

ISTANBUL TECHNICAL UNIVERSITY ★ GRADUATE SCHOOL OF SCIENCE
ENGINEERING AND TECHNOLOGY

**A PROBABILISTIC DYNAMIC SECURITY ASSESSMENT OF LARGE
POWER SYSTEMS USING ARTIFICIAL NEURAL NETWORKS**



M.Sc. THESIS

Sevda JAFARZADEH

Department Of Electrical Engineering

Electrical Engineering Program

JUNE 2017

ISTANBUL TECHNICAL UNIVERSITY ★ GRADUATE SCHOOL OF SCIENCE
ENGINEERING AND TECHNOLOGY

**A PROBABILISTIC DYNAMIC SECURITY ASSESSMENT OF LARGE
POWER SYSTEMS USING ARTIFICIAL NEURAL NETWORKS**



M.Sc. THESIS

**Sevda JAFARZADEH
(504151038)**

Department Of Electric Engineering

Electrical Engineering Program

Thesis Advisor: Assoc. Prof. Dr. Veysel Murat Istemihan GENÇ

JUNE 2017

İSTANBUL TEKNİK ÜNİVERSİTESİ ★ FEN BİLİMLERİ ENSTİTÜSÜ

**BÜYÜK GÜÇ SİSTEMLERİNİN YAPAY SİNİR AĞLARI KULLANARAK
OLASILIKSAL DİNAMİK GÜVENLİĞİNİN DEĞERLENDİRİLMESİ**



YÜKSEK LİSANS TEZİ

**Sevda JAFARZADEH
(504151038)**

Elektrik Mühendisliği Anabilim Dalı

Elektrik Mühendisliği Programı

Tez Danışmanı: Doç. Dr. Veysel Murat İstemihan GENÇ

HAZİRAN 2017

Sevda JAFARZADEH, a M.Sc. student of İTÜ Graduate School of Science Engineering and Technology student ID 504151038, successfully defended the thesis entitled “**A PROBABILISTIC DYNAMIC SECURITY ASSESSMENT FOR LARGE POWER SYSTEMS USING ARTIFICIAL NEURAL NETWORKS**”, which she prepared after fulfilling the requirements specified in the associated legislations, before the jury whose signatures are below.

Thesis Advisor : **Assoc.Prof. V. M. İstemihan Genç**
Istanbul Technical University

Jury Members : **Assoc.Prof. Ramazan Çağlar**
Istanbul Technical University

.....
Prof.Dr. Mehmet Bayrak
Sakarya University

Date of Submission : 5 May 2017
Date of Defense : 9 June 2017



To my Family,





FOREWORD

I would like to express my deep appreciation and thanks for my advisor, Asso.Prof.Veyssel Murat Istemihan Genc. This was impossible without his warm guidance, support and encouragement during the entire master study.

I would like to thank my friend Meric Alkar for his kind and effective helps and guidance.

At the end, I would like to thanks my family for their warm support and love in my life.

June 2017

Sevda JAFARZADEH

TABLE OF CONTENTS

	<u>Page</u>
FOREWORD	ix
TABLE OF CONTENTS	xi
ABBREVIATIONS	xiii
NOMENCLATURE	xv
LIST OF TABLES	xvii
LIST OF FIGURES	xix
1. INTRODUCTION	1
1.1 Literature Review	1
1.2 Power system stability	3
1.2.1 Overview	3
1.2.2 Small signal angle stability	2
1.2.3 Transient stability	4
1.2.3.2 Critical Clearing Time(CCT)	4
1.3 Risk analysis	5
1.4 Optimal power flow	6
1.4.1 Overview	6
1.4.2 Standard ACOPF	6
1.5 Scope of thesis	8
2. MACHINE LEARNING APPROACHES	9
2.1 Monte Carlo sampling	9
2.1.1 The law of large numbers	9
2.1.2 Central limit theorem	9
2.2 MCMC sampling method	9
2.2.1 Gibbs sampling approach	10
2.2.2 LHS approach	12
2.2.3 Determine the number of required samples for Monte-Carlo approach ...	12
2.3 Artificial neural networks	13
2.3.1 MLP ANNs	14
2.4 Feature selection	16
2.4.1 Relief feature selection	17
2.5 GP distribution	17
3. PROPOSED METHODOLOGY	19
3.1 Create a dataset for training neural network	21
3.2 Construct a good predictor for system performance	22
3.3 Generation of instances (OPs) Using Gibbs sampling method	23
3.4 Generation of the exceedance sample	24
3.5 Fitting the distribution	26
3.6 Computation of the risk	27
4. SIMULATION RESULTS	29

4.1 Test system	29
4.2 Results	33
4.2.1 Result of small signal stability analysis	34
4.2.2 Results for transient stability analysis	39
5. CONCLUSION.....	56
REFERENCES	59
CURRICULUM VITAE	63



ABBREVIATIONS

ANN	: Artificial Neural Network
CCT	: Critical Clearing Time
GP	: Generalized Pareto
LHS	: Latin Hypercube Sampling
MC	: Monte-Carlo
MCMC	: Markov-Chain Monte-Carlo
MLP	: Multi-Layer Perceptron
OP	: Operating point
OPF	: Optimal Power Flow
PDF	: Probability Density Function
SSSA	: Small Signal Stability Analysis
TSA	: Transient Stability Analysis



NOMENCLATURE

P_g	: Active power of generator
Y_N	: Admittance matrix
δ	: Angle difference between two generators
Q	: Cumulative density function of normal distribution
I	: Current vector
ζ	: Damping ratio
w	: Imaginary part of eigenvalue
B	: Input matrix for power system
n_b	: Number of buses
n_g	: Number of generators
C	: Output coefficient matrix
η	: Power angle based stability margin
Q_g	: Reactive power of generator
T_s	: Settling time
A	: State matrix for power system
V	: Voltage vector



LIST OF TABLES

	<u>Page</u>
Table 4.1: Correlation coefficient between load and wind generation unit uncertainties	30
Table 4.2: Forecasted value for active power of each load and its standard deviation	31
Table 4.3: Forecasted value for active power of each wind generation unit and its standard deviation	31
Table 4.4: Cost coefficient of each generator.....	32
Table 4.5: Parameters for wind generation unit models	32
Table 4.6: Active powers of generators obtained by OPF for the forecasted operating point	33
Table 4.7: Id number of features	34
Table 4.8: Weights for active power of wind generation units with respect to small signal stability index	35
Table 4.9: Weights for active power of loads with respect to small signal stability index.....	35
Table 4.10: Performance of MLP neural network with different number of features	36
Table 4.11: Total Risk for small signal stability of the forecasted operating condition	39
Table 4.12: Critical contingencies and corresponded CCT values	40
Table 4.13: Weights for active power of wind generation units contingency 78.....	41
Table 4.14: Weights for active power demand of loads contingency 78	41
Table 4.15: Weights for active power of wind generation units contingency 76.....	42
Table 4.16: Weights for active power demand of loads contingency 76	43
Table 4.17: Weights for active power of wind generation units contingency 62.....	44
Table 4.18: Weights for active power demand of loads contingency 62	44
Table 4.19: MLP performance of different number of features for contingency 78.	45
Table 4.20: MLP performance of different number of features for contingency 76.	45
Table 4.21: MLP performance of different number of features for contingency 62.	46
Table 4.22: Required number of MC instances for three critical contingencies	48
Table 4.23: Total number of produced instances to obtain exceedance sample	48
Table 4.24: Total risk of forecasted operating point for each contingency with different methods	52
Table 4.25: Length of time duration taken for each step of the methodology for contingency 78	53
Table 4.26: Length of time duration taken for each step of the methodology for contingency 76	54
Table 4.27: Length of time duration taken for each step of the methodology for contingency 62	55

Table 4.28: Required time for exceedance sample generation with machine learning tools and without machine learning tools.....	55
---	-----------



LIST OF FIGURES

	<u>Page</u>
Figure 2.1: Flowchart of gibbs sampling method	11
Figure 2.2: Division of a distribution for LHS approach	12
Figure 2.3: Three layered neural networks	14
Figure 2.4: Basic perceptron	14
Figure 2.5: GP distribution for different k values	18
Figure 3.1: : Flowchart of proposed methodology	21
Figure 3.2: Procedure to create a dataset to train neural network	22
Figure 3.3: Procedure to train ANN	23
Figure 3.4: Procedure to produce operating points with Gibbs method.....	24
Figure 3.5: Flowchart for exceedance sample production	25
Figure 3.6: Area under the probability density function of a distribution.....	26
Figure 3.7: Severity function.....	27
Figure 4.1: Modified 16 generator-68 bus test system.....	29
Figure 4.2: Simple structure for DFIG model(TSAT user manual, 2015).....	33
Figure 4.3: Weights of features based on small signal stability performance of the system.....	36
Figure 4.4: Performance of MLP for whole features corresponding to small signal stability system performance	37
Figure 4.5: Distribution of damping ratio	38
Figure 4.6: Product of probability density function and severity function for small signal angle stability.....	39
Figure 4.7: Weights of feature for contingency number 78.....	41
Figure 4.8: Weights of feature for contingency number 76.....	42
Figure 4.9: Weights of feature for contingency number 62	43
Figure 4.10: Best MLP performance with 50 number of features for contingency number 78	46
Figure 4.11: Best MLP performance with 50 number of features for contingency number 76	47
Figure 4.12: Best MLP performance with 50 number of features for contingency number 62	47
Figure 4.13: Distribution of CCT for contingency number 78	49
Figure 4.14: Distribution of CCT for contingency number 76	49
Figure 4.15: Distribution of CCT for contingency number 62	50
Figure 4.16: Product of probability density function and severity function for contingency number 78	50
Figure 4.17: Product of probability density function and severity function for contingency number 76	51

Figure 4.18: Product of probability density function and severity function for contingency number 62	51
--	-----------



A PROBABILISTIC DYNAMIC SECURITY ASSESSMENT OF LARGE POWER SYSTEMS WITH USING ARTIFICIAL NEURAL NETWORKS

SUMMARY

Due to extensive utilization of intermittent energy sources in recent years, probabilistic dynamic security assessment has attracted much more attention. Conventional deterministic studies may not provide accurate security assessments for the power systems under uncertainties. Therefore, probabilistic approaches are necessary for decision makers to take appropriate preventive control actions based on more reliable assessments.

In this thesis, a new methodology is proposed to perform probabilistic security assessment for one forecasted operating point of a large power system. This approach can be applied for both transient and small-signal angle stability analyses.

The main uncertainty in a power system originates from the uncertainties of loads and wind generation units; therefore, the active power output of wind generation units and active power demand of loads are considered as random variables.

The main purpose of the proposed methodology for dynamic security assessment of a large power system is to compute the risk for one forecasted operating condition. In order to compute the risk, probability distribution of the system's dynamic performance is required. To obtain the probability distribution of the system performance, Monte-Carlo simulation method is used with some additional efforts to increase the accuracy and to decrease the computational burden of the process. To obtain the probability distribution of the system performance more precisely, left tailed region of the distribution is fitted by a Generalized Pareto (GP) distribution and the remaining part is fitted by a normal distribution.

To decrease the computational burden of the exceedance sample generation process, machine learning approaches such as neural-networks and feature selection methods are utilized.

The procedure starts with the creation of a dataset for training a neural network. The trained neural network is utilized to predict the system performance of an operating point for an exceedance sample generation process. The exceedance sample involves instances which result in system performances below a predefined threshold. Due to the fact that most of these instances have a low probability to be sampled, a large number of instances must be examined in order to get a reasonably large sample. However, using dynamic simulations to directly distinguish all the instances in such a large sample would not be suitable for online applications. Instead of the extensive dynamic simulations requiring an excessive computational effort, the trained neural network is used to generate an exceedance sample with a proper size as it predicts the system performance of each operating point more efficiently.

