Machine Learning Based Real-Time Monitoring of Long-Term Voltage Stability Using Voltage Stability Indices

by

Don Kalana Dulanjith Dharmapala

A thesis submitted to the Faculty of Graduate Studies of

The University of Manitoba

in partial fulfilment of the requirements for the degree of

MASTER OF SCIENCE

Department of Electrical and Computer Engineering

University of Manitoba

Winnipeg.

Abstract

Voltage instability manifests as a progressive voltage drop leading to a voltage collapse. It has been associated with many recent blackouts occurred at various parts of the world. Ability to predict the proximity to voltage instability in real-time using measurements can improve the security of power systems.

Among the various techniques that are available to assess the proximity to voltage instability of a power system, Voltage Stability Indices (VSIs) are the widely considered as suitable for generating real-time quantitative parameters that indicate the voltage security level. The majority of the VSIs are theoretically coherent at the proximity of the point of system voltage collapse where they reach some critical value. However, their ability to accurately indicate the voltage stability margin varies with power system operating conditions and under different contingencies: It becomes difficult for the system operators to differentiate voltage instabilities from normal operating conditions using simple rules such as thresholds.

Loadability Margin (LM) is a direct representation of proximity to long-term voltage stability and it is also an easily understandable indicator of voltage stability. However, it is difficult to determine the LM of a large power system through direct computations in real-time. As a solution, a new machine learning based approach is suggested to predict the LM. The proposed technique uses different VSIs computed from synchrophasor measurements as inputs to several Machine Learning Models (MLMs) which forms an ensemble that collectively predict the LM. The VSIs used for each MLM is carefully chosen to include those based on different principles. A procedure that include automated training data generation, correlation based input feature selection, and performance based choice of

machine learning algorithm is proposed to develop the proposed system. The studies conducted on the IEEE-14 and -118 bus systems showed that the Random Forest Regression gives the best performance in terms of the accuracy and robustness.

The effects of practical aspects such as synchrophasor measurement errors on the LM prediction accuracy were analyzed. Finally, a real-time version of the proposed LM prediction system was implemented on PhasorSmart® synchrophasor application platform and was tested using the IEEE 14-bus system and phasor measurement units simulated on RTDS® real-time simulator.

Acknowledgment

This research was completed with the enormous support from various individuals and institutes. I would like to express my sincere gratitude to all those who helped me throughout the research work.

First and foremost, I would like to acknowledge my supervisor Prof. Athula Rajapakse. I profoundly admire his encouragement, guidance and dedication throughout this research.

Then I would like to be thankful for the MITACs Accelerate programme collaborated with RTDS Inc. which provided the financial support for this research.

I would also like to convey my gratitude to Dr. Krish Narendra from Electric Power Group, Pasadena, U.S.A for his valuable advises and providing the PhasorSmart® software for the real-time implementation.

I'm also sincerely thankful for Dr. Yi Zhang and Mr. Paul Forsyth from RTDS Inc., Winnipeg, Manitoba for their advises and inputs provided throughout the research.

Special thank must be conveyed to Mr. Shrimal Koruwage for providing the technical and IT support whenever necessary.

Finally, I would like to thank my wife, parents, family members, friends and alma maters for constantly nurturing, supporting, motivating and bringing me up to be the human who I am today.

Dedication

To my beloved wife, parents, family members teachers and friends.

Contents

List of Tables

List of Figures

List of Abbreviation

ANNR	artificial neural network regression
CPF	continuation power flow
CSF	contingency severity factor
DSA	dynamic security assessment
DWT	discrete wavelet transform
ERPR	reactive power reserve index
GBR	gradient boost regression
GPS	global positioning system
ISI	impedance matching stability index
LM	loadability margin
LR	linear regression
LTVS	long term voltage stability
MLM	machine learning model
MWE	mean wavelet energy
PDC	phasor data concentrator
RED	relative electrical distance
RFR	random forest regression
RMSE	root mean square error
ROI	region of interest
SCADA	supervisory control and data acquisition
SDC	S difference criterion
STVS	short term voltage stability
SVSI	simplified voltage stability index
TVE	total vector error
UCT	universal coordinated time
VCPI_1	voltage collapse proximity index

VSI	voltage stability index
VSIbus	bus voltage stability index
VSLBI	voltage stability load bus index
VSLI	voltage stability load index
VSM	voltage stability margin
WALI	wide area loss index
WAMS	wide-area monitoring system

Chapter 1

Introduction

1.1 Background

Modern day power systems are operated at stressed conditions due to multitude of reasons such as growth of demand, diverse load patterns, change of the nature of generation due to integration of renewable energy, and optimization of the use of power system infrastructure. Therefore, they are exposed to conditions that can set off frequency, oscillatory, voltage, rotor angle, and harmonic instabilities [1]. Out of all the different types of instabilities, voltage instability draws special attention because it has been associated with or responsible for major blackouts occurred at various parts of the world: examples include Canada-United States blackout (2003), Sweden-Denmark blackout (2003), India blackout (2012), Turkey blackout (2015) [3]-[6]. The voltage instability manifests as a progressive voltage drop leading to the collapse of system voltages, resulting in cascaded tripping of transmission elements on the way. Progressive voltage drops could also be associated with rotor angle instabilities which can very quickly lead to total or partial blackouts [2].

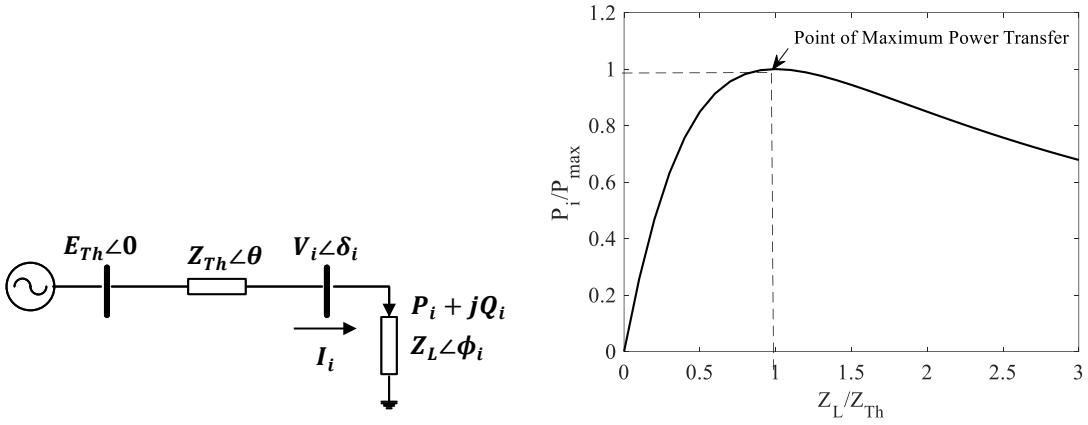
The maximum loadability limits of transmission lines, generator voltage or reactive power control limits, transmission network strength, limits of reactive power compensation devices and the load characteristic are the main factors which affect the voltage stability.

However with the increasing renewable penetration levels, the capability limitations of distributed resources also govern the voltage stability margin of power systems [4].

Different approaches have been used to assess the power system voltage stability. These approaches include modal analysis, P-Q and Q-V curve analysis, singular value decomposition, Voltage Stability Indices (VSI) and continuation power flow [5]. The use of VSIs have gained the attention of researchers since they are able to provide a real-time quantitative parameter to determine the proximity to the voltage instability [6]. The voltage stability in a Region of Interest (ROI) can be considered as wide area phenomenon. Therefore, the VSIs which are calculated using real-time wide-area measurements provides a better insight into the voltage stability of a particular ROI [7]. Phasor Measurement Units (PMUs) and high speed communication infrastructure that facilitate the synchronized wide-area measurements are being increasingly deployed in the power systems. Synchrophasor measurements can provide a high-resolution dynamic snapshot across a wide area of power system [8],[9]enabling a wide range of voltage stability monitoring and controlling potentials.

1.2 Voltage Stability Fundamentals

Voltage stability is commonly explained considering a load (PQ bus) with the rest of the power system represented using a Thévenin equivalent circuit as shown in Figure 1.1 (a). The inductive network elements in the power system limit the voltage support at the load bus, as the power flow increases [4]. This leads to a point where more power cannot be transferred through the transmission network.



(a) (b)

Figure 1.1. (a) Radial power system, (b) Variation of real power with load impedance

Considering the simple radial system in Figure 1.1 (a), an expression for the active power consumption at bus-i (P_i) can be obtained using the known parameters:

$$I_i = \frac{|\overline{E_{Th}}|}{|\overline{Z_{th}} + \overline{Z_L}|} \tag{1.1}$$

$$V_i = Z_L I_i \tag{1.2}$$

$$P_i = V_i I_i cos\phi_i = Z_L I_i^2 cos\phi_i \tag{1.3}$$

$$P_i = \frac{Z_L}{\eta}\left(\frac{E_{Th}}{Z_{th}}\right)^2 cos\phi_i \tag{1.4}$$

where $\eta = \left(\frac{Z_L}{Z_{Th}}\right)^2 + 2\frac{Z_L}{Z_{Th}}\cos(\theta - \phi_i) + 1$

Figure 1.1. (b) illustrates the relative active power ($\frac{P_i}{P_{max}}$) variation as function of relative load impedance, when the $\theta = 84.3^0$ and $\phi_i = 18.2^0$. As the load demand increases ($Z_L$decresres) the value of P_i will increase gradually and reach the maximum when $\frac{Z_L}{Z_{Th}}$ equals to 1. This point is known as the point of maximum power flow. This point is also

defined as the voltage collapse point in voltage stability related literature. Typically, power systems are designed to operate well below the voltage stability limits under normal operating conditions. However, these systems can become vulnerable to voltage instabilities under N-1 or more severe contingency situations.

The additional power that can be transmitted before reaching the voltage collapse point from the current point of operation is usually referred to as the Voltage Stability Margin or the Loadability Margin (LM). This is typically illustrated using the P-V curve as shown in Figure 1.2. LM is an intuitive and easily understandable indicator of proximity to voltage instability. Operators can be alerted and automatic remedial actions can be initiated when the LM drops below the pre-determined critical values.

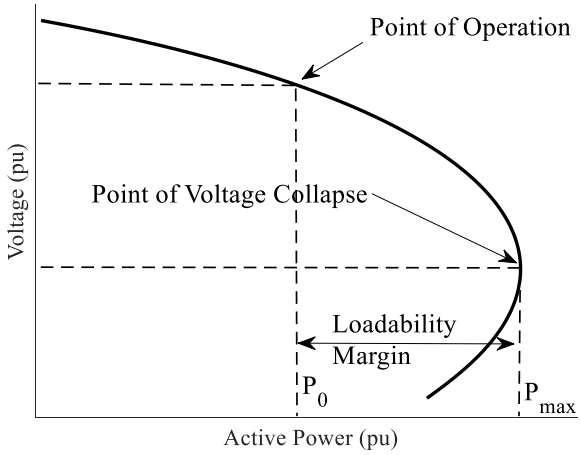


Figure 1.2. P-V curve and Loadability Margin

However, it is impossible to calculate the LM using traditional power flow solution methods since they diverge near the vicinity of the voltage collapse point due to ill conditioned Jacobian matrix. In order to avoid singularity of the Jacobian matrix, the Continuation Power Flow (CPF) algorithm [3] can be used to trace the P-V curve beyond the voltage collapse point. Although it is a good indicator, it is difficult to compute LM in

real-time using the CPF for a large network, due to iterative computations involved. Furthermore, the topology and parameters of the network in real-time is required for CPF. Therefore an alternative approach such as machine learning is required to predict LM in real-time using power system measurements.

1.3 Voltage Stability Indices

Voltage Stability Indices are typically derived considering the theoretical conditions at the voltage collapse point or considering the effects of influential factors such as reactive power reserves. The advantage of VSIs is that they can be easily adapted to provide real-time quantitative parameters to determine the proximity of voltage instability [7]. Calculation of some of the indices require the knowledge of network topology to calculate the Thévenin equivalent circuit. Another class of indices depend on the local synchronized voltage and current measurements while there are indices based on other considerations such as the complex power balance or reactive power reserves in a region of interest. More details of the definitions and principles of selected VSIs can be found in the comprehensive review provided in section 2.5.

1.4 Problem Statement

Many VSIs based on different theoretical approaches have been proposed in literature with the aim of identifying impending voltages instabilities in power systems. Majority of the VSIs are theoretically coherent at the proximity of the power system voltage collapse point, however, their levels of accuracy vary with the power system conditions and under contingencies [6].Furthermore, the sensitivity and the characteristics of some VSIs change

under different system operating conditions. Hence it becomes difficult to define the threshold levels that can help systems operators to differentiate system voltage instabilities from normal system operating conditions. It is almost impossible to depend on a particular VSI to accurately predict voltage instabilities under all circumstances. In addition, other practical aspects such as measurement errors, incorrect network topology information, etc. can also affect VSIs in different ways and it is difficult to single out one VSI which is robust under all circumstances. Hence there is need to improve the way the VSIs are used to determine the proximity to voltage instability of a power system, especially in real-time applications.

1.5 Research Motivation

In order to take fast remedial actions during a disturbance in a power system, the system operators have to process a large volume of information within a short period of time to determine the security state of the power system. This is practically impossible when the system size becomes larger. This is why the necessity of a real-time voltage stability security assessment technique arises. Such a system should assess the voltage stability and alert the system operators when the system security level is low, so that they can make timely decisions to initiate remedial actions based on the assessed security level.

Even though well-established off-line security assessment tools such as VSAT [54], PSAT [51], CPFlow [55] and PSS®E [56] are available, they are neither suitable nor designed for real-time application due to the required computational time. Among them, VSAT is compatible to integrate with an energy manegment system for on-line security assessment, but it is not intended or capable of performing the analyses in real-time. However, there

are wide range of VSIs which can be computed in real-time from the measurements, specifically using wide-area synchrophasor measurements, with much less computational burden. Typical approach has been to use a single VSI to determine the system voltage stability. Many of these VSIs are derived considering equivalent circuits that represent sections of the power system, which are subjected to change, especially under stressed conditions. Therefore, various VSIs show different levels of accuracies under different power system conditions and contingencies. There is a potential for improving the voltage stability assessment by combining different VSIs which are based on different principles, because they are affected differently by the system conditions and changes. However, this requires a real-time intelligent system which can learn from situations previously encountered and apply that knowledge to new situations to get better assessments of voltage stability. Modern machine learning algorithms are highly capable and their use is becoming more accepted in the power industry. Therefore, investigation of the usage of machine learning models (MLMs) and combinations of VSIs to enhance the accuracy and robustness of the on-line voltage security assessment is a worthwhile endeavour.

1.6 Research Objectives

Main objective of the proposed research is to explore the feasibility of developing more accurate and robust voltage stability predictor by combining different types VSIs proposed in literature as input features. The following sub-objectives are proposed as steps to achieve the main objective.

1. To critically review research contributions in the area of voltage stability analysis and to identify several VSIs with distinct advantages and features while being amenable for real-time calculation using synchrophasor measurements.
2. To evaluate the selected VSIs using a test power system suitable for demonstrating the voltage stability through off-line simulations.
3. To develop a voltage security level predictor by combining the most promising VSIs using a suitable MLM.
4. To implement the developed voltage security level predictor on a real-time platform and verify the algorithms under real-time conditions using a laboratory test setup based on a real-time power system simulator.

1.7 Thesis Organization

In Chapter 2, a literature review conducted on voltage stability assessment is presented. This includes categories of power system stability, different traditional techniques to assess the power system voltage stability margin, descriptions of various VSIs used evaluate voltage stability and some background on wide area monitoring using the synchrophasor technology. Finally, different machine learning approaches presented in literature are discussed.

In Chapter 3, overview of the proposed machine learning strategy to predict voltage stability margin using VSIs is presented.

In Chapter 4, the method of generating training data for machine learning is discussed. Generation of random operating points and screening them through automated PSS®E

runs, and evaluation of credible N-1 system contingencies is presented in this chapter. This chapter includes work related to bus voltage phasors and LM computation up to the voltage collapse point by running Continuation Power Flow (CPF). -Finally the process of calculating VSIs and generating a database with VSIs as input features and LM as the output feature is presented.

In Chapter 5, a brief introduction of the multi variant MLMs used in this thesis is presented. Furthermore, the processes of feature selection, MLM training and validation are explained.

In Chapter 6, the development of a real-time voltage stability prediction testbed is presented.

In Chapter 7 presents, conclusions, contributions, and some suggestions for future research in the area of real-time voltage stability monitoring.

Chapter 2

Literature Review

2.1 Real-time Power System Security Assessment

Modern power systems spread over a large geographical terrain and interconnect a large number of equipment through transmission network. The dynamic nature of the system and the advanced control algorithms of equipment makes the power system operation even more challenging for power system operators. In order to address the issue of uncertainty of predicting the subsequent operating condition, on-line Dynamic Security Assessment (DSA) approach has been proposed and implemented in various parts of the world.

On-line DSA approaches require centralized control centers where a snapshot of the power system is taken through near real-time measuring systems such as Supervisory Control and Data Acquisition (SCADA) system, Phasor Measurement Units (PMU) networks, and disturbance recorders. Then using the aforementioned information a system model is generated to perform the security assessment. Computation of security assessment can be done using an intelligent system which consist of logical algorithm or trained MLM. Afterwards the results of assessment is reported to raise the alarms when security issues are detected. Figure 2.1 illustrate an on-line DSA architecture which is reproduce from [53].

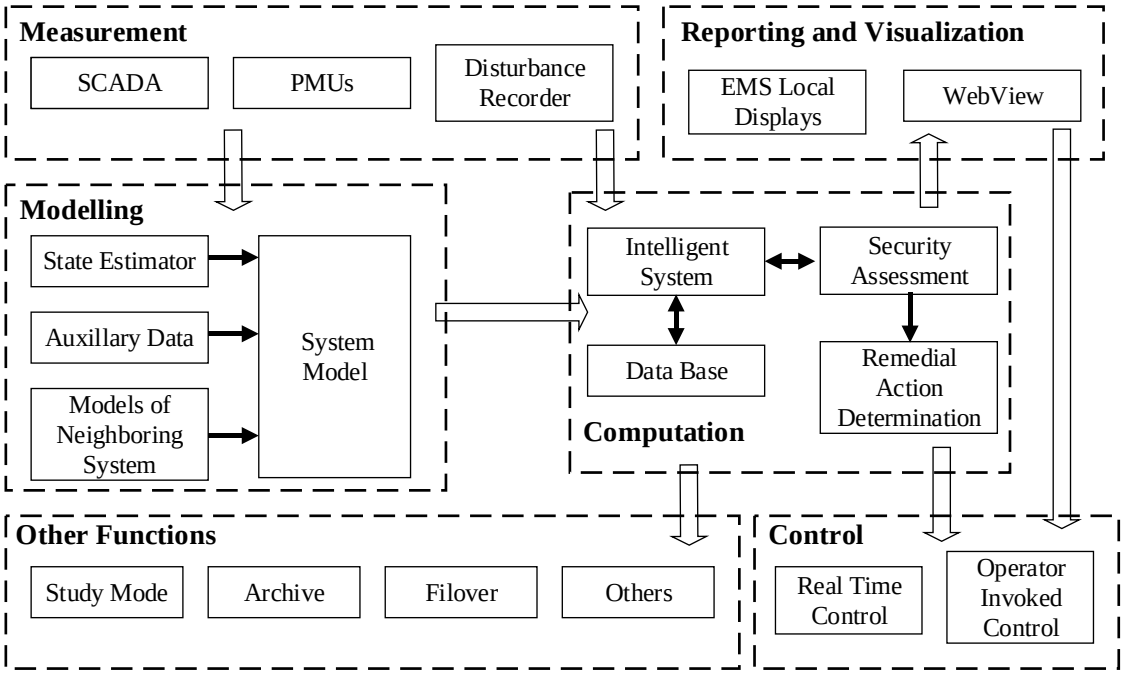


Figure 2.1. Components of online Dynamic Security Assessment System [53]

2.2 Power System Operating States

Operating states of a power system are classified according to the security level. This allows the power system operators to take remedial actions accordingly. Basically the system operating states can be categorised into three states as normal, emergency and restorative. This is further extended to five states as normal, alert, emergency, in-extremis and restorative in [53].

In the normal state all power system constrains are satisfied and the system operates at an adequate security margin. In the alert state, which typically results in due to some component outages, the system security state drops below a certain level. At this state, even though all the system constrains are satisfied, some of the network equipment would still be overloaded. Therefore it cannot be guaranteed that there will be an adequate amount of

reserve margin to withstand a severe contingency. In the emergency state, the power system will still be intact but the violation of system constraints is severe. Therefore most of the equipment will be overloaded under this condition. Such a system can be restored to the normal state if suitable corrective actions are taken. In the extreme state both the system equality and the inequality constraints are violated. At this stage, the system will lose it synchronism and cascade outages will occur and lead to a system wide blackout. In the restoration state, system operators try to reconnect the power system equipment and restore the system either under normal or alert state, depending on the conditions.

2.3 Power System Stability

Power system stability can be defined as the ability of a power system to regain an operational equilibrium after being subjected to a physical disturbance without collapsing and be within the operation limits. The dynamic response of a modern power system after a disturbance is governed by the devices which are intact within the power system and has various characteristics and response times. A power systems remains at stability due to the equilibrium of opposing forces. These opposing forces may be subjected to substantial amount of imbalance during a disturbance leading the power system to different forms of instabilities. This is valid to power systems with Converter Interfaced Generation (CIG) as well, as long as the dynamic response of a disturbance only affects an individual CIG without causing a cascading instability in the main system [4]. Figure 2.2, which is reproduced from [4], represents different categories of power system instabilities.

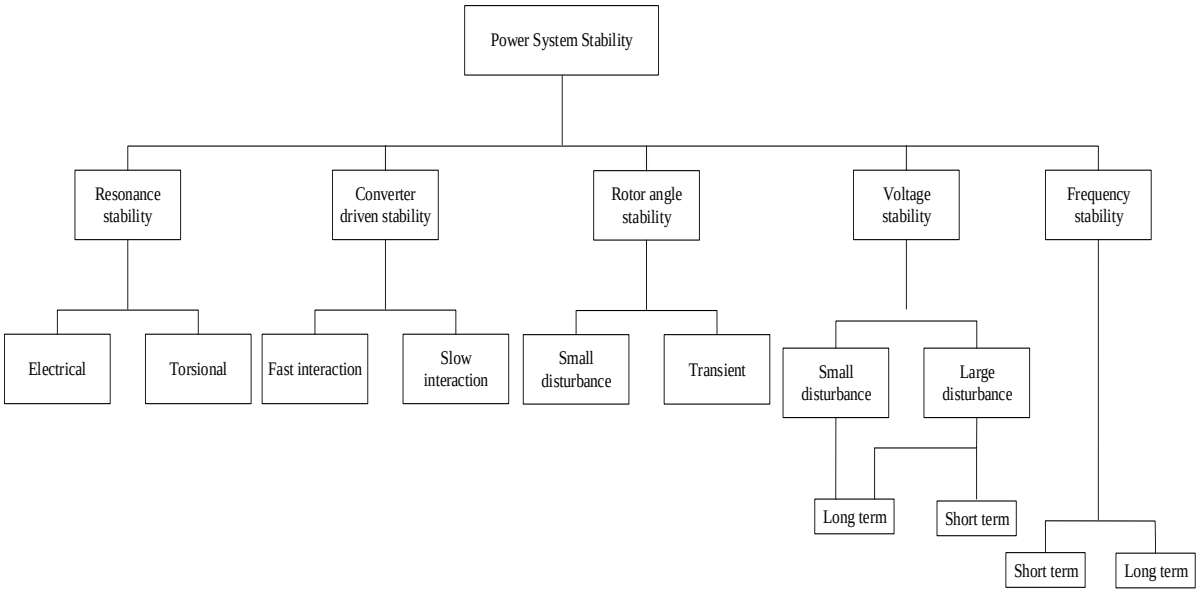


Figure 2.2. Power system stability categorization [4]

2.4 Power System Voltage Stability

The voltage stability of a power system can be defined as the ability of the system to maintain an acceptable steady voltage at all the buses after the system has been subjected to a disturbance from an initial steady operating point. Power system voltage instabilities may result in the form of progressive fall or rise of voltages of some buses. The possible outcomes of such progressive voltage drops are tripping of transmission lines and other equipment by their protective intelligent electronic devices (IEDs), loss of load in an area and tripping of synchronous generators by their protective systems due to pole slipping. Several such simultaneous events may lead to cascading outages of the entire network [2],[4]. Voltage stability can be classified in to two categories: small disturbance voltage stability and large disturbance voltage stability [4]. Both small and large disturbances can be further categorized into short term and long term voltage stability based on the

timeframe. Usually short term voltage stability is defined for a time range of 0-10 seconds and long term voltage stability is defined for a time range of few minutes [2].

2.4.1 Small-disturbance Voltage Stability

The ability of a power system to maintain steady voltages throughout the system when subjected to a small perturbation such as incremental system load change is referred to as the small–disturbance voltage stability. This form of voltage stability is mostly influenced by the load characteristics, and the response of continuous and discrete controllers at the point of operation [2].

2.4.2 Large-disturbance Voltage Stability

The ability of a power system to maintain steady voltages throughout the system when subjected to a large perturbation such as a system fault, loss of generator unit or network contingencies is referred to as the large–disturbance voltage stability. This form of voltage stability is mostly influenced by slow responding equipment such as tap-changing transformers, generator excitation current limiters and thermostatically controlled loads [2].

2.4.3 Voltage Stability Assessment Techniques

Several approaches have been proposed in literature to assess the voltage stability of a power system. Some of these traditional approaches are based on the analysis of power system mathematical models, and the others are based on curve analysis and sensitivity analysis based on eigenvalues of the power flow Jacobian. These traditional voltage stability assessment methods can be categorized as follows:

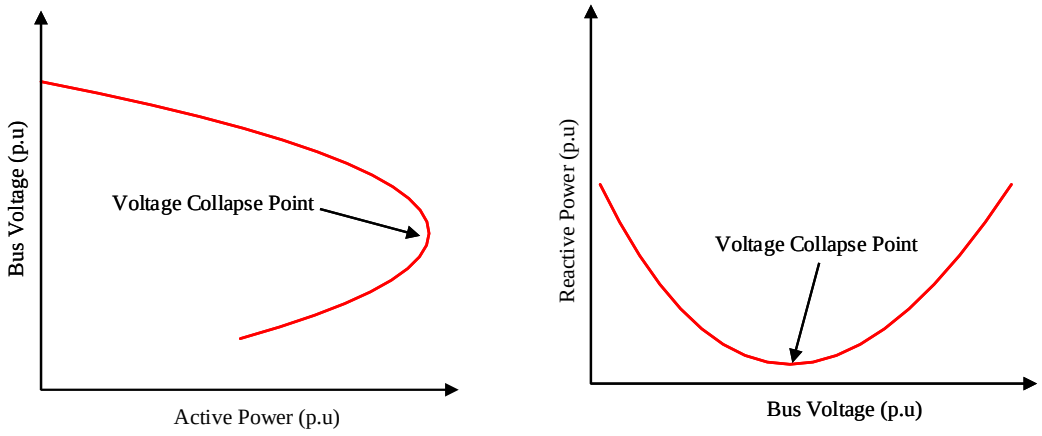
- P-V and V-Q curve analysis method
- Modal analysis method
- Sensitivity analysis method
- Thévenin equivalent methods
- Voltage stability index based methods

2.4.3.1. Curve Analysis Method

The P-V and V-Q curves of the selected buses can be used to analyze the capability of a transmission network and the reactive power compensation requirement of the respective load buses. The P-V curve represents voltage response at a particular bus due to load increase at the bus or at the other system buses. Similarly, V-Q curve represents the voltage response of a system bus for the reactive power support provided by the sources [45],[46].

This can be further explained by using a two bus system where the load is connected to a source through a transmission line with Z impedance. The P-V curve is obtained by gradually increasing the active power demand of the load. Similarly the V-Q curve is obtained by gradually increasing the reactive power demand of the loads. However, reactive power injection of the source is taken into account when plotting the V-Q curve. The characteristics of the P-V and V-Q curves are shown in Figure 2.3.(a) and (b) respectively. These curves can be computed theoretically for a simple radial system.

