this goal. An investigation of how sensitive these models are for unforeseen operating conditions can also be done.

The investigation of MLRM performance on unforeseen contingencies can also be object of future research. In this work, a large but limited set of contingencies has been used to generate the MLRMs. Although a large number of contingencies have been considered, it would be interesting to analyze the performance of these models for unforeseen contingencies.

Since the accuracy of the MLRMs will depend on the proximity of the test cases to the current SCADA/EMS condition, a mechanism to compare the cases used to train the MLRMs and the current operating conditions would be helpful. Such mechanism could determine whether the current SCADA conditions have been used during the MLRM training phase. We did not have the opportunity to make such comparison and it is believed that they can be helpful in creating more accurate models.

Another direction of further research can focus on the investigation of the effects of a high penetration of wind turbines on the amount of VSM. The inclusion of wind variability with time can provide a more realistic effect on how the amount of margin varies for different wind scenarios. Variation of the capability curve of wind generators from constant power factor to the actual capability curve can also be included in the simulations for a more realistic assessment.

A further investigation of other machine learning techniques and their performance on the studied scenarios can also be object of future research.

Regarding the optimal control method, other variables can be investigated as potential sources of control to increase the amount of RPRs and VSM. The methodology has shown good results on a mid-sized system and could also be tested on larger networks.

A method to select the control variables that will be used in the optimization problem can also be established. As of now, the control variables with highest sensitivities on critical generators are used.

PUBLICATIONS

– Journals

B. Leonardi and V. Ajjarapu, “Development of multi-linear regression models for online voltage stability margin estimation”, *IEEE Trans. on Power Systems*, vol. 26, no. 1, 2011, pp. 374-383.

B. Leonardi and V. Ajjarapu, “Reactive Power Reserve Sensitivities and Control: A Linearized Approach towards Real Time Applications”, to be *submitted to IEEE Trans. on Power Systems*.

– Conferences

B. Leonardi and V. Ajjarapu, “Investigation of various generator reactive power reserve (GRPR) definitions for online voltage stability/security assessment”, *IEEE PES General Meeting*, Pittsburgh, USA, 2008.

B. Leonardi and V. Ajjarapu, and M. Djukanovic, and P. Zhang, “Investigation of multi-linear regression models and artificial intelligence techniques for on line voltage stability margin estimation”, *VIII IREP symposium - Bulk Power System Dynamics and Control*, Rio de Janeiro, Brazil, 2010.

B. Leonardi and V. Ajjarapu, and M. Djukanovic, “A practical two-Stage online voltage stability margin estimation method for utility-scale systems”, *IEEE PES General Meeting*, Detroit, USA, 2011.

ACKNOWLEDGEMENTS

This dissertation is a material proof that education CAN and WILL provide a brighter future for every person in the world, all it takes is a dream and an opportunity.

Firstly, I would like to thank god for giving me the chance to better understand the world he has created. I am aware that thousands of students would like to have an opportunity like this and I will be eternally grateful for being the chosen one.

I have no words to express all my gratitude and appreciation to my advisor Dr. Venkataramana Ajjarapu. His patience, generosity, guidance, friendship and immense knowledge have been critical in my PhD studies and I would be lost without his help. I also gratefully appreciate the opportunity he has given to me and the financial support for all these years.

A special thanks to Dr. Umesh Vaidya, Dr. Sigurdur Olafsson, Dr. Manimaram Govindarasu and Dr. Cheng-Ching Liu for everything that I learned from them, inside and outside of the classroom. All the skills gained from your courses and talks will certainly guide me in making decisions throughout my professional life.

I would also like to thank my fellow graduate students and the ECpE staff, with whom I have had valorous and helpful discussions regarding everything one can possibly imagine.

As stated in the first pages of this document, this thesis is dedicated to my parents, Jamil Leonardi and Vitória Leonardi, for their efforts, dedication, support and guidance in providing me with everything that I needed to have access to education. Mom and Dad, thank you so much for your words of encouragement along the years.

Last and definitely not least, I would like to thank my fiancée Andrea Bekić, who has been with me along the way. You have given me so much during these busy years and yet have never complained or questioned any of my dreams. You have kept my hopes high during moments of difficulty and we never let it go. Although I have no words to thank you enough for all that you did, I commit myself to give you more of my free time so that I can start “paying off” some of the dedication, care and love that you have provided all these years.