Latin Hypercube Sampling technique is used to sample the space of random variables evenly and produce instances of the dataset for training the neural network. For each instance, Optimal Power Flow solution is found and new operating points are obtained. For each operating point produced, a series of dynamic simulations is applied and the system performance is computed. Since the initial dataset created for training the neural network will most likely involve features which are not relevant to the system performance, a reliable and fast feature selection algorithm, namely Relief, is used to find the relevant features. Thus, with this approach the accuracy of the neural network can be improved as its training time is reduced.

The next step of the methodology is to generate instances by an appropriate Monte-Carlo sampling technique and to obtain an initial probability distribution of the system performance. Gibbs sampling technique, which is a type of Markov-Chain Monte-Carlo sampling method, is used to generate the operating conditions which are used to fit a normal distribution for the system performance. Considering the existence of correlation between the random variables, Gibbs sampling is selected as an appropriate method to produce new operating conditions. For each newly produced operating condition, an Optimal Power Flow (OPF) solution is found and the system performance for the resulting operating point is computed by dynamic simulation.

The required number of operating points for Gibbs sampling is determined according to the sample standard deviation of the system performance and specified error and confidence level. The sample standard deviation of the system performance is calculated based on the previously generated training dataset for the neural network. In this study, the maximum tolerable error and the confidence level are specified as 5% and 95%, respectively. It means that we are 95% sure that error we make is less than 5%.

The exceedance sample which is required for Generalized Pareto (GP) distribution fitting is produced with the help of trained neural network. First, a large number of instances is generated by LHS method and, then, we randomly select one instance at a time, without a replacement, as a candidate exceedance instance. The performance of each selected instance is predicted by the trained neural network. If the system performance is less than the specified threshold, this candidate is included in the exceedance sample; otherwise it is discarded. The system performance of each exceedance instances are also computed through dynamic simulation, which also confirm their exceedance. This process continues until the required number of exceedance instances are collected. The system performance at these operating points are used to fit the GP distribution.

The overall probability distribution of the system performance is obtained by combining the two fitted (normal and GP) distributions above for a given forecasted operating condition. A risk analysis has been done to give a probabilistic security assessment for the forecasted operating condition. In order to perform the risk analysis, a severity function related to the system performance is selected. In this study, the severity function is the combination of a unit step and a quadratic function. The product of the probability density function of the system performance and the corresponding severity function is used to calculate the risk of the forecasted operating condition.



BÜYÜK GÜÇ SİSTEMLERİNİN YAPAY SİNİR AĞLARI KULLANARAK OLASILIKSAL DİNAMİK GÜVENLİK DEĞERLENDİRMESİ

ÖZET

Son yıllarda kentsintisiz olarak işletilmesi mümkün olmayan enerji kaynaklarının yoğun kullanımı nedeniyle, olasılıksal dinamik güvenlik değerlendirme büyük önem kazanmıştır. Konvansiyonel deterministik çalışmalar, belirsizlikler altında güç sistemleri için doğru güvenlik değerlendirmeleri sağlamayabilir. Bu nedenle, karar vericilere daha güvenilir değerlendirmelere dayanılarak uygun önleyici kontrol faaliyetleri için olasılıksal yaklaşımlar gereklidir.

Bu tezde, büyük bir güç sisteminin öngörülen bir çalışma noktası için olasılıksal güvenlik değerlendirmesini gerçekleştirmek için yeni bir metodoloji önerilmiştir. Bu yaklaşım hem geçici hem de küçük sinyal açısı kararlılığı analizleri için uygulanabilir.

Bir güç sistemindeki temel belirsizlik, yüklerin ve rüzgar enerjisi üretim birimlerinin belirsizliklerinden kaynaklanmaktadır; Bu nedenle, rüzgar enerjisi üretim ünitelerinin aktif güç çıkışı ve yüklerin aktif güç talebi rastgele değişken olarak kabul edilir. Büyük bir güç sisteminin dinamik güvenlik değerlendirmesi için önerilen metodolojinin ana amacı, öngörülen bir çalışma koşulunun riskini hesaplamaktır. Riski hesaplamak için sistemin dinamik performansının olasılık dağılımı gereklidir.

Sistem performansının olasılık dağılımını elde etmek için Monte Carlo simülasyon yöntemi, işlemin hesaplama yükünü azaltmak ve doğruluğunu artırmak için bazı ek çabalarla kullanılır. Sistem performansının olasılık dağılımını daha doğru bir şekilde elde etmek için dağılımın sol kuyruk bölgesi, bir Genelleştirilmiş Pareto (GP) dağılımı ile yerleştirilir ve geri kalan kısım normal bir dağılım ile yerleştirilir.

GP dağılımı uydurmak için, bir aşırılık örneğinin elde edilmesi gereklidir, ancak söz konusu örneğin elde edilmesinde sadece dinamik simülasyonlardan yararlanılması, önemli miktarda uzun bir zaman gerektirir.

Aşan örnek oluşturma sürecinin hesaplama yükünü azaltmak için, sinir ağı ve nitelik seçme yöntemleri gibi makine öğrenme yaklaşımları kullanılır.

İşlem, sinir ağının eğitimi için bir veri kümesinin oluşturulmasıyla başlar. Eğitilen sinir ağı, aşan bir numune oluşturma sırasında, ele alınan bir çalışma noktasındaki sistem performansını tahmin etmek için kullanılır. Aşım örneği, sistem performanslarının önceden tanımlanmış bir eşiğin altına düşmesine neden olan örnek elemanları içerir. Bu örneklerin çoğunun örnekleme olasılığının düşük olması nedeniyle, makul büyüklükte bir örnek elde etmek için çok sayıda örnek elemanı incelenmelidir. Bununla birlikte, dinamik simülasyonlar kullanarak büyük bir örnekteki tüm elemanları doğrudan ayırt etmek, çevrimiçi uygulamalarda uygun olmayacaktır. Aşırı bir hesaplama çabası gerektiren kapsamlı dinamik simülasyonların yerine, eğitilen sinir ağı, her çalışma noktasının sistem performansını daha verimli bir şekilde kestirmesi nedeniyle, uygun bir büyüklüğe sahip aşım örneği oluşturmak için kullanılır.

Latin Hiperküp Örnekleme tekniği, rastgele değişkenlerin alanını eşit olarak örnekleme ve sinir ağının eğitimi için veri kümesinin örneklerini üretmek için kullanılır. Her bir örnek elemanı için Optimal Güç Akışı çözümü bulunmuş ve yeni çalışma noktaları elde edilmiştir. Üretilen her bir çalışma noktası için bir dizi dinamik simülasyon uygulanır ve sistem performansı hesaplanır. Sinir ağının eğitimi için başlangıçta oluşturulan veri kümesi büyük olasılıkla sistem performansı ile alakalı olmayan nitelikleri de içerdiğinden, ilgili nitelikleri bulmak için güvenilir ve hızlı bir nitelik seçimi algoritması olan Relief kullanılabilir. Böylece, bu yaklaşımla, eğitme zamanının azaltılmasıyla birlikte, sinir ağının performansı artırılmış olur.

Metodolojinin bir sonraki adımı uygun Monte-Carlo örnekleme tekniği ile örnek elemanları üretmek ve sistem performansına ait bir başlangıç olasılık dağılımını elde etmektir. Markov-Zinciri Monte-Carlo örnekleme yönteminin bir türü olan Gibbs örnekleme tekniği, sistem performansı normal dağılımında kullanılan çalışma koşullarını oluşturmak için kullanılır. Rastgele değişkenler arasındaki korelasyon varlığı göz önüne alındığında, Gibbs örnekleme, yeni çalışma koşulları üretmek için uygun bir yöntem olarak seçilmiştir. Yeni üretilen her çalışma durumu için Optimal Güç Akışı (OPF) çözümü bulunur ve elde edilen çalışma noktası için sistem performansı, dinamik simülasyon ile hesaplanır.

Gibbs örneklemesi için gereken çalışma noktası sayısı, sistem performansına ait örnek standart sapmasına, kabul edilebilir hataya ve güven seviyesine göre belirlenir. Sistem performansının örnek standart sapması, sinir ağını eğitmek için daha önce oluşturulmuş olan veri kümesine dayanarak hesaplanır. Bu çalışmada, maksimum tolere edilebilir hata ve güven seviyesi sırasıyla, % 5 ve % 95 olarak belirlenmiştir. Bu kabul, yaptığımız hatanın % 5'ten az olduğundan % 95 emin olduğumuz anlamına gelmektedir.

Genelleştirilmiş Pareto (GP) dağılım uydurması için gereken aşan numune eğitilmiş olan sinir ağı yardımıyla üretilir. İlk olarak, LHS yöntemiyle çok sayıda örnek oluşturulur ve daha sonra, aşma örnek elemanı adayı olarak, yerine koymaksızın, tek tek rastgele bir örnek elemanı seçeriz. Her bir seçilen örnek elemanı için sistem performansı, sinir ağı tarafından öngörülür. Sistem performansı belirtilen eşik değerden düşükse, bu aday aşım örneğine dahil edilir; yoksa atılır. Her aşım örneğinin sistem performansı ayrıca, aşırılıklarını da doğrulayan bir dinamik simülasyon yoluyla hesaplanır. Bu işlem, aşım örneği toplanana kadar devam eder. Bu çalışma noktalarındaki sistem performansının GP dağılımına uydurulması için kullanılır.

Ele alınan tahmini bir çalışma koşulu için sistem performansının tüm olasılık dağılımı, yukarıdaki iki uydurulmuş (normal ve GP) dağılımlar birleştirilerek elde edilir. Tahmini bir çalışma koşuluna ait olasılıksal bir güvenlik değerlendirmesi yapmak için, bir risk analizi yapılmıştır. Risk analizini gerçekleştirmek için, sistem performansı ile ilişkili bir şiddet fonksiyonu seçilir. Bu çalışmada, şiddet fonksiyonu bir birim adım ve bir ikinci dereceden fonksiyonun birleşimi olarak seçilmiştir. Sistem performansının olasılık yoğunluk fonksiyonunun ve ilgili şiddet fonksiyonunun çarpımı, tahmini çalışma koşullarının riskinin hesaplanmasında kullanılır.



1. INTRODUCTION

1.1 Literature Review

Reliability of large power systems requires sufficient power generation and adequate security [1]. In other words, the bulk power system should provide adequate power for the required load and it must be secure to be a reliable power system. Security is the ability of a power system to remain stable when it is subjected to a set of credible contingencies.

In recent years dynamic security assessment has been recognized as an important issue for large power systems. A dynamic security assessment is performed to identify the current security level of a power system. The assessment provides recommendations for preventive control actions which can be really useful. Dynamic security assessment for a power system is divided into 3 categories. The first category discusses the transient stability of the power system which involves large disturbances such as a short circuit on transmission lines. The second category explores the small signal stability of a power system. The third section discusses the voltage stability of the power system.

Power systems are set off under growingly inconsistent conditions due to rapid reproduction of intermittent renewable energies. These variable conditions integrate uncertainty to the power systems therefore a conventional deterministic security assessment cannot be able to present the precise overview for the security of the power systems. In order to take the right predictive control actions in an emergency situation, decision-makers should be aware of the precise security level of the power system however, deterministic N-1 security criterion cannot assure the security level of the power system [1]. The necessities of applying probabilistic dynamic security assessment are studied in [2,3].

Probabilistic security assessment was first introduced by Felix Wu and Yu-kun Tsai [4]. The online probabilistic security assessment gives the security level of the system for near future. There are two types of methods for probabilistic security assessment.

The first one is analytic method [5,6]. One of the difficulties of analytic method, arises in cases with large number of uncertainties which causes dimension explosion. The other problem with this method raises when the uncertainty of loads and generations are considered. Analytic method needs to compute security translation rates in the case of uncertain loads and generations. Some efforts have been done in [1] to overcome the difficulties related to the analytic method.