For a networked power system, voltages need to be computed through power flow analysis. However, it is impossible to perform a power flow analysis up to the voltage collapse point using traditional power flow methods. Traditional power flow equations which are solved using iterative techniques are prone to divergence at the vicinity of the voltage collapse point. In order to avoid the divergence caused by the singularity of Jacobian matrix, the power flow equations are reformulated by applying a locally parameterized continuation technique. This method is called the Continuation Power Flow (CPF) [3],[38].and it is used to obtain the voltage profiles of the system buses up to and beyond the voltage collapse point.



(b)

Figure 2.3 (a). PV curve for 2 bus system and (b) VQ curve for 2 bus system

2.4.3.2. Modal Analysis Method

The modal analysis method assesses the power system voltage stability by computing the eigenvalues and eigenvectors of a reduced Jacobian matrix (J_R).The Jacobian matrix used in power flow analysis is given as

$$\begin{bmatrix}\Delta P\\ \Delta Q\end{bmatrix} = \begin{bmatrix}J_{11} & J_{12}\\ J_{21} & J_{22}\end{bmatrix}\begin{bmatrix}\Delta\delta\\ \Delta V\end{bmatrix} = \mathrm{J}\begin{bmatrix}\Delta\delta\\ \Delta V\end{bmatrix} \quad . \tag{2.1}$$

Where ΔP and ΔQ represent the incremental change of active and reactive power injection respectively. $\Delta V$and $\Delta\delta$ represent the incremental change of bus angles and voltages respectively. The Jacobian matrix is represented by J.

Vector equation in (2.1) is reduced by letting $\Delta P = 0$:

$$\Delta Q = J_R \Delta V \tag{2.2}$$

where $J_R = [J_{22} - J_{21}J_{11}^{-1}J_{12}]$

Therefore the reduced matrix J_R represents the linear relationship between the incremental changes in bus voltages and incremental change in injecting reactive power. The voltage

stability of the power system can be determined by the sign of the eigenvalues of the J_R matrix. Since the J_R matrix can be taken as a symmetrical matrix the eigenvalues of J_R matrix is close to be purely real. The positive eigenvalues represent a stable system in terms of voltage stability. Whereas at least one negative eigenvalue represents an unstable system in terms of voltage stability. A zero eigenvalue represent a marginally stable system. Furthermore, the eigenvalues of J_R matrix provide the proximity to voltage instability, the smaller the positive magnitude of the eigenvalues closer the system is to voltage instability [47].

2.4.3.3. Sensitivity Analysis Method

The V-Q sensitivity of a particular system bus can be determined by the slope of the V-Q curve at a particular operating point. The system is said to be stable when the sensitivity parameter is positive. When the sensitivity parameter becomes zero, the system is said to be at the voltage collapse point. A negative sensitivity indicates that the system is unstable. The proximity to voltage collapse can be estimated by the singularity vector of the power flow Jacobian. The singular vector can also be used to identify the most critical buses of the system which are more prone to voltage instability [38],[48].

2.4.3.4. Thévenin Equivalent Method

When determining the voltage stability of a particular bus, the rest of the network can be represented as a Thévenin equivalent. This reduces the network to a two bus system, which can be easily analyzed as described in Section 1.2 Some of the VSIs are based on the Thévenin equivalent. Thévenin equivalent can be obtained both in online and off-line methods. Figure 2.4 illustrates the concept of reducing a meshed network to a Thévenin equivalent. If the load demand at the ith bus is $S_i = P_i + jQ_i$ the current flow can be expressed in terms of the Thévenin equivalent parameters.

$$\frac{P_i + jQ_i}{\bar{V}_i} = \bar{I}_i = \left(\frac{\bar{E}_{Th} - \bar{V}_i}{\overline{Z_{th_i}}}\right)^* \tag{2.3}$$

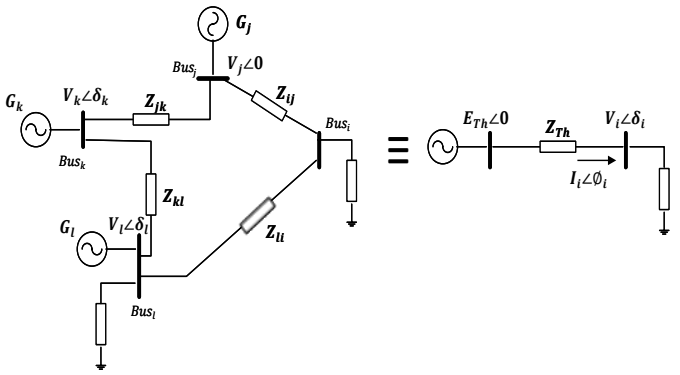


Figure 2.4. Thévenin equivalent of a meshed network

It can be shown that the conditions in (2.4) and (2.5) hold at the voltage collapse point;

$$(\bar{E}_{Th} - \bar{V}_i)^* = \bar{V}_i \tag{2.4}$$

$$\bar{Z}_{Th} = \bar{Z}_{Load_i} \tag{2.5}$$

Hence these conditions can be used to assess the proximity to voltage collapse of a meshed power system if it can be represented using the Thévenin equivalent. The methods to evaluate the Thévenin equivalent will be discussed in Section 2.5

2.4.3.5. Voltage Stability Index Method

Even though there are various types of VSIs proposed in literature [7],[10]-[12] ,[25]-[32] all these VSIs are based on theoretical conditions at the point of voltage collapse as described in Section 1.3. VSIs provide a quantitative parameter which reflects the proximity to voltage collapse. However, characteristic and sensitivity of each VSI is unique. VSIs have been used for online voltage stability assessment applications due to the simplicity in terms of calculation.

2.4.4 Thévenin Equivalent

Consider two sub systems (a and b) connected as shown in Figure 2.5 The impact of one subsystem on another can be evaluated using a Thévenin equivalent circuit consisting of

an ideal voltage source and an impedance connected in series. The validity of the equivalent circuit will persist, until the network topology remains unchanged and not subjected to conditions such as generators reaching their excitation limits or changing of transformer tap positions [49],[50]. Thévenin equivalent of a system can be determined in both online and offline methods.

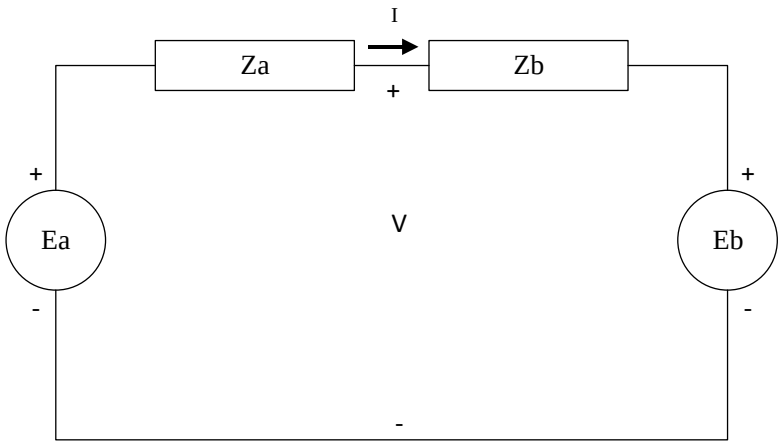


Figure 2.5. Interconnected sub systems a and b

2.4.4.1. Off-line Thévenin Equivalent

A multiport Thévenin equivalent circuit can be represented as a coupled single port circuit [24]. Current and voltage relationship of the above multi-port system can be represented as given in (2.6).

$$\begin{bmatrix} -\bar{I}_L \\ 0 \\ \bar{I}_G \end{bmatrix} = \begin{bmatrix} \overline{Y_{LL}} & \overline{Y_{LT}} & \overline{Y_{LG}} \\ \overline{Y_{TL}} & \overline{Y_{TT}} & \overline{Y_{TG}} \\ \overline{Y_{GL}} & \overline{Y_{GT}} & \overline{Y_{GG}} \end{bmatrix} \begin{bmatrix} \bar{V}_L \\ \overline{V_T} \\ \overline{V_G} \end{bmatrix} \tag{2.6}$$

where *V* and *I* are the voltage and current vectors respectively and the subscripts *L,T* and *G* represent load bus, tie bus and generator bus respectively. The admittance matrix *Y* is divided into nine sub matrices. The bus voltage vector V_L can be obtained as

$$\bar{V_L} = \overline{Z_{LL}}(\overline{Y_{LT}}\bar{Y}_{TT}^{-1}\bar{Y}_{TG} - \bar{Y}_{LG})\overline{V_G} - \bar{Z}_{LL}\bar{I}_L = \bar{E}_{Th} - \bar{Z}_{LL}\bar{I_L} \tag{2.7}$$

where $\overline{Z_{LL}} = (\overline{Y_{LL}} - \overline{Y_{LT}}\bar{Y}_{TT}^{-1}\bar{Y}_{TL})^{-1}$

In order to represent the coupling impact of other loads on the considered i^{th} bus, (2.7) has been modified as below. The Thévenin equivalent voltage and impedance of bus *i* can be obtained as

$$\bar{E}_{Th} = \bar{V}_{Li} + \bar{Z}_{Th}\bar{I}_{Li} + \bar{E}_{coupledi} \tag{2.8}$$

where $\bar{E}_{coupled_i} = \sum_{i=0, i\neq j}^{N} \bar{Z}_{LLji}\bar{I}_{Li}$

This equivalent circuit remains valid as long as the topology and the loads remains unchanged.

2.4.4.2. On-line Thévenin Equivalent

Online or nodal measurement based Thévenin equivalent determination is based on the assumption that the system is at steady state or its variation is slow enough not to affect the nodal measurements used for a calculation step [50]. The voltage current relationship at the nodal point can be represented as

$$\bar{V} = \bar{E}_{Th} + \bar{Z}_{Th}\bar{I} \tag{2.9}$$

Therefore, two distinct measurement pairs $(\bar{V}_1, \bar{I}_1)$ and $(\bar{V}_2, \bar{I}_2)$ can be used to calculate the Thévenin equivalent based on Tellegen's theory [49]:

$$\bar{Z}_{Th} = \frac{\bar{V}_1 - \bar{V}_2}{\bar{I}_1 - \bar{I}_2} \tag{2.10}$$

$$\bar{E}_{Th} = \frac{\bar{I}_1\bar{V}_2 - \bar{I}_2\bar{V}_1}{\bar{I}_1 - \bar{I}_2} \quad . \tag{2.11}$$

2.5 Compatible Voltage Stability Indices

Eleven candidate VSIs are considered in this thesis for the purpose of evaluating each VSI characteristic and combining them to obtain a composite VSI which has more LM prediction accuracy than individual VSIs. These candidate VSIs have been derived based on diverse principles and are calculable using synchrophasor measurements. All these VSIs are presented in the Subsections 2.5.1-2.5.11.

2.5.1 Voltage Stability Load Index (VSLI)

The voltage stability load index is derived from the voltage equation used for power flow in [25]. Consider two nodes of the power system as shown in Figure 2.6 where, P_i, Q_i, P_j and Q_j are the active and reactive power injected at the nodes *i* and *j* respectively.

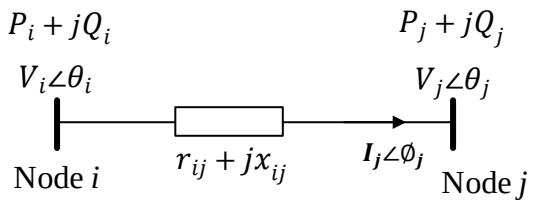


Figure 2.6. Two nodes of a power system

The current flowing through the transmission line can be expressed using the defined parameters as

$$\left|\bar{I}_J\right|^2 = \frac{P_j^2 + Q_j^2}{\bar{V}_j^2} \quad . \tag{2.12}$$

The active and reactive power flows at node *j* can be denoted as in (2.13) and (2.14):

$$P_j = P_i - P_{loss} \tag{2.13}$$

$$Q_j = Q_i - Q_{loss.} \tag{2.14}$$

The active and reactive power losses due to complex power flow through the transmission line can be computed as in (2.15) and (2.16):

$$P_{loss} = \left(\frac{P_j^2 + Q_j^2}{V_j^2}\right) r_{ij} \tag{2.15}$$

$$Q_{loss} = \left(\frac{P_j^2 + Q_j^2}{V_j^2}\right) x_{ij} \tag{2.16}$$

The current at node *i* can be expressed as in (2.17).

$$I_i{}^2 = \frac{P_i^2 + Q_i^2}{V_i^2} \tag{2.17}$$

$$I_i{}^2 = \frac{\left[P_j + \left(\frac{P_j^2 + Q_j^2}{V_j^2}\right) r_{ij}\right]^2 + \left[Q_j + \left(\frac{P_j^2 + Q_j^2}{V_j^2}\right) x_{ij}\right]^2}{V_i^2} \tag{2.18}$$

By substituting to (2.18) from (2.12), a quadratic equation of V_j^2 can be obtained:

$$V_j^4 + \left[2\left(P_j r_{ij} + Q_j x_{ij}\right) - V_i^2\right] V_j^2 + \left(P_j^2 + Q_j^2\right)\left(r_{ij}{}^2 + x_{ij}{}^2\right) = 0 \tag{2.19}$$

The quadratic equation in (2.19) have real roots if $b^2 - 4ac \geq 0$, i.e. :

$$8P_jQ_jr_{ij}x_{ij} - 4V_i^2\left(P_jr_{ij} + Q_jx_{ij}\right) + V_i^4 - 4\left(P_j^2{x_{ij}}^2 + Q_j^2{r_{ij}}^2\right) \geq 0 \quad (2.20)$$

The condition in (2.20) can be further simplified to

$$\frac{4\left[V_i^2\left(P_jr_{ij} + Q_jx_{ij}\right) + (P_jx_{ij} - Q_jr_{ij})^2\right]}{V_i^4} \leq 1 \quad (2.21)$$

The left hand side of the inequality in (2.21) is defined as the Voltage Stability Load Index (VSLI), and the VSLI should be less than or equal to 1 for a feasible solution for the voltage magnitude to exist.

$$\text{VSLI} = \frac{4\left[V_i^2\left(P_jr_{ij} + Q_jx_{ij}\right) + (P_jx_{ij} - Q_jr_{ij})^2\right]}{V_i^4} \quad (2.22)$$

Since

$$V_iV_j\cos\left(\theta_i - \theta_j\right) - V_j^2 = P_jr_{ij} + Q_jx_{ij} \quad (2.23)$$

and

$$V_iV_j\sin\left(\theta_i - \theta_j\right) = P_jx_{ij} + Q_jr_{ij} \quad . \quad (2.24)$$

The index VLSI can be expressed in terms of the voltage phasors at buses *i* and *j* as

$$\text{VSLI} = \frac{4\left[V_iV_j\cos\left(\theta_i - \theta_j\right) - V_j^2\cos\left(\theta_i - \theta_j\right)^2\right]}{V_i^2} \quad . \quad (2.25)$$

The derived voltage stability index is applied to the Thévenin equivalent circuit as shown in Figure 1.1. (a) Radial power system, (b) Variation of real power with load impedance with respect to a load bus of interest.

$$\text{VSLI}_i = \frac{4[V_i E_{Th}\cos(\delta_i) - V_i^2\cos(\delta_i)^2]}{E_{Th}^2} \tag{2.26}$$

where $\theta_i - \theta_j = \delta_i$

Hence the VSLI corresponding to the concerned load bus in the power system should be less than 1 in order to be able to maintain the voltage stability. If the VSLI becomes greater than 1, then the voltage of the concerned load bus theoretically becomes imaginary and bus voltages will collapse.

2.5.2 Voltage Collapse Proximity Index-1 (VCPI_1)

This index is derived considering the maximum power transfer limit among two buses through a transmission line [26]. When the rest of the system is represented using a Thévenin equivalent as shown in Figure 1.1. (a) with respect to a load bus of interest, then the power flow into the i^{th} bus can be expressed as

$$\bar{S}_i = \bar{V}_i \bar{I}_i^{\,*} = \bar{V}_i \left(\frac{\bar{E}_{Th} - \bar{V}_i}{\bar{Z}_{Th}} \right)^* \quad . \tag{2.27}$$

By considering the derivative of (2.27), the condition at the maximum power transfer can be obtained as

$$\frac{d\bar{S}}{d\bar{v}_i} = (\bar{E}_{Th} - \bar{V}_i)^* - \bar{V}_i \quad . \tag{2.28}$$

Therefore at the maximum power transfer, the voltage magnitude of the Thévenin source (E_{Th}) becomes twice the real part of the voltage (V_i) at the load bus of interest. Considering this fact, VCPI_1 index is defined as:

$$VCPI\_1_i = V_i cos\delta_i - 0.5E_{Th} \quad . \tag{2.29}$$

Hence when the system approaches the voltage collapse point, the VCPI_1 index will approach to zero

2.5.3 Voltage Stability Load Bus Index (VSLBI)

Voltage Stability Load Bus Index is proposed in [27] by considering the Thévenin equivalent circuit as seen from the load bus of interest. In this equivalent circuit, voltage instability occurs at the maximum power transfer point. Under maximum power flow, the voltage drop across the Thévenin equivalent impedance $\Delta\bar{V}$ is equal to the load bus voltage:

$$\bar{V}_i = \Delta\bar{V} = \bar{E}_{Th} - \bar{V}_i \tag{2.30}$$

Therefore, to assess the proximity to the voltage collapse point in the presence of a constant power load, the VSLBI is defined as

$$VSLBI_i = \frac{|\bar{E}_{Th} - \bar{V}_i|}{|\bar{V}_i|} \quad . \tag{2.31}$$

Therefore when the system approaches the point of voltage collapse, the $VSLBI$ index will approach 1.

2.5.4 S Difference Criterion (SDC)

This voltage stability index is derived considering the rate of change of apparent power [28]. The rate of change is calculated using two different pairs of measurements taken at the load bus of interest. At proximity of the voltage collapse point, transmission line losses start to grow rapidly. Hence the rate of change of apparent power start to reduce. This can be demonstrated by considering two consecutive voltage and current measurements at a

bus. The difference of two consecutive voltage and current measurements at the i^{th} bus can be represented respectively as

$$\Delta\bar{V}_i^{n+1} = \bar{V}_i^{n+1} - \bar{V}_i^n \tag{2.32}$$

$$\Delta\bar{I}_i^{n+1} = \bar{I}_i^{n+1} - \bar{I}_i^n \quad . \tag{2.33}$$

The complex power demand of the i^{th} load bus of interest at the n^{th} measurement can be expressed as

$$\bar{S}_i^n = \bar{V}_i^n(\bar{I}_i^n)^* \quad . \tag{2.34}$$

Similarly, the complex power demand of the i^{th} load bus of interest at the n+1th measurement can be represented as

$$\bar{S}_i^{n+1} = \bar{S}_i^n + \Delta\bar{S}_i^{n+1} = (\bar{V}_i^n + \Delta\bar{V}_i^{n+1})(\bar{I}_i^n + \Delta\bar{I}_i^{n+1})^* \tag{2.35}$$

The term $\Delta\bar{V}_i^{n+1} \cdot \Delta\bar{I}_i^{n+1^*}$ can be neglected when the differences are small, and hence can be rewritten as

$$\bar{S}_i^{n+1} \approx \bar{V}_i^n I_i^{n^*} + \bar{V}_i^n \Delta\bar{I}_i^{n+1^*} + \Delta\bar{V}_i^{n+1}\bar{I}_i^{n^*} \quad . \tag{2.36}$$

The change of power flow of the line can be expressed as

$$\Delta\bar{S}_i^{n+1} = \bar{V}_i^n \Delta\bar{I}_i^{n+1^*} + \Delta\bar{V}_i^{n+1}\bar{I}_i^{n^*} \quad . \tag{2.37}$$

At the point of voltage collapse, the additional complex power injected from the source end will be lost as transmission losses, and therefore ΔS_i^{n+1} becomes 0. Hence the SDC is defined as

$$SDC = \left| 1 + \frac{\Delta \bar{V}_j^{n+1} \bar{I}_i^{n^*}}{\bar{V}_j^n \Delta \bar{I}_i^{n+1^*}} \right| \quad . \tag{2.38}$$

Therefore the SDC becomes 0 at the voltage collapse point.

2.5.5 Impedance matching Stability Index (ISI)

The magnitude of the load impedance of the i^{th} load bus is equal to the magnitude of the Thévenin impedance at i^{th} bus. At the maximum power transfer limit of a transmission line:

$$|\bar{Z}_i| = |\bar{Z}_{Th}| \tag{2.39}$$

By applying the difference form of Tellegen's theorem to the i^{th} bus of interest, the expression in (2.39) can be obtained.

$$\bar{I}_i^* \Delta \bar{V}_i - \bar{V}_i \Delta \bar{I}_i^* = 0 \tag{2.40}$$

Where those parameters have the same definitions as in (2.32) and (2.33). The apparent load impedance at the i^{th} bus can be obtained from measurements as:

$$\overline{Z_\iota} = \frac{\bar{V}_i}{\bar{I}_\iota} \quad . \tag{2.41}$$

The Thévenin impedance of the adjoined network with respect to the i^{th} bus can be expressed in-terms of the Thévenin equivalent voltage

$$\overline{Z_{Th}} = \frac{\bar{E}_{Th} - \bar{V}_i}{\bar{I}_i} \quad . \tag{2.42}$$

Substituting to (2.42) from (2.40)

$$\overline{Z_{Th}}^* = \frac{(\bar{E}_{Th} - \bar{V}_i)^*}{\bar{V}_i \Delta \bar{I}_i^*} \Delta \bar{V}_i \tag{2.43}$$

At the maximum power transfer

$$\bar{V}_i = (\bar{E}_{Th} - \bar{V}_i)^* \tag{2.44}$$

Therefore, from (2.43) and (2.44), at the maximum power transfer,

$$\bar{Z}_{Th} = \frac{\Delta \bar{V}_i^*}{\Delta \bar{I}_i} \tag{2.45}$$

At the normal loading conditions (2.43) holds,

$$|\bar{Z}_{Th}| < |\bar{Z}_i| \quad . \tag{2.46}$$

At the proximity of voltage instability, the voltage difference between the two impedances approaches to zero and impedance magnitudes becomes equal as given in (2.39) at the point of collapse.

Hence a simple normalized impedance based voltage stability index can be defined as

$$ISI = \frac{|\bar{Z}_i| - |\bar{Z}_{Th}|}{|\bar{Z}_i|} \quad . \tag{2.47}$$

Substituting to (2.47) from (2.41) and (2.45), ISI can be expressed as

$$ISI = 1 - \frac{|\Delta \bar{V}_i \bar{I}_i|}{|\bar{V}_i \Delta \bar{I}_i|} \quad . \tag{2.48}$$

Hence, when the system reaches the voltage collapse point, ISI will reach 0. This index is proposed in [31].

2.5.6 Voltage Stability Index (VSI_b)

Voltage Stability Index (VSI_b) is proposed in [29] and it is based on the same concept of SDC which identifies the proximity to voltage collapse by the rate of change of complex power. However VSI_b is defined as the SDC raised to the power of an index α (>1), as given in (2.48).

$$\mathrm{VSI}_b = SDC^{\alpha} \tag{2.49}$$

$$\mathrm{VSI}_b = \left(\left| 1 + \frac{\Delta \bar{V}_j^{n+1} \bar{I}_i^{n*}}{\bar{V}_j^{n} \Delta \bar{I}_i^{n+1*}} \right| \right)^{\alpha} \tag{2.50}$$

As shown in Figure 2.7 when the value of α increases, the variation of VSI_b index with increasing load power becomes linear. However, the exact value of α that linearizes the characteristics depends on the system.

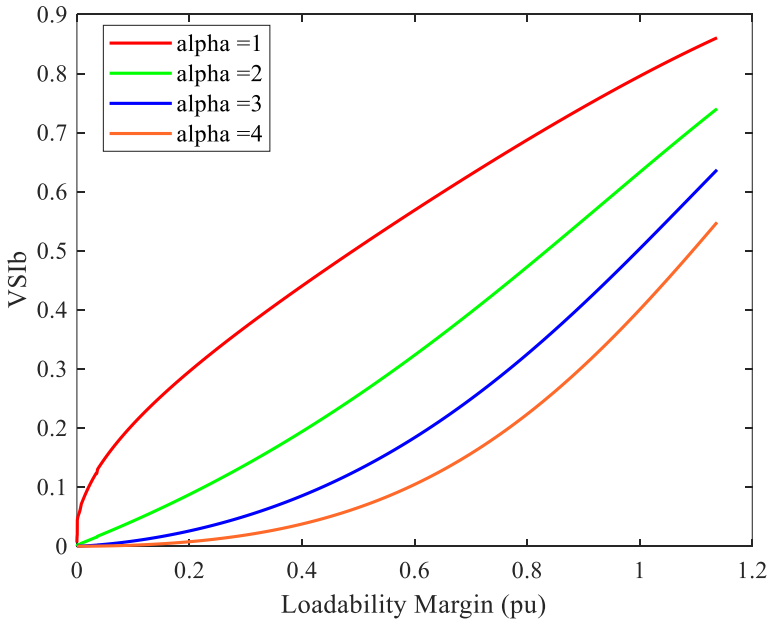


Figure 2.7. VSIb characteristic under different α

2.5.7 Simplified Voltage Stability Index (SVSI)

The equivalent circuit for this index is based on the Relative Electrical Distance (RED) rather than Thévenin equivalent circuit and it is proposed in [11]. The generator bus located at the minimum RED from the load bus of interest is identified as the voltage source equivalent to the rest of the network sources. RED can be obtained for a power system from its nodal equation. For a network with combination of load (denoted by subscript L) and generator (denoted by subscript G) buses, the nodal equation can be written as

$$\begin{bmatrix} \bar{I}_G \\ \bar{I}_L \end{bmatrix} = \begin{bmatrix} \bar{Y}_{GG} & \bar{Y}_{GL} \\ \bar{Y}_{LG} & \bar{Y}_{LL} \end{bmatrix} \begin{bmatrix} \bar{V}_G \\ \bar{V}_L \end{bmatrix} \tag{2.51}$$

where $[\bar{Y}_{GG}], [\bar{Y}_{GL}], [\bar{Y}_{LG}]$ and$[\bar{Y}_{LL}]$ are the corresponding partitions of admittance matrix when it is categorised as generator and load buses. The relationship between voltage and current in (2.51) can be rearranged as

$$\begin{bmatrix} \bar{V}_L \\ \bar{I}_G \end{bmatrix} = \begin{bmatrix} \bar{Z}_{LL} & \bar{F}_{LG} \\ \bar{K}_{GL} & \bar{Y}_{GG} \end{bmatrix} \begin{bmatrix} \bar{I}_L \\ \bar{V}_G \end{bmatrix} \quad . \tag{2.52}$$

In (2.52) $\bar{F}_{LG} = -[\bar{Y}_{LL}]^{-1}[\bar{Y}_{LG}]$ denotes the relationship between the load and source voltages. RED of all the loads buses in the network can be obtained as

$$RED_{LG} = [A] - abs[\bar{F}_{LG}] \tag{2.53}$$

where A is a (N-N_G) × N_G unity matrix, N is the total number of buses of the network and N_G is the number of generator buses. Hence the rows of RED_{LG} represents the load buses and the columns represents the generator buses. The respective column which contain the

minimum value of a particular row represent the generator bus with the lowest RED for that particular load bus.

Once the nearest generator bus to a particular load bus is identified, the voltage drop between these two buses is calculated and it is assumed to be the voltage drop between the Thévenin impedance ($\Delta\bar{V}_i$):

$$\Delta\bar{V}_i \approx \bar{V}_{g(minRED)} - \bar{V}_i \quad . \tag{2.54}$$

Similar to VSLBI, SVSI considers the voltage difference ($\Delta\bar{V}_i$) with respect to the load bus of interest, the compensation factor β is included to increase the sensitivity of the index with respect to the other buses of the system. SVSI for the i^{th} bus of interest is defined as

$$SVSI_i = \frac{|\Delta\bar{V}_i|}{\beta V_i} \quad . \tag{2.55}$$

where $\beta = 1 - (\max(|V_m| - |V_l|))^2$.

In this expression, V_m is the maximum bus voltage in the network, and V_l is the minimum bus voltage in the network.