BIBLIOGRAPHY

- Aha, D., Kibler, D. and Albert, M. (1991) Instance-based learning algorithms. *Machine Learning*, vol. 6, no. 1, pp: 37-66.
- Aha, D. (1992) Tolerating noisy, irrelevant and novel attributes in instance-based learning algorithms. *International journal of Man-Machine Studies*, vol. 36, no. 2, pp: 267-287.
- Aly, M. (2005). Survey on multiclass classification methods. November, 2005
- Anand, R., Mehrotra, K., Mohan, C. and Ranka, S. (1995). Efficient Classification for Multiclass Problems Using Modular Neural Networks. *IEEE Trans. on Neural Networks*, vol. 6, no. 1, 177-124.
- Ajjarapu, V. (2006). *Computational techniques for voltage stability assessment and control*. New York: Springer.
- Ajjarapu, V. and Lee, B. (1992). Bifurcation theory and its application to nonlinear dynamical phenomena in an electrical power system. *IEEE Trans. Power Systems*, vol. 7, no. 1, 424–431.
- Anderson, P. and Geckil, I. (2003). Northeast Blackout Likely to Reduce US Earnings by \$6.4 Billion. *AEG Working Paper*. Available at: <http://www.andersoneconomicgroup.com/Portals/0/upload/Doc544.pdf>.
- Begovic, M. and Phadke, A. (1992). Control of voltage stability using sensitivity analysis. *IEEE Trans. on Power Systems*, vol. 7, no. 1, pp. 114 – 123.

- Bishop, C. M., (1995). *Neural Networks for Pattern Recognition*. Oxford University Press: New York.
- Bishop, C. M. (2006). *Pattern recognition and machine learning*. Spring: New York.
- Bao, L., Huang, Z. and Xu, W. (2003). On-line voltage stability monitoring using var reserves. *IEEE Trans. on Power Systems*, vol. 18, no. 4, pp. 1461 – 1469.
- Breiman, L., Friedman, J., Olshen, R. A. and Stone, C. J. (1984). *Classification and Regression Trees*. Chapman & Hall: Boca Raton, Florida.
- Burges, C. J. (1998). A tutorial on support vector machines for pattern recognition. *In Data Mining and Knowledge Discovery*, pp 1–47.
- Box, G. E. P. and Cox, D. R. (1964). An analysis of transformations, *Journal of the Royal Statistical Society*, series B 26 (2): pp. 211–252.
- Breiman, L., Friedman, J.H., Olshen, R.A. and Stone, C.J. (1984). *Classification and Regression Trees, 1st Ed.* New York: Chapman & Hall.
- Chen, Y., Crawford, M. and Ghosh. J. (2004). Integrating support vector machines in a hierarchical output space decomposition framework. *In Proceedings of Geosciences and Remote Sensing Symposium*, vol 2, pp. 949–952.
- Cortes, C. and Vapnik, V. (1995). Support-vector networks. *Machine Learning*, pp 273–297.
- Diao, R., Sun, K., Vittal, V., O'Keefe, R.J., Richardson, M.R., Bhatt, N. Stradford, D. and Sarawgi, S.K. (2009). Decision Tree-Based Online Voltage Security Assessment Using PMU Measurements, *IEEE Trans. on Power Systems*, vol. 24, no. 2, pp:832-839.

- Duda, R. O., Hart, P. E. and Stork, D. G. (2000). *Pattern Classification*, 2nd edition. John Wiley and sons: New York, USA.
- Dong, F., Chowdhury, B. H., Crow, M. L. and Acar, L. (2005). Improving voltage stability by Reactive Power Reserve Management. *IEEE Transactions on Power Systems*, vol. 20, no. 1, 338 – 345.
- EISA (2007). US Senate and House of Representatives, *Energy Independence and Security Act of 2007*.
- El-Keib, A. A. and Ma, X. (1995) Application of artificial neural networks in voltage stability assessment. *IEEE Trans. on Power Systems*, 10, 1890-1896.
- Friedman, J. (1996). Another approach to polychotomous classification. *Department of Statistics, technical report*. Stanford University.
- Freund, Y. and Schapire, R. E. (1996). Experiments with a new boosting algorithm. *Thirteenth International Conference on Machine Learning*, San Francisco, USA, 148-156.
- FERC (2006). Federal Energy Regulatory Commission Staff Preliminary Assessment of the North American Electric Reliability Council's Proposed Mandatory Reliability Standards. Available at: <http://www.ferc.gov/legal/staff-reports/05-11-06-nerc-assessment.pdf>
- Fürnkranz, J. (2007). Round robin classification. *The Journal of Machine Learning Research*, pp.721-747.
- Frank, I. and Witten, E. (1999). Data mining: practical machine learning tools and techniques with java implementation, 1st ed. San Diego: Academic press.