The second method is simulation [8,7]. In this practice, each item of the power system can be characterized uncertain easily. Monte-Carlo is the most famous simulation method for the probabilistic security assessment. One of the difficulties with Monte-Carlo method is its heavy computational burden for high precision therefore it is not appropriate for cases which are rare. Some research has been taken MC simulation practice for probabilistic security assessment and overcome the difficulties with this method. In [9,10], efficient sampling techniques have been used to cut down on computational burden of the procedure. The problem about these methods is the limitation of symmetric output.

1.2 Power System Stability

1.2.1 Overview

Power system stability is an important issue for the security assessment of a power system since the 1920s [11]. Major blackouts which were the result of instability of the power system revealed the significance of the phenomenon [12]. The stability of a system refers to the ability of a system to return back to its stable state when it is disturbed. Large disturbances are categorized as a transient stability issue and small disturbances are categorized as a small signal angle stability problem.

1.2.2 Small signal angle stability

Small signal angle stability refers to the disturbances which are small. The power system is stable if the small variation in system 's steady state variables cannot cause major deviation on the system variables for a long period of time [13]. The damping of oscillations inherent in power system is considered a crucial problem. These oscillations are categorized into 4 groups: control mode oscillations, inter-area mode oscillations, torsional mode oscillations, local plant mode oscillations [14].

The most troublesome of the oscillations are the local plant mode ones which originate often from the performing of AVR's for a large amount of output. In addition, these oscillations exacerbated by integrating speedy excitation systems. The frequency of local plant mode oscillations is in the range of 1 to 2 Hz. Power System Stabilizers (PSS) are effective enough to damp these oscillations.

Closely coupled machines are the main reason of the inter-area mode oscillations. The frequency range for these oscillations is 0.1 to 1 Hz. Control mode oscillations are engendered due to the control of generation units and other elements. The reciprocal action between components of turbine-generators caused torsional mode oscillations [15].

For a power system, dynamic behavior can be presented by a set of differential equations and algebraic equations [14].

$$\dot{x} = f(x, u), y = g(x, u) \quad (1.1)$$

$x = [x_1, x_2, \dots, x_n]^T$ and $y = [y_1, y_2, \dots, y_n]^T$ represent the vector of state variables and system outputs respectively. System inputs is entitled by u . f and g are the sets of nonlinear differential and algebraic functions respectively. To perform small signal stability analysis, modal analysis technique can be utilized. In this practice, system is linearized around one operating point and equation (1.2) gives the linearized equations.

$$\Delta \dot{x} = A \Delta x + B \Delta u, \Delta y = C \Delta x + D \Delta u \quad (1.2)$$

A is state matrix and B represents input matrix. C denotes output coefficient matrix and D explains the relationship between input and output vectors. Eigenvalues of the state matrix declare the small signal angle stability status of the system. If the real parts of all eigenvalues are negative, this system is stable. Damping ratio is determined by equation (1.3) as a criteria for small signal angle stability analysis. α_i and ω_i are the real and imaginary parts of i^{th} eigenvalue respectively.

$$\xi_i = \frac{-\alpha_i}{\sqrt{\alpha_i^2 + \omega_i^2}} \quad (1.3)$$

1.2.3 Transient stability

Transient stability is characterized with the ability of a power system to preserve its synchronism when exposed to a large disturbance such as a disconnection of a generator or a short circuit on transmission lines.

Initial operating condition, fault type, severity of the fault and the fault location affect the transient stability of the power system. Various approaches exist for transient stability analysis. One of the common methods for transient stability analysis is time domain simulation which is used in this study. Time domain simulation solves a set of nonlinear differential and algebraic equations to find the transient stability status of the network.

$$\dot{x} = f(x, V), I(x, V) = Y_N V \quad (1.4)$$

V and I are the voltage and current vectors respectively. x presents the state vector and Y_N describes the admittance matrix of the network. In this thesis, DSAtools program [42] is used for transient stability analysis of the power system. This method is not able to present the severity of the contingency therefore different indices have been introduced to determine the severity of the contingency. The indices which are used in this study are introduced in the next section.

1.2.3.1 Power angle based stability margin(AM)

Power angle index is defined based on the maximum angle difference of any two generator (δ_{max}). This index is given in equation (1.5).

$$\eta = \frac{360 - \delta_{max}}{360 + \delta_{max}} \times 100 \quad (1.5)$$

The system is secure if η is greater than zero otherwise it is insecure [42].

1.2.3.2 Critical Clearing Time(CCT)

Critical Clearing Time (CCT) is the maximum time duration in which power system can tolerate the subjected contingency [16]. The system is secure if the computed CCT for the subjected contingency is greater than the length of time which circuit breakers require to take action [42].

1.3 Risk Analysis

Risk provides information about the probability of the occurrence of a contingency or an event and the severity of its further consequences. Some studies have performed risk analysis for an uncertain power system [3,17]. In [17] the risk of small signal instability is computed based on the uncertainty of forecasted loading values and power generation from intermittent energy sources and power flows through voltage-source direct current (VSC-HVDC) lines. Settling time for the power system oscillations is utilized to determine the severity. For these oscillations the settling time T_S is dependent on the tolerance (tol) and damping (ξ) (1.6).

$$T_S = \frac{\ln(tol)}{\xi} \quad (1.6)$$

Different continuous severity functions can also be used. Some examples of continuous severity functions are presented in equations (1.7-1.9).

$$\text{Stepped: } S(\xi) = \begin{cases} 0.25 & \text{for } -0.05 \leq \xi < 0 \\ 1 & \xi \geq 0 \end{cases} \quad (1.7)$$

$$\text{Linear: } S(\xi) = 20(\xi + 0.05) \quad \text{for } \xi \geq -0.05 \quad (1.8)$$

$$\text{Quadratic: } S(\xi) = 400(\xi + 0.05)^2 \quad \text{for } \xi \geq -0.05 \quad (1.9)$$

Risk matrix can also be used to quantify the risk for cases with discrete probability ranges and discrete severity ranges [18].

Risk can be obtained from simple equation (1.10) for continuous probability ranges and continuous severity ranges.

$$Risk = \int P(\xi)S(\xi) d\xi \quad (1.10)$$

$P(\xi)$ represents the probability of damping ξ .

There is another explanation for risk index for cases with discrete probability ranges and discrete severity ranges which is an expectation of the severity computed by summing over all possible outcomes the product of outcome probability and its severity [19]. The probability of each terminal state is the product of the probabilities of each branch, severity values are assigned to each terminal state. The severity is the sum of product of the probability and corresponded severity (1.11).

$$Risk(X_{t,f}) = \sum_i \Pr(E_i) \left(\sum_j \Pr(X_{t,f}|X_{t,j}) \times Sev(E_i, X_{t,j}) \right) \quad (1.11)$$

$X_{t,f}$, $X_{t,j}$, E_i represent forecasted condition at time t, j^{th} possible loading condition and i^{th} contingency respectively. $Sev(E_i, X_{t,j})$ quantifies the severity of i^{th} contingency occurring under j^{th} possible operating condition.

1.4 Optimal Power Flow

1.4.1 Overview

In this study, in order to determine the output power of generators for each loading values and output power of wind generation units, Optimal power flow (OPF) analysis is required. In this thesis Matpower program is used to solve the OPF problem.

The optimal power flow problem was first introduced by Carpentier in 1960's [20]. Optimal power flow analysis is used to find the optimum solution for a power system with subject to the specified constraints such as power flow equation and other operational constraints. There are different types of OPF problem based on an objective function, constraints and solution methods. One of the OPF problem types is AC optimal power flow which uses exact AC power flow equation. Various objective functions are used for ACOPF problem such as minimizing generation cost, maximizing marketing surplus, minimizing losses, minimizing generation and maximizing transfers [21]. Minimizing generation cost and maximizing marketing surplus are equivalent if demand functions are discarded. Minimizing cost includes minimizing fuel cost of the generators, minimizing output of the generators and minimizing losses. In this study standard AC OPF is selected to find the optimum solution. Objective function and constraints are defined in the next section.

1.4.2 Standard ACOPF

In this section, objective function and constraints are defined. Equation (1.12) presents the general objective function for ACOPF.

$$\min_{\theta, V_m, P_g, Q_g} \sum_{i=1}^{n_g} f_p^i(P_g^i) + f_Q^i(Q_g^i) \quad (1.12)$$

In this study, simply the objective function is the summation of quadratic cost functions for each generator (1.13).

$$\min_{\theta, V_m, P_g, Q_g} \sum_{i=1}^{n_g} C_{P,i}^2 (P_g^i)^2 + C_{P,i}^1 (P_g^i) + C_{P,i}^0 + C_{Q,i}^2 (P_Q^i)^2 + C_{Q,i}^1 (P_Q^i) + C_{Q,i}^0 \quad (1.13)$$

Cost coefficients of each generator is given by C_i .

Equality constraint of ACOPF is power flow equation and inequality constraints are given in equations (1.14-1.19.).

Inequality constraints:

$$\theta_i^{ref} \leq \theta_i \leq \theta_i^{ref} \quad i \in I_{ref} \quad (1.14)$$

$$v_m^{i,min} \leq v_m^i \leq v_m^{i,max} \quad i = 1 \dots n_b \quad (1.15)$$

$$P_g^{i,min} \leq P_g^i \leq P_g^{i,max} \quad i = 1 \dots n_g \quad (1.16)$$

$$Q_g^{i,min} \leq Q_g^i \leq Q_g^{i,max} \quad i = 1 \dots n_g \quad (1.17)$$

$$h_f(\theta, v_m) = |F_f(\theta, v_m)| - F_{max} \leq 0 \quad (1.18)$$

$$h_t(\theta, v_m) = |F_t(\theta, v_m)| - F_{max} \leq 0 \quad (1.19)$$

v_m^i is voltage magnitude for bus i and P_g^i, Q_g^i represent the active power and reactive power for generator i respectively. θ_i is voltage angle of bus i . n_b and n_g are number of buses and number of generators in a power system. Equations (1.18) and (1.19) are inequality constraints for branches flow limits. The first for flow limits from end and the second one is for to end. $h_f(\theta, v_m)$ and $h_t(\theta, v_m)$ are defined in equations (1.20-1.21).

$$F_f(\theta, v_m) = \begin{cases} S_f(\theta, v_m) & \text{apparent power} \\ P_f(\theta, v_m) & \text{real power} \\ I_f(\theta, v_m) & \text{current} \end{cases} \quad (1.20)$$

$$F_t(\theta, v_m) = \begin{cases} S_t(\theta, v_m) & \text{apparent power} \\ P_t(\theta, v_m) & \text{real power} \\ I_t(\theta, v_m) & \text{current} \end{cases} \quad (1.21)$$

Various types of solvers are available in Matpower program [43] to solve the OPF problem. In this thesis Interior Point solver is used to find an optimum solution.

1.5 Scope Of Thesis

The current chapter was an introduction of this thesis. As it is mentioned in previous sections, dynamic security assessment is an important issue for large power systems and uncertainties inherent in power system cause uncertainty in the stability of the power system therefore probabilistic methods are required to obtain a realistic view for the security level of one forecasted operating point of the large power systems.

Required background, detailed discussion of sampling techniques, machine learning approaches have been presented in the next chapter. Proposed methodology for probabilistic security assessment which is implemented in this thesis, is covered in the third chapter. Chapter 4 presents the simulation results and the details of the test system which is used in this thesis. Last chapter of this study covers the conclusion of the whole study.

2. MACHINE LEARNING APPROACHES

2.1 Monte-Carlo Sampling

Monte-Carlo is a method to estimate the properties of an output distribution by extracting information from the distribution of inputs. Monte-Carlo practice is developed based on two fundamental mathematical theories [22]. The first one is the law of large numbers and the second one is central limit theorem.