2.5.8 L-Index (L)

Considering (2.52), the voltage of the i^{th} load bus ($\bar{\mathrm{V}}_i$) can be expressed as

$$\bar{\mathrm{V}}_{\mathbf{i}} = \sum_{i=1}^{N_L} \bar{Z}_{LL\,ji}.\,\bar{\mathrm{I}}_{\mathbf{i}} + \sum_{j=1}^{N_G} F_{LG\,ji}.\,\bar{\mathrm{V}}_{\mathbf{j}} \quad . \tag{2.56}$$

.

This can be expressed in the form

$$\overline{\mathrm{V}}_{\mathbf{i}}^{\,\mathbf{2}} + \overline{\mathrm{V}}_{0\mathrm{i}}\overline{\mathrm{V}}_{\mathbf{i}}^{\,*} = \frac{\bar{S}_i^{\,*}}{\bar{Y}_{ii}} \tag{2.57}$$

with the substitution $$\overline{\mathrm{V}}_{0\mathrm{i}} = -\sum_{j=1}^{N_G} F_{LG\,ji}.\overline{\mathrm{V}}_{\mathrm{j}} \tag{2.58}$$

L index is defined in [30] considering the voltage at the sending end ($\overline{\mathrm{V}}_{0\mathrm{i}}$) and the voltage at the receiving end load ($\overline{\mathrm{V}}_{\mathrm{i}}$) as

$$L_i = \left|1 - \frac{\overline{\mathrm{V}}_{0\mathrm{i}}}{\overline{\mathrm{V}}_{\mathrm{i}}}\right| \quad . \tag{2.59}$$

The L-index considers an equivalent voltage source obtained by taking the summation of the matrix multiplication of F_{ji} matrix and generator voltages, where F_{ji} represents the admittance relationship between generators (j=1, 2…N_G) and loads (i=1, 2…N_L). Combining (2.58) and (2.59), the L-index can be expressed as

$$L_i = \left|1 - \frac{\sum_{j=1}^{N_G} F_{ji}.\overline{\mathrm{V}}_{\mathrm{j}}}{\overline{\mathrm{V}}_{\mathrm{i}}}\right| \quad . \tag{2.60}$$

where $\overline{\mathrm{V}}_{\mathrm{j}}$ denotes the voltages of the jth generator bus and $\overline{\mathrm{V}}_{\mathbf{i}}$ denotes the voltage of the load bus of interest.

2.5.9 Wide Area Loss Index (WALI)

Wide Area Loss Index (WALI) is based on the power balance principle. In this method a particular region of interest (ROI) is considered. In order to satisfy the power balance, the summation of total power imports into the ROI and the total power generation within this ROI should be equal to the summation of the total power consumption and the losses with

in the ROI. Criticality of a power system with respect to a voltage collapse can be clearly visualize by observing the change of system variables, hence for WALI, the power balance is computed for the changes between two consecutive measurement times as:

$$\overline{\Delta S_{GEN}^{n+1}} = \overline{\Delta S_{LOAD}^{n+1}} + \overline{\Delta S_{LOSS}^{n+1}} \tag{2.61}$$

where

$\overline{\Delta S_{GEN}^{n+1}} = \overline{S_{GEN}^{n+1}} - \overline{S_{GEN}^{n}}$

$\overline{\Delta S_{LOAD}^{n+1}} = \overline{S_{LOAD}^{n+1}} - \overline{S_{LOAD}^{n}}$

$\overline{\Delta S_{LOSS}^{n+1}} = \overline{S_{LOSS}^{n+1}} - \overline{S_{LOSS}^{n}}$

In the vicinity of a voltage collapse, the additional generation will be lost within the ROI as the losses in the transmission line and other equipment rather supplying to the loads. Therefore, in the vicinity of voltage collapse

$$\overline{\Delta S_{GEN}^{n+1}} \approx \overline{\Delta S_{LOSS}^{n+1}} \tag{2.62}$$

Based on (2.61) WALI is defined in [7]as

$$WALI = \frac{\left|\overline{\Delta S_{GEN}^{n+1}} - \overline{\Delta S_{LOSS}^{n+1}}\right|}{\max(\left|\overline{\Delta S_{GEN}^{n+1}}\right|, \left|\overline{\Delta S_{LOSS}^{n+1}}\right|)} \quad . \tag{2.63}$$

In a healthy system WALI should be around 1, but at the vicinity of a voltage collapse in the ROI it becomes significantly smaller than 1.

2.5.10 Voltage Stability Margin (VSM)

Most VSI characteristics are not linear, therefore there are some situations where VSI characteristic will not show significant changes for a wide range of loads. A linear index can solve this issue and can be used to accurately identify the proximity to a voltage collapse. VSM is a linear index which is derived by considering the Thévenin equivalent at the i^{th} load bus of interest. The active and reactive power transfer through the Thévenin impedance can be computed as

$$P_i^2 = \frac{E_{Th}^2 V_i^2 sin^2\theta_i}{x^2} \tag{2.64}$$

$$Q_i^2 = \frac{E_{Th}^2 V_i^2 cos^2\theta_i - 2E_{Th}V_i^3 cos\theta_i + V_i^4}{x^2} \quad . \tag{2.65}$$

Assuming that the power factor doesn't change during the load increment, and substituting for V_i^2 in (2.64) and (2.65) from its real and imaginary parts as

$$V_i^2 = {V_{i_{re}}}^2 + {V_{i_{im}}}^2 \quad . \tag{2.66}$$

.

The ratio between (2.64) and (2.65) can be expressed as

$$\frac{P_i}{Q_i} = \frac{-E_{Th}V_{i_{im}}}{E_{Th}V_{i_{re}} - ({V_{i_{re}}}^2 + {V_{i_{im}}}^2)} \quad . \tag{2.67}$$

The expression in (2.67) can be rearranged to obtain (2.68).

$${V_{i_{re}}}^2 - E_{Th}V_{i_{re}} + {V_{i_{im}}}^2 - E_{Th}V_{i_{im}}\left(\frac{Q_i}{P_i}\right) = 0 \quad . \tag{2.68}$$

$$\left(V_{i_{re}} - \frac{E_{Th}}{2}\right)^2 + \left(V_{i_{im}} - \frac{E_{Th}}{2}tan\theta\right)^2 - (1 + tan^2\theta)\frac{E_{Th}^2}{4} = 0 \quad (2.69)$$

$$\left(V_{i_{re}} - \frac{E_{Th}}{2}\right)^2 + \left(V_{i_{im}} - \frac{E_{Th}}{2}tan\theta\right)^2 = \left(\frac{E_{Th}}{2cos\theta}\right)^2 \quad (2.70)$$

The equation in (2.70) corresponds to the locus of a circle with the radius $\frac{E_{Th}}{2cos\theta}$ with the center at $(\frac{E_{Th}}{2}, \frac{E_{Th}}{2}tan\theta)$. Figure 2.8 is recreated from [12] to illustrate the locus of the voltage at different load power factors (p.f = 0.9,1 and -0.9).

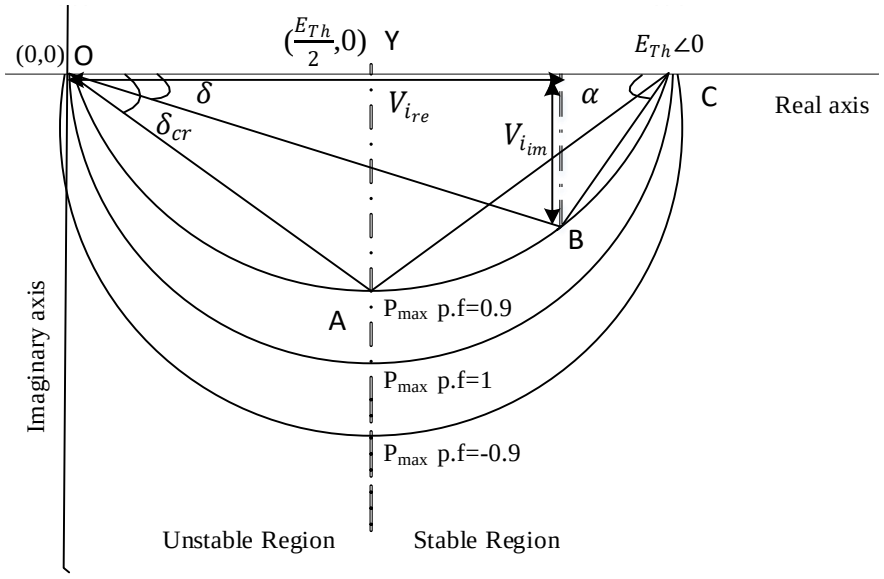


Figure 2.8. Voltage locus at load power factor 0.9,1 and -0.9 [12]

Considering the voltage phasor diagram shown in Figure 2.9, δ can be determined using the load power factor angle (θ) and the transmission line impedance angle ($\emptyset$) as given in (2.71).

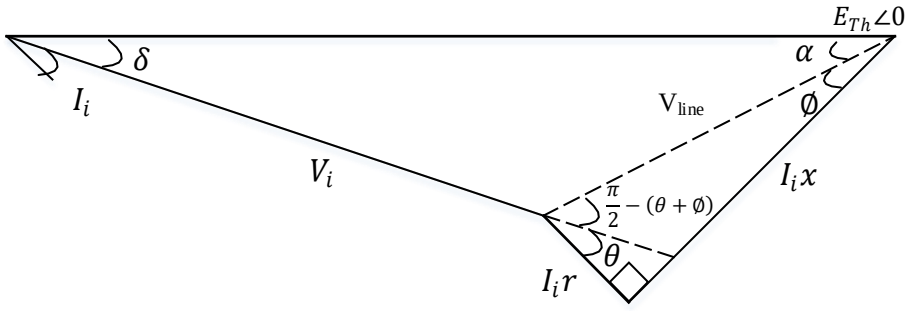


Figure 2.9. Voltage phasor of a 2 bus system

At the maximum power transfer $\delta = \delta_{cr} = \alpha$. Therefore

$$\frac{\pi}{2} - (\theta + \emptyset) = 2\delta_{cr} \quad . \tag{2.71}$$

For a particular load bus of interest *i*, the critical loadability angle (δ_{cri}) is defined as

$$\delta_{cr_i} = \frac{\pi}{4} - \frac{(\theta_i + \emptyset_i)}{2} \quad . \tag{2.72}$$

where

$$\theta_i = tan^{-1}\left(\frac{Q_i}{P_i}\right) \tag{2.73}$$

$$\emptyset_i = tan^{-1}\left(\frac{Re(Z_{Th})}{Im(Z_{Th})}\right) \quad . \tag{2.74}$$

Therefore, the critical voltage at the load bus of interest *i* can be given as

$$V_{cr_i} = \frac{E_{Th}}{2cos\delta_{cr_i}} \quad . \tag{2.75}$$

The M-index or the VSM is defined in [12] as the power margin relative to the critical power as:

$$VSM_i = 1 - \frac{P}{P_{max}} \quad . \tag{2.76}$$

Since the load power is proportional to the imaginary part of the voltage ($V_{i_{im}}$), VSM_i can be represented as

$$VSM_i = 1 - \frac{\overline{BX}}{\overline{AY}} \quad . \tag{2.77}$$

Considering the area of the triangles in the circumscribed, BX and AY can be represented as:

$$\overline{BX} = \frac{\overline{OB}.\overline{BC}.sin\beta}{\overline{OC}} \tag{2.78}$$

$$\overline{AY} = \frac{\overline{OA}.\overline{AC}.sin\beta}{\overline{OC}} \quad . \tag{2.79}$$

Therefore VSM can be reformulated as

$$VSM_i = 1 - \frac{\overline{OB}.\overline{BC}}{\overline{OA}.\overline{AC}}. \tag{2.80}$$

$$VSM_i = 1 - \frac{|\bar{V}_i||\bar{E}_{Th} - \bar{V}_i|}{\left|\bar{V}_{cr_i}\right|^2} \quad . \tag{2.81}$$

2.5.11 Reactive Power Reserve Index (ERPR)

Reactive power reserve is one of the most important factors which govern the degree of power system voltage stability. Stressed power systems without adequate reactive power are more prone to voltage instabilities. Available reactive power reserve relative to the total

reactive power capacity of the grid connected generators can be used as an indicator to predict voltage instability. Therefore the reactive power reserve index, ERPR, is formulated as given in (2.82) in terms of the ratio between the summation of all injected reactive power (Q_{op}) and the summation of the reactive power capacities (Q_{lim}) of the generators that are connected to the system [25].

$$ERPR = 1 - \frac{\sum_{i=1}^{N_G} Q_{op,i}}{\sum_{i=1}^{N_G} Q_{lim,i}} \quad . \tag{2.82}$$

2.6 Application of Machine Learning for Voltage Stability Assessment

2.6.1 Machine learning

The process of converting experience to expertise or knowledge is known as learning. People are prone to make mistakes when trying to generate a relationship between multiple features and analysis. Some of these problems can be solved using Machine Learning techniques with higher accuracy and robustness. Machine learning algorithms are trained using different kinds of data sets such as continuous, binary and categorical. If the data instances contain labels (the corresponding output) then it is called supervised learning. When the data instances don't contain labels, unsupervised learning or clustering can be applied to discover useful unknown relationships. Semi-supervised learning is used when a small amount of data instances contain labels while a large amount of data does not. There is another category of learning which is known as reinforced learning where an

algorithm learn the rules by acting according to the feedback received from the environment. Every action has an impact on the environment, hence a feedback is provided to guide the learning to gain the maximum yield [52]. Examples for each learning category are tabulated in Table 2-1. However learners such as artificial neural network and deep neural network belong to all three learning categories [60],[61].

TABLE 2-1 MACHINE LEARNING ALGORITHMS FOR DIFFERENT LEARNING CATEGORIES [60], [61]

Supervised learning	Decision Tree
	K-Nearest Neighbours
	Multinomial Naïve Bayesian
	Rule System
	Discriminative Multinomial Naïve Bayesian
	Support Vector Machine
Unsupervised learning	Latent Dirichlet Allocation
	K-mean
	Hierarchical Agglomerative Clustering
	Bi-term Topic Modelling
Semi- supervised learning	Expectation Maximization
	Self-Training
	Active Learning
	Random Subspace Method for Co-training
	Relevant Random Subspace Method for Co-training
Reinforced learning	Q Learning
	Temporal Difference Learning
	Deep Q Network

Supervised learning can be categorised under regression analysis and classification depending on the desired output of a problem. Regression analysis is a curve fitting techniques where a MLM is trained to obtain a value as the output for the corresponding

input feature/s. In machine learning classification, a MLM is trained to classify the output into one of the provided classification classes, depending on the values of input features. A suitable machine learning algorithm must be selected based on the application.

These trained MLMs, especially the regression models, could be under-fit or over-fit depending on the training data set. Therefore, the generally machine learning data set is randomly split for the purpose of training and testing. After training the MLM using the training data set, the testing data set is used to obtain the output from the trained MLM. Then the results can be evaluated using a suitable performance measure such as:

- Root Mean Square Error (RMSE)
- Prediction correlation (R^2)

A low value of RMSE and a prediction correlation close to ±1would indicate an accurate model.

2.6.2 Application for assessing on-line voltage stability

Application of computational intelligence and data mining for voltage stability assessment has been investigated throughout the recent past years. Online Long Term Voltage Stability (LTVS) assessment using an Artificial Neural Network (ANN) with voltage phasors is proposed in [16]. A decision Tree (DT) based classification approaches are proposed in [13] and [14] for assessing system voltage security, while [15] used a DT classification for LTVS assessment under small disturbances. In order to optimize the prediction accuracy and speed, Extreme Learning Machine (ELM) technique is introduced for Short Term Voltage Stability (STVS) assessment in [17], [18]. Voltage stability margin prediction using Convolutional Neural Network (CNN) is proposed in [19]. Support Vector Machine

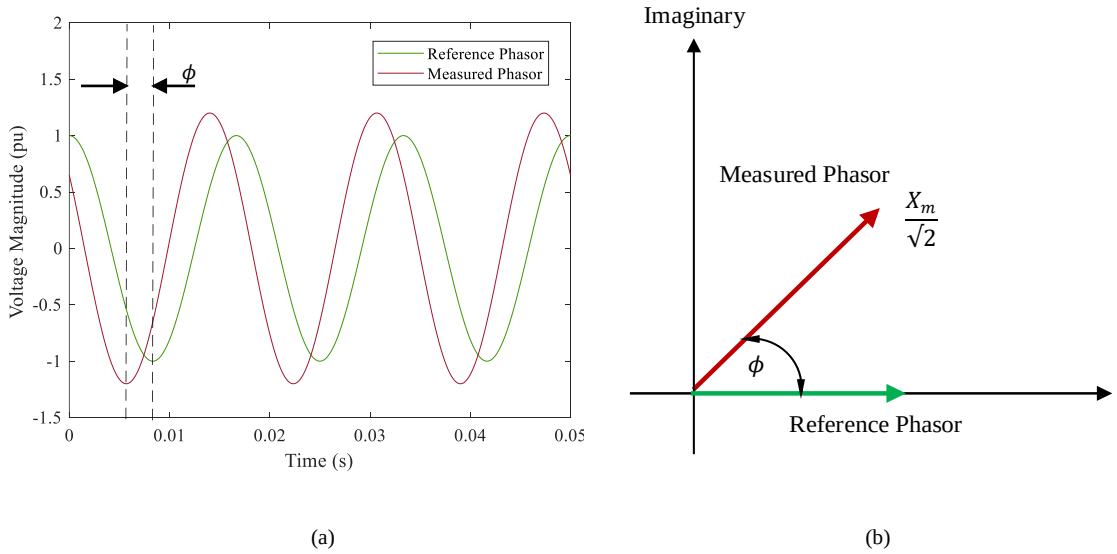
(SVM) based regression has been presented in [20], [21] to assess LTVS. These work has shown a good potential, however, are not widely applied in practice. There is a rapid progress in machine learning technology and data science, while the traditional drawbacks of machine learning approaches such as computational power requirements and the engineering effort required for training data generation when applying to large power systems are becoming more manageable with cloud computing [22], [23] and automation tools for power system analysis software.

2.7 Wide Area Monitoring Systems

Wide Area Monitoring Systems (WAMS) are capable of near real-time monitoring of power systems extending over a large geographical terrain. Supervisory Control and Data Acquisition (SCADA) systems which are used in wide area monitoring are associated with several drawbacks such as impossibility of capturing dynamic events and incapability of providing synchronized data [8]. Development of the Synchrophasor technology along with the computational, networking and communicational advancements has mitigated those drawbacks using higher sampling rates and precise time tagging.

Furthermore, synchrophasor measurements provide dynamic snapshots across a wide area of the power system, allowing a range of potential wide area voltage stability monitoring and controlling algorithms. In WAMS, synchrophasors are stationed at optimal locations which provide voltage and current magnitudes and their phase angles. VSIs which use the magnitudes and angles of the currents and voltages are compatible with synchrophasor data and can be conveniently applied for wide area voltage stability assessment.

Synchrophasor measurement represents the voltage or current phasors at a given time. The phase angle of these measurements are expressed relative to a reference cosine signal. This reference cosine signal has the nominal system frequency and it is synchronized to the Universal Coordinated Time (UCT) at the beginning of each second [27]. The UCT is based on the International Atomic Time which is highly precise and, this time reference is obtained for synchrophasor calculations through GPS based satellite clocks. Therefore all the measurements taken at different geographical locations are synchronized with each other.



(a) (b)

Figure 2.10. (a) Waveform of reference cosine wave and the measured voltage. (b) Phasor diagram showing the reference and measured phasors.

Chapter 3

Scheme for Predicting Voltage Stability Margin

3.1 Voltage Stability Margin

The additional power that can be transmitted before reaching the voltage collapse point from the current point of operation is usually referred to as the Voltage Stability Margin or the LM. Since the LM is easily understandable and because it reflects the proximity to voltage instability in terms of a measurable and controllable quantity, the system operators can be alerted and automatic remedial actions can be initiated when the LM drops below pre-determined set of critical values.

Although the LM is a good indicator, it is difficult to compute the LM in real-time using the CPF for a large network, due to iterative computations involved. Furthermore, topology and parameters of the network at current time is required for CPF. Therefore, an alternative strategy is required to predict the LM using real-time power system measurements. The approaches based on machine learning techniques are potential candidates, as previously shown in literature [13]-[21]. Most of these machine learning based approaches have used homogenies set of data such as voltage magnitudes, phase angles, reactive power injection, etc. as inputs to the MLMs. In this thesis, it is hypothesized that (i) by using VSIs which

already encode specific information on voltage stability status as inputs to the MLMs, and (ii) by combining multiple MLMs on different categories of inputs, the accuracy and robustness of on-line LM predictions can be improved. In order to investigate this possibility, a new framework is proposed in Section 3.2 to predict the LM as a composite VSI.

3.2 Framework for Real-Time Prediction of Loadability Margin

It is proposed in this thesis to express the LM as a function of a set of VSIs which can be computed using synchrophasor measurements on a real-time basis. The particular function that relates the input VSIs to LM can be learned from data using a suitable machine learning technique. It is also proposed to use an ensemble of MLMs, each of which uses a different group of input VSIs or machine learning principle. The underlying hypothesis is that, the LM can be predicted in a robust manner by combining models that use VSIs based on different principles. The same concept is illustrated in Figure 3.1.

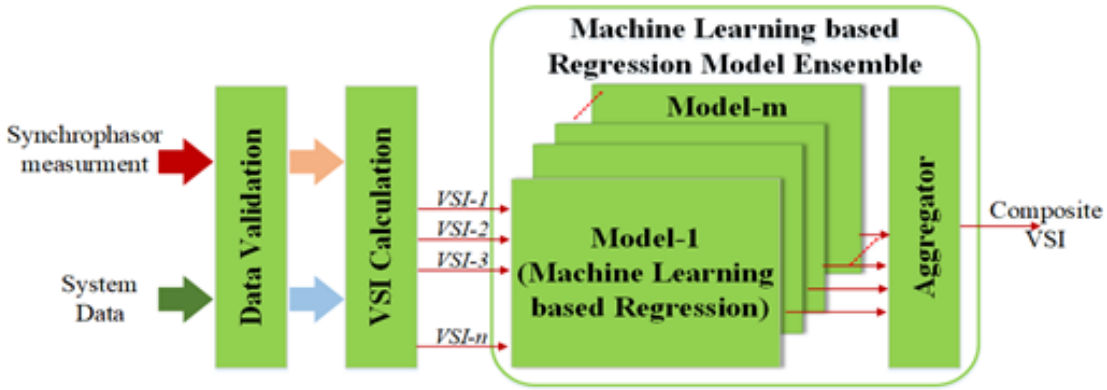


Figure 3.1. Proposed approach for computing LM in real-time

In the proposed approach, real-time synchrophasor data is input to a module that compute a set of VSIs. System data such as the network topology and the admittance matrix are also

input to this module at the beginning and updated when the system changes. Since, these data may contain noises, bias errors and commutation errors the data is validated prior to the VSI calculation. In the data validation step, any existing measurement errors are identified and corrected. The VSIs computed by this module are input to the machine learning based regression models. These regression models may use different machine learning techniques to train the inputs or different inputs to train each MLM. Finally, the results from these MLMs are aggregated to obtain a composite VSI

3.3 Design Methodology

The methodology proposed to develop a real-time LM predictor is shown in a flow diagram in Figure 3.2. The development process starts by generating a database of examples of input-output data pairs to be used in training and testing of MLMs. Each data pair should include a set of VSIs computed at a particular operating point and the LM at that point structured as [VSI_1, VSI_2, …..VSI_n: LM]. Then a statistical analysis is performed to remove the less relevant and redundant inputs. Thereafter the selected input data are categorized into different types and the potential MLMs are trained. In this process, the particular function that relates the input VSIs to LM is learned from the data, and different machine learning algorithms can be explored. After the training, the accuracy of the MLMs can be tested and the less accurate and feeble MLMs can be discarded. Finally a suitable method to aggregate the outputs of the ensemble of selected MLMs is identified. More details about these steps are described with validation examples in Chapter 4 and Chapter 5. This theoretical methodology is implemented experimentally using real-time data from a real-

time digital simulator. The theoretical results are then validated using the results obtained using the experimental setup.

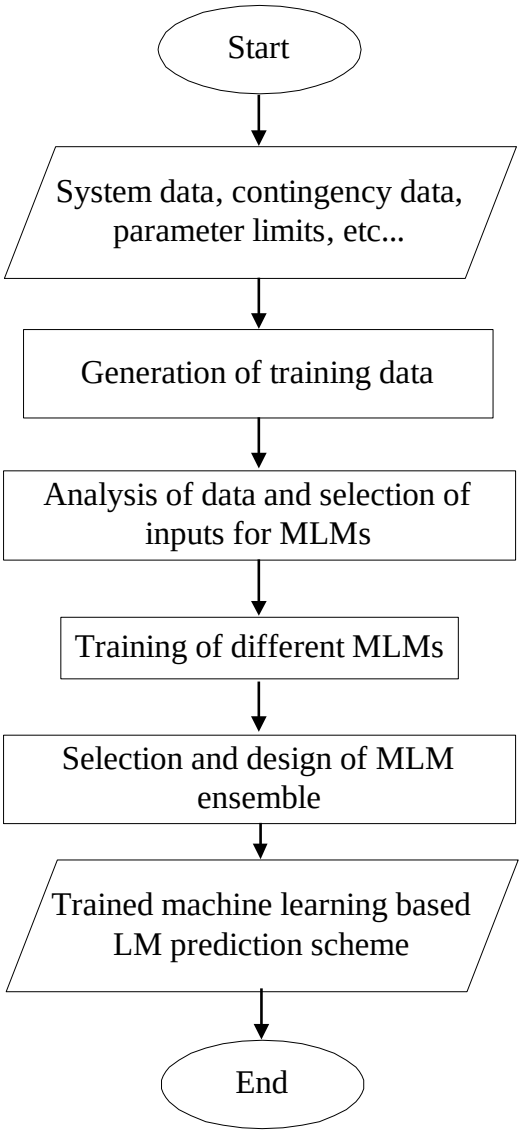


Figure 3.2. Methodology for the development of machine learning based LM predictor.

3.4 Summary

The proposed methodology suggests a new framework to predict the LM using an ensemble of MLMs that take VSIs as inputs. An outline of the said methodology to develop the proposed real-time LM predictor is presented.

Chapter 4

Preparation of Training and Testing Data

4.1 Introduction

A comprehensive database of training examples is an essential requirement for the development of the proposed machine learning based on-line load margin predicting system. It is not practical to obtain these data from historical measurements as voltage stability events which in reality changed a power system to an unstable state are very rare. Therefore, training data need to be invariably generated through simulations.

4.2 Overview of Data Generation Process

The process which is used to generate training and testing data is illustrated in the flowchart in Figure 4.1. In order to obtain data sets corresponding to a diverse set of operating conditions, the process is initiated by generating a large set of initial operating points. These operating points are obtained through a random process and then validated with the conventional power flow to identify the feasible random initial operating points. The continuation power flow is performed to trace the P-V curve up to the voltage collapse point, stating from each validated initial operating point.

The CPF results are sampled to obtain a series of operating points. The electrical variables obtained as CPF results in this process are represented by synchrophasor measurements in

the real-time application. After that, the VSIs described in Section 2.5 and the corresponding LMs are calculated at each operating point. These calculations start from the initial operating point and end at the voltage collapse point. In order to generate a learning data base to train MLMs, this process is repeated for different initial operating points under different power system contingencies.