- Freund, Y. and Schapire, R. E (1997). A Decision-Theoretic Generalization of On-Line Learning and an Application to Boosting. *Journal of Computer and System Sciences*, vol. 55, no. 1, pp. 119-139.
- Gao, B., Morison, K., Kundur, P. (1992). Voltage stability evaluation using modal analysis. *IEEE Trans. on Power Systems*, vol. 7, no. 4, pp. 1529 – 1542.
- Goubin, M. (1996). Contingency Severity Assessment for Voltage Security using Non-parametric Regression Techniques, *IEEE Trans. on Power Systems*, vol. 11, no. 1, pp. 101-111.
- Hastie, T. and Tibshirani, R. (1998) Classification by pairwise coupling. *The Annals of Statistics*, vol. 26, no. 2, pp. 451–471.
- Hall, M., Frank, E., Holmes, G., Pfahringer, B., Reutemann, P. and Witten, I. (2009). The WEKA Data Mining Software: An Update. *SIGKDD Explorations*, vol. 11, no. 1.
- Henry, S., Pompee, J., Devatine, L., Bulot, M. and Bell, K. (2004). New trends for the assessment of power system security under uncertainty. *In Proc. of Power Systems Conference and Exposition*, New York, USA.
- Jeyasurya, B. (2000). Artificial neural networks for on-line voltage stability assessment. *IEEE PES General meeting, Seattle, USA*.
- Jimenez, C.A. and Castro, C.A. (2005). Voltage stability security margin assessment via artificial neural networks. *IEEE Power tech conference, St Petersburg, Russia*.
- Kundur, P. (1994). *Power System Stability and Control*, 1st ed. New York: McGraw-Hill, 1994.

- Kleinbaum, D., Kupper, L., Muller, K. and Nizam, A. (1998). *Applied Regression Analysis and Other Multivariable Methods*, 3rd edition. Duxbury Press: Pacific Grove, 1998.
- Kundur, P. (2004). Paserba, J., Ajarapu, V., Andersson, G., Bose, A., Canizares, C., Hatziargyriou, N., Hill, D., Stankovic, A., Taylor, C., Van Cutsem, T., Vittal, V. Definition and classification of power system stability: IEEE/CIGRE joint task force on stability terms and definitions. *IEEE Trans. on Power Systems*, vol. 19, no 3, pp. 1387–1401.
- Kumar, S., Ghosh, J., Crawford, M. M. (2002). Hierarchical fusion of multiple classifiers for hyperspectral data analysis. *Pattern Analysis and Applications*, vol 5, pp.210–220.
- Kutner, M., Nachtsheim, C., Neter, J. and Li, W. (2004). *Applied linear statistical models*, 5th edition. New York: McGrall-Hill/Irwin.
- Leonardi, B. and Ajarapu, V. (2008). Investigation of Various Generator Reactive Power Reserve (GRPR) Definitions for Online Voltage Stability/Security Assessment. *IEEE PES General Meeting*, Pittsburgh, USA.
- Leonardi, B. and Ajarapu, V. (2011). Development of Multi-linear Regression Models for Online Voltage Stability Margin Estimation. *IEEE Trans. on Power Systems*, vol. 1, no. 1, pp 275-279.
- Leonardi, B., Ajarapu, V., Djukanovic, M. and Zhang, P. (2010). Application of multi-linear regression models and machine learning techniques for online voltage stability margin estimation. *In Proc. VIII IREP Symposium*, Buzios, Brazil.

- Lin, Z. C. and Wu, W. J. (1999). Multiple linear regression analysis of the overlay accuracy model. *IEEE Trans. on Semiconductor Manufacturing*, vol.12, no. 2, pp.229-237.
- Morgan, J.N. and Sonquist, J.A. (1963). Problems in the analysis of Survey Data, and Proposal, *Journal of the American Statistical Association*, vol. 58, pp. 415-434.
- Morison, K., Wang, X., Moshref, A. and Edris, A. (2008). Identification of voltage control areas and reactive power reserve: An advancement in on-line voltage security assessment. *In IEEE Power and Energy Society General Meeting*, Pittsburgh, USA.
- Makarov, Y. V., Guttromson, R. T., Huang, Z., Subbarao, K., Etingov, P. V., Chakrabarti, B. B. and Ma, J. (2010). Incorporating Wind Generation and Load Forecast Uncertainties into Power Grid Operations. *Pacific Northwest National Laboratory, technical report*. Available at: http://www.pnl.gov/main/publications/external/technical_reports/PNNL-19189.pdf.
- Nirenberg, S. A., McInnis, D. S., and Sparks, K. D. (1992). Fast acting load shedding. *IEEE Trans. on Power Systems*, vol. 7, no. 2, pp. 63–67.
- NERC (2004). Final Report on the August 14th, 2003 Blackout in the United States and Canada: Causes and Recommendations, *Technical report from U.S. and Canada Power System Outage Task Force*. Available at <https://reports.energy.gov/>.
- NERC recommendations (2004). NERC Final Blackout Recommendations. Available at: <http://www.nerc.com/docs/docs/blackout/section5.pdf>
- NERC-TPL-001-0.1 (2008). System performance under normal (no contingency) conditions (category A). Available at: http://www.nerc.com/files/TPL-001-0_1.pdf.