2.1.1 The law of large numbers

The law of large numbers presents that as number of experiments increases, sampled average of the output of the experiments converge to the expected value of the experiment. Equation (2.1-2.2) depicts the law of large numbers.

$$\bar{X}_n = \frac{1}{n} (X_1 + X_2 + \dots + X_n) \quad (2.1)$$

$$\lim_{n \rightarrow \infty} \bar{X}_n = E[X] = \mu \quad (2.2)$$

X_i is i^{th} iteration of an experiment with X is a random variable and $E[X]$ is an expected value for the random variable X .

2.1.2 Central limit theorem

Central Limit Theory (CLT) is a statistical theory which establishes that the distribution of sufficiently large number of repeated identical independent experiments will converge to a normal distribution. Equation (2.3) presents the Central Limit Theorem.

$$\lim_{n \rightarrow \infty} Pr \left\{ \alpha \leq \frac{|\bar{X}_n - \mu|}{\sigma/\sqrt{n}} \leq \beta \right\} = \frac{1}{\sqrt{2\pi}} \int_{\alpha}^{\beta} e^{-u^2/2} du \quad (2.3)$$

2.2 MCMC Sampling Method

MCMC is an increasingly popular Monte-Carlo sampling method which is the combination of Markov-Chain and Monte-Carlo properties [23]. Generating new

instance based on the previous instance is a sequential process which is the Markov-Chain property of MCMC sampling method. Each generated random instance is considered as a step to generate the next random instance by comparing the probabilities of both instances (chain property). Each generated instance is based on only the one previous instance which is the Markov property of the chain.

The idea of MCMC sampling procedure is simple. Steps 1 to 4 present the MCMC approach:

- Step1. Determine a distribution function represent all the input variables
- Step2. Sample each input variable independently
- Step3. Calculate the output from the specified function
- Step4. Repeat many times to obtain the output distribution

There are many types of MCMC sampling approaches with their own advantages and disadvantages. Some of these methods cannot be used for highly correlated inputs such as Metropolis-Hasting approach. Due to the fact that there some correlations in our uncertain inputs in this thesis, simple Gibbs sampling method has been utilized.

2.2.1 Gibbs sampling approach

For a given multivariate distribution, each parameter is treated separately. The process of Gibbs sampling is presented in below steps [24].

- Step1. Choose staring values for all the parameters
- Step2. Start from the first variable $i=1$
- Step3. If $i \leq n$ go to step 3.a otherwise go to step 3.b (n is total number of parameters)
- Step3.a. Generate a proposal for i^{th} parameter
- Step3.b. one iteration of the sampling is completed and go to step2 for a new iteration
- Step4. Accept the new proposal if the probability of new proposal given the other parameters is higher than the previous value of it otherwise accept the new proposal with a probability equal to a

$$prob(X_{new}) = prob(X_1, X_2, \dots, X_i = X_{i,proposal}, \dots, X_n) \quad (2.4)$$

$$prob(X_{old}) = prob(X_1, X_2, \dots, X_n)$$

$$a = \frac{prob(X_{new})}{prob(X_{old})}$$

- Step5. $i=i+1$ and return to step 3

The process is simple and efficient. In this study MATLAB program [41] is used to generate new instances. Figure 3.1 presents the flowchart of the process.

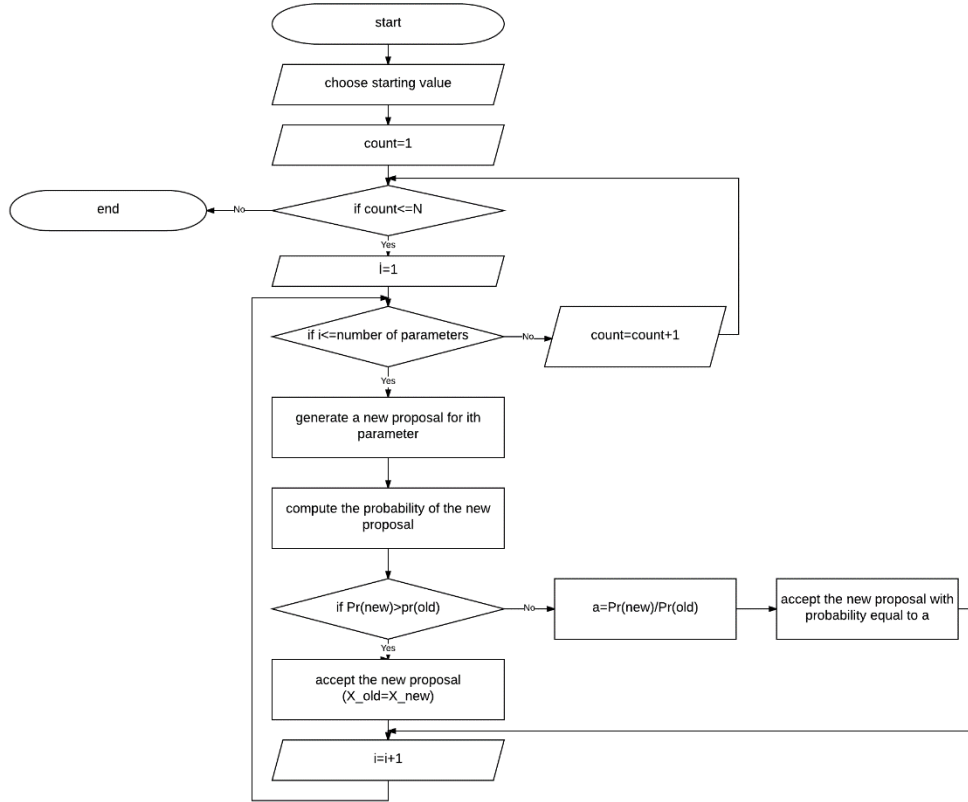


Figure 2.1: Flowchart of Gibbs sampling method

The mentioned practice is useful for Monte-Carlo simulation and in this thesis, the instances which are obtained with this method are implemented for Monte-Carlo simulation part of the proposed methodology, however in the first part of the methodology limited number of operating points should be sampled from the whole search space evenly to train the neural network therefore Latin Hypercube Sampling (LHS) method is used for the mentioned case. Detailed explanation of LHS technique is presented in the next section.

2.2.2 LHS approach

Latin Hypercube Sampling was first introduced by McKay in 1979. LHS is one of the Monte-Carlo sampling methods which is able to sample the search space evenly. The method is started with division of the distribution into n sections in a way that all sections have equal probability. Figure 3.2 depicts the section division of the distribution for a normal distribution. In the next step, predefined number of instances are sampled randomly from each section with uniform probability [25]. The procedure is so simple that makes the method easy to implement.

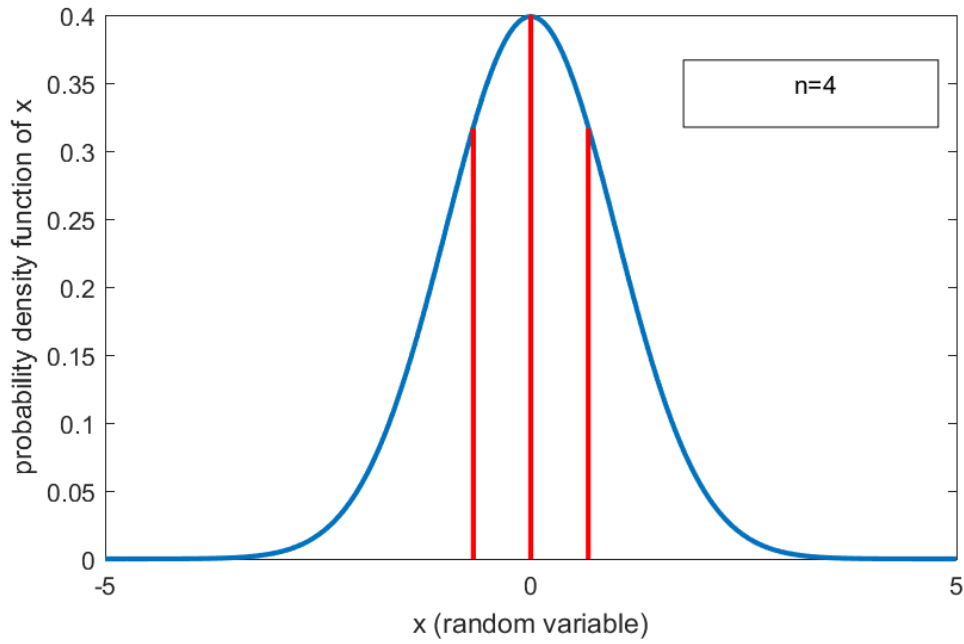


Figure 2.2: Division of a distribution for LHS approach

2.2.3 Determine the number of required samples for Monte-Carlo approach

In order to obtain the acceptable error with predefined confidence level, the number of required Monte-Carlo instances should be determined. In [26] the basic idea is presented to determine the number of required instances. If central limit theorem in equation (2.3) is applied for a specified absolute error with the predefined confidence level, the number of required instances will be obtained.

$$Pr \left\{ |\bar{X}_n - \mu| < z_{\alpha/2} \frac{\sigma}{\sqrt{n}} \right\} = 1 - \alpha \quad (2.5)$$

$$n = (z_{\alpha/2} \times \sigma / \text{err})^2 \quad (2.6)$$

err presents the specified error and σ is the standard deviation for output distribution. Predefine confidence level is denoted by α . $z_{\alpha/2}$ describes the upper percentile for the standard gaussian distribution which can be computed from equation (2.7).

$$z_{\alpha/2} = Q^{-1}\left(\frac{\alpha}{2}\right) \quad (2.7)$$

Q^{-1} represents the inverse function of normal cumulative density function.

The problem about the presented practice is that, the standard deviation of the output distribution is also unknown. In this thesis, the standard deviation of the instances which are produced by LHS sampling method to train the neural network is used to determine the number of required instances.

2.3 Artificial Neural Networks

Artificial neural network (ANN) is a mathematical based model which is able to learn the connection between inputs and outputs of a system.

In recent years ANNs have been used a lot in different contexts. ANNs can be used both for classification and regression problems. Due to the fact that, dynamic simulation of a large power system and computing its performance for a large number of operating points is really time consuming, ANNs can be an appropriate alternative in order to compute system performance of different operating points. Many researchers have been used ANNs in power system. In [27] ANN has been used for dynamic security assessment. In [28] Critical clearing time (CCT) is estimated by ANN.

There are different types of ANNs for regression problems such as Multi-Layer Perceptron (MLP), Radial Basis Function neural network (RBF), Support Vector Machine (SVM). Figure 3.3 presents a three layered neural network.

As shown in figure 3.3, there are some nodes in each layer. The first layer is called input layer and the last layer is output layer. The layers between input layer and output layer are called hidden layers. For input layer, number of nodes is the number of inputs and for output layer the number of nodes is equivalent to number of outputs. The weights are the connection between layers and updated until a desired error is reached [29].