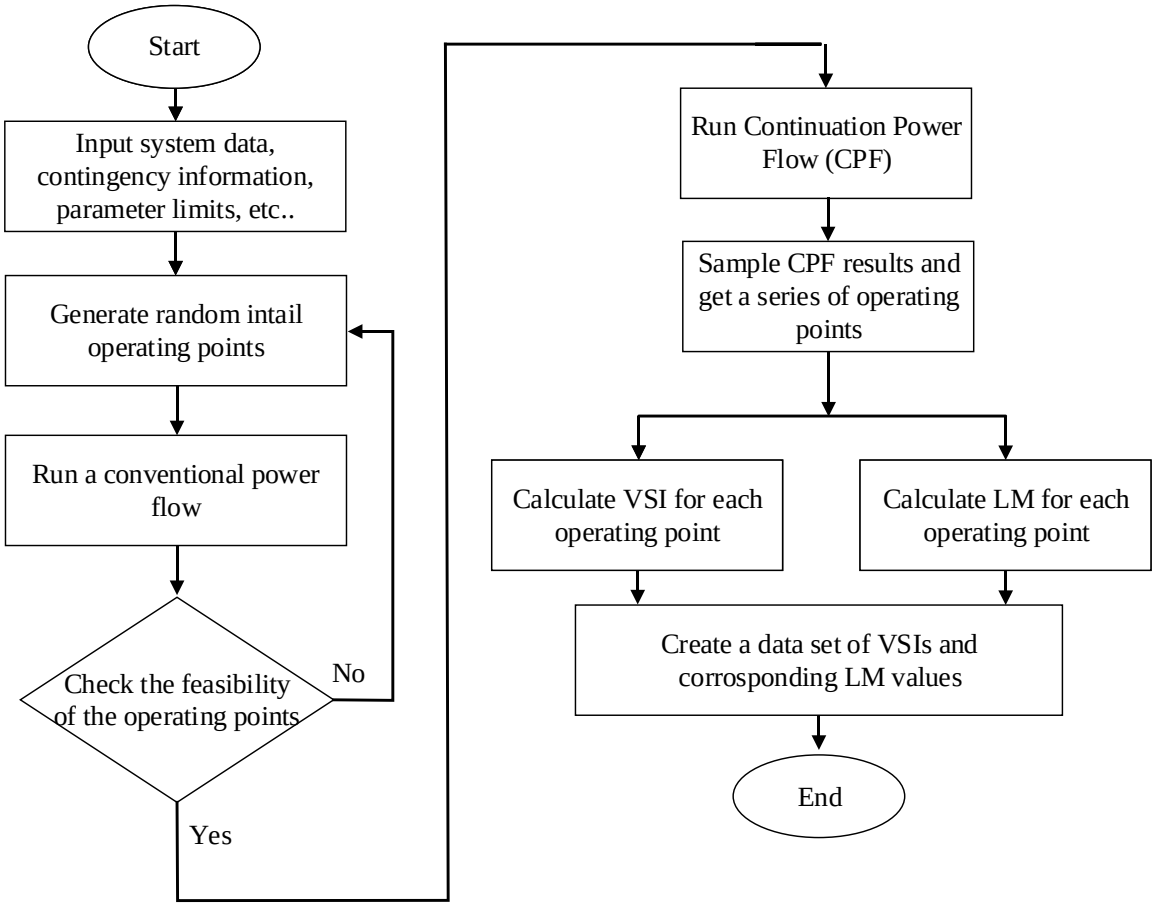


Figure 4.1. Training and testing data generation process

4.3 Initial Power Flow Data Generation

The approach used to calculate the training data is to perform a continuation power flow, starting from a given operating point. The values of VSIs and the corresponding LMs are then calculated using CPF results. The training data set should properly cover the expected region of operation under both normal and contingency situations. Therefore, the base case scenario along with credible N-1 (and N-2, etc. as required) contingency scenarios such as tripping of generator units, transmission lines, fixed shunts, transformers and loads should be taken in to account when generating initial operating points for the CPF. Random deviations are introduced to load and generator settings of the base case and other contingency scenarios in order to have different initial operating points. It is more desirable to generate these random deviations using a normally distributed random variable to follow the natural trend of decrease in the frequency of occurrence with the increase of magnitude of deviation, rather than using a uniformly distributed random variable. Therefore active and reactive power perturbation of the load are generated using (4.1) and (4.2) respectively; where $P_{LO}^{(i)}$ denotes the original active power of the i^{th} bus and $\varepsilon_{PL}^{(i)}(k)$ represents the normally distributed k^{th} random variable with a mean of 0 and a standard deviation of 0.1 for the i^{th} bus. The load reactive power and generator active power deviations are generated in a similar fashion with the same mean and the standard deviation using (4.3). A standard deviation of 0.01 is considered when computing the generator voltage set point deviations using (4.4). This complete process of generating initial operating points is illustrated in Figure 4.2. Figure 4.2. Flowchart of initial operating point generation

$$P_L^{(i)}(k) = P_{LO}^{(i)}\left\{1 + \varepsilon_{PL}^{(i)}(k)\right\} \tag{4.1}$$

$$Q_L^{(i)}(k) = Q_{LO}^{(i)}\left\{1 + \varepsilon_{QL}^{(i)}(k)\right\} \quad . \tag{4.2}$$

$$P_G^{(i)}(k) = P_{G0}^{(i)}\left\{1 + \varepsilon_{PG}^{(i)}(k)\right\} \tag{4.3}$$

$$V_G^{(i)}(k) = V_{GO}^{(i)}\left\{1 + \varepsilon_{VG}^{(i)}(k)\right\} \tag{4.4}$$

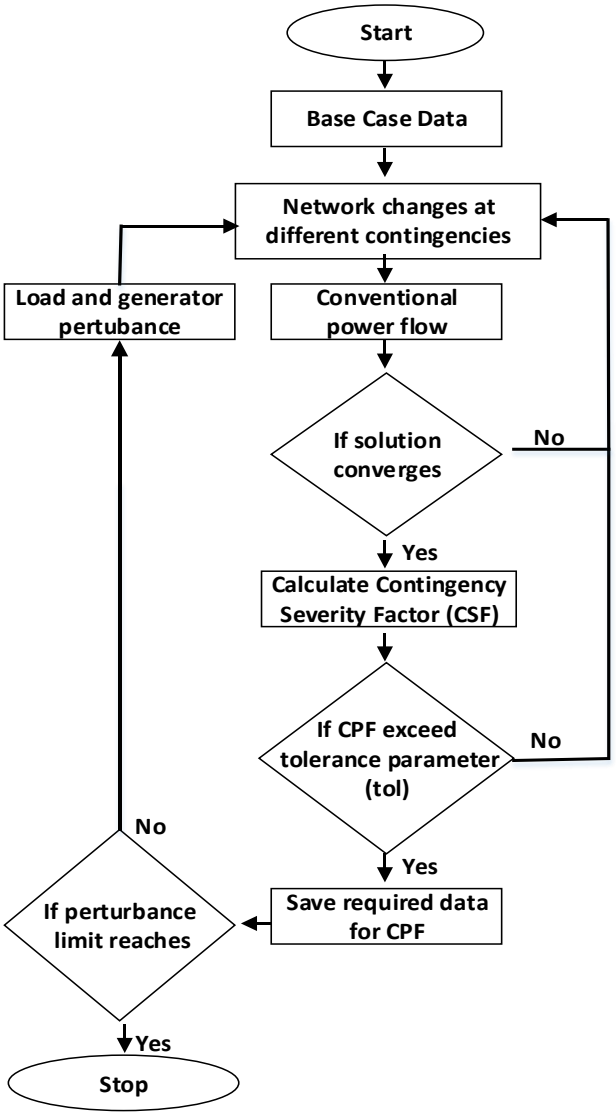


Figure 4.2. Flowchart of initial operating point generation process

However, all randomly generated operating points are not realistic, and their feasibility should be checked by running a power flow analysis. The well-known power system analysis software PSS®E [37] is automated through psspy library to conduct power flow studies for the feasibility check in this work. The operating point is considered feasible if the power flow converges before reaching the iteration limit.

4.3.1 Severe Case Identification

In order to limit the number of data points, some non-influential contingencies (for voltage stability) are eliminated by considering the value of Contingency Severity Factor (CSF) represented in (4.5) where $\bar{V}_i$ represents the steady-state voltage of the i^{th} bus after the contingency, $\bar{V}_i^{BC}$ represents the base case voltage and ΔV_i^{tol} denotes the defined voltage deviation tolerance which is equal to 0.025 in this study. If CSF exceeds 1 for at least one of the load buses after the contingency, it is considered as a severe contingency.

$$CSF_i = \frac{|\bar{V}_i| - |\bar{V}_i^{BC}|}{\Delta V_i^{tol}} \tag{4.5}$$

4.4 Continuation Power Flow

A power system contains different types of nodes such as PQ nodes where P and Q are known values; PV nodes where P and the magnitude of V are specified; and the slack bus where V is a known value (both magnitude and angle). These different type of nodes impose different boundary conditions and hence the problem of solving power flow equations becomes nonlinear. Therefore power flow equations are solved iteratively using techniques such as Gauss-Seidel or Newton-Raphson algorithms. However, conventional

iterative power flow solution methods diverge at the vicinity of the voltage collapse. This happens because the Jacobian matrix corresponding to the power flow becomes singular at the vicinity of a voltage collapse. The power flow equations can be reformulated by applying a locally parameterized continuation technique in order to avoid the singularity of the Jacobian matrix. This method is known as Continuation Power Flow (CPF) [3],[38]. This algorithm can be used to trace the P-V curve beyond the voltage collapse point.

CPF technique uses iterative predictor-corrector scheme to trace the solution path for a specific loading pattern as shown in Figure 4.3

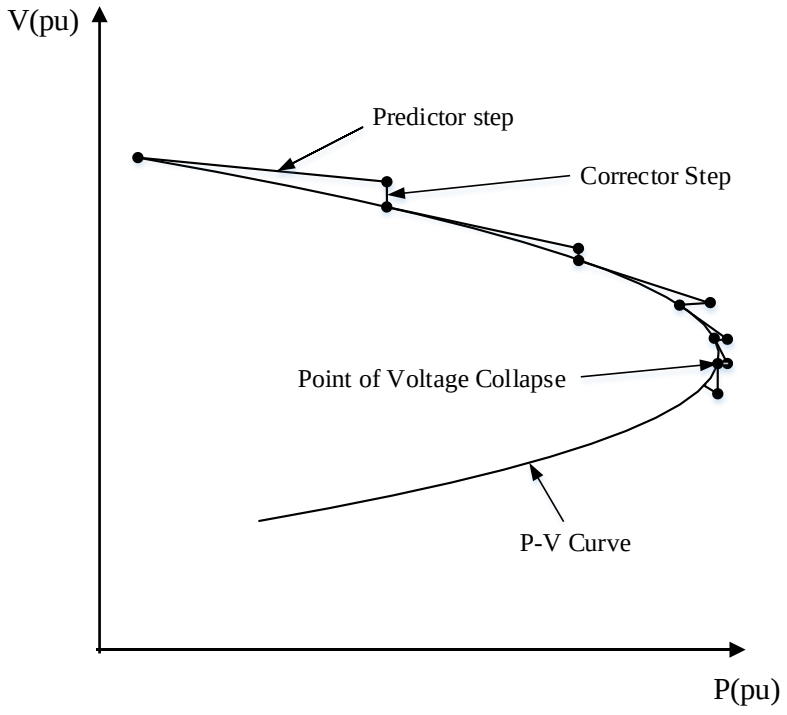


Figure 4.3. Predictor Corrector scheme of continuation power flow

4.4.1 CPF formulation

Traditional power flow equation is reformulated to avoid singularity in the Jacobian matrix. Load parameter (λ) is inserted in to the equation in order to apply the continuation technique as shown in (4.6) and (4.7).

Considering active and reactive power balance following equations can be derived as

$$0 = P_{gi0}(1 + \lambda k_{Gi}) - P_{Li0} + \lambda(k_{Li}S_{base}cos\psi_i) - \sum_{j=1}^{n} V_i V_j y_{ij}\cos(\delta_i - \delta_j - \nu_{ij}) \quad (4.6)$$

$$0 = Q_{Gi} - Q_{Li0} + \lambda(k_{Li}S_{base}sin\psi_i) - \sum_{j=1}^{n} V_i V_j y_{ij}\sin(\delta_i - \delta_j - \nu_{ij}) \quad (4.7)$$

where P_{gi0} represent scheduled initial active power injection at the $i^{th}bus$. S_{base} represents the base apparent power and ψ_i is the power factor of the i^{th} bus .The parameters k_{Li}, k_{Gi} represent the multipliers (constants) to designate rate of change of load and generation at the i^{th} bus. V_i, δ_i and V_j, δ_j: represent the voltage magnitude and voltage angle at the i^{th} and j^{th} buses respectively. y_{ij}, ν_{ij} represent the magnitude and angle of the transmission line admittance between the i^{th} and j^{th}buses respectively. Scheduled load active and reactive powers at the i^{th} bus is represented by P_{Li0}, Q_{Li0} respectively.

This parameterized power flow equation can be expressed as in (4.8).

$$F(\delta, V, \lambda) = 0 \quad (4.8)$$

where V represent a vector of bus voltage magnitudes, δ represent a vector of voltage angle and λ represent the load change parameter or the loading factor.

This parameterized power flow equation can be solved using the predictor corrector method. Predictor estimate the value through the tangential direction and the corrector step provides the correct value on the curve.

4.4.2 Prediction step

Prediction step is the main step of determining the possible solution set of the continuation power flow. Initially, the conventional power flow solution is fed to the algorithm, then the next prediction can be obtained by considering an appropriately sized step in the direction of the tangent of solution path. The tangent vector of the parameterized power flow equation can be determine by differentiating (4.8):

$$d[F(\delta, V, \lambda)] = [F_\delta{}' \quad F_v{}' \quad F_\lambda{}'] \begin{bmatrix} d\delta \\ dv \\ d\lambda \end{bmatrix} = [0] \tag{4.9}$$

Since there is an additional variable than the number of equations, the Jacobian matrix of (4.9) should be modified. The Jacobian matrix is modified by inserting a row vector e_k with all elements equal to zero except the last element which equal to 1. When the tangent direction is increasing, the right side matrix of (4.10) should be $[0\ 1]^T$ and when the tangent direction is decreasing, the right side matrix of (4.10) should be$[0\ -1]^T$.

$$\begin{bmatrix} F_\delta{}' & F_v{}' & F_\lambda{}' \\ & e_k & \end{bmatrix} \begin{bmatrix} d\delta \\ dv \\ d\lambda \end{bmatrix} = \begin{bmatrix} 0 \\ \pm 1 \end{bmatrix} \tag{4.10}$$

where $\begin{bmatrix} F_\delta{}' & F_v{}' & F_\lambda{}' \\ & e_k & \end{bmatrix} = \begin{bmatrix} \frac{\partial f_1}{\partial \delta} & \frac{\partial f_1}{\partial V_2} & P_0 \\ \frac{\partial f_2}{\partial \delta} & \frac{\partial f_2}{\partial V_2} & Q_0 \\ 0 & 0 & 1 \end{bmatrix} = J_{aug}$

The augmented Jacobian matrix is represented as J_{aug}. After calculating the tangent vector, predictor step can be conducted as

$$\begin{bmatrix} \delta^* \\ V^* \\ \lambda^* \end{bmatrix} = \begin{bmatrix} \delta \\ V \\ \lambda \end{bmatrix} + \sigma \begin{bmatrix} d\delta \\ dV \\ d\lambda \end{bmatrix} \quad . \tag{4.11}$$

Step size (σ) is selected in a way that the prediction must be within the boundary of convergence of the corrector, hence a small step size should be used.

4.4.3 Corrector step

After calculating the predictor, corrector step is performed using the predictor solution as the initial guess. However, parameterization is the most vital step of the corrector process. In this thesis local parameterization is considered. Hence, the original set of equations represented in (4.8) is altered as shown in (4.12) by adding an equation with the state variable x_k and prediction value of the state variable $x_k{}^*$. This state variable is known as the continuation parameter and it is determined as explained in section 4.4.4

$$\begin{bmatrix} F(\delta, V, \lambda) \\ x_k - x_k{}^* \end{bmatrix} = [0] \tag{4.12}$$

The Jacobian matrix can be obtained by using (4.12). The mismatches (Δf_1 and Δf_2) can be obtained by substituting the predictor values to (4.6) and (4.7). Then the mismatch values are substituted to (4.13). Afterwards the corrector coefficients can be obtained by (4.14).

$$\begin{bmatrix}\Delta\delta\\ \Delta V\\ \Delta\lambda\end{bmatrix} = J_{aug}^{-1} * \begin{bmatrix}\Delta f_1\\ \Delta f_2\\ 0\end{bmatrix} \quad (4.13)$$

Hence the corrected solution is obtained as

$$\begin{bmatrix}\delta\\ V\\ \lambda\end{bmatrix} = \begin{bmatrix}\delta^*\\ V^*\\ \lambda^*\end{bmatrix} + \sigma \begin{bmatrix}\Delta\delta\\ \Delta V\\ \Delta\lambda\end{bmatrix} \quad . \quad (4.14)$$

4.4.4 Selecting the Continuation Parameter

The best continuation parameter is selected by analysing the tangent vector component of each state variable. The state variable with the largest tangent vector component is selected as

$$x_k := \max\{[x_1', x_2', \dots, x_m']\} \quad . \quad (4.15)$$

Where x_k is the k^{th} state variable and x_m' represents the tangent of the state variable x_m. Each part of the P-V curve has a different continuation factor.

- Upper part of the P-V curve (prior to the critical point): Load change (λ) is taken as the continuation factor.
- Near the critical point of P-V curve: a voltage change (V) is taken as the continuation factor.
- Lower part of the P-V curve (after the critical point): Load change is taken as the continuation factor.

4.5 Factors Impacting the Power System Voltage Stability

Voltage stability issues generally occur in heavily stressed systems [1]. However there are other factors which contributes to the voltage instability [2].They are:

- Generator reactive power capability limit
- Automatic transformer tap changing
- Reactive power compensation device operation and control
- Load characteristics
- Distributed resources operation and control

In this thesis, the impact of generator reactive power capability limit and automatic transformer tap changing is taken in to consideration.

4.5.1 Generator reactive power capability limit

When the generator reactive power reaches it's over excitation level, the generator cannot inject reactive power to the system. Hence a generator bus which has reached the reactive power limit is considered as a PQ bus rather than a PV bus. Reactive power (Q) remains constant even when the system load increases. This will change the trajectory of the voltage profile when system loads increase gradually. There is a considerable change in the LM when the generators reach the reactive power limit.

4.5.2 Automatic transformer tap changing

Automatic on-load transformers operate to regulate the secondary side voltages of the power systems. Transformer tap increases or decreases the voltages, according to the voltage regulation requirement and this will lead to several changes in the system admittance matrix.

4.6 Case Study

In this section, the proposed method of training and testing data generation is applied to two test systems, and the accuracy of the methodology and computational routines are validated when possible.

4.6.1 Test Systems

The IEEE 14 bus system and the IEEE 118 bus system are considered when evaluating the proposed voltage stability margin assessment approach. The IEEE 14 bus system consists of 5 generators, 9 load buses with static loads and one fixed shunt. The generators at bus-3, -6 and -8 are operated as synchronous condensers. Analysis of the voltage profile of the system under severe contingency scenarios showed that bus-14 is the most critical bus in terms of voltage instability for this system.

The IEEE 118 bus system consists of 19 generators, 35 synchronous condensers, 56 load buses with static loads and 14 fixed shunts. It was found that bus-45 is the most vulnerable bus for voltage instability in the IEEE 118 test system. The data of IEEE -14 and -118 bus test systems are given in Appendix. A and Appendix. B respectively.

4.6.2 Validation of the developed CPF Program

A CPF program implemented in Matlab® with all the required features. A commercial transmission planning and analysis programme Siemens Power System Simulator for Engineers (PSS®E) [56] is used to validate the CPF program. This program allows to perform wide verity of analysis functions such as power flow, power system dynamics, short circuit, contingency analysis, optimal power flow, voltage stability and transient

stability. Repetitive power flow analysis is conducted using PSS®E to validate the bus voltage magnitudes and phase angles obtained from the developed CPF routine. Only the initial parts of the voltage trajectories can be tested, as PSS®E power flow solution diverges when the voltages drops well below the respective nominal values.

The IEEE 14 bus system is used for this comparison. The scheduled loads and the active power injections of PV buses of PSS®E raw data file are increased by 12.5 % w.r.t the initial scheduled load. After each incremental power flow solution is obtained, the bus voltage magnitudes and the phase angles are recorded. The power mismatch after the power flow solution is maintained below 0.1 MW. As the voltages deviate, the magnitude of the total mismatch of PSS®E power flow solutions increase and finally diverges. Therefore, a comparison between PSS®E and CPF can be only done up to the point of divergent of the power flow. The results obtained from CPF and PSS®E are presented in Figure 4.4 to Figure 4.7.The data points obtained from PSS®E coincide with the CPF results, confirming the accuracy of the CPF results.

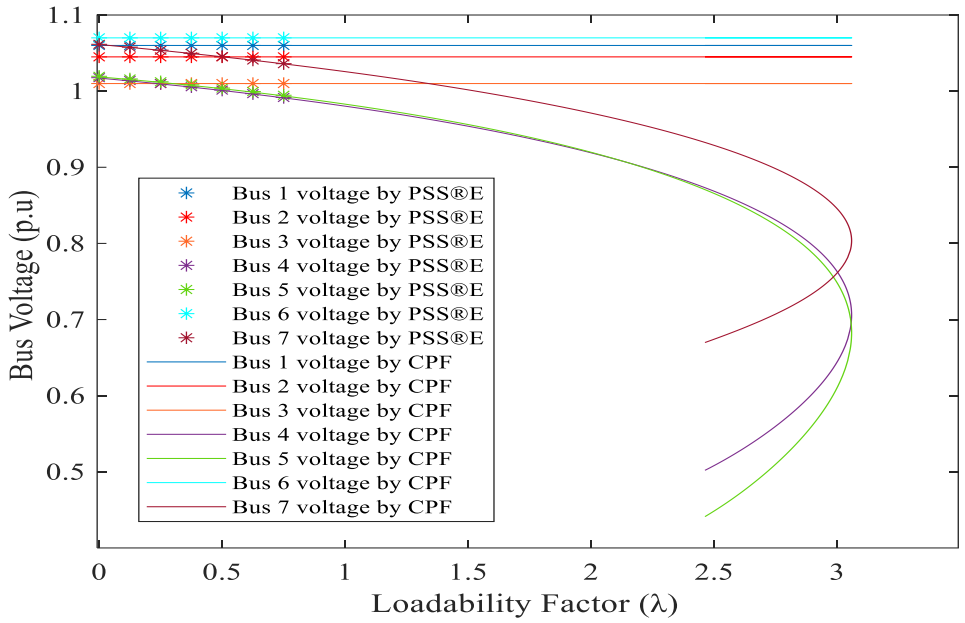


Figure 4.4. Comparison of P-V curves obtained by CPF and PSS®E power flow results for bus-1 to -7 of IEEE 14 bus system

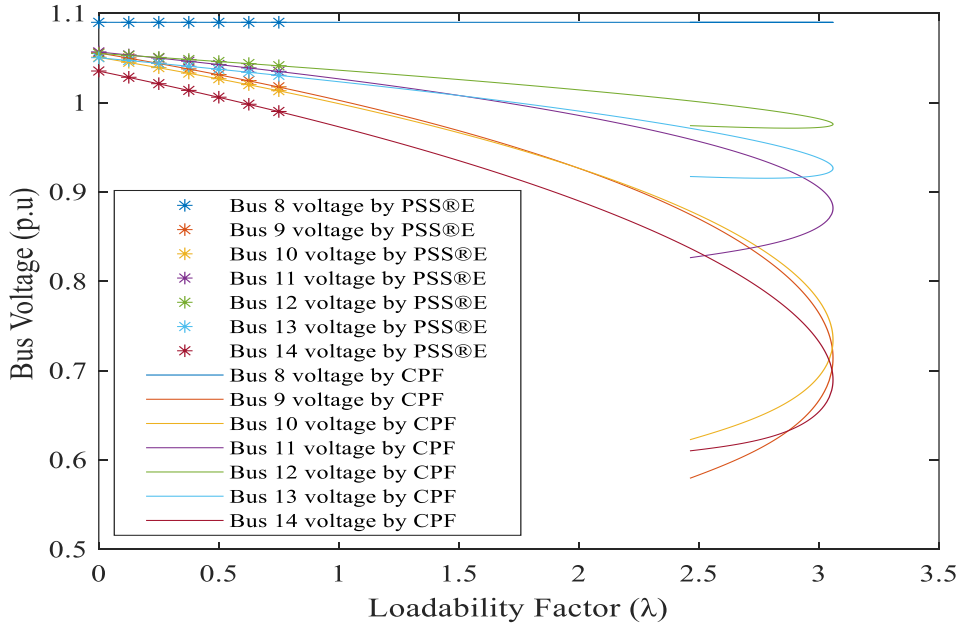


Figure 4.5. Comparison of P-V curves obtained by CPF and PSS®E power flow results for buses-8 to -14 of IEEE 14 bus system

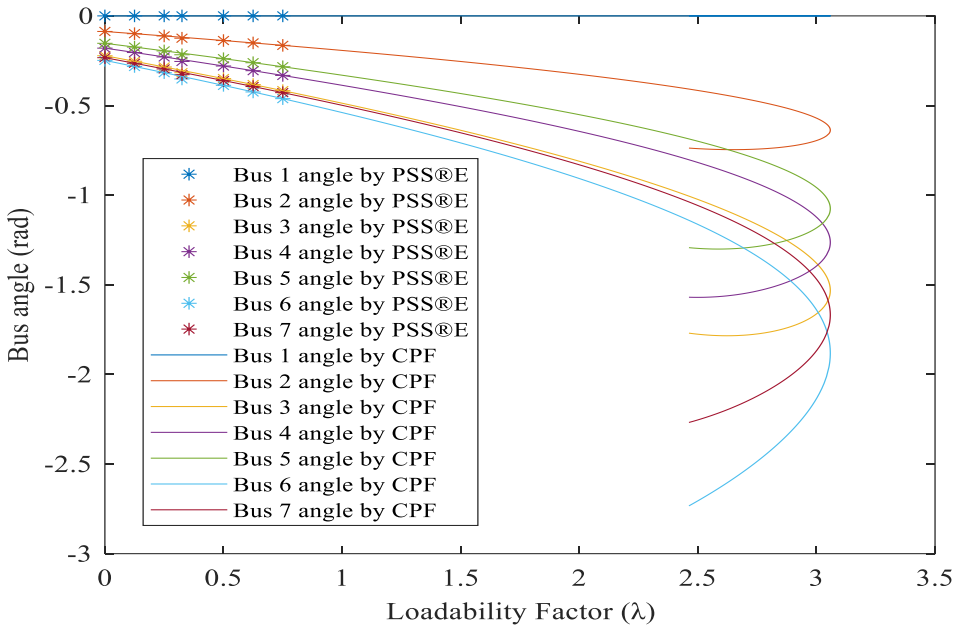


Figure 4.6. Comparison of P-δ curves obtained by CPF and PSS®E power flow results for buses-1 to -7 of IEEE 14 bus system

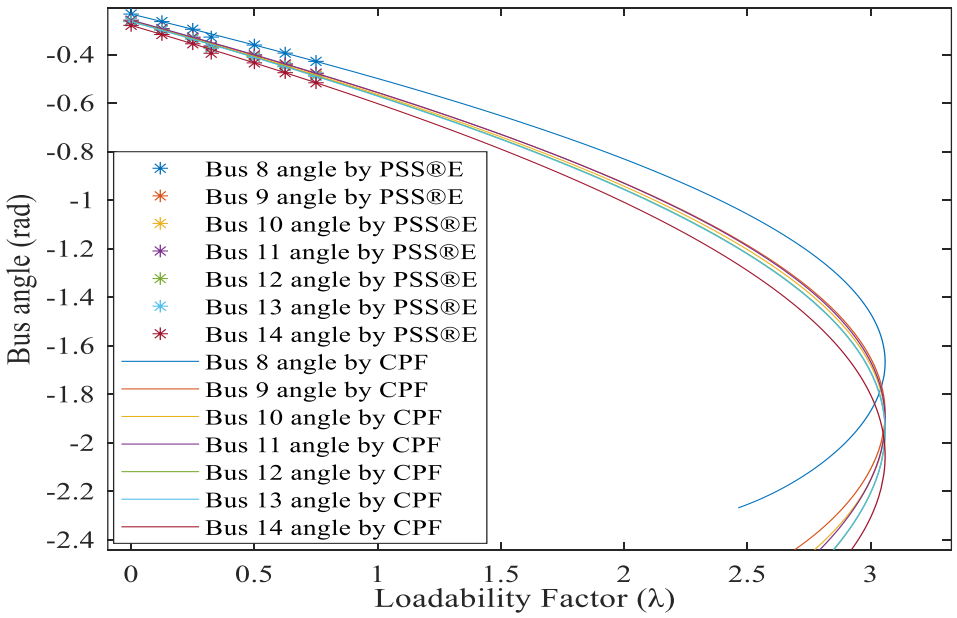


Figure 4.7. Comparison of P- δ curves obtained by CPF and PSS®E power flow results for buses-8 and -14 of IEEE 14 bus system

Figure 4.8 shows the validation results of the CPF with PSS®E data of the IEEE 14 bus system considering the reactive power capability of the generator buses.