- NERC TSD. (2010). *Total summer Demmand (TSD)*. Available at: <http://www.nerc.com/page.php?cid=4|38|41> (Total_Demand_Summer.xls)
- Nocedal, J, and Wright, S. (2006). *Numerical optimization*, 2nd edition. New York: Srping-Verlag.
- Quinlan, J. R. (1993). *C4.5: Programs for Mahine Learning*. Morgan Kaufmann: San Mateo, California.
- Quinlan, J.R. (1986). Induction of Decision Trees, *Machine Learning*, vol. 1, pp. 81-106.
- Ruiz, P. A. (2006), Sauer, P. W. Reactive power reserve issues. *The 38th North American Power Symposium*, pp. 439 – 445.
- Rifkin, R. and Klautau. A. (2004). Parallel networks that learn to pronounce English text. *Journal of Machine Learning Research*, pages 101–141.
- Sandberg, L., Roudén, K. and Ekstan, L. (1994). Security assessment against voltage collapse based on real-time data including generator reactive capacity. *CIGRÉ, paper 39/11-03*.
- Studenmund, A. H. (2006). *Using econometrics: a practical guide*. Boston: Addison Wesley Pearson.
- Schlueter, R.A. (1998). A voltage stability security assessment method. *IEEE Trans. on Power Systems*, vol. 13, no. 4, pp. 1423 – 1438.
- Song, H., Lee, B., Kwon, S. and Ajjarapu, V. (2003). Reactive reserve-based contingency constrained optimal power flow (RCCOPF) for enhancement of voltage stability margins. *IEEE Trans. on Power Systems*, vol. 18, no. 4, pp. 1538-1546.

Taylor, C. (1994). *Power System Voltage Stability*. New York: McGrall-Hill.

Taylor, C. W. and Ramanathan, R. (1998). BPA reactive power monitoring and control following the august 10, 1996 power failure. *In proc. of the VI symposium of specialists in electric operational and expansion planning, Salvador, Brazil*.

TOP-006-1 (2006) – Monitoring System Conditions. Available at <http://www.nerc.com/files/TOP-006-1.pdf>.

TOP-006-2 (2008) – Monitoring System Conditions. Available at <http://www.nerc.com/files/TOP-006-2.pdf>.

USGPO (2007). United States Government Printing Office (USGPO): “Energy Independence and security act of 2007”. Available at: http://frwebgate.access.gpo.gov/cgi-bin/getdoc.cgi?dbname=110_cong_bills&docid=f:h6enr.txt.pdf.

USDOE (2007). United States Department of Energy. *The smart grid initiative*. Available at: <http://www.smartgrid.gov/>

USDOE (2008). United States Department of energy. *The smart grid: An introduction*.

Vaahedi, E., Mansour, Y., Fuchs, C., Granville, S., Latore, M., and Hamadanzadeh, H. (2001). Dynamic security constrained optimal power flow/VAr planing,” *IEEE Trans. Power Systems*, vol. 16, pp. 38–43.

Van Cutsem, T., Wehenkel, L., Pavella, M. and Heilbronn, B. (1993). Decision tree approaches to voltage security assessment, *In IEE Proc. Generation, Transmission and Distribution*, vol. 140, no. 3, pp. 189 – 198.

Van Cutsem, T. (1998). Vournas, C. *Voltage stability of electrical power systems*. New York: Spring.

VAR-001-1 (2006) – Voltage and Reactive Power Standard. Available at <http://www.nerc.com/files/VAR-001-1.pdf>.

VAR-001-2 (2010) – Voltage and Reactive Power Standard. Available at <http://www.nerc.com/files/VAR-001-2.pdf>.

Zhong, J., Nobile, E., Bose, A., and Bhattacharya, K. (2004). Localized reactive power markets using the concept of voltage control areas. *IEEE Trans. on Power Systems*, vol. 19, no. 3, pp. 1555-1561.