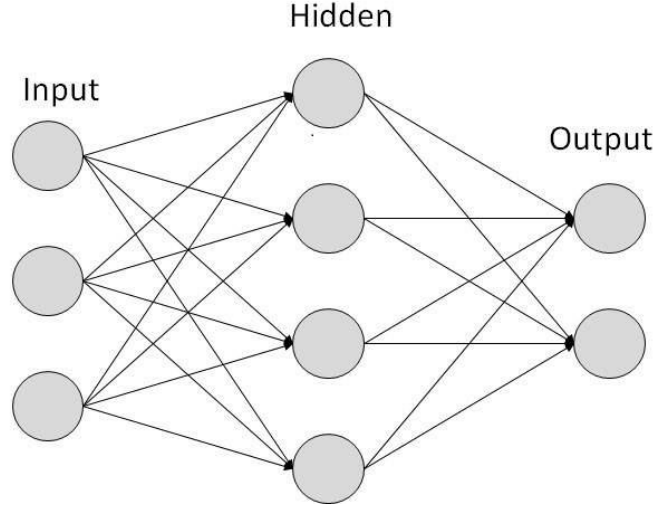


Figure 2.3: Three layered neural networks

2.3.1 MLP ANNs

In this study MLP neural network is used to estimate system performance and it is implemented by MATLAB toolbox [41]. Multi-layered perceptron which is one of the famous types of feedforward neural network, is modified practice of linear perceptron [30,31]. Perceptron is defined as a simple element which has inputs and an output [32]. Inputs of each perceptron can be the outputs of other perceptrons or inputs of associated system. Each input of one perceptron is multiplied by its weight and in the simplest case the output of perceptron is the weighted sum of inputs. A basic perceptron is shown in figure 3.4.

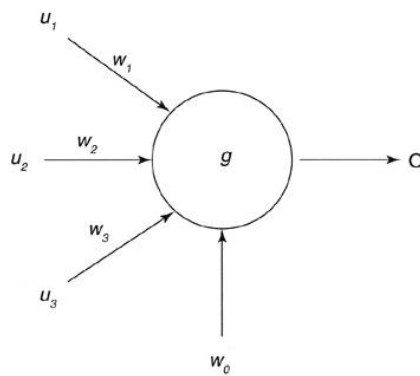


Figure 2.4: Basic perceptron

w_i s are the weights and u_i s are inputs and g is a basic perceptron. w_0 is a bias term. Basic formula of the perceptron is given in equation (2.8).

$$O = w_0 + \sum_{i=1}^d w_i u_i \quad (2.8)$$

d is the number of inputs of perceptron. The mentioned simple function can be used for classification of linearly separable classes by checking the sign of output O .

$$choose \begin{cases} class\ 1\ if\ O > 0 \\ class2\ if\ O < 0 \end{cases} \quad (2.9)$$

However for nonlinearly separable classes, more complex functions are required. The function which exists at the end of each perceptron such as mentioned sign function in equation (2.9) is called activation function. For nonlinear separable cases various types of activation functions are used such as sigmoid and tangent hyperbolic (2.10-2.11).

$$y = \frac{1}{1+e^{-x}} \quad (2.10)$$

$$y = \tanh x \quad (2.11)$$

The next step is training the neural network and obtain the weights. First weights are initialized randomly and error is calculated. Weights are updated at each iteration in a way that minimizes the error function. Equation (2.16) gives a regression error for one instance. r^t is the target output for input x^t .

$$y^t = w_0 + \sum_{i=1}^d w_i^t x_i^t \quad (2.12)$$

$$X^t = \begin{bmatrix} x_1^t \\ x_2^t \\ \vdots \\ x_d^t \\ 1 \end{bmatrix}_{(d+1) \times 1} \quad (2.13)$$

$$W = \begin{bmatrix} w_1 \\ w_2 \\ \vdots \\ w_d \\ w_0 \end{bmatrix}_{(d+1) \times 1} \quad (2.14)$$

$$y^t = W^T X^t \quad (2.15)$$

$$E_t(w|x^t, r^t) = \frac{1}{2} (r^t - y^t)^2 = \frac{1}{2} [r^t - (W^t x^t)]^2 \quad (2.16)$$

Stochastic gradient descent algorithm is used to update the weights to minimize the error function [32]. This structure is only able to estimate the linear functions. In order to generalize the model, more than one perceptron is used and its name is Multi-Layer Perceptron. Neurons of one hidden layer are fully connected to the neurons of next

hidden layer and the connections are the weights. The whole neurons of hidden layers have a sigmoid or tangent hyperbolic activation function. In this study Levenberg-Marquardt algorithm is used to train the MLP neural network. Details of this algorithm can be found in [33,34].

2.4 Feature Selection

Dimensionality reduction is necessary for some applications such as data mining. Feature selection is a kind of dimensionality reduction technique which is able to find the most relevant subset of features. Dimensionality reduction means cut down on number of dimensions in order to have simpler model and higher accuracy [35]. There are a lot of reasons to do feature selection. In real world applications, too many features are candidate to represent the domain, however, most of these features are irrelevant to the target. Also the relevant features can have correlation with each other and knowing one of the features is sufficient to estimate the target and the other relevant features cannot add any information [36]. Therefore feature selection is a necessary step for preprocessing. By selecting the most relevant features, training neural network takes less time and the model of the system will be simpler which drastically increases the accuracy of the neural network. Feature selection methods are formed from two phases. The first one is generation procedure and the second one is evaluation function. Various types of feature selection techniques differ in the methods for the mentioned phases. Generation procedure is a phase that searches for the candidate subset of features and evaluation function defines a criterion to evaluate and compare different candidate subset of features. Feature selection methods can be categorized based on their evaluation function into 3 groups [37]. The first group is called filter methods. Filter methods use statistical measure to rank the features based on their relevance. Due to the fact that, Classifiers and estimators are not required in these methods, they are drastically fast in comparison to other groups. The second group is wrapper method which uses the classifier or estimator to find the best subset of features. Embedded methods are considered in the third group. Embedded techniques select the features based on how they affect the accuracy of the model while it is being created. In this thesis, a novel methodology for probabilistic security assessment is proposed which is online therefore, the selected feature selection technique must be fast enough for online assessment. Relief feature selection practice which is a kind of filter method, is

selected to find the most relevant features. The next section gives an introduction to relief feature selection method.

2.4.1 Relief feature selection

Relief feature selection is first introduced by K. Kira and L. A. Rendell in 1992 [38]. This practice assigns weight for each instance. Relief feature selection technique uses heuristic method for generation procedure and distance measure as an evaluation function. The method was firstly proposed for classification problems and then it is extended to regression problems also. Weights of features are initialized zero and it is updated at each iteration. The generation process starts with randomly picking samples from a training set. For each instance near hit and near miss instances are found. Near hit is an instance which has minimum Euclidean distance from the selected sample among all the instances which belongs to the class of selected sample. Near miss is an instance which has minimum Euclidean distance from the selected sample among the instances of other classes. The relevance of features are determined based how it distinguishes the instance and its near miss. The feature which is able to distinguish the instance and its near miss is considered relevant feature and it has more weight and a feature which distinguished the instance and its near miss is considered as irrelevant feature and is has less weights. The best features are selected based on their weights. The limitation of this practice is the ignorance of redundant features. In this thesis, MATLAB programming language is used for feature selection codes [41].

2.5 GP Distribution

To model the tailed regions of the distributions, Generalize-Pareto is used which is combination of Exponential and Pareto distribution [18,9]. GP distribution was first introduced by Pickands in 1975 [39].

$$f(x; \lambda) = \begin{cases} \lambda e^{-\lambda x} & x \geq 0 \\ 0 & x < 0 \end{cases} \quad (2.17)$$

$$f_X(x; x_m, a) = \begin{cases} \frac{ax_m^a}{x^{a+1}} & x \geq x_m \\ 0 & x < x_m \end{cases} \quad (2.18)$$

Three parameters μ , σ and k are sufficient to define the GP distribution. μ is location or mean parameter and σ . Represents the scale parameter. k is shape parameter.

Cumulative density function and probability density function of GP distribution is presented in equation (2.19-2.20) respectively.

$$f(x; k, \mu, \sigma) = \frac{1}{\sigma} \left(1 + k \frac{(x-\mu)}{\sigma}\right)^{\left(-\frac{1}{k}-1\right)} \text{ for } k \neq 0 \quad (2.19)$$

$$F(x; k, \mu, \sigma) = \begin{cases} 1 - \left(1 + k \frac{(x-\mu)}{\sigma}\right)^{-\frac{1}{k}} & \text{for } k \neq 0 \\ 1 - e^{-\left(\frac{x-\mu}{\sigma}\right)} & \text{for } k = 0 \end{cases} \quad (2.20)$$

The Generalized-Pareto distribution can be characterized effectively by exceedance samples. In order to define exceedance samples, a threshold should be specified for the output first. The samples whose outputs are greater the threshold are considered as exceedance samples. The distribution of the exceedance samples is approximated by GP and in this case, right tail of the original distribution is modeled. To model the left tailed region of the original distribution, it is sufficient to only set another threshold and discard the samples whose outputs are greater than the threshold. It is obvious from equations (2.19-2.20) that if shape parameter is set to zero, GP distribution will be equivalent to exponential distribution. Positive shape factor, depicts polynomial decrease in tailed region and negative shape factor outlines finite tailed distribution. Figure 2.5 illustrates GP distribution for different shape parameters.

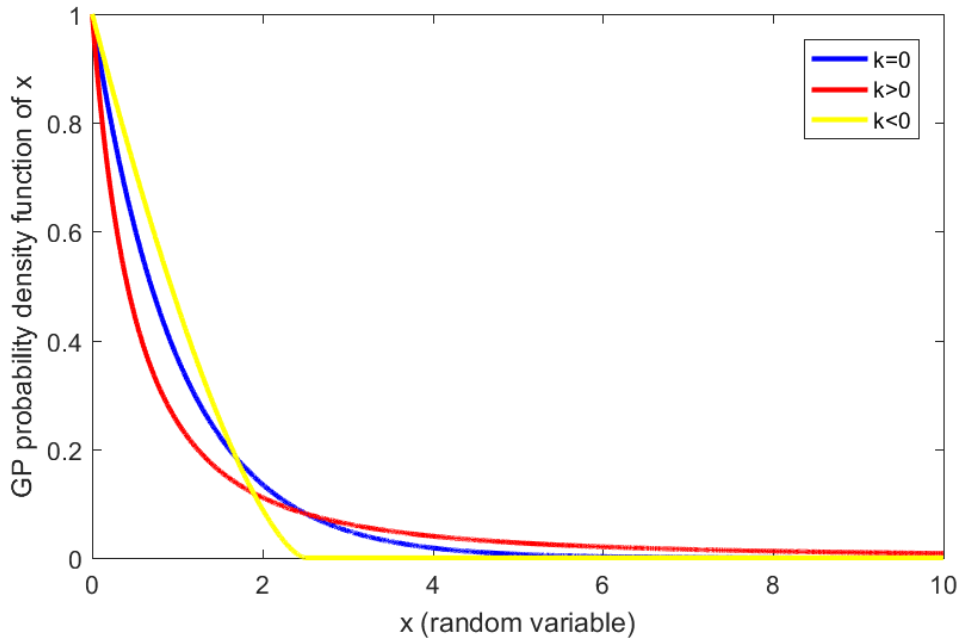


Figure 2.5: GP distribution for different k values

3. PROPOSED METHODOLOGY

In this study, an efficient methodology is proposed to make a probabilistic security assessment for small signal stability and transient stability of a bulk power system. It is assumed that a near future operating condition is forecasted. This operating condition cannot be fully known because of the uncertainties inherent in the power system with a variable generation and load demand. A sufficiently small period of time is considered in order to assume a normal or Gaussian distribution with mean equal to the forecasted value of each uncertainty. In this study, loading conditions and output active power of wind generations are the cause of uncertainty. Correlation between loads and wind generations are also considered to construct the complete multi-variate normal distribution.

In this thesis, the term, operating condition, refers to the condition that is defined by the specific values of the active power outputs of wind generation units and load demand at a given instant. On the other hand, operating point refers to the optimal power flow solution for a given operating condition. Each operating point is obtained by the optimal solution that can be defined by the resulting bus voltages of the network and the power output of all other generators in addition to the power generated by the wind generation units and the power consumed by the loads.

Figure 3.1 depicts the flowchart of proposed methodology for both transient security assessment and small signal security assessment. In general, system performance is used for both transient stability index and small signal stability index. For transient security assessment, the proposed methodology must be repeated for each critical contingency, therefore contingency scanning is required to determine the critical contingencies.