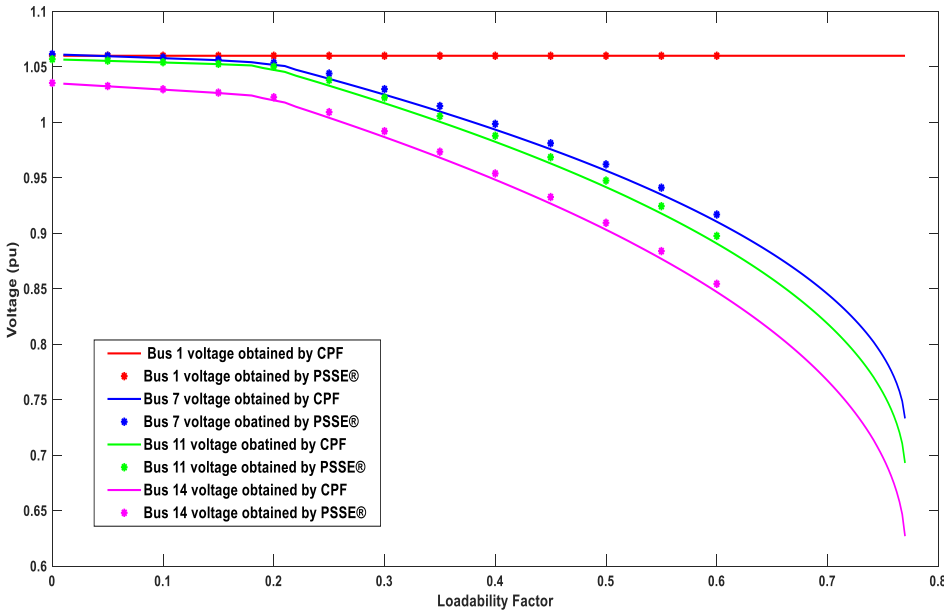


Figure 4.8. Comparison between PV curves of generator buses of IEEE 14 bus system when generators reactive power limits are considered

The in-house developed CPF contains a controller model to change the taps of the system transformers to regulate the secondary voltage at 0.9 p.u. When the bus voltage reaches below this threshold, transformer taps will rise until the transformer reaches its tap limit. This operation raises the P-V curve further along with the voltage critical point as shown in Figure 4.9, which depicts an example of a PV curve of a load bus (Bus14) in the IEEE 14 test system. The curve shows the impact of a generator reaching its over-excitation limit as well as the action of an automatic tap changing operations of a transformer. The generator connected to bus-2 reaches its over-excitation limit and changes the trajectory of the voltage profile when the system loads increase gradually as shown in Figure 4.9. The

automatic transformer tap changing operation of the transformer between bus-4 and -7 raises the P-V curve and pushes the critical point further as shown in Figure 4.9.

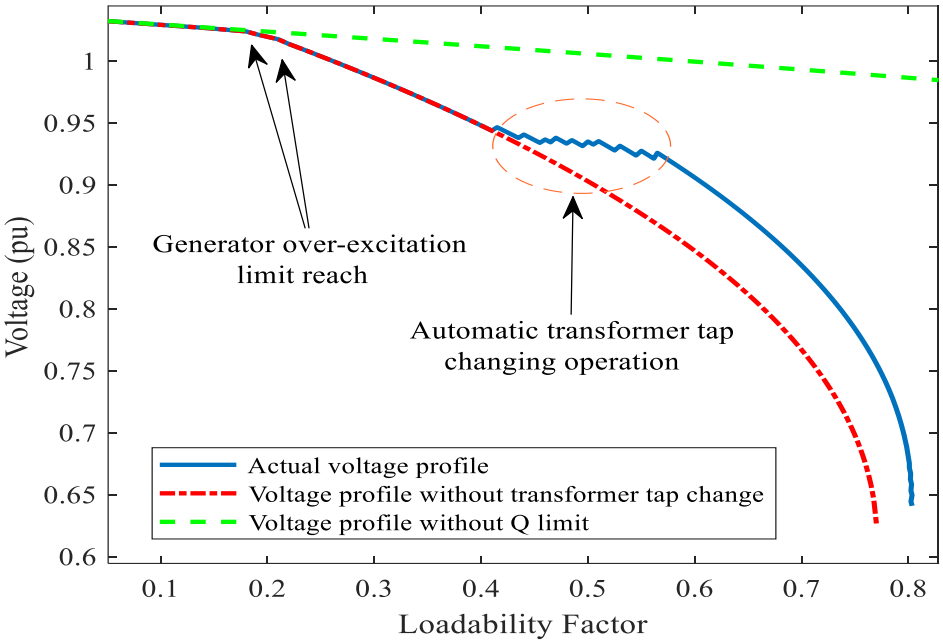


Figure 4.9. PV curve of bus 14 of IEEE 14-bus system under different operating conditions

4.7 Loadability Margin and VSI Calculation

In order to generate voltage and current phasors, and LM data up to the voltage collapse point, the conventional CPF technique [3],[38] is used. Important aspects such as generator over-excitation limit reach and automatic transformer tap changing operations are integrated to the CPF algorithm. The load is increased uniformly for all load buses in this work, but it is possible to introduce more specific or randomized load change patterns if desired. Once the CPF is performed, the trajectory is resampled, and for each sample point, VSIs and the corresponding LMs are computed, and saved to a database. In this work, 11 VSIs were computed using (2.25), (2.29), (2.31), (2.38), (2.48),(2.50), (2.55), (2.60), (2.63), (2.81) and (2.82) .

4.8 Data Base Preparation

The bus voltage phasors, VSI data sets and the LM were generated through the approach described at the beginning of this chapter for both test systems. Therefore, each column of the database represent the bus voltage magnitudes, phase angles, VSIs, and the LMs. In the data generation process, cases with different initial operating points and contingencies are considered. Therefore rows of the database are generated at each operating point for each of these cases. The factors which define the created data base size such as; number of initial operational points considered, number of contingency cases used in addition to base case and the total number of data points generated through resampling of CPF are tabulated in Table 4-1.

TABLE 4-1.DATA BASE SIZE

Test System	Number of initial operating points	Number of contingencies considered	Total number of data points
IEEE 14 bus system	3,196	32	338,366
IEEE 118 bus system	5,424	55	632,161

4.9 Summary

A database of possible power system operating scenarios is created by randomly generating different operating points and verifying them through power flow analysis under normal operation as well as under feasible contingencies. Starting from each initial operating point, continuation power flow was performed to determine the voltage collapse point. CPF was incorporated with generator reactive power capability limits and automatic tap changing

operation to obtain more realistic off-line simulation data. Part of the CPF data are then validated using the results from a commercial power system analysis tool to confirm the validity. Finally, after computing VSIs and LMs at different operating points, a database is generated with a sufficient amount of data to train and test the proposed machine learning algorithms.

Chapter 5

Development and Validation of Machine Leaning Model Ensemble

5.1 Introduction

Using the data generated as described in Chapter 4, a set of candidate MLMs are trained to predict the LM. Since every feature in the MLM contributes for the computation burden, it is important to evaluate the input features prior to train the MLM. After the feature selection, MLM training is performed using different machine learning algorithms. Each model trained through different learning algorithms are then tested in order to select the best predictor model. Finally, the best MLMs are combined together to improve the robustness of the predictor model.

5.2 Background of Multi-Variant Machine Learning Regression Algorithms and Ensemble Methods

5.2.1 Multi-Variant Machine Learning Regression Algorithms

Predicting the system LM using multiple inputs require multi-variant regression models. Among different regression techniques available, the following techniques were

investigated:

- Linear Regression
- Polynomial Regression
- Gradient Boost Regression
- Artificial Neural Network Regression
- Random Forest Regression

5.2.1.1. Linear Regression (LR)

Linear regression represents the linear relationship between the input features and output features. LR learner assigns optimum values to set of coefficients $[\beta_0,\beta_1, \dots \beta_p]$ of the corresponding set of input features including a bias term $[1, x_1, \dots\ x_p]$ to obtain a linear relationship as in (5.1), with the objective of minimizing the residual value.

$$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 \dots\dots \beta_p x_p \tag{5.1}$$

where p indicates the number of input features. When the input features have a linear or near-liner relationship with the dependant (output) feature, the trained LR model shows lower error and high correlation. This method of determining coefficients is given in Appendix. C.1.

5.2.1.2. Polynomial Regression (PR)

Polynomial regression attempts to capture the non-linear relationship between the input features and outputs assuming that the non-linear relationship can be approximated by a

polynomial. A single variable PR fitted curve is represented in (5.2), where the degree of the fitted curve is *d* and the coefficient of the k^{th} power of the input feature x_k is β_{k} .

$$y = \sum_{k=0}^{d} \beta_{\text{k}}\, x^k \tag{5.2}$$

However, for multi-variant polynomial regression the fitting curve will include products of *d*-1 powers of the input features. For example, the polynomial fitting curve of a second order two variable regression problem can be represented as in (5.3).

$$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_1 x_2 + \beta_4 x_1^2 + \beta_5 x_2^2 \tag{5.3}$$

Similar to the LR learner, the PR learner assigns optimum values to each coefficient of the fitting curve with the objective of minimizing the residual [58]. If the input features have a non-linear relationship with the dependant (output) feature, the trained PR model shows a lower error and a higher correlation.

5.2.1.3. Gradient Boost Regression (GBR)

J.H. Friedman proposed the Gradient Boosting machine in [32] which is capable of doing both regression and classification, especially for mining applications with less clean data [34]. This learning algorithm contains *M* number of small decision trees which are trained using the pseudo residual of each data instance. Each decision tree is scaled using a learning rate (ν) to lower the variance of the prediction. Predictions can be obtained as represented in (5.4) where $F_i(x)$ represents the prediction from each tree. A further description on the training methodology and the assumptions is given in Appendix. C.2.

$$y = F_0(\text{x}) \; + \nu \sum_{i=1}^{M} F_i(\text{x}) \tag{5.4}$$

5.2.1.4. Artificial Neural Network Regression (ANNR)

Artificial neural network was inspired by the operation of biological neural networks. It was proposed by Hebb in 1940 [33]. In ANN, each neuron in a layer receives signals from successive downstream nodes multiplied by the corresponding weightage. These are summed up with a bias unit, the total summation is fed to an activation function and the output is fed to the successive upstream hidden layer neurons. This process continues through all the hidden layers until it converge to one neuron at the output layer. The output of j^{th} neuron can be repented as

$$y_j = \frac{1}{1 + e^{-net_j}} \tag{5.5}$$

where $net_j = \sum_{i=0}^{N_d} x_i \omega_{ij} + \omega_{j0}$.

N_d represents the number of neurons connected from the downstream hidden layer. This model and learning procedure is further elaborated in Appendix. C.3.

5.2.1.5. Random Forest Regression (RFR)

Random forest was proposed by L. Breiman in 2001 [35]. Random forest is a collection of decision trees trained using independently sampled random vectors. The generalization error of the RF depends on the strength of each tree and the correlation between each tree. A random forest consisting of *N* number of decision trees which is trained using random combinations of boost trap data instances. Average of the predictions (F_i(x)) of each tree from the forest is considered as the output (y) as given in (5.6). Additional descriptions on the methodology and assumptions are discussed in Appendix. C.4.

$$y = \frac{1}{N}\sum_{i=1}^{N} F_i(\mathrm{x}) \tag{5.6}$$

5.2.2 Machine Learning Model Ensembles

The main objective of a machine learning model ensemble is to integrate several predictions of the base MLMs built with a specific learning algorithm and a set of input features in order to improve the generalizability and predictor robustness over a single MLM [59]. There are two types of main ensemble methods:

- Averaging method: The basic principle of this method is to obtain the average prediction from several independent MLMs. Generally the averaged or weighted averaged prediction is better than the individual MLMs since it reduce the variance.
- Boosting method: This method is used to combine several weak MLMs to build a powerful ensemble. The base MLMs are built sequentially while trying to reduce the bias of the final prediction model.

5.3 Multi-Variable Regression Training

5.3.1 Training Feature Selection and Categorization

Since feeding irrelevant and redundant input features into a MLM increases the computational burden and contributes to increased errors and overfitting, it is important to evaluate the relevance of input features to select only the most important features. This can be performed prior to MLM training or as a part of the training process. In this thesis, the former approach is considered.

It is proposed to select the most relevant inputs through Spearman's rank correlation coefficient method which provides correlation coefficients equal or close to 1 or -1 when two data sets have a close monotonic relationship. When there is no monotonic relationship, the correlation coefficients are close to 0. Only the features with a very high positive or negative correlation should be selected, in order to minimize the number of input features and improve the accuracy. As a rule of thumb used in statistics [57], Spearman's rank correlation coefficients which are equal to or higher than 0.9 or equal or less than -0.9 are considered to have a very high correlation. Therefore, 0.9 can be used as the magnitude threshold for selection of features with a very high correlation.

Recursive feature elimination is a backward feature selection method. In this technique the model is created using all the available feature sets and the importance score is computed for each feature. The least important feature is removed and the model is retrained and the importance score is calculated again, until the desired number of features is reached [62].

5.3.2 Machine Learning Model Training and Testing Methodology

The training of MLMs with a given set of data must be performed in a systematic manner to achieve a model with the best possible generalization capabilities. The k-fold cross validation has emerged as an acceptable method to realize this purpose in machine learning literature. A flow chart summarizing this methodology is presented in Figure 5.1.

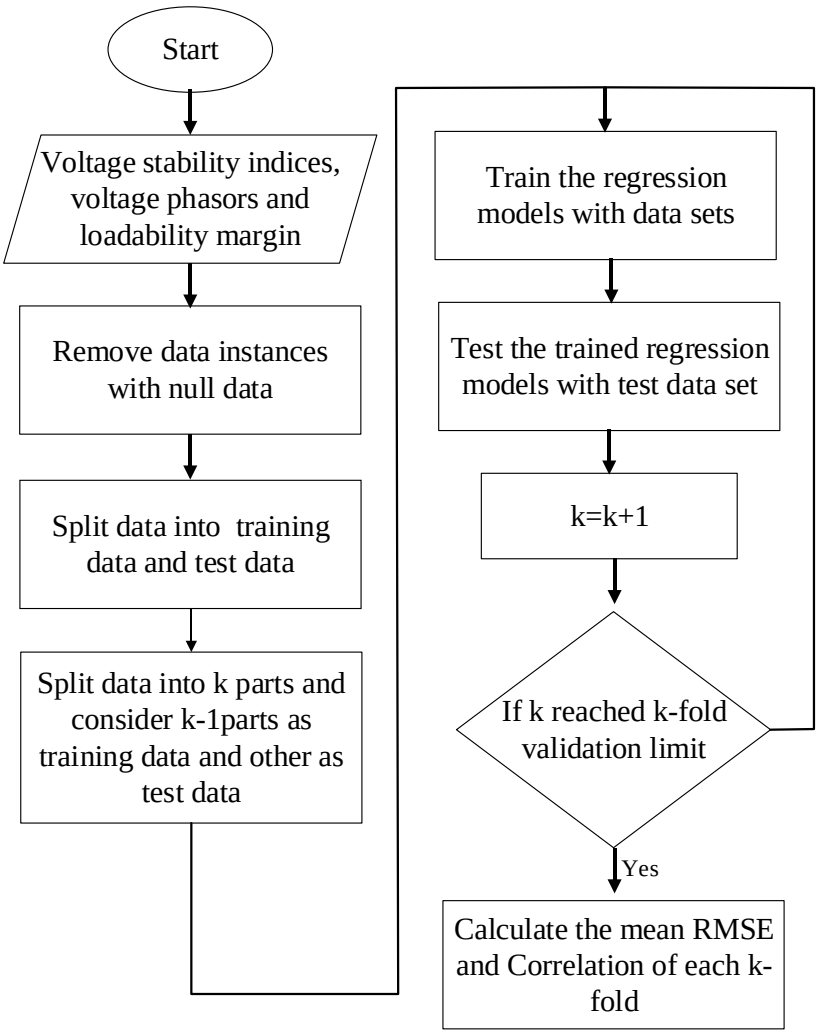


Figure 5.1. Machine learning model training and testing methodology

After categorizing and selecting the features, it is recommended to generate separate databases which consist of the selected feature sets. Each of these databases is conditioned by removing the null data instances and are fed individually to the learners, and split in the ratio of 4:1 as training and testing data. For k-fold cross validation, the training dataset is again split into k consecutive folds and each split is then used once as the validation set

while the remaining k-1 folds are used for training. The best cross validation split is used to obtain the best generalized estimate. Root Mean Square Error (RMSE) and the Prediction correlation (R^2) are considered as the measurements to evaluate each ML model in this thesis.

5.4 Application to IEEE 14 and IEEE 118 Bus Systems

The training data bases developed in Chapter 4 using IEEE 14 and IEEE 118 bus test systems are used to train the MLMs. Four different MLMs that use different categories of input features are developed for identifying the best feature set:

MLM-1: Uses a set of bus voltage magnitudes and phase angles as inputs. The most relevant buses are selected using the recursive feature elimination method [41]. This MLM is similar to the approach proposed in [16] and will be used as a reference model.

MLM-2: Uses all of the candidate VSIs as input features.

MLM-3: Uses a set of VSIs which are calculated using only the local measurements of the considered bus. The VSIs used in MLM-3 are SDC, ISI, VSI_b, VSLI, VSLBI, VCPI_1 and VSM.

MLM-4: Uses a set of most relevant VSIs as inputs.

The recursive feature elimination procedure was applied in selecting features for MLM-1, as there are a large number of available candidate features. Feature selection via recursive feature elimination for the IEEE 14 bus system phasor data resulted in the selection of voltage phasors corresponding to the buses 1, 2, 8, 10 and 14 as the inputs for MLM-1. The

corresponding voltage phasors for the MLM-1 for IEEE 118 bus system are those of the buses 2, 3, 5, 11, 13, 18, 19, 29, 32, 56, 57, 80, 81, 88 and 95.

The Spearman's rank correlation coefficients method is applied for selecting input features for MLM-4. The Spearman's rank correlation coefficients heat map for the IEEE 14 bus system, computed with the database created in Chapter 4, are shown in Figure 5.2.

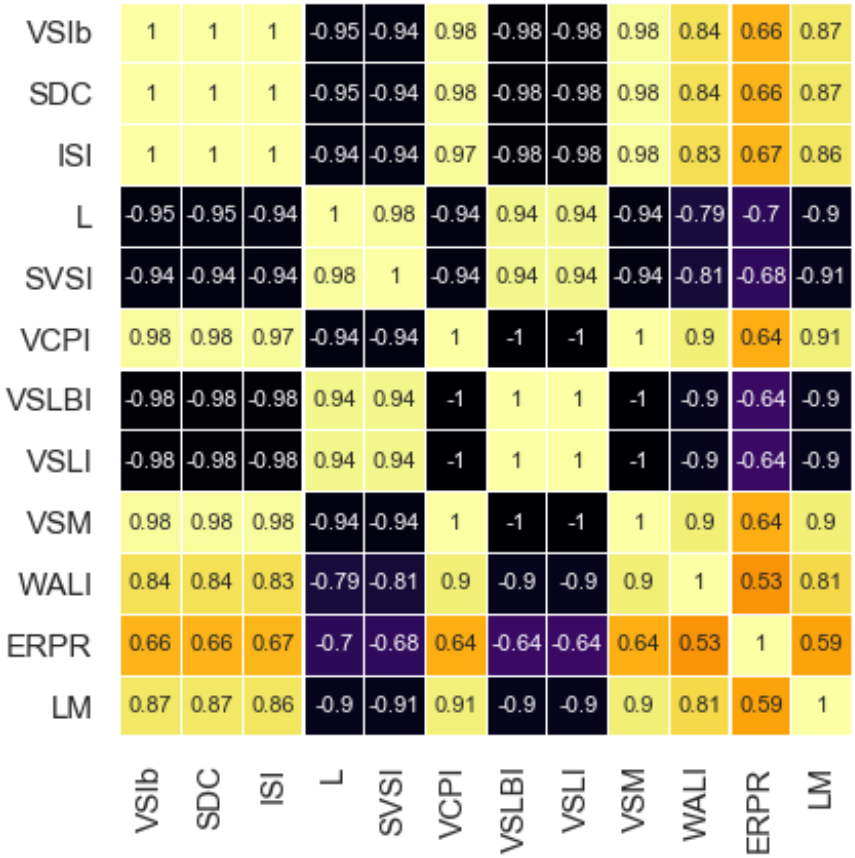


Figure 5.2. Correlation heat map of VSIs and loadability margin

The indices SVSI, VCPI, VSLBI, VSLI, VSM and L indices can be recognized as the most important features (correlation > 0.9). However the correlation between VSM, VSLI, VCPI and VSLBI is 1 and, hence representing only one of these is sufficient. Preliminary trials run to confirm this showed that the MLMs taking more than one VSI from the group having

perfect correlation as inputs do not display significant improvement in the LM prediction accuracy. Thus the index VCPI was selected to represent the above four indices and the inputs to MLM-4 were finalized as VCPI, SVSI, and L.

With the selected features, four categories of MLMs were trained using k-fold cross validation approach, with k = 10. For each MLM category, the five different machine learning approaches presented in Section 5.2.1 are applied. Each of these trained MLMs are evlauted in terms of the Root Mean Square Errror (RMSE) and the coefficient of multiple determination (R^2) for both IEEE 14-bus sysem and IEEE 118-bus system, and the results are presented in Table 5-1 to Table 5-4.

TABLE 5-1. MLM 1 VALIDATION AND TESTING

		IEEE 14		**IEEE 118**	
Regression Type		**RMSE**	**R^2**	**RMSE**	**R^2**
LR	Train	0.0428	0.9047	12.643	0.1445
	Test	0.0428	0.9065	12.569	0.1373
PR	Train	0.0326	0.9455	10.514	0.4556
	Test	0.0332	0.9437	11.003	0.3963
GBR	Train	0.0472	0.9467	11.685	0.4883
	Test	0.0473	0.9465	11.823	0.4377
ANNR	Train	0.0088	0.9929	3.949	0.9497
	Test	0.0088	0.9923	562.64	0.6136
RFR	Train	0.0027	0.9999	2.265	0.9853
	Test	0.0063	0.9981	5.296	0.9431

TABLE 5-2. MLM 2 VALIDATION AND TESTING

		IEEE 14		IEEE 118	
Regression Type		**RMSE**	**R^2**	**RMSE**	**R^2**
LR	Train	0.0329	0.9425	10.552	0.0596
	Test	0.0329	0.9456	10.455	0.0418
PR	Train	0.0287	0.9563	11.573	0.0371
	Test	0.0362	0.9304	12.037	0.0261
GBR	Train	0.0463	0.9498	3.675	0.9997
	Test	0.0463	0.9496	3.666	0.9986
ANNR	Train	0.0134	0.9811	8.508	0.5863
	Test	0.0133	0.9816	1.3e12	0.4430
RFR	Train	0.0027	0.9996	0.142	0.9999
	Test	0.0061	0.9981	0.286	0.9993

TABLE 5-3. MLM 3 VALIDATION AND TESTING

		IEEE 14		IEEE 118	
Regression Type		**RMSE**	**R^2**	**RMSE**	**R^2**
LR	Train	0.0370	0.9260	10.553	0.0585
	Test	0.0371	0.9260	10.456	0.0384
PR	Train	0.0358	0.9311	10.450	0.0371
	Test	0.0391	0.9169	12.037	0.0261
GBR	Train	0.0507	0.9190	10.357	0.2155
	Test	0.0507	0.9185	10.392	0.0746
ANNR	Train	0.0294	0.8527	8.506	0.5258
	Test	0.0294	0.8525	6.5e12	0.3634
RFR	Train	0.0374	0.9992	1.881	0.9841
	Test	0.0084	0.9963	4.216	0.9097

TABLE 5-4. MLM4 VALIDATION AND TESTING

		IEEE 14		IEEE 118	
Regression Type		**RMSE**	**R^2**	**RMSE**	**R^2**
LR	Train	0.0421	0.8360	9.986	0.3107
	Test	0.0422	0.8432	9.907	0.3132
PR	Train	0.0358	0.9312	9.855	0.3355
	Test	0.0356	0.9313	11.38	0.3261
GBR	Train	0.0367	0.9482	8.128	0.6952
	Test	0.0368	0.9476	8.220	0.5409
ANNR	Train	0.0205	0.9279	8.505	0.5698
	Test	0.0206	0.9274	22.820	0.4866
RFR	Train	0.0049	0.9980	2.359	0.9762
	Test	0.0104	0.9907	5.384	0.9450

The results clearly indicate that ANNR and RFR are the better algorithms for this problem. Although the linear regression performance is somewhat statisfactorily for the IEEE 14 bus system, its performance is poor when the system becomes larger. Both the LR and PR are generic parametric functions where the parameters are determined with the objective of minimizing the error between the fitted curve estimation and the training data. It can be observed that these parametric curves will predict the LM with an acceptable accuracy when the system is small; however when the system size increases, the predictions from these generic parametric functions show higher errors and lower correlation. Therefore, the generic parametric functions are not very satisfactory for this problem.

The ANNR learner shows much higher testing error for MLM-2 and MLM-3. Analysis of the learning curves shows that these two trained ANNR models have a high variance in

testing error [43], although the training error decreases with each iteration. Since the RFR shows a better accuracy than the ANNR model in both IEEE 14 and IEEE 118 bus systems, and exhibit good generalization compared to ANNR, the RFR is clearly the most appropriate algorithm for this problem.

When considering the results from both IEEE 14 and 118 bus test systems, predictions from MLM-2 which uses VSIs as the input features shows a better and promising overall performance (especially for the larger 118 bus system) than MLM-1 which directly uses voltage phasors, confirming the initial hypothesis. Moreover, MLM-2 achieves this with a lower number of inputs features; MLM-1 of 118 bus system uses 30 features (magnitudes and angles of 15 phasors) while MLM-2 uses 11 input features. Although the calculation of 11 features need all of phasor measurements, the intermediate step of calculating VSIs makes the MLM much simpler and more accurate. MLM-3 is better suited for practical implementation as it requires monitoring only the synchrophasor data from the most critical buses, but comes with a little sacrifice with regard to accuracy. MLM-4 and MLM-3 are comparable, but MLM-4 has a marginally better accuracy than MLM-3 for the 118 bus system.

5.4.1 Synchrophasor Measurement Errors

Noises, bias errors and communication errors (ex: PMU data packet losses) are general errors associated with field measurements. Noise and bias errors in synchrophasor measurements can be identified and corrected by data validation as proposed in [47]. Detailed investigation of data validation is out of the scope of this thesis. However

synchrophasor measurement errors cannot be corrected using data validation. Therefore the sensitivity of synchrophasor measurement errors to the LM prediction is analysed

Errors in Synchrophasor measurements may induce errors in the VSIs which could affect the LM prediction. According to IEEE standard C37.118-2011 [42] synchrophasors must maintain a total vector error (TVE) less than 1% under steady state conditions. The effect of synchrophasor measurement errors to LM prediction is analyzed by inputting erroneous voltage and current phasors which are generated by introducing random errors to magnitudes and phase angles such that TVE ≤1%. Voltage input for the i^{th} bus of interest with measurement errors can be represent as in (5.7) where ΔV_i and $\Delta\theta_i$ are the per unit magnitude and angle errors.

$$\begin{aligned} V_i^* &= V_i(1 + \Delta V_i) \\ \theta_i^* &= \theta_i(1 + \Delta\theta_i) \end{aligned} \tag{5.7}$$

The erroneous current phasors are generated in a similar fashion, and are used to calculate the VSIs which are fed to different MLMs trained using the RFR algorithm. Figure 5.3 compares the RMSE of different MLMs with and without the influence of synchrophasor measurement errors. MLM-3 shows a significant prediction error under the influence of synchrophasor measurement errors. The others show a relatively lower prediction error even with synchrophasor measurement errors.