The process starts with construction of the multivariate probability distribution of uncertainties. Due to the fact that small period of time is considered for security assessment, normal distribution is assumed for each of uncertainty with a mean equal to its forecasted value and a small standard deviation. In order to generalize the model,

the correlation between the uncertainties are also considered. This step of the process is necessary for sampling.

The next step is creation of the database for training the neural network. Details of this step is given in section 3.1.

The step is followed by the process of developing a good performance estimator, which is an MLP neural network in this study.

Details of this process is presented in section 3.2. The trained neural network is used to speed up the exceedance sample generation. Conventional Monte-Carlo instances are obtained in the next step. Gibbs sampling method is used to produce the instances based on the multivariate distribution.

The obtained sample is used for a distribution fitting process. This step is explained in section 3.3. In order to increase the accuracy of the output distribution, a GP distribution is used to fit the left tailed region of the output distribution. For the GP distribution, exceedance sample is required.

Exceedance instances are produced by means of the trained neural network. Section 3.4 presents the details of exceedance sample generation.

The exceedance instances and MC instances are used to fit the distribution of the system performance. The step is followed by a risk analysis. Section 3.5 and 3.6 present the details of fitting process and risk analysis, respectively.

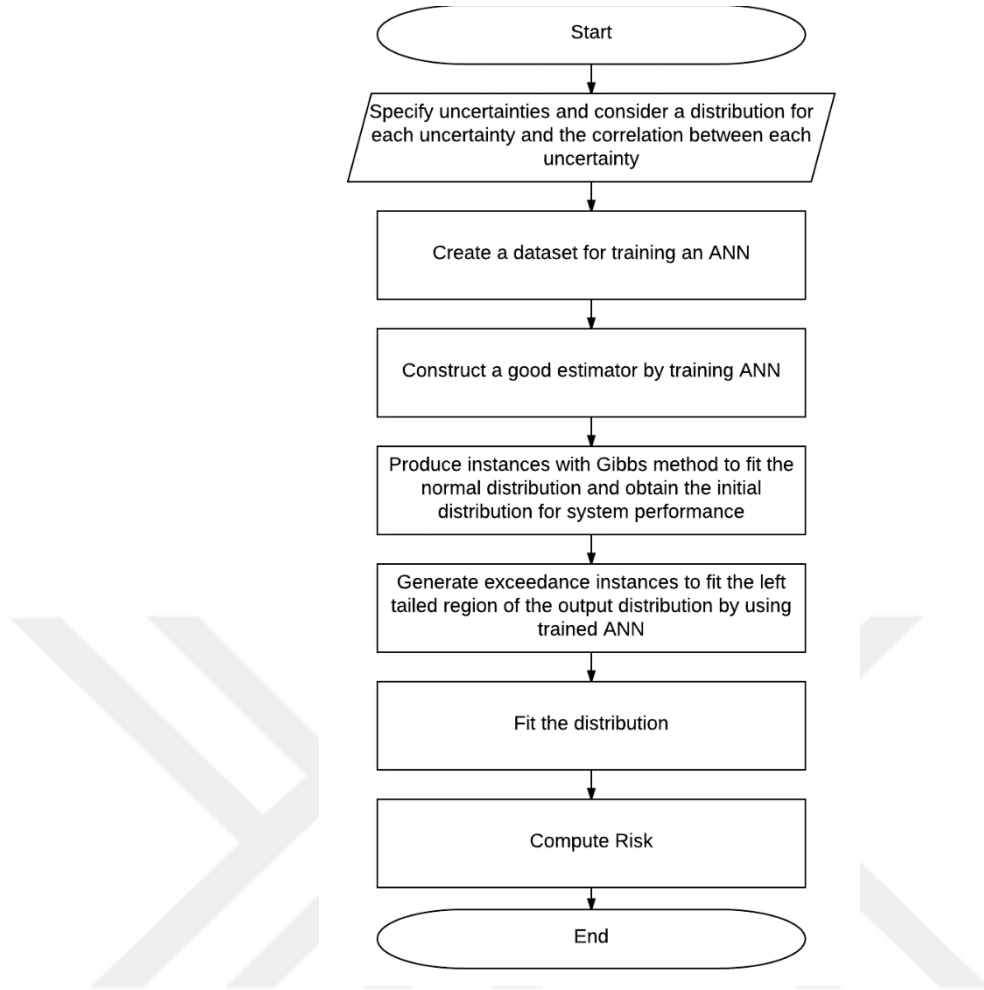


Figure 3.1: : Flowchart of proposed methodology

3.1 Create A Dataset For Training Neural Network

A limited number of different operating conditions specified by the power outputs of wind generation units and load demands are produced considering the uncertainties in their forecasted values. The created dataset will be used to train a neural network which will predict the performance of any operating point to distinguish and obtain an exceedance sample. Without using a neural network, the generation of an exceedance sample having a reasonably large size will take a long time, because an exceedance instance has a low probability to be sampled. In this case, a large number of instances will be examined through a time consuming time-domain simulation and modal analysis. Instead of a these computations, the neural network is used to predict the system's performance at the operating points, thus, it will speed up the process of obtaining the exceedance sample. A large dataset for the training the neural network

would not be appropriate, since the dynamic performance of the instances in this set can only be obtained by the time consuming dynamic simulation.

LHS sampling method is used to sample the search space evenly. For each operating condition produced, an optimal power flow solution is found and new operating points are generated. A full deterministic study of time-domain simulations and modal analyses is performed and the system's dynamic performance at each operating point is evaluated. Figure 3.2 illustrates the flowchart of the procedure.

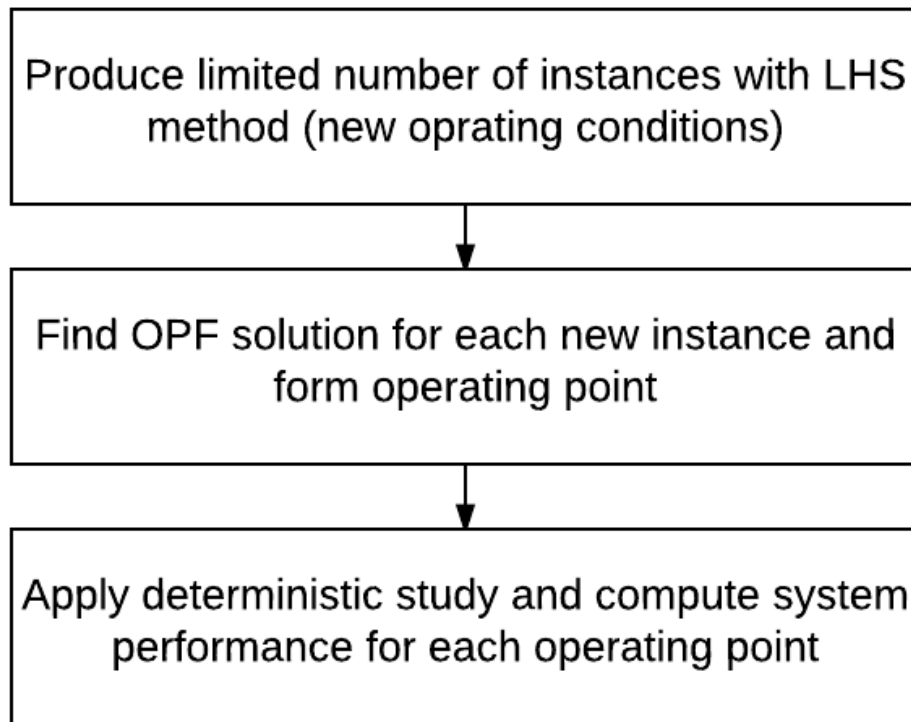


Figure 3.2: Procedure to create a dataset to train neural network

3.2 Construct A Good Predictor For System Performance

A feature selection method should be used to cut down on the number of features before training an MLP neural network. The MLP neural network is trained using the dataset covering a limited number of operating points which are generated in the previous step.

To reduce the training time and to enhance the performance of the MLP, Relief, which is a fast and a reliable type of feature selection method, is used to determine the relevant features to the regression problem of interest. The flowchart of the procedure is given in figure 3.3.

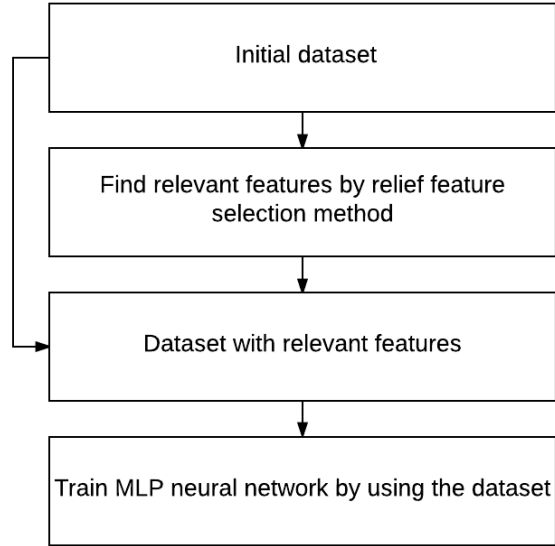


Figure 3.3: Procedure to train ANN

3.3 Generation Of Instances (OPs) Using Gibbs Sampling Method

The required number of operating points for Gibbs sampling is determined according to the sample standard deviation of the system performance and specified error and confidence level. The sample standard deviation of the system performance is calculated based on the previously generated training dataset for the neural network. In this study, the maximum tolerable error and the confidence level are specified as 5% and 95%, respectively. It means that we are 95% sure that error we make is less than 5%.

A new dataset of the probable operating points for the power system is to be produced to represent the initially guessed distributions for the dynamic performances. The Gibbs sampling method, a type of Monte-Carlo sampling, is used to generate this dataset that includes the optimal operating points distributed around the forecasted operating condition which is subjected to the changes due to the uncertainties both in generation and demand.

For each newly produced operating condition, an OPF solution is found and thus an operating point is formed. Through a series of time-domain simulations and modal analysis, the system's performance is evaluated for each operating point. The computed system performances of the generated operating points are used to fit the normal distribution of system performance. The tailed region of the obtained distribution at this step, is not accurate enough, therefore it is not appropriate to use

the generated instances to fit the tailed regions of the distribution of the system performance. For the left tailed region of the probability distribution for system performance, a GP distribution is utilized, as explained in section 3.5. The flowchart of the overall procedure is illustrated in figure 3.4.