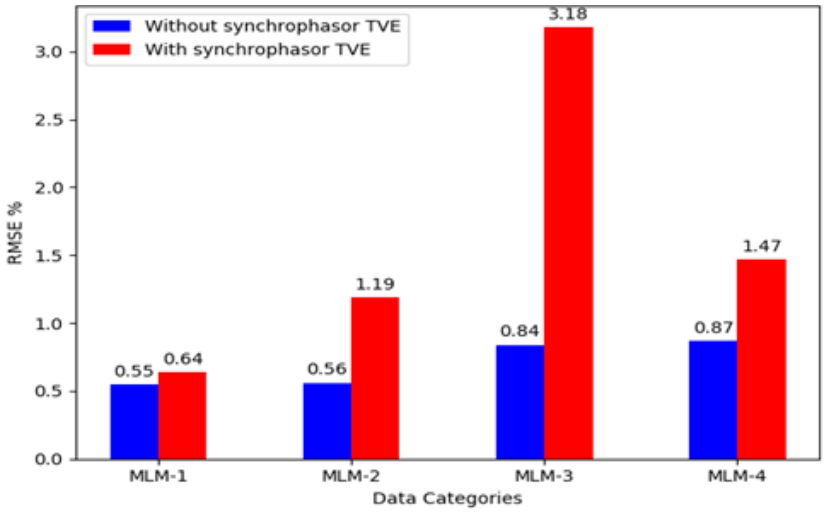


Figure 5.3. MSE of prediction from synchrophasor measurements from IEEE 14 bus system with and without synchrophasor measurement errors

5.4.2 Machine Learning Model Ensemble

In an MLM ensemble, multiple diverse models are used to predict a common outcome in order to improve the overall performance and robustness. Even though MLM-3 is more amenable for a practical implementation, MLM-1 shows higher resilience to synchrophasor measurement errors. MLM-4 on the other hand is more accurate for larger systems. Hence the strengths of different MLMs can be yielded through an ensemble modeling. The weighted average ensemble method is used to obtain the final prediction Y_{en} as shown in (5.8) using the predictions of individual MLM-s.

$$Y_{en} = \sum_{i \in E} w_i Y_{MLM-i} \tag{5.8}$$

where E is the set of models included in the ensemble, and w_i is the weightage which can be determined either empirically or using a grid search optimization method.

A LM prediction from an ensemble consisting of MLM-1, MLM-3, and MLM-4 is attempted. The results of the limited grid search conducted to find a good set of weights

for MLM-1, MLM-3, and MLM-4 are shown in Table 5-5. The best weightages, 0.5, 0.25 and 0.25, were selected considering the RMSE and R^2 for the testing data set of IEEE 14 and IEEE 118 bus systems.

TABLE 5-5 ENSEMBLE MODELS WITH DIFFERENT WEIGHTS

Test System	**Weightage ratio MLM1: MLM3: MLM4**	**RMSE**	**R^2**
IEEE 14	1:1:1	0.0057	0.9983
	2:1:1	0.0053	0.9985
	1:2:1	0.0061	0.9981
	1:1:2	0.0062	0.9980
	2:2:1	0.0056	0.9984
	2:1:2	0.0056	0.9983
	2:1:2	0.0062	0.9980
IEEE 118	1:1:1	2.8252	0.9725
	2:1:1	2.6935	0.9752
	1:2:1	2.8551	0.9717
	1:1:2	3.0425	0.9663
	2:2:1	2.7216	0.9748
	2:1:2	2.8537	0.9713
	2:1:2	2.9715	0.9686

It can be seen that the ensemble model shows a better performance than individual MLMs. Figure 5.4 shows an example case of an LM prediction for the IEEE 14 bus system under *n*-1 contingency situation (transmission line between bus-2 and -4 out of service). The predictions of different MLMs from the initial operating point to the voltage collapse point are shown. The scenario also includes a generator hitting its reactive power limit (when LM≈0.13) The predictions of the ensemble model are much closer to the target compared to the predictions of individual MLMs, some of which clearly deviate at the point of generator reaching its reactive power limit.

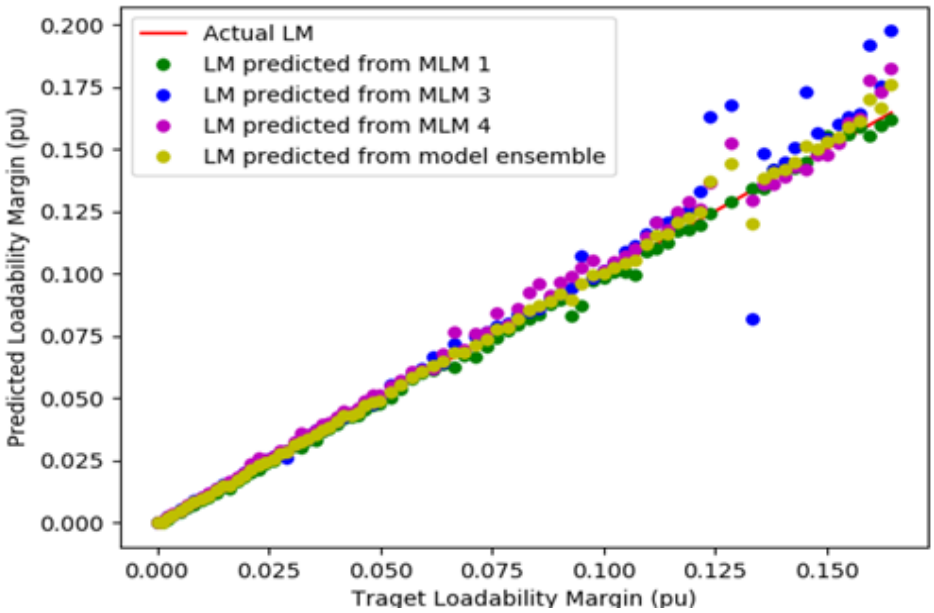


Figure 5.4. LM prediction from different MLMs and the model ensemble under the contingency of line tripping between bus 2 and 4 of IEEE 14 bus system

5.5 Summary

The generated data bases of IEEE 14 and IEEE 118 bus systems are used to train and test a set of MLMs to predict the LM. The most feasible input features for the MLMs are selected considering their correlation with the LM and through recursive feature elimination. Different multi-variant regression algorithms are evaluated and the, random forest regression was identified as the most appropriate predictor model to prediction the LM. The effects of synchrophasor measurement errors on the LM prediction are also evaluated. Finally, three different MLMs (MLM-1, MLM-3 and MLM-4) which were trained using different features are combined to form an ensemble to improve the overall performances and robustness.

Chapter 6

Online Voltage Stability Margin Prediction Testbed

6.1 Introduction

The proposed framework for predicting the LM using VSIs and voltage phasors of significant buses is implemented on an experimental testbed to demonstrate its applicability in real-time. This chapter presents details of this implementation and addresses some issues pertaining to real-time implementation.

6.2 Testbed Architecture

The experimental setup is implemented on the real-time synchrophasor application framework PhasorSmart®, and tested using an experimental setup consisting of RTDS® real-time simulator and a laboratory synchrophasor network. The RTDS system simulated the IEEE 14 bus system model in real-time and published the synchrophasor measurements. The loads are gradually increased during the real-time simulation to push the system towards voltage instability. VSIs are calculated and fed to the trained RFR model to predict the LM at each steady state operation condition. The results are displayed on the Grafana® visualization tool, which is a part of the PhasorSmart package.

6.2.1 Hardware Architecture

A schematic of the hardware architecture of the test bed is shown in Figure 6.1. The real-time simulator consist of a RTDS rack containing PB5 processor cards. The RTDS simulator emulate the power grid and generate voltage and current signals. These signals are converted to synchrophasor data frames compatible with the IEEE Std. C37.118.2 [42] using a GTNET_PMU communication card. Time stamps for these data frames are generated using a GTSYNC card connected to a SEL® 2407 satellite clock. A local area network (implemented using a RuggedCom utility grade Ethernet switch) connects the GTNET output with a PC that runs the PhasorSmart platform.

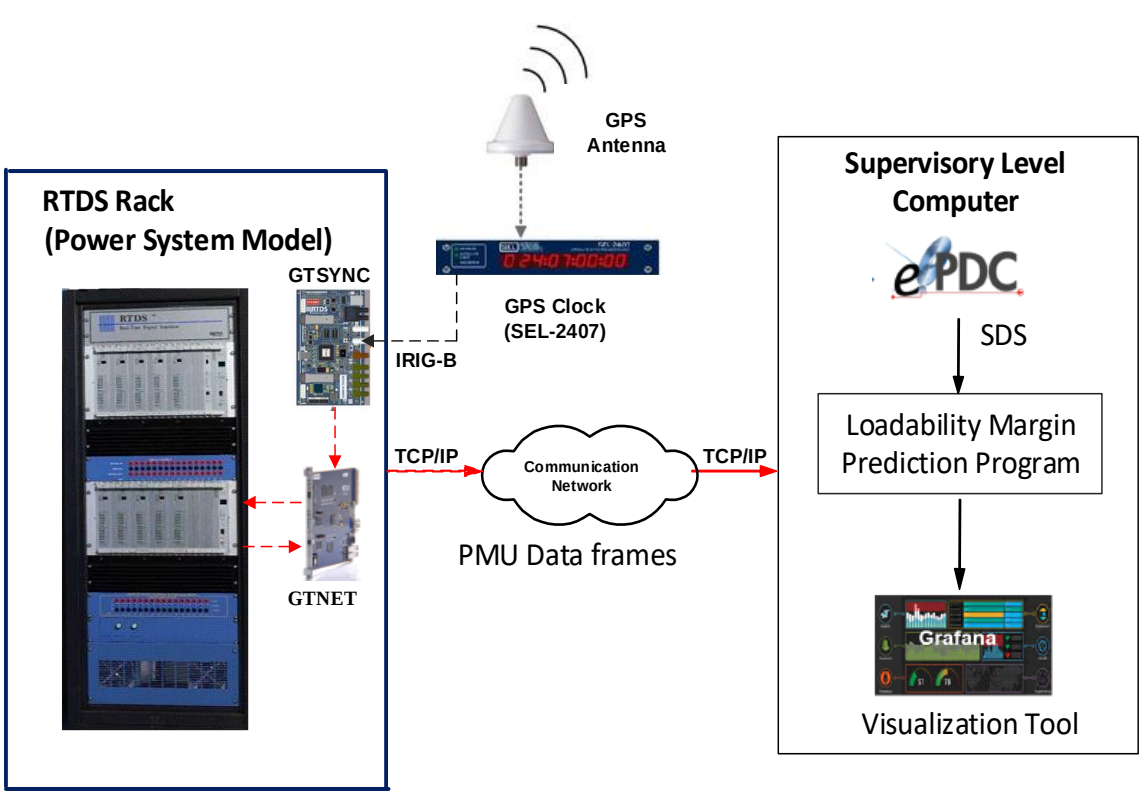


Figure 6.1. Real-time testbed arrangement

6.2.2 Software Architecture

Software architecture of the supervisory level computer contains three platforms; ePDC is the first platform which is a virtual Phasor Data Concentrator (PDC). It concentrates the PMU data published by the RTDS and feeds them to the Synchronous Data Server (SDS) running on the same computer, with a sampling rate of 30 samples per second. The second platform is the C++ programme which accesses the SDS and calculate the VSIs from those data. The trained random forest regression data is stored within the same programme which predict the LM from each individual MLM. Next, these individual MLM predictions are fed to the ensemble model to obtain the final prediction. This prediction along with other data is then distributed to the web based Grafana visualization tool. Grafana provides the operator a user friendly monitoring interface along with the alarming features when the system reaches the verge of voltage stability. The complete arrangement of the real-time voltage assessment testbed is shown in Figure 6.1

6.3 Transient Identification Using Wavelet Analysis

Power systems are dynamic in nature. However, it is required to avoid transient measurements when calculating VSIs to assess the long term voltage stability. In addition, local measurement based VSIs require voltage and current phasors of both the previous and the current steady state and hence, transient identification can be used to differentiate between the new and previous steady state. System transients may occur due to load changes, power system equipment tripping, faults and various other reasons. Wavelet analysis has been used as an efficient technique to analyses the transient phenomena. The wavelet transform decomposes a signal into basis functions that are localized in scale and

time. These decomposed signals can be used to identify the transients both online and offline studies. In this study the online Discrete Wavelet Transform (DWT) is implemented using Mallet tree algorithm as denoted in Figure 6.2. This algorithm consists of a series of low pass filters and their dual high pass filters as presented in [26]. The filter coefficients are defined based on three decomposition levels with "Haar" mother wavelet based Discrete Wavelet analysis.

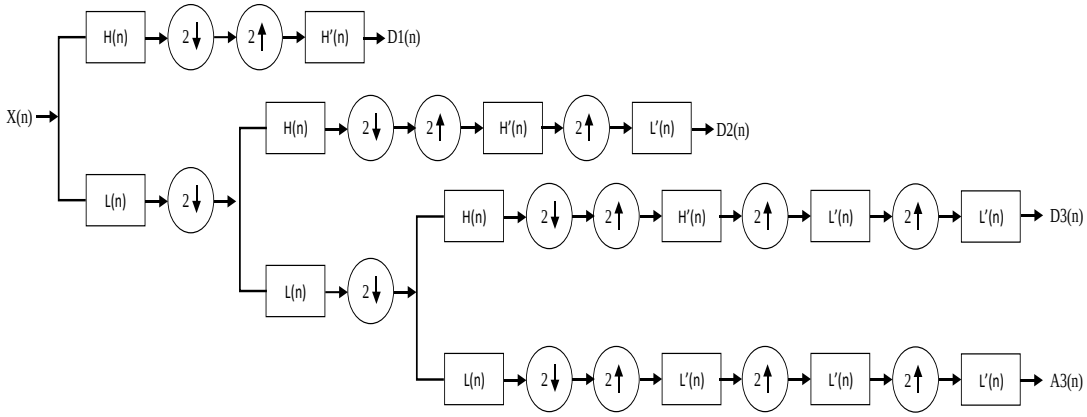


Figure 6.2. Mallet tree algorithm

Analysing the decomposed coefficients which are obtained from the DWT is not adequate to identify the transient period of the signal. Therefore in order to identify the transient period, the Mean Wavelet Energy (MWE) has been calculated as proposed in [44].The MWE is calculated using 8 sample data windows as mentioned in (6.1) where $d_{i,k}$ denotes the decomposed wavelet coefficient of the k^{th} sample of the data window with a total N number of samples of the i^{th} decomposition level.

$$\text{Mean Wavelet Energy} = \sqrt{\frac{1}{N}\sum_{k}^{N}\sum_{i}^{L}\left(d_{i,k}\right)^2} \tag{6.1}$$

The calculated MWE of the moving window is compared against a threshold to identify any transients within the window. A data window of 8 measurements (N=8) and a threshold of 0.005 pu is used in this study. When the measurements are free of transients, the local measurement based VSIs are calculated in the intervals of 128 cycles (2.13s).

At the beginning, the voltage and current phasors are measured to ensure if the system is at the steady state. Afterwards if the system changes from one steady state to another steady state the voltages and current phasors will be measured and saved to V_1 and I_1. The previous V_1 and I_1 values will be saved to V_0 and I_0. Hence it avoids calculating VSIs during transients. Therefore the LM prediction shows the previous LM during the transient and it will update the value when the transient is over. This process of transient identification and voltage and current measurement is shown in a flow diagram in Figure 6.3.

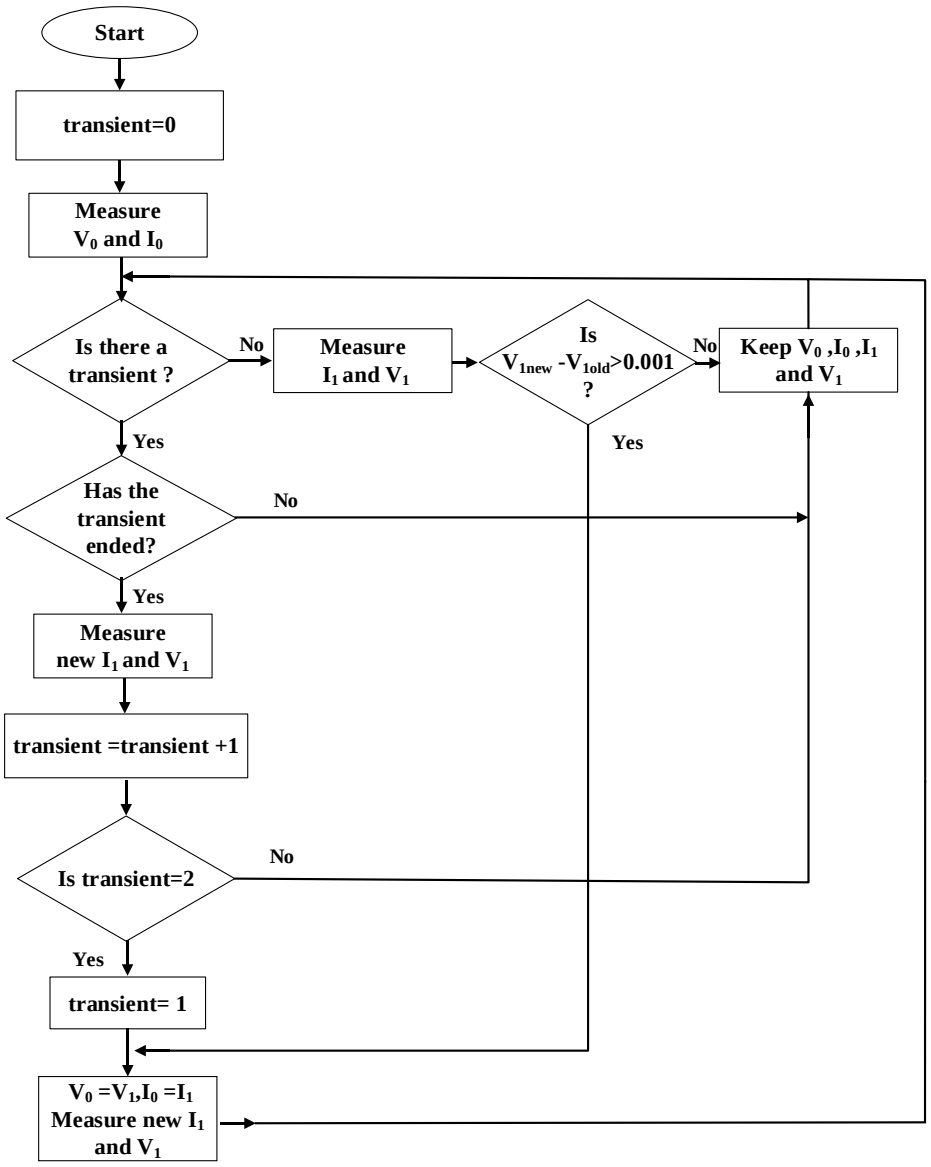


Figure 6.3. Flow chart of new steady state identification method

This process can be further explained using two scenarios. There is a gradual voltage reduction during time interval of 220s to 480s in Figure 6.4. However, it is not identified as a transient since the MWE spikes do not exceed the defined threshold. In such a scenario the voltage magnitude is monitored and if there is a difference exceeding 0.001 pu the

voltages and current phasors (V_1 and I_1) are updated to new values. The previous V_1 and I_1 values are saved to V_0 and I_0.

In the second scenario, the system voltage drops suddenly due to a sudden load increment at around 675s. For this event, the MWE spike has exceeded the defined threshold hence the system transient status monitor updates its output bit value from 0 to 1 and keeps it at 1 until the voltage magnitude becomes steady. When the system changes from one steady state to another steady state, the transient count increases. When the transient count reaches 2 the voltages and current phasors (V_1 and I_1) are updated to new values. The previous V_1 and I_1 values are saved to V_0 and I_0.

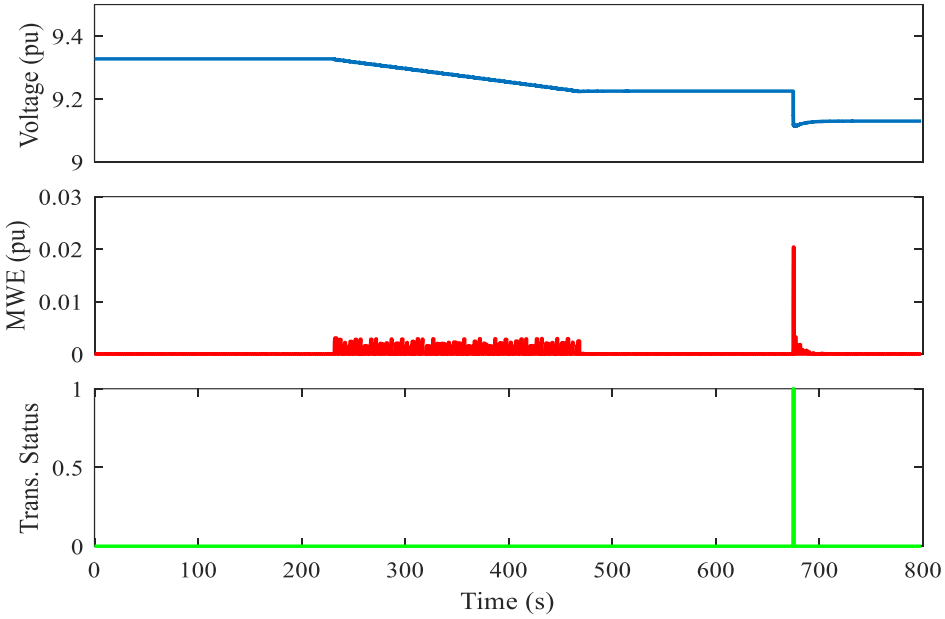


Figure 6.4. Transient identification using Wavelet analysis

6.4 Predicting the LM via trained ML Model

The trained RFR model is stored in to six separate text files, and when the program start running this data is retrieved to six separate arrays; Tolerance Value (T_{val}), Right child

(R_{child}), Left child (L_{child}), node feature (N_f), Leaf or non-leaf (L_n) and Leaf value (L_{Val}). Each array contains data that is required for a random forest to predict the LM. Each column of an array contains the values of each decision tree in the random forest. After the VSIs are calculated it is fed to the trained RFR model as shown in Figure 6.5. The predicted LM is obtained as the leaf values of the end node of each tree in the random forest. Then the average of the results from all the trees in the forest is taken as the predicted LM.

Table 6-1 compares the execution time of two techniques to obtain the LM: the proposed machine learning based method and the direct use of CPF. LM prediction through CPF is implemented in Matlab® using conventional continuation power flow algorithm and LM predictions through trained MLMs are obtained through a Python program. The execution times are computed by considering 10 random operating points because the computing time for CPF method is dependent on the operating point. The same 10 operating points were used for the both LM determination methods and the times reported in Table 6-1 are the average computing times. It can be observed that the proposed method is more than 5,000 times faster than the CPF method for IEEE 14-bus system, and it is more than 50,000 times faster for IEEE 118-bus system. These results also show the impracticality of using the CPF for on-line calculation of LM in case of large power systems. In this case, it takes 23 s to compute the LM of the IEEE 118-bus system. When the size of the system increases, time to obtain the LM also increases. However, the time taken by the proposed methodology still remains very low.

TABLE 6-1 EXECUTION TIMES OF DIFFERENT LM DETERMINATION METHODS WITH DIFFERENT SYSTEM SIZES

	IEEE 14 bus system	**IEEE 118 bus system**
Calculate LM trough CPF	0.5879 s	3.3079 s
Predict LM from trained MLM	131.4 μs	400 μs

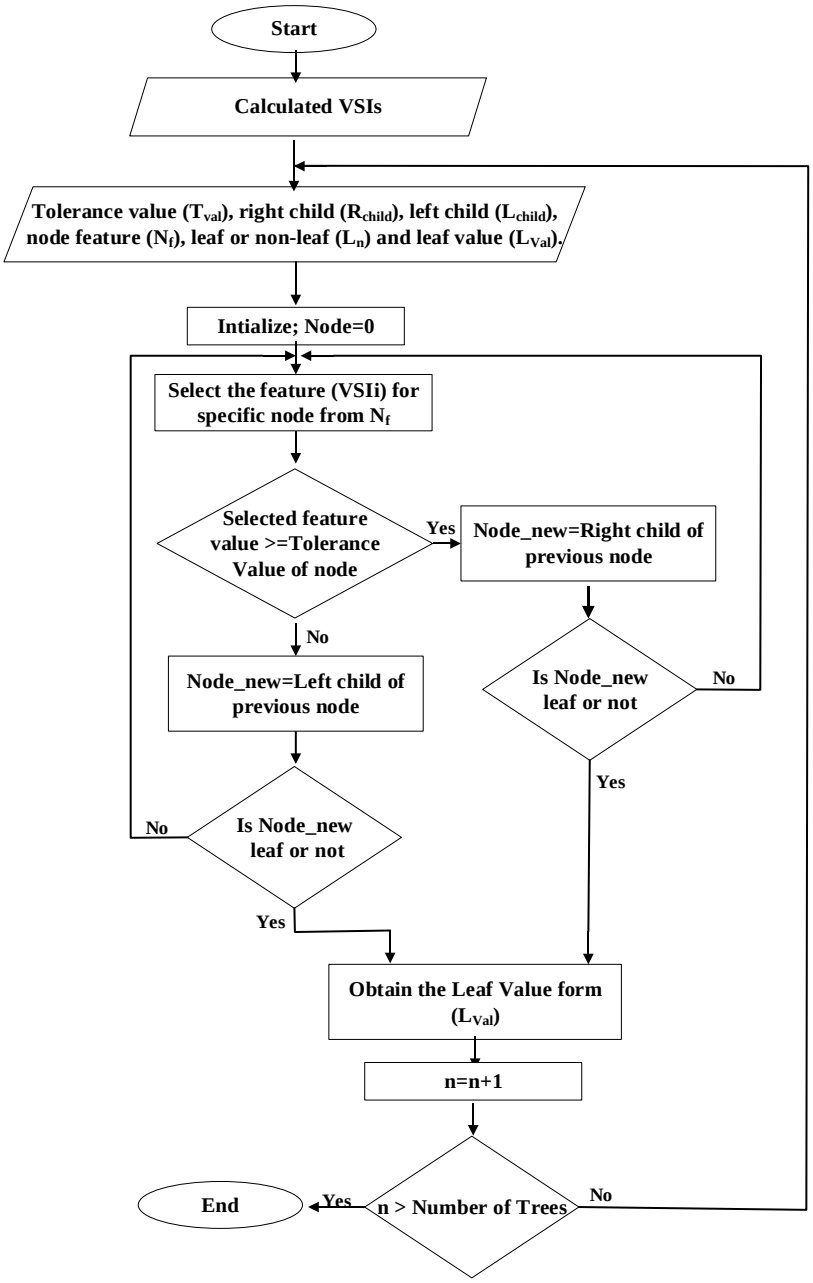


Figure 6.5. Flow chart of LM prediction using trained RFR model

6.5 Voltage Monitoring Dash-Board

A sample of the system voltage stability monitoring dash-board implemented on PhasroSmart® platform is shown in Figure 6.6. This provides a simple overview of the current status of the system voltage stability and its trend over a selected time window in the immediate past to the system operator. The visualization is customizable; the window shown in Figure 6.6 shows the real-time voltage magnitudes of the most critical buses with alarms for under and over voltage conditions. Furthermore, the LM predictions are visualized for the two most critical buses, bus 14 and bus 8. The dials showing the LM are divided in to three regions; normal region, alarm region and critical region. The set of graphs in the middle shows the recent trend of the predicted LM, and the bottom right graph shows the recent transient events. Multiple windows can be created for monitoring different regions in the network.

Figure 6.6. Real-time voltage stability monitoring dash-board

6.6 Real-Time Monitoring Results

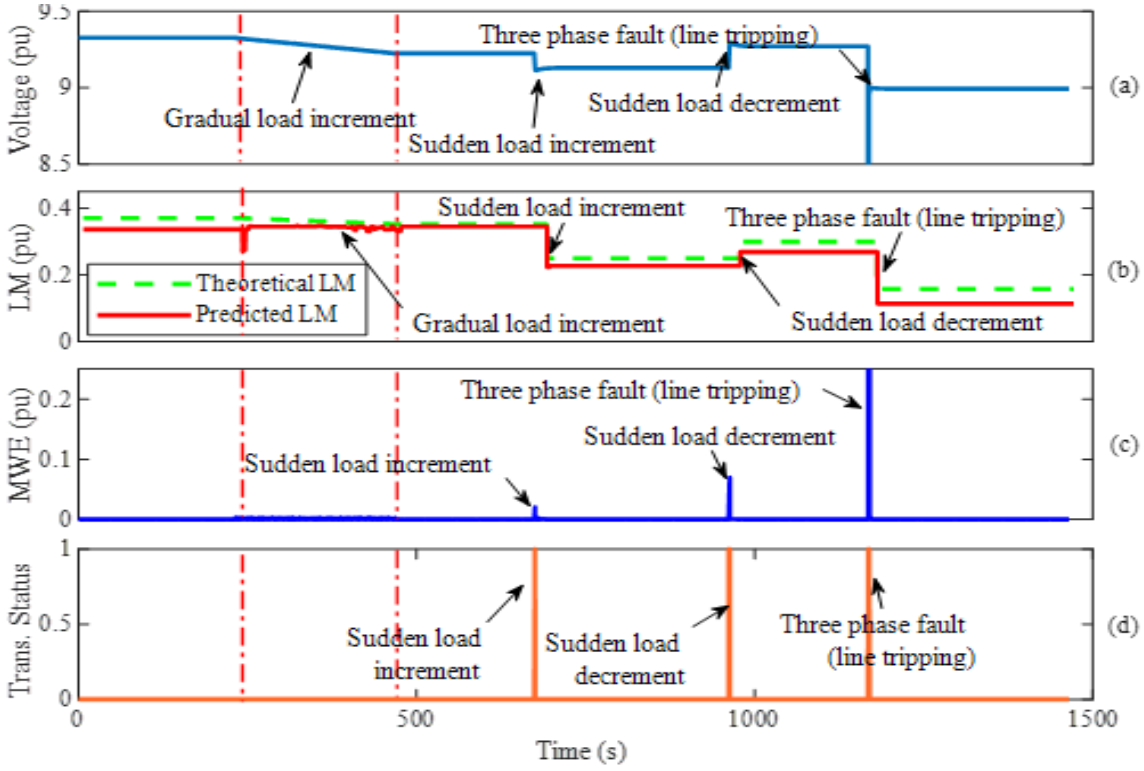


Figure 6.7. (a) Voltage profile of bus-14 of the IEEE 14 bus system (b) LM prediction form the ensemble ML model, (c) MWE variation (d) Transient status under different power system operational situations

Figure 6.7(a) shows the voltage profile of bus-14 of the IEEE 14 bus system. Figure 6.7(b) shows the LM prediction from the ensemble ML model together with the theoretical LM under the considered power system operating condition. The simulated scenario starts in a steady state, and then goes through a period of gradual load increment. After a brief period of steady operation, the system sees a sudden load increment followed by a sudden load decrement. The last part of the simulation correspond to a three phase fault with the consequent tripping of a transmission line.

The loads were increased or decreased by a percentage of scheduled active and reactive powers of the buses. For the gradual power increment, active and reactive power is

increased by steps of 0.001 pu in every 5s. In the case of sudden power increment, a 20% rise in the active and reactive power at bus 9 is introduced. When there is a sudden power decrement, a power decrement of 10% is introduced to the same bus. A three phase fault is applied on the line between bus-6 and bus-13 in close proximity to bus-6. The fault is set to persist for 4 cycles before it is cleared by tripping the line. The predicted LM is plotted with respect to the theoretical LM under all the operational conditions in Figure 6.7(b).

The spikes in MWE in Figure 6.7(c) corresponds to the transient changes such as sudden changes in voltage or faults. When MWE is above the defined threshold of 0.005, the bit value of the transient status changes from 0 to 1, and remains at 1 until the voltage magnitude becomes steady as shown in Figure 6.7(d). The predicted LM varies responding to the changes in the load. The voltage of bus-14 drops slightly after removing the line after the fault, but the predicted LM drops significantly, which agrees with the correct behavior of the voltage stability monitoring system.

6.7 Summary

The real-time monitoring results shows that the proposed method of predicting the power system LM through ensemble MLMs provide a prediction within an acceptable level of accuracy and time. The developed graphical user interface provides a voltage security level indication using the proposed methodology, enabling the system operators to clearly observe the LM and its trend, and then initiate the remedial actions if necessary.

Chapter 7

Conclusions and Future Work

7.1 Conclusions

A novel machine learning based approach to predict the long term voltage stability, as represented by the loadability margin, is presented in this thesis. The proposed method utilizes an ensemble of machine learning based regression models which uses a selected sets of voltage stability indices as input features for the MLMs. MLMs that use Random Forest regression has proven to be the most robust and accurate irrespective of the system size. Although the loadability margin can be predicted with a random forest regression model that directly uses voltage magnitudes and phase angles with a reasonable accuracy, the predictions obtained with voltage stability indices as inputs have a higher accuracy, especially for the larger systems. Furthermore, the accuracy and the robustness can be increased under both normal operating conditions and N-1 contingencies when an ensemble of machine leaning models that use different input features is used. With the ensemble approach, the correlation between the predicted and actual loadability margin could be improved to 99.7 % and 97.6% for the IEEE 14 bus and 118 bus test cases respectively. The practicality of the proposed method is demonstrated by implementing the scheme on PhasorSmart® real-time platform and testing it on a laboratory testbed based

on RTDS® real-time simulator. However, there are practical issues such as managing transients that need to be addressed in real-time implementation.

7.2 Contributions

The main research contribution of this thesis is an improved approach for real-time estimation of the loadability margin in selected locations of a power system using wide area synchrophasor measurements. While the proposed method has more complexity than previously published machine learning based approaches, it is better in terms of performance and feasibility when implemented as a real-time voltage stability assessment system. A methodical approach to develop the proposed loadability margin predictor is presented in this work. Over the course of research, a number of useful results, tools and methodologies were developed. They include:

1. A CPF program implemented using Matlab® to calculate the bus voltages of power systems up to the voltage collapse point. This CPF program take into account important aspects such as generator reactive power limits and automatic transformer tap changing, and is sufficiently generic to apply for any power system with conventional generation and power flow control devices.
2. A systematic methodology for generation of large datasets for machine learning and tools to automate commercial grade power system analysis software and the developed CPF program to compute the data. The large datasets generated for IEEE 14 bus and IEEE 118 bus test power system considering various operating conditions and contingencies are also available for future use for research.

3. Analysis of the correlations between various VSIs proposed by different authors and the LM in terms of Spearman's rank correlation coefficient. While individual VSIs have been proposed to serve as measures of closeness to voltage instability, their relationship to LM and other VSIs have not been analyzed or compared broadly. Some of the VSIs were found to be equivalent despite the differences in their definitions.
4. Ensemble approach to improve the accuracy and robustness of MLMs where MLMs with different input feature categories are combined. A data driven approach to select the weightages of different MLMs is also presented. Although not considered in this thesis, the same framework can be used to combine MLMs based on different machine learning algorithms to improve the robustness.
5. The real-time test system implemented in the intelligent power grid laboratory of the University of Manitoba, complete with a synchrophasor network, phasor data concentrator and data server, computational modules for estimating LM and visualization tools. The RTDS® real-time simulator with synchrophasor streaming capability and PhasorSmart® package that included a PDC, synchrophasor data server and the Graphana® visualization tool were found to be highly effective for testing new real-time synchrophasor applications such as the voltage stability assessment system proposed in this thesis. The setup developed can be used in future research.

7.3 Future Work

Further work is recommended as an extension to the wide range of studies presented in this thesis.

1. Further opportunities exist to improve this methodology to adapt for an hourly dispatch schedule according to the load prediction of the power system. Parallel computing can be used to simulate the power system under probable operational and contingency scenarios and generate a data set to train the fast ML models. These dedicated trained models can be used in real-time monitoring.
2. Modern day power system dynamic characteristic are becoming more and more affected by power converter outputs and advance control algorithms associated with the renewable energy sources. The effect of high renewable energy penetration to the power system towards the voltage collapse can be further studied. Robustness and the accuracy of the proposed voltage assessment methodology can be validated under high renewable energy penetration scenarios such as integration of a wind farm, solar PV array, battery storage etc.
3. This proposed voltage stability assessment technique can be used to initiate automatic remedial control actions such as shedding the loads to avoid cascaded tripping of transmission lines which may lead to system wide blackouts.

Appendix. A IEEE 14 bus system

This appendix provides the system parameters and steady state data of the system

Figure A.1 shows the IEEE 14 bus system network [39]. Table A-1 to Table A-4 contain the system data.

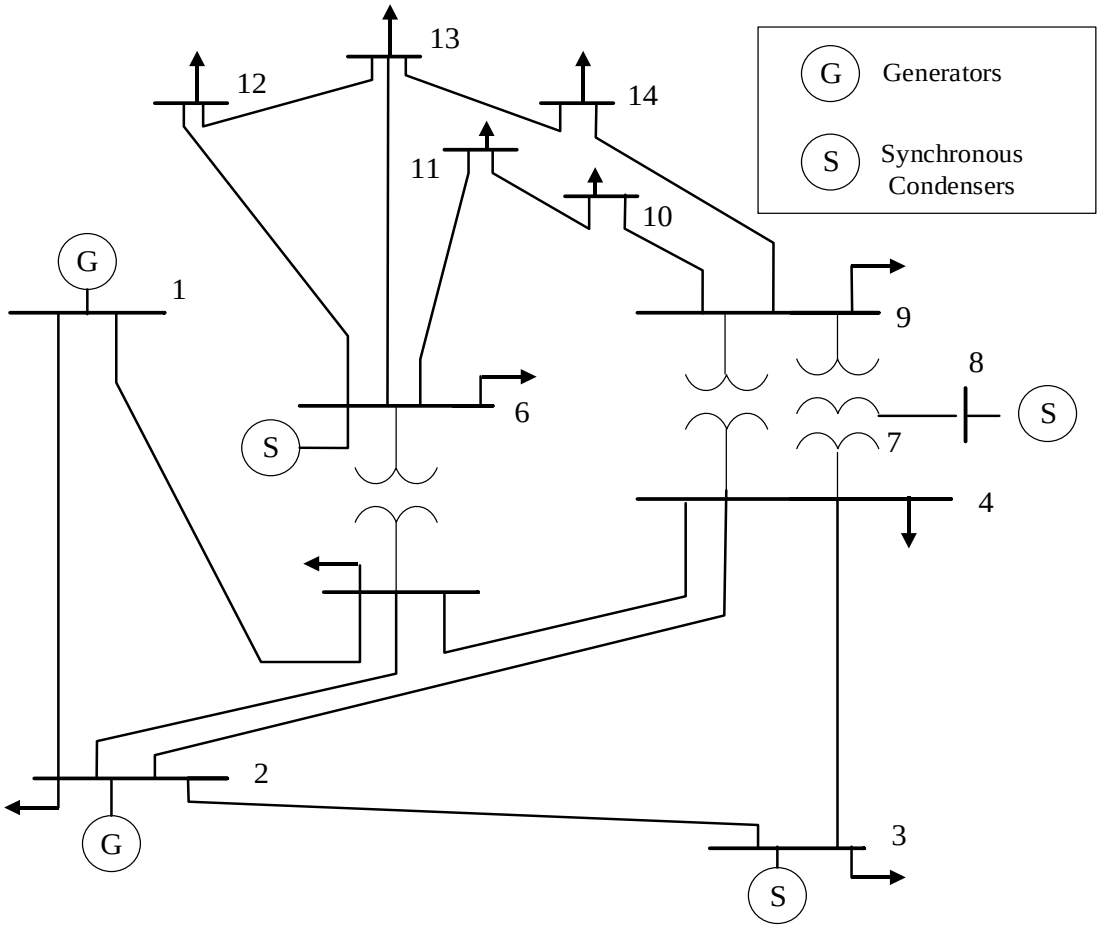


Figure A.1. IEEE 14 bus system

TABLE A-1. IEEE 14 BUS DATA

Bus Number	Type	$\|V\|$ (p.u)	Gen (MW)	Shunt (MVAr)	Load	
					P (MW)	Q (MVAr)
1	Slack	1.060	-	-	-	-
2	PV	1.045	40	-	21.7	12.7
3	PV	1.010	0	-	94.2	19.0
4	PQ	-	-	-	47.8	-3.9
5	PQ	-	-	-	7.6	1.6
6	PV	1.070	0	-	11.2	7.5
7	PQ	-	-	-	-	-
8	PV	1.090	0	-	-	-
9	PQ	-	-	190	29.5	16.6
10	PQ	-	-	-	9.0	5.8
11	PQ	-	-	-	3.5	1.8
12	PQ	-	-	-	6.1	1.6
13	PQ	-	-	-	13.5	5.8
14	PQ	-	-	-	14.9	5.0

TABLE A-2. IEEE 14 BUS SYSTEM GENERATOR REACTIVE POWER CAPABILITY

Bus Number	$\|V\|$ (p.u)	Q_{max} (MVAr)	Q_{min} (MVAr)
2	1.045	-40	50
3	1.010	0	40
6	1.070	-6	24
8	1.090	-6	24

TABLE A-3. IEEE 14 BUS SYSTEM TRANSFORMER DATA

From Bus	To Bus	R (p.u)	X (p.u)	Tap Ratio
4	7	0.0	0.20912	0.978
4	9	0.0	0.55618	0.969
5	6	0.0	0.25202	0.932
7	8	0.0	0.17615	1.00
7	9	0.0	0.11001	1.00

TABLE A-4. IEEE 14 BUS SYSTEM TRANSMISSION LINE DATA

From Bus	To Bus	R (p.u)	X (p.u)	B (p.u)
1	2	0.01938	0.05917	0.0264
1	5	0.05403	0.22304	0.0246
2	3	0.04699	0.19797	0.0219
2	4	0.05811	0.17632	0.0187
2	5	0.05695	0.17388	0.0170
3	4	0.06701	0.17103	0.0173
4	5	0.01335	0.04211	0.0064
6	11	0.09498	0.19890	-
6	12	0.12291	0.25581	-
6	13	0.06615	0.13027	-
9	10	0.03181	0.08450	-
9	14	0.12711	0.27038	-
10	11	0.08205	0.19207	-
12	13	0.22092	0.19988	-
13	14	0.17093	0.34802	-

Appendix. B IEEE 118 bus system

This appendix provides the system parameters and steady state data of the system

Figure B.1 shows the IEEE 118 bus system [40]. Table B-1 to Table B-4 contain the system data.

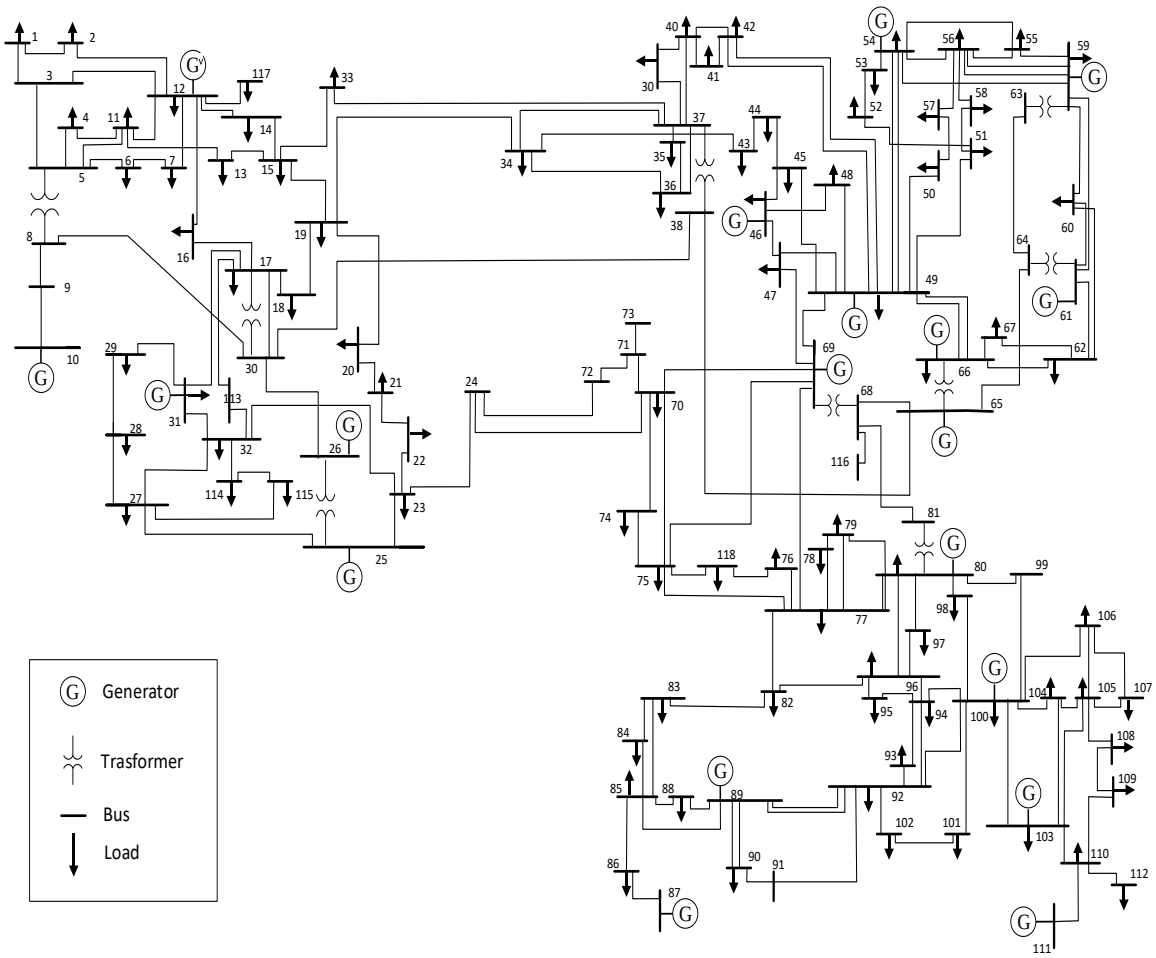


Figure B.1. IEEE 118 bus system

TABLE B-1. IEEE 118 BUS DATA

Bus Number	Type	$\|V\|$ (p.u)	Gen (MW)	Shunt (MVAr)	Load	
					P (MW)	Q (MVAr)
1	PV	0.9550	0	-	-	-
2	PQ	-	-	-	20	9
3	PQ	-	-	-	39	10
4	PV	0.9980	-	-	39	12
5	PQ	-	-	-40	-	-
6	PV	0.9900	0	-	52	22
7	PQ	-	-	-	19	2
8	PV	1.0150	0	-	28	0
9	PQ	-	-	-	-	-
10	PV	1.0500	450	-	-	-
11	PQ	-	-	-	70	23
12	PV	0.9900	85	-	47	10
13	PQ	-	-	-	34	16
14	PQ	-	-	-	14	1
15	PV	0.9700	0	-	90	30
16	PQ	-	-	-	25	10
17	PQ	-	-		11	3
18	PV	0.9730	0	-	60	34
19	PV	0.9634	0	-	45	25
20	PQ	-	-	-	18	3
21	PQ	-	-	-	14	8
22	PQ	-	-	-	10	5
23	PQ	-	-	-	7	3
24	PV	0.9920	0	-	13	0
25	PV	1.0500	220	-		
26	PV	1.0150	314	-		
27	PV	0.9680	0	-	71	13
28	PQ	-	-	-	17	7
29	PQ	-	-	-	24	4
30	PQ	-	-	-		
31	PV	0.9670	7	-	43	27
32	PV	0.9635	0	-	59	23
33	PQ	-	-	-	23	9
34	PV	0.9854	0	14	59	26
35	PQ	-	-	-	33	9
36	PV	0.9800	0	-	31	17
37	PQ	-	-	-25		

38	PQ	-	-	-		
39	PQ	-	-	-	27	11
40	PV	0.9700	0	-	66	23
41	PQ	-	-	-	37	10
42	PV	0.9850	0	-	96	23
43	PQ	-	-	-	18	7
44	PQ	-	-	10	16	8
45	PQ	-	-	10	53	22
46	PV	1.0050	19	10	28	10
47	PQ	-	-	-	34	0
48	PQ	-	-	15	20	11
49	PV	1.0250	204	-	87	30
50	PQ	-	-	-	17	4
51	PQ	-	-	-	17	8
52	PQ	-	-	-	18	5
53	PQ	-	-	-	23	11
54	PV	0.9550	48	-	113	32
55	PV	0.9520	0	-	63	22
56	PV	0.9540	0	-	84	18
57	PQ	-	-	-	12	3
58	PQ	-	-	-	12	3
59	PV	0.9850	155	-	277	113
60	PQ	-	-	-	78	3
61	PV	0.9950	160	-	-	-
62	PV	0.9980	0	-	77	14
63	PQ	-	-	-	-	-
64	PQ	-	-	-	-	-
65	PV	1.0050	391	-	-	-
66	PV	1.0500	392	-	39	18
67	PQ	-	-	-	28	7
68	PQ	-	-	-	-	-
69	Slack	1.0350	-	-	-	-
70	PV	0.9840	0	-	66	20
71	PQ	-	-	-	-	-
72	PV	0.9800	0	-	12	0
73	PV	0.9910	0	-	6	0
74	PV	0.9580	0	12	68	27
75	PQ	-	-	-	47	11
76	PV	0.9430	0	-	68	36
77	PV	1.0028	0	-	61	28
78	PQ	-	-	-	71	26
79	PQ	-	-	20	39	32
80	PV	1.0400	477	-	130	26
81	PQ	-	-	-	-	-

82	PQ	-	-	20	54	27
83	PQ	-	-	10	20	10
84	PQ	-	-	-	11	7
85	PV	0.9850	0	-	24	15
86	PQ	-	-	-	21	10
87	PV	1.0150	4	-	-	-
88	PQ	-	-	-	48	10
89	PV	1.0050	607	-	-	-
90	PV	0.9850	0	-	163	42
91	PV	0.9800	0	-	10	0
92	PV	0.9799	0	-	65	10
93	PQ	-	-	-	12	7
94	PQ	-	-	-	30	16
95	PQ	-	-	-	42	31
96	PQ	-	-	-	38	15
97	PQ	-	-	-	15	9
98	PQ	-	-	-	34	8
99	PV	1.0100	0	-	42	0
100	PV	1.0170	252	-	37	18
101	PQ	-	-	-	22	15
102	PQ	-	-	-	5	3
103	PV	1.0007	40	-	23	16
104	PV	0.9710	0	-	38	25
105	PV	0.9660	0	20	31	26
106	PQ	-	-	-	43	16
107	PV	0.9520	0	6	50	12
108	PQ	-	-	-	2	1
109	PQ	-	-	-	8	3
110	PV	0.9730	0	6	39	30
111	PV	0.9800	36	-	-	-
112	PV	0.9750	0	-	68	13
113	PV	0.9930	0	-	6	0
114	PQ	-	-	-	8	3
115	PQ	-	-	-	22	7
116	PV	1.0050	0	-	184	0
117	PQ	-	-	-	20	8
118	PQ	-	-	-	33	

TABLE B-2. IEEE 118 BUS SYSTEM GENERATOR REACTIVE POWER LIMITS

Bus Number	$\|V\|$ (p.u)	Q_{max} (MVAr)	Q_{min} (MVAr)
1	0.9550	-3.1057	15.0000
4	0.9980	-15.0637	300.0000
6	0.9900	15.9308	50.0000
8	1.0150	64.6537	300.0000
10	1.0500	-51.0423	200.0000
12	0.9900	91.3676	120.0000
15	0.9700	4.0217	30.0000
18	0.9730	26.0961	50.0000
19	0.9620	-8.0000	24.0000
24	0.9920	-15.7812	300.0000
25	1.0500	50.1066	140.0000
26	1.0150	11.1032	1000.0000
27	0.9680	3.0903	300.0000
31	0.9670	32.1872	300.0000
32	0.9630	-14.0000	42.0000
34	0.9840	-8.0000	24.0000
36	0.9800	1.2989	24.0000
40	0.9700	33.0461	300.0000
42	0.9850	60.7410	300.0000
46	1.0050	-4.2639	100.0000
49	1.0250	123.2477	210.0000
54	0.9550	31.1925	300.0000
55	0.9520	4.9692	23.0000
56	0.9540	14.2346	15.0000
59	0.9850	72.9174	180.0000
61	0.9950	-38.0212	300.0000
62	0.9980	2.4615	20.0000
65	1.0050	87.7531	200.0000
66	1.0500	-15.1923	200.0000
69	1.0350	-87.3266	300.0000
70	0.9840	11.7162	32.0000
72	0.9800	-11.0497	100.0000
73	0.9910	9.6874	100.0000
74	0.9580	-4.4850	9.0000
76	0.9430	7.5765	23.0000
77	1.0060	70.0000	70.0000
80	1.0400	60.4292	280.0000
85	0.9850	3.4550	23.0000

87	1.0150	11.0216	1000.0000
89	1.0050	-5.8333	300.0000
90	0.9850	81.7106	300.0000
91	0.9800	-4.3001	100.0000
92	0.9900	9.0000	9.0000
99	1.0100	-17.1197	100.0000
100	1.0170	123.2887	155.0000
103	1.0100	40.0000	40.0000
104	0.9710	5.6519	23.0000
105	0.9650	-8.0000	23.0000
107	0.9520	5.6951	200.0000
110	0.9730	4.8604	23.0000
111	0.9800	-1.8438	1000.0000
112	0.9750	41.5117	1000.0000
113	0.9930	7.0447	200.0000
116	1.0050	51.8787	1000.0000

TABLE B-3. IEEE 118 BUS SYSTEM TRANSFORMER DATA

From Bus	To Bus	R (p.u)	X (p.u)	Tap Ratio
5	8	0.0	0.026700	0.9850
17	30	0.0	0.038800	0.9600
25	26	0.0	0.038200	0.9600
37	38	0.0	0.037500	0.9350
59	63	0.0	0.038600	0.9600
61	64	0.0	0.026800	0.9850
65	66	0.0	0.037000	0.9350
68	69	0.0	0.037000	0.9350
80	81	0.0	0.037000	0.9350

TABLE B-4.IEEE 118 BUS SYSTEM TRANSMISSION LINE DATA

From Bus	To Bus	R (p.u)	X (p.u)	B (p.u)
1	2	0.0303	0.0999	0.0254
1	3	0.0129	0.0424	0.01082
2	12	0.0187	0.0616	0.01572
3	5	0.0241	0.108	0.0284
3	12	0.0484	0.16	0.0406
4	5	0.00176	0.00798	0.0021
4	11	0.0209	0.0688	0.01748
5	6	0.0119	0.054	0.01426

5	11	0.0203	0.0682	0.01738
6	7	0.00459	0.0208	0.0055
7	12	0.00862	0.034	0.00874
8	9	0.00244	0.0305	1.162
8	30	0.00431	0.0504	0.514
9	10	0.00258	0.0322	1.23
11	12	0.00595	0.0196	0.00502
11	13	0.02225	0.0731	0.01876
12	14	0.0215	0.0707	0.01816
12	16	0.0212	0.0834	0.0214
12	117	0.0329	0.14	0.0358
13	15	0.0744	0.2444	0.06268
14	15	0.0595	0.195	0.0502
15	17	0.0132	0.0437	0.0444
15	19	0.012	0.0394	0.0101
15	33	0.038	0.1244	0.03194
16	17	0.0454	0.1801	0.0466
17	18	0.0123	0.0505	0.01298
17	31	0.0474	0.1563	0.0399
17	113	0.00913	0.0301	0.00768
18	19	0.01119	0.0493	0.01142
19	20	0.0252	0.117	0.0298
19	34	0.0752	0.247	0.0632
20	21	0.0183	0.0849	0.0216
21	22	0.0209	0.097	0.0246
22	23	0.0342	0.159	0.0404
23	24	0.0135	0.0492	0.0498
23	25	0.0156	0.08	0.0864
23	32	0.0317	0.1153	0.1173
24	70	0.00221	0.4115	0.10198
24	72	0.0488	0.196	0.0488
25	27	0.0318	0.163	0.1764
26	30	0.00799	0.086	0.908
27	28	0.01913	0.0855	0.0216
27	32	0.0229	0.0755	0.01926
27	115	0.0164	0.0741	0.01972
28	29	0.0237	0.0943	0.0238
29	31	0.0108	0.0331	0.0083
30	38	0.00464	0.054	0.422
31	32	0.0298	0.0985	0.0251
32	113	0.0615	0.203	0.0518
32	114	0.0135	0.0612	0.01628
33	37	0.0415	0.142	0.0366
34	36	0.00871	0.0268	0.00568