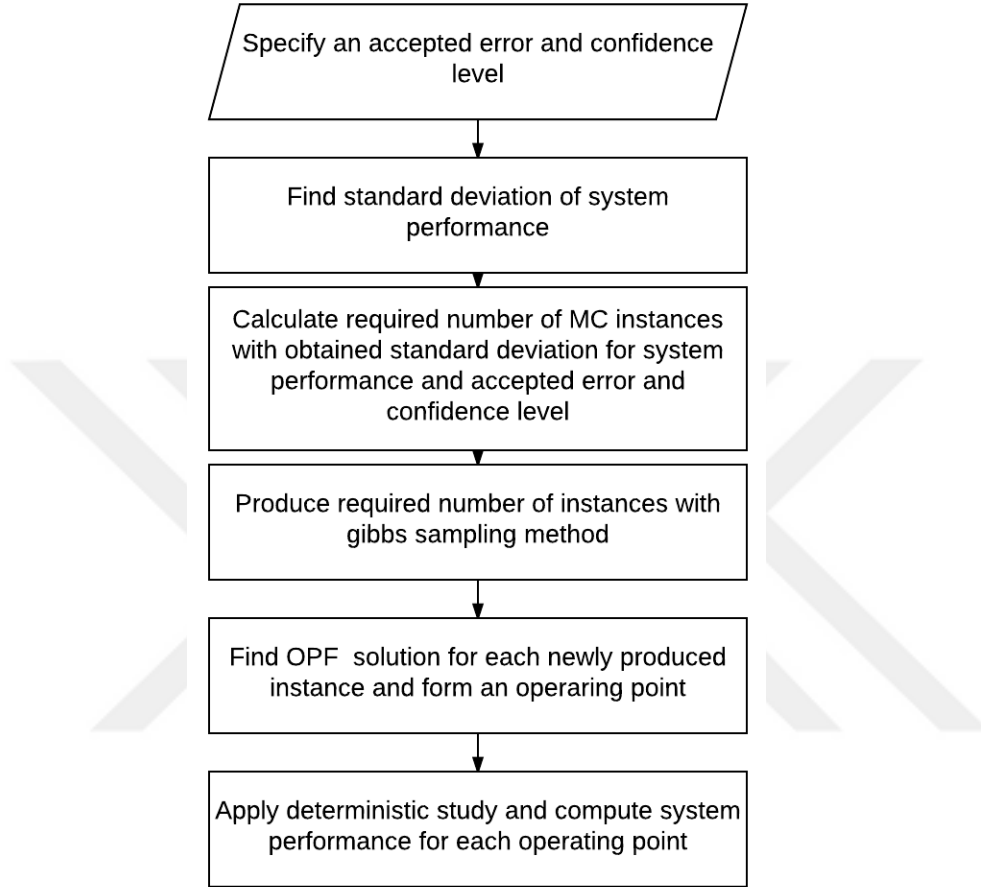


Figure 3.4: Procedure to produce operating points with Gibbs method

3.4 Generation Of the Exceedance Sample

In order to fit the tailed region of the distribution for system performance, another procedure is followed. A GP distribution is used to fit the left tailed region of the system performance distribution. For the GP distribution, an exceedance sample that includes the instances resulting in a system performance below a predetermined threshold is required [17].

Most of the instances which were produced by the Monte-Carlo sampling method were not exceedance instances due to the fact that each exceedance instance has a low

probability to be sampled. New operating conditions are produced and for each newly generated operating point, system performance is predicted by the trained neural network. If the system performance of the operating point is less than a threshold, it is considered as an exceedance case, otherwise another operating point is to be produced until a required number of exceedance instances are collected. For each exceedance instance, a deterministic dynamic security assessment is made to compute the system's performance. The system performance of the exceedance instances will be used to fit the left tailed regions of the PDF's associated with the system performance. Figure 3.5 depicts the flowchart of exceedance sample generation.

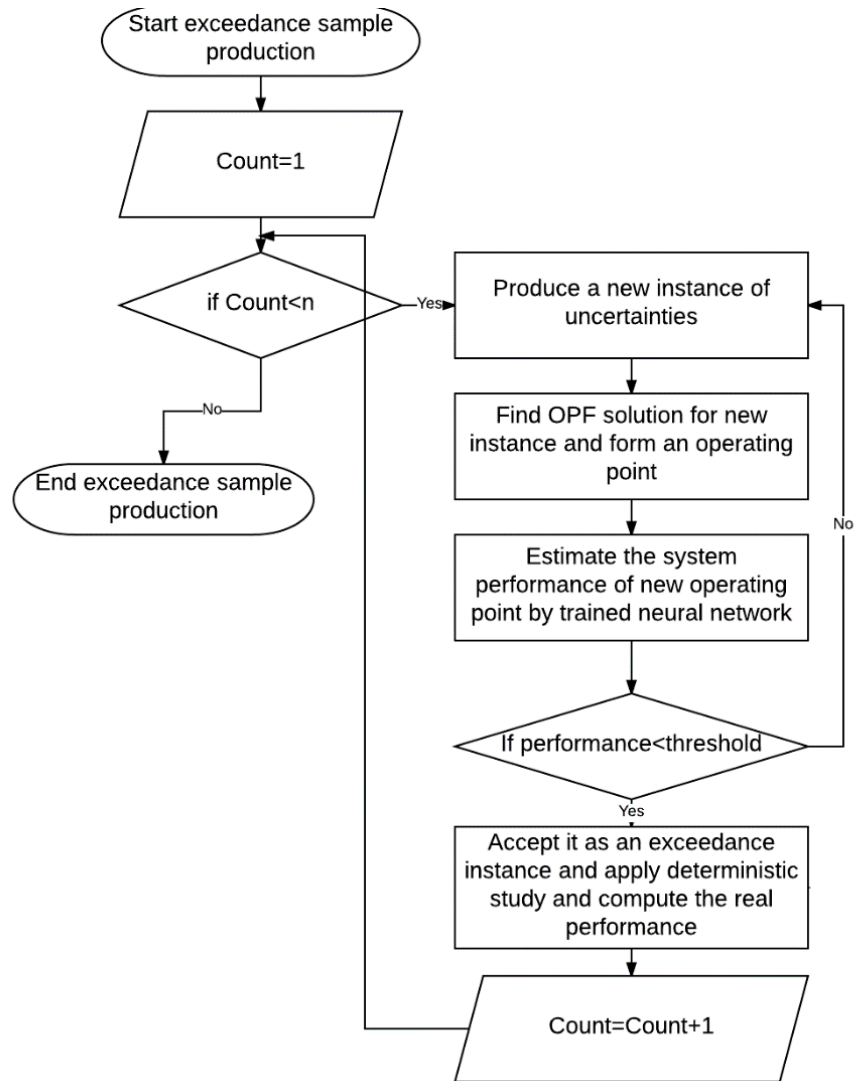


Figure 3.5: Flowchart for exceedance sample production

In figure 3.5, n represents the specified number of exceedance instances.

3.5 Fitting The Distribution

Exceedance instances are used to fit a GP distribution and Monte-Carlo instances are used to fit the normal distribution. Combination of these two distributions will give a more accurate distribution for the overall range of operating points around a forecasted one. The area under a probability density function must be 1 (Figure 3.6). Therefore, the modifications below must be done to obtain the final distribution.

GP distribution must be multiplied by $n_{exc}/$ Exceedance instances are used to fit a GP distribution and Monte-Carlo instances are used to fit the normal distribution. Combination of these two distributions will give a more accurate distribution for the overall range of operating points around a forecasted one.

A GP distribution is used for the left tailed region of the system performance distribution and a normal distribution is used for the remaining part of the distribution. It is obvious that some modifications are required to make the resulting distribution reasonable.

The area under a probability density function must be 1 (Figure 3.6). Therefore, the modifications below must be done to obtain the final distribution.

1. n_t , where n_{exc} is number of exceedance instances and n_t is total number of instances which are examined.
2. Probability of performance must be divided by a constant s , which ensures that the area under the probability density function is equal to one. Therefore, the constant s is computed as follows:

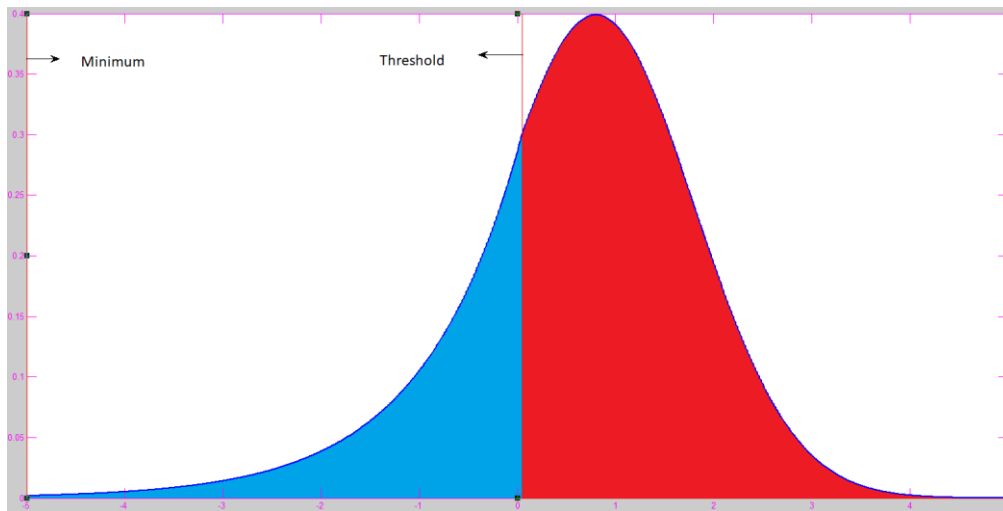


Figure 3.6: Area under the probability density function of a distribution

$$s = b + a \quad (3.1)$$

a is blue area and b is red area in figure 3.6. a and b can be computed by equations (3.2-3.6)

$$x = F_{GP}(\text{threshold}) \frac{n_{exc}}{n_t} \quad (3.2)$$

$$z = F_{GP}(\min) \frac{n_{exc}}{n_t} \quad (3.3)$$

$$y = F_N(\text{threshold}) \quad (3.4)$$

$$a = x - z \quad (3.5)$$

$$b = 1 - y \quad (3.6)$$

F_N is cumulative density function for normal distribution and F_{GP} is cumulative density function for GP distribution. *threshold* and *min* are the specified threshold and minimum of system performance respectively.

3.6 Computation Of The Risk

In order to compute the risk, a suitable severity function should be defined. In this study, a combination of a quadratic and a stepped severity function is used. Equation (3.7) and figure 3.7 represent such a severity function.

$$S(\text{perf}) = \begin{cases} 0 & \text{perf} < -0.05 \\ (perf + 0.05)^2 - 0.05 & -0.05 < \text{perf} < 0.05 \\ 1 & \text{perf} > 0.05 \end{cases} \quad (3.7)$$

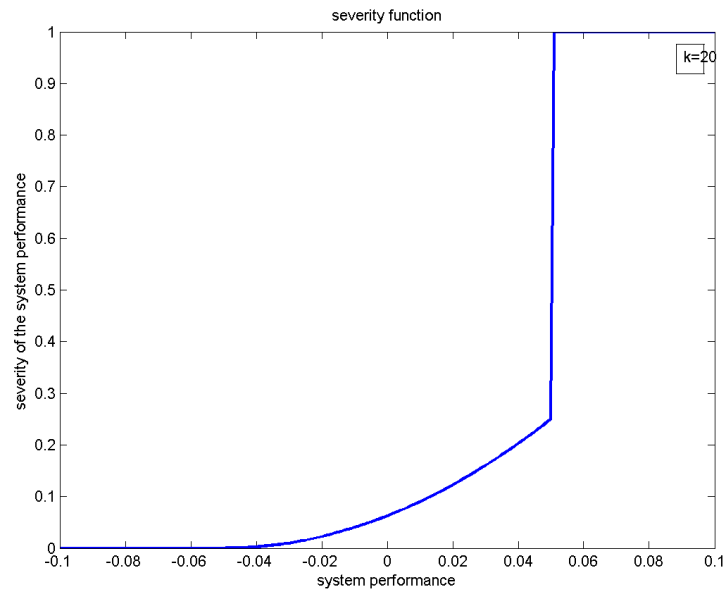


Figure 3.7: Severity function

The system's dynamic performance $perf$ for an operating point is computed for both transient and small signal angle stability assessments. For transient security assessment, the system's performance, $perf_{TS}$ is computed as follows:

$$perf_{TS} = CT - CCT \quad (3.8)$$

where CCT is the clearing time of a critical fault and CCT is its critical clearing time, both in cycles.

The system's performance associated with small signal angle stability, $perf_{SS}$ is defined as:

$$perf_{SS} = \xi_{threshold} - \xi \quad (3.9)$$

Using the severity functions defined, the risk of the operating condition for the two types of system's dynamic performances are calculated. The risk related to transient stability can be calculated with equation (3.10).

$$Risk_{TS,i} = Prob(contingency_i) \int P(perf_{TS,i}) S(perf_{TS,i}) dperf_{TS} \quad (3.10)$$

$Prob(contingency_i)$ represent the probability of the occurrence of contingency i .