34	37	0.00256	0.0094	0.00984
34	43	0.0413	0.1681	0.04226
35	36	0.00224	0.0102	0.00268
35	37	0.011	0.0497	0.01318
37	39	0.0321	0.106	0.027
37	40	0.0593	0.168	0.042
38	65	0.00901	0.0986	1.046
39	40	0.0184	0.0605	0.01552
40	41	0.0145	0.0487	0.01222
40	42	0.0555	0.183	0.0466
41	42	0.041	0.135	0.0344
42	49	0.0715	0.323	0.086
43	44	0.0608	0.2454	0.06068
44	45	0.0224	0.0901	0.0224
45	46	0.04	0.1356	0.0332
45	49	0.0684	0.186	0.0444
46	47	0.038	0.127	0.0316
46	48	0.0601	0.189	0.0472
47	49	0.0191	0.0625	0.01604
47	69	0.0844	0.2778	0.07092
48	49	0.0179	0.0505	0.01258
49	50	0.0267	0.0752	0.01874
49	51	0.0486	0.137	0.0342
49	54	0.0869	0.291	0.073
49	66	0.018	0.0919	0.0248
49	69	0.0985	0.324	0.0828
50	57	0.0474	0.134	0.0332
51	52	0.0203	0.0588	0.01396
51	58	0.0255	0.0719	0.01788
52	53	0.0405	0.1635	0.04058
53	54	0.0263	0.122	0.031
54	55	0.0169	0.0707	0.0202
54	56	0.00275	0.00955	0.00732
54	59	0.0503	0.2293	0.0598
55	56	0.00488	0.0151	0.00374
55	59	0.04739	0.2158	0.05646
56	57	0.0343	0.0966	0.0242
56	58	0.0343	0.0966	0.0242
56	59	0.0803	0.239	0.0536
59	60	0.0317	0.145	0.0376
59	61	0.0328	0.15	0.0388
60	61	0.00264	0.0135	0.01456
60	62	0.0123	0.0561	0.01468
61	62	0.00824	0.0376	0.0098

62	66	0.0482	0.218	0.0578
62	67	0.0258	0.117	0.031
63	64	0.00172	0.02	0.216
64	65	0.00269	0.0302	0.38
65	68	0.00138	0.016	0.638
66	67	0.0224	0.1015	0.02682
68	81	0.00175	0.0202	0.808
68	116	0.00034	0.00405	0.164
69	70	0.03	0.127	0.122
69	75	0.0405	0.122	0.124
69	77	0.0309	0.101	0.1038
70	71	0.00882	0.0355	0.00878
70	74	0.0401	0.1323	0.03368
70	75	0.0428	0.141	0.036
71	72	0.0446	0.18	0.04444
71	73	0.00866	0.0454	0.01178
74	75	0.0123	0.0406	0.01034
75	77	0.0601	0.1999	0.04978
75	118	0.0145	0.0481	0.01198
76	77	0.0444	0.148	0.0368
76	118	0.0164	0.0544	0.01356
77	78	0.00376	0.0124	0.01264
77	80	0.0294	0.105	0.0228
77	82	0.0298	0.0853	0.08174
78	79	0.00546	0.0244	0.00648
79	80	0.0156	0.0704	0.0187
80	96	0.0356	0.182	0.0494
80	97	0.0183	0.0934	0.0254
80	98	0.0238	0.108	0.0286
80	99	0.0454	0.206	0.0546
82	83	0.0112	0.03665	0.03796
82	96	0.0162	0.053	0.0544
83	84	0.0625	0.132	0.0258
83	85	0.043	0.148	0.0348
84	85	0.0302	0.0641	0.01234
85	86	0.035	0.123	0.0276
85	88	0.02	0.102	0.0276
85	89	0.0239	0.173	0.047
86	87	0.02828	0.2074	0.0445
88	89	0.0139	0.0712	0.01934
89	90	0.0238	0.0997	0.106
89	92	0.0393	0.1581	0.0414
90	91	0.0254	0.0836	0.0214
91	92	0.0387	0.1272	0.03268

92	93	0.0258	0.0848	0.0218
92	94	0.0481	0.158	0.0406
92	100	0.0648	0.295	0.0472
92	102	0.0123	0.0559	0.01464
93	94	0.0223	0.0732	0.01876
94	95	0.0132	0.0434	0.0111
94	96	0.0269	0.0869	0.023
94	100	0.0178	0.058	0.0604
95	96	0.0171	0.0547	0.01474
96	97	0.0173	0.0885	0.024
98	100	0.0397	0.179	0.0476
99	100	0.018	0.0813	0.0216
100	101	0.0277	0.1262	0.0328
100	103	0.016	0.0525	0.0536
100	104	0.0451	0.204	0.0541
100	106	0.0605	0.229	0.062
101	102	0.0246	0.112	0.0294
103	104	0.0466	0.1584	0.0407
103	105	0.0535	0.1625	0.0408
103	110	0.03906	0.1813	0.0461
104	105	0.00994	0.0378	0.00986
105	106	0.014	0.0547	0.01434
105	107	0.053	0.183	0.0472
105	108	0.0261	0.0703	0.01844
106	107	0.053	0.183	0.0472
108	109	0.0105	0.0288	0.0076
109	110	0.0278	0.0762	0.0202
110	111	0.022	0.0755	0.02
110	112	0.0247	0.064	0.062
114	115	0.0023	0.0104	0.00276

Appendix. C Machine Learning Models

C.1 Linear Regression

The Linear regression model describes the linear relationship between the dependent variable and the predictor variable. Since many features influence the final prediction, a multi-variant linear model is considered and it can be represented by (C.1) where optimal values for $\beta_0, \beta_1, \ldots \beta_p$ can be obtained by solving the objective function (C.2).

$$y_i = \beta_0 + \beta_1 x_{1i} + \beta_2 x_{2i} \ldots\ldots \beta_p x_{pi} \tag{C.1}$$

$$min: \sum_{i=1}^{n} \left(y_i - \beta_0 + \beta_1 x_{1i} + \beta_2 x_{2i} \ldots\ldots \beta_p x_{pi}\right)^2 \tag{C.2}$$

There are many approaches developed to estimate parameters in linear regression. These methods differ from each other depending on the simplicity of the algorithm, robustness and efficacy. Among different forms of parameter estimators such as least squares estimator, maximum-likelihood estimator and Bayesian estimator, the most common estimation technique for linear regression is linear least square estimation. The following algorithm denotes main steps of the least square estimation technique [33],[63].

Algorithm:

Input: Data $\left\{\{x_{ik}\}_{k=1}^{p}, y_i\right\}_{i=1}^{n}$

Step 1: Input data is written in matrix form as $X\beta = y$

Where

$$X = \begin{bmatrix} 1 & X_{11} & X_{12} & \cdots & X_{1p} \\ 1 & X_{21} & X_{22} & \cdots & X_{2p} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & X_{n1} & X_{n2} & \cdots & X_{np} \end{bmatrix} \quad , \quad \beta = \begin{bmatrix} \beta_0 \\ \beta_1 \\ \vdots \\ \beta_p \end{bmatrix} \quad , \quad y = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix}$$

Step 2: This minimization problem has unique solutions. This method minimizes the sum of squared residuals which leads to estimate the unknown parameter vectorβ.

$$\hat{\beta} = (X^T X)^{-1} X^T y$$

C.2 Gradient Boost Regression

Gradient boost regression is an ensemble of decision trees. The structure of the decision tree is shown in Figure C.1. Similar to other boosting techniques the gradient boost regression is built in a stage wise manner and it generalizes them by using a differentiable loss function. The loss function contains the desired response or the target output (y_i) and the predicted value (F(x)). The following algorithm contains the procedure of gradient boost regression learner [34],[64].

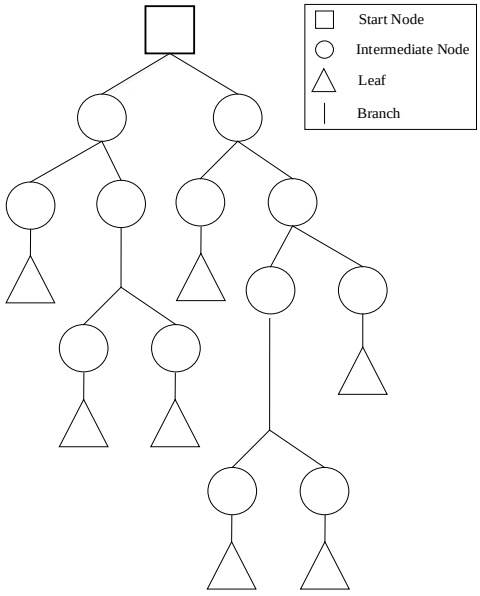


Figure C.1. Decision tree architecture

Algorithm:

Input: Data $\left\{\{x_{ik}\}_{k=1}^{p}, y_i\right\}_{i=1}^{n}$ and differentiable loss function $L\left(y_i, F(x)\right)$

Step 1: Initiate model with a constant value $F_0(x) = argmin \sum_{i=1}^{n} L(y_i, \gamma)$ where γ is the prediction

Step 2: for m=1 to M, where M is the number of trees

(a) Compute residual $r_{im} = -\left[\frac{\partial L(y_i, F(x_i))}{\partial F(x_i)}\right]_{F(x)=F_{m-1}(x)}$

(b) Fit the regression tree to the r_{im} values and create terminal regions R_{jm} for j=1…J_m

(c) For j=1…J_m compute $\gamma_{jm} = argmin \sum_{x \in R_{ij}} L(y_i, F_{m-1}(x) + \gamma)$

(d) Update $F_m(x) = F_{m-1}(x) + \upsilon \sum_{j=1}^{J_m} \gamma_{jm} I(x \in R_{jm})$ where υ is the learning rate

Step 3: Output $F_m(x)$

C.3 Artificial Neural Network Regression

Artificial neural network is the most widely used machine learning technique. There are numerous types of ANN in usage. Among them, the multi-layer feed forward network is the most commonly used ANN technique. In this technique there are three main layers of neurons, namely input layer, hidden layer and output layer as shown in Figure C.2. Every neuron in a particular layer is interconnected to all the neurons of the consecutive layer. Every interconnection hold a weightage factor which is fine tuned using back-propagation method [35],[65].

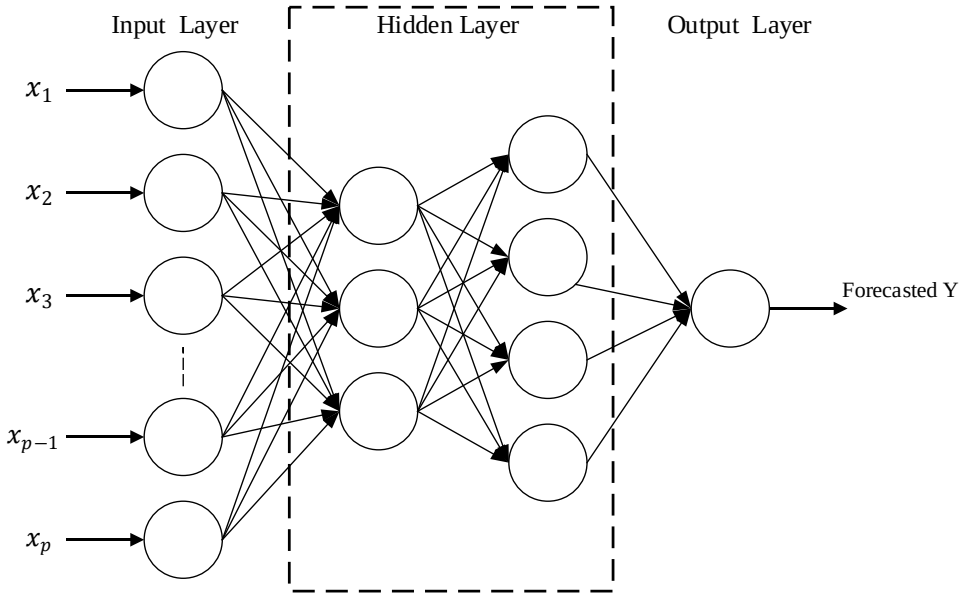


Figure C.2. Multilayer neural network architecture

In this technique the sum of all input to a specific neuron can be expressed as (C.3)

$$net_j = \sum_{i=1}^{N_d} x_i \, \omega_{ji} + \omega_{j0} \tag{C.3}$$

Since every neuron in the ANN contain an activation function the output of the j^{th} neuron of the hidden layer can be expressed as denoted in (C.4)

$$y_j = \frac{1}{1 + e^{-net_j}} \tag{C.4}$$

The learning algorithm of the ANN is based on an iterative technique which adjust the weightages and biases iteratively. Back-propagation is the most commonly used learning algorithm for ANN and it is explained in the section below.

Algorithm

Input: Data $\left\{\{x_{ik}\}_{k=1}^{p}, y_i\right\}_{i=1}^{n}$

Step 1: Initialize the weights in the network

Step 2: for i=1 to n

(a) Forward pass

Feed the input data to the neural network and using the forecasted Y, calculate the error (E) using differentiable loss function: $L\big(y_i, F(x)\big)$

(b) Backward pass

Compute $\Delta\omega$ for all the weights of the hidden layer based on gradient descent $\Delta\omega = -\eta \frac{\partial \Delta E}{\partial \Delta \omega}$ where η is the learning rate.

(c) Update the weights in the network $\omega' = \omega + \Delta\omega$

Step 3: Output

C.4 Random Forest Regression

Random forest regression is an ensemble learning method formed by a collection of decision trees. The inhabit nature of overfitting (low bias however, high variance) in decision trees is mitigated in this technique. Random Forest Regression average the value of several deep decision trees which are trained using different subsets of the same data set [36],[66].

The training algorithm for random forest regression includes bootstrap aggregating. If there is a training set of $\{\{x_{ik}\}_{k=1}^{p}, y_i\}_{i=1}^{n}$ bagging repeatedly will select a set of random sample instances from the training data set. Main steps of the RFR learner is as follows.

Algorithm

Input: Data $\{\{x_{ik}\}_{k=1}^{p}, y_i\}_{i=1}^{n}$

Step 1: for b=1 to B where B is the number of bagging

(d) Sample with replacement and obtain bootstrap data set with vector X_b and Y_b
(e) Train the regression tree on X_b, Y_b

Step 2: Output $\hat{f} = \frac{1}{B}\sum_{b=1}^{B} f_b(x')$

References

[1] C. V. T. Van Cutsem, Voltage Stability of Electric Power Systems. 1998.

[2] P. Kundur, Power System Stability and Control. McGraw-Hill, INC, 1994.

[3] V. Ajjarapu and C. Christy, "The continuation power flow: a tool for steady state voltage stability analysis," [Proceedings] Conference Papers 1991 Power Industry Computer Application Conference, Baltimore, MD, USA, 1991, pp. 304-311

[4] N. Hatziargyriou, J. Milanovi´c, C. Rahmann, V. Ajjarapu, C. Ca˜nizares,I. Erlich, D. Hill, I. Hiskens, I. Kamwa, B. Palet al., "Stabilitydefinitions and characterization of dynamic behavior in systems withhigh penetration of power electronic interfaced technologies," IEEEPower & Energy Society, Tech. Rep., 2020..

[5] H. W. K. M. Amarasekara, L. Meegahapola, A. P. Agalgaonkar, and S. Perera, "Impact of Renewable Power Integration on VQ Stability Margin", Australian Universities Power Eng. Conf.,AUPEC 2013,Hobart,TAS, Australia, 2013.

[6] Chandra, A. K. Pradhan, and A. K. Sinha, "A comparative study of voltage stability indices used for power system operation," Int. Conf. 21st dy Energy Needs - Mater. Syst. Appl. ICTFCEN 2016, pp. 1–4, 2017.

[7] H. Liu, "Real-Time Dynamic Voltage Stability Assessment Through a Wide-Area Loss Index," IEEE Elect. Power and Energy Conf. 2018,Toronto,Canada,2018.

[8] Y. M. Sangwook Han, Byongjun Lee, Sangtae Kim, "Real Time Wide Area Voltage Stability Index in the Korean Metropolitan Area," Electr. Eng. Technol., vol. 4, no. 4, pp. 451–456, 2009.

[9] M. Glavic and T. Van Cutsem, "Wide-Area Detection of Voltage Instability From Synchronized Phasor Measurements. Part II: Simulation Results," IEEE Trans. Power Syst., vol. 24, no. 3, pp. 1417–1425, 2009.

[10] J. Modarresi, E. Gholipour, and A. Khodabakhshian, "A comprehensive review of the voltage stability indices," Renew. Sustain. Energy Rev., vol. 63, pp. 1–12, 2016.

[11] Pérez-Londoño, L. F. Rodríguez, and G. Olivar, “A simplified voltage stability index (SVSI),” Int. J. Electr. Power Energy Syst., vol. 63, pp. 806–813, 2014.

[12] R. R. Matavalam and V. Ajjarapu, “Calculating the long term voltage stability margin using a linear index,” IEEE Power Energy Soc. Gen. Meet., vol. 2015-Septe, no. 2, 2015.

[13] H. Mohammadi and M. Dehghani, “PMU based voltage security assessment of power systems exploiting principal component analysis and decision trees,” Int. J. Electr. Power Energy Syst., vol. 64, pp. 655–663, 2015.

[14] V. Vittal et al., “Decision Tree-Based Online Voltage Security Assessment Using PMU Measurements,” IEEE Trans. Power Syst., vol. 24, no. 2, pp. 832–839, 2009.

[15] S. M. Shahrtash and H. Khoshkhoo, “On-line dynamic voltage instability prediction based on decision tree supported by a wide-area measurement system,” IET Gener. Transm. Distrib., vol. 6, no. 11, pp. 1143–1152, 2012

[16] D. Q. Zhou, U. D. Annakkage, and A. D. Rajapakse, “Online monitoring of voltage stability margin using an artificial neural network,” IEEE Trans. Power Syst., vol. 25, no. 3, pp. 1566–1574, 2010.

[17] Ren, Y. Xu, Y. Zhang and R. Zhang, "A Hybrid Randomized Learning System for Temporal-Adaptive Voltage Stability Assessment of Power Systems," in IEEE Transactions on Industrial Informatics, vol. 16, no. 6, pp. 3672-3684, June 2020

[18] Y. Zhang, Y. Xu, Z. Y. Dong and R. Zhang "A Hierarchical Self-Adaptive Data-Analytics Method for Real-Time Power System Short-Term Voltage Stability Assessment" in IEEE Transaction on Industrial Informatics. vol. 15, no. 1, pp. 74–84, 2020

[19] S. Li and V. Ajjarapu, "Real-time monitoring of long-term voltage stability via convolutional neural network," 2017 IEEE Power & Energy Society General Meeting, Chicago, IL, 2017, pp. 1-5, PESGM.2017

[20] M. Suganyadevi, Babulal,"Support vector regression model for the prediction of loadability margin of a power system". Appl Soft Comput 24:304–315

[21] K. Sajan, K. Vishal , Barjeev,"Genetic algorithm based support vector machine for on-line voltage stability monitoring". Electr Power Energy Syst 73:200–208

[22] Q. Huang, M. Zhou, Y. Zhang and Z. Wu, "Exploiting cloud computing for power system analysis," 2010 International Conference on Power System Technology, Hangzhou, 2010, pp. 1-6

[23] F. Ma, X. Luo and E. Litvinov, "Cloud Computing for Power System Simulations at ISO New England—Experiences and Challenges," in IEEE Transactions on Smart Grid, vol. 7, no. 6, pp. 2596-2603, Nov. 2016

[24] J. H. Liu and C. C. Chu, "Wide-area measurement-based voltage stability indicators by modified coupled single-port models," IEEE Trans. Power Syst., vol. 29, no. 2, pp. 756–764, 2014.

[25] T. K. Abdul Rahman and G. B. Jasmon, "New technique for voltage stability analysis in a power system and improved loadflow algorithm for distribution network," Proc. Int. Conf. Energy Manag. Power Deliv. EMPD, vol. 2, pp. 714–719, 1995.

[26] L. Wang, Y. Liu, and Z. Luan, "Power transmission paths based voltage stability assessment" . In: Proceedings of the transmission and distribution conference and exhibition: Asia and Pacific,. pp. 1–5, 2005.

[27] Milǒsević and M. Begović, "Voltage-stability protection and control using a wide-area network of phasor measurements," IEEE Trans. Power Syst., vol. 18, no. 1, pp. 121–127, 2003

[28] G. Verbič and F. Gubina, "A new concept of protection against voltage collapse based on local phasors," PowerCon 2000 - 2000 Int. Conf. Power Syst. Technol. Proc., vol. 2, no. 3, pp. 965–970, 2000.

[29] M. H. Haque, "Use of local information to determine the distance to voltage collapse," 8th Int. Power Eng. Conf. IPEC 2007, pp. 407–412, 2007.

[30] P. Kessel and H. Glavitsch, "Estimating the Voltage Stability of a Power System," IEEE Transactions on Power Delivery, Vol. -1, No. 3, July 1986.

[31] Smon, S. Member, and G. Verbiˇ, "Local Voltage-Stability Index Using Tellegen ' s Theorem" IEEE Trans. Power Syst., vol. 21, no. 3, pp. 1267–1275, 2006.

[32] B. Leonardi, V. Ajjarapu, M. Djukanovic and P. Zhang,"Application of Multi-linear Regression Models and Machine Learning Techniques for Online Voltage Stability Margin Estimation",2010 IREP Symposium--Bulk Power System Dynamics and Control- VIII (IREP) : Buzios, Rio de Janeiro, Brazil : August 1-6, 2010. IEEE, 2010.

[33] S. Weisberg,"Applied Linear Regression-Third Edition", John Wiley & Sons,Inc 2005

[34] J.H. Friedman, "Greedy Function Approximation: A Gradient Boosting Machine" The Annals of Statistics. Vol- 29, No5,pp. 1189-1232,2001

[35] K. Suzuki, "Artificial Neural Network: Architectures and Applications" InTech,2013

[36] L. Breman, "Random Forest", Machine Learning ,Kluwer Academic Publishers, pp.5-32, 2001

[37] Siemens Industy. Inc. "PSSE 34 Application Program Interface Manual",Mar 2015

[38] V. Ajjarapu, Computational Techniques for Volatge Stability Assessment and Control, Springer Science-Business Media, LLC,2005

[39] L. L. Freris and A. M. Sasson, "Investigation of the load-flow problem," in Proceedings of the Institution of Electrical Engineers, vol. 115, no. 10, pp. 1459-1470, October 1968,piee.1968.0260.

[40] R. Christie, "Power systems test case archive," Univ. Washington, Seattle,WA,USA.1993.[Online].Available: https://www.ee.washington.edu/research/pstca/pf118/pg_tca118bus.htm

[41] M. A. Hall, "Correlation-based Feature Selection for Machine Learning" Doctor of Philosophy thesis ,Department of Computer Science University of Waikato, Hamilton, New Zealand ,1999

[42] IEEE Standard for Synchrophasor Measurements for Power Systems, IEEE Standard C37.118.1-2011 (Revision of IEEE Standard C37.118-2005), 2011

[43] N.Perera, A.D Rajapakse and T.E Buchholzer "Isolation of Faults in Distribution Network With Distributed Generators" IEEE Trans. Power Delivery, vol. 23, no. 4, pp. 2347–2355, 2008

[44] Kim, T. Y. Chun, S. Yoon, G. Lee and Y. Shin, "Wavelet-based event detection method using PMU data," 2017 IEEE Power & Energy Society General Meeting, Chicago, IL, 2017, pp. 1-1,PESGM.2017

[45] Z. de Souza, A.C., Monn, F., Borges, I.F. 2011. Using PV and QV Curves with the Meaning of Static Contingency Screening and planning. Electric Power System Research, vol. 81, pp. 1491-1498

[46] Overbye, T.J., Dobson, I., DeMarco, I. 1994. Q-V Curve Interpretations of Energy Measures for Voltage Security. IEEE Transactions on Power Systems, vol. 9, no. 1, pp. 331-340.

[47] Morison, G.K., Gao., B., Kundur, P. 1993. Voltage Stability Analysis Using Static and Dynamic Approaches. IEEE Transactions on Power Systems, vol., 8, no. 3, pp.1159-1171.

[48] J. Barquin, T. Gomez, F.L Pagola. "Estimating the Loading Limit Margin Taking into Account Voltage Collapse Areas". IEEE Transactions on Power Systems, vol. 10, no. 4, pp. 1952-1962.

[49] S. M. Abdelkader and D. J. Morrow, "Online Tracking of Thévenin Equivalent Parameters Using PMU Measurements," in *IEEE Transactions on Power Systems*, vol. 27, no. 2, pp. 975-983

[50] S. M. Abdelkader and D. J. Morrow, "Online Thévenin Equivalent Determination Considering System Side Changes and Measurement Errors," in IEEE Transactions on Power Systems, vol. 30, no. 5, pp. 2716-2725

[51] F. Milano, "An open source power system analysis toolbox," in *IEEE Transactions on Power Systems*, vol. 20, no. 3, pp. 1199-1206,TPWRS.2005

[52] A. Smola and S.V.N. Vishwanathan, Introduction to Machine Learning, Cambridge Press,2008.

[53] Working Group o1 od Advisory Group 6 of Study Committee C4,"Review of On-line Dynamic Security Assessment Tools and Techniques", CIGRE Thechnical Brochure No. 325,December 2004

[54] Powertech Labs Inc., Voltage Security Assessment Tool (VSAT) User Manual, April,2008

[55] Hsiao-Dong Chiang, A.J. Flueck, K.S Shah and N.Balu, "CPFLOW: A Practical Tool for Tracing Power System Steady-state Stationary Behavior due to Load and Genaratortion Variation", IEEE Transaction on Power Systems, vol. 10 no.2pp. 623-634, May 1995.

[56] Shaw Power Technologies,Inc, "PSS/E™ 30 User Manual" ,August, 2004

[57] W.W. Hinkle DE, Wiersma W, S.G. Jurs, "Applied Statistics for the Behavioral Sciences". 5th ed. Boston: Houghton Mifflin; 2003

[58] G. Kleinbaum, L. Kupper, E. Muller and A. Nizam, "Applied Regression Analysis and Multivariable Methods" 3rd Edition. California: Thomson Publication , pp: 281-301

[59] T. Hastie, R. Tibshirani and J, Friedman, "The Elements of Statistical Learning: Data Mining, Interface and Prediction". 2nd Edition. California Springer Publication, pp 261-293

[60] M. Binkhonain, L. Zhao, "A review of machine learning algorithms for identification and classification of non-functional requirements", Expert Systems with Applications: X, Volume 1, 2019.

[61] M. Bowles, "Machine Learning in Python: Essential Techniques for Predictive Analytics", John Wiley & Sons Inc.

[62] I. Guyon, J Weston, S Barnhill, and V Vapnik.. "Gene Selection for Cancer Classification Using Support Vector Machines." *Machine Learning* 46 (1), pp 389–422, 2002.

[63] https://scikitlearn.org/stable/modules/generated/sklearn.linear_model.LinearRegression.html

[64] https://scikitlearn.org/stable/modules/generated/sklearn.ensemble.GradientBoostingRegressor.html#sklearn.ensemble.GradientBoostingRegressor

[65] https://scikitlearn.org/stable/modules/generated/sklearn.neural_network.MLPRegressor.html#sklearn.neural_network.MLPRegressor

[66] https://scikitlearn.org/stable/modules/generated/sklearn.ensemble.RandomForestRegressor.html