The risk related to small signal angle stability can be computed by equation (3.11).

$$Risk_{SS} = \int P(perf_{SS}) S(perf_{SS}) dperf_{SS} \quad (3.11)$$

4. SIMULATION RESULTS

4.1 Test System

A modified version of the IEEE 68-bus-16-generator test system [40] is selected as the power system where all the studies in this thesis are performed. To study the efficiency of the proposed methodology of probabilistic security assessment, three wind generation units are added into the system at the buses numbered 7, 30, 37. Figure 4.1 depicts the test system with the wind generation units included.

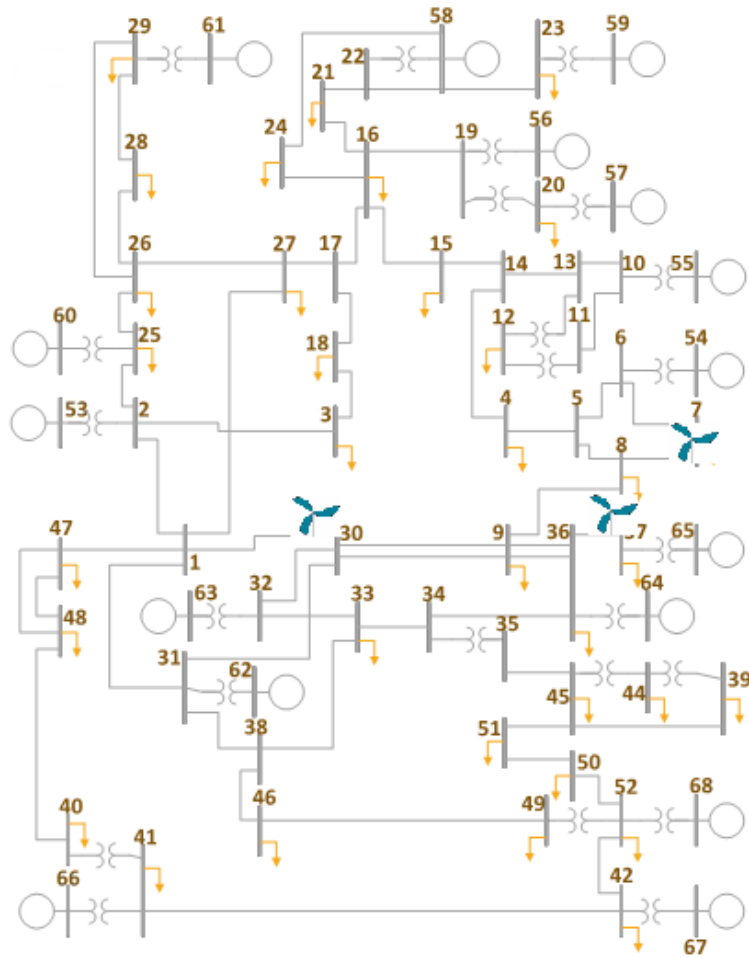


Figure 4.1: Modified 16 generator-68 bus test system

The outputs of the 3 wind generation units and the demand in 35 loads in the test system are presumed to have some uncertainty within a time frame of interest.

Since an online probabilistic security assessment has been proposed for small period of time, for each variable, a Gaussian distribution is adopted.

The correlation between loads and wind generation units are also considered in this study. The correlation coefficients are presented in table 4.1.

Table 4.1: Correlation coefficient between load and wind generation unit uncertainties

Correlation between wind generation units	Correlation between loads	Correlation between load and wind generation units
0.2	0.4	0.3

For a given operating condition that is forecasted within a small time window of interest, a probabilistic security assessment is to be performed.

Table 4.2 and Table 4.3 give the forecasted values and the standard deviations associated with the loads and the wind generation units. For all variables, the standard deviation is assumed to be 3% of their forecasted values.

The coefficients in the operating cost of each generator is given in Table 4.4. It is obvious that cost coefficient of wind generation units are zero due to the fact that generation of wind power does not have any cost. Equation (4.1) presents the cost function for each generator.

$$f_{cost,i} = C_{2,i}P_{g,i}^2 + C_{1,i}P_{g,i} + C_{0,i} \quad (4.1)$$

Equality constraint of ACOPF is power flow equation and inequality constraints are given in equations (1.14-1.19.).

Various types of solvers are available in Matpower program [43] to solve the OPF problem. In this thesis Interior Point solver is used to find an optimum solution.

OPF solution is required to change the operating condition to an appropriate operating point and determine the active output power of generators which is needed for power flow solutions.

Table 4.2: Forecasted value for active power of each load and its standard deviation

Load no	Active power demand	Standard deviation
L1	252.70	7.581
L2	322	9.66
L3	500	15
L4	234	7.02
L5	522	15.66
L6	104	3.12
L7	9	0.27
L8	320	9.60
L9	329	9.87
L10	158	4.74
L11	680	20.40
L12	274	8.22
L13	248	7.44
L14	309	9.27
L15	224	6.72
L16	139	4.17
L17	281	8.43
L18	206	6.18
L19	284	8.52
L20	112	3.36
L21	102	3.06
L22	6000	180
L23	267	8.01
L24	65.63	1.9689
L25	1000	30
L26	1150	34.5
L27	267.55	8.0265
L28	208	6.24
L29	150.70	4.521
L30	203.12	6.0936
L31	241.20	7.236
L32	164	4.92
L33	100	3
L34	337	10.11
L35	2470	74.10

Table 4.3: Forecasted value for active power of each wind generation unit and its standard deviation

Wind-generation-no	P	Standard deviation
W_{P1}	100	3
W_{P2}	200	6
W_{P3}	150	4.5

While computing an OPF solution, a quadratic operating cost function of generation units is minimized subject to the constraints as given in (1.14) - (1.19).

Table 4.4: Cost coefficient of each generator

Bus-no	Gen-no	C_2	C_1	C_0
7	1	0	0	0
30	2	0	0	0
37	3	0	0	0
53	4	0.04	20	0
54	5	0.0194	20	0
55	6	0.0149	20	0
56	7	0.0165	20	0
57	8	0.0194	20	0
58	9	0.0143	20	0
59	10	0.0194	20	0
60	11	0.0194	20	0
61	12	0.0125	20	0
62	13	0.0194	20	0
63	14	0.0099	20	0
64	15	0.0067	20	0
65	16	0.0032	20	0
66	17	0.0062	20	0
67	18	0.0099	20	0
68	19	0.0024	20	0

In this thesis, each synchronous generator is modeled with a two-axis model of order 6. For each wind generation unit, a simplified model of a Doubly Fed Induction Generator (DFIG) is used. The simplified model adopted is the one named as WGNC in TSAT [43] with the parameters listed in Table 4.5. Figure 4.2 depicts the structure of the selected DFIG model.

Table 4.5: Parameters for wind generation unit models

Bus-no	model	I	MVA	X''	L_{uplsw}	R_{LVPL}	$V_{\text{lvp}2}$	$V_{\text{lvp}1}$
7	WGNC	1	100	0.8	1	5	0.9	0.6
30	WGNC	1	100	0.8	1	5	0.9	0.6
37	WGNC	1	100	0.8	1	5	0.9	0.6

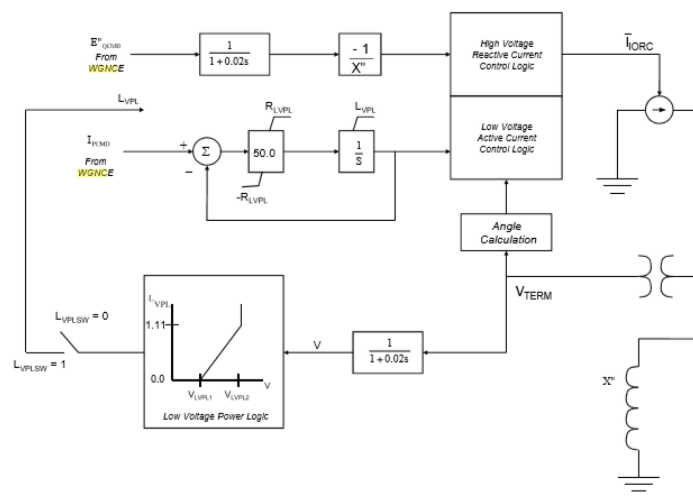


Figure 4.2: Simple structure for DFIG model [42]

4.2 Results

The proposed methodology is applied to the test system described in the previous section.

Table 4.6: Active powers of generators obtained by OPF for the forecasted operating point

Gen no	Bus no	P	Gen no	Bus no	P
1	53	253.2096	9	61	760.2177
2	54	525.1822	10	62	515.5013
3	55	674.3725	11	63	989.7671
4	56	592.1824	12	64	1578.772
5	57	505.3201	13	65	3436.831
6	58	691.1258	14	66	1504.875
7	59	505.7540	15	67	961.2510
8	60	498.6201	16	68	3920.552

Considering the nature of the problem and the computational power, a limited number of instances, as compared to the sizes of the other datasets required in the proposed methodology, are generated by Latin Hypercube Sampling technique to train the neural network. 190 instances are generated through the OPF solutions obtained by using Matpower [43]. Table 4.6 presents the output of synchronous generators which are determined by the OPF solution for the forecasted operating condition.

Time-domain simulations and modal analysis for each operation point produced have been applied and the system performance at each operating point is obtained. The Relief feature selection technique is implemented and the weight of each feature is determined. The Id numbers of the features included in the feature selection algorithm are given in Table 4.7.

Table 4.7: Id number of features

Id number of each feature	feature
1-19	Active power of generation units
20-38	Reactive power of generation units
39-73	Active power demanded of loads
74-108	Reactive power demanded of loads
109-176	Voltage of buses
177-244	Voltage angle of buses

4.2.1 Result of small signal stability analysis

In this section, probabilistic small signal stability analysis of forecasted operating condition is presented. For 190 operating point, small signal stability analysis has been performed and damping ratio of corresponded operating points are computed. For produced dataset, relief feature selection technique has been implemented. Table 4.8 and 4.8 present the weights of the features which are related to the active output power of wind generation units and active power demand of loads respectively.

Table 4.8: Weights for active power of wind generation units with respect to small signal stability index

The bus number of wind generation unit	Weight of the feature
7	-0.0014915
30	-0.0026741
37	-0.0021014

Table 4.9: Weights for active power of loads with respect to small signal stability index

Load no	Weight	Load no	Weight	Load no	Weight	Load no	Weight	Load no	Weight
L1	-0.00270	L8	-0.00201	L15	-0.0049	L22	-0.0029	L29	-0.0017
L2	-0.00218	L9	-0.00193	L16	-0.0013	L23	-0.0046	L30	-0.0023
L3	-0.00175	L10	-0.00276	L17	-0.0018	L24	-0.0021	L31	-0.0010
L4	-0.00037	L11	-0.00256	L18	-0.0012	L25	-0.0021	L32	-0.0020
L5	-0.00195	L12	-0.00177	L19	-0.0027	L26	-0.0018	L33	-0.0015
L6	-0.00204	L13	-0.00180	L20	-0.0020	L27	-0.0028	L34	-0.0024
L7	-0.00265	L14	-0.00202	L21	-0.009	L28	-0.0039	L35	-0.0026

Figure 4.3 depicts the weights of each feature which are obtained by relief approach for the small signal angle stability system performance.

Negative weights of above-mentioned features in table 4.9 and 4.8 reveal that, output of wind generation units and load demands have less relevance to the system performance. It can be concluded that, output of wind generation units and power demand of loads affect the optimal power flow solution, and the output of synchronous generators, voltage of buses and voltage angle of buses which are the solution for OPF, are more relevant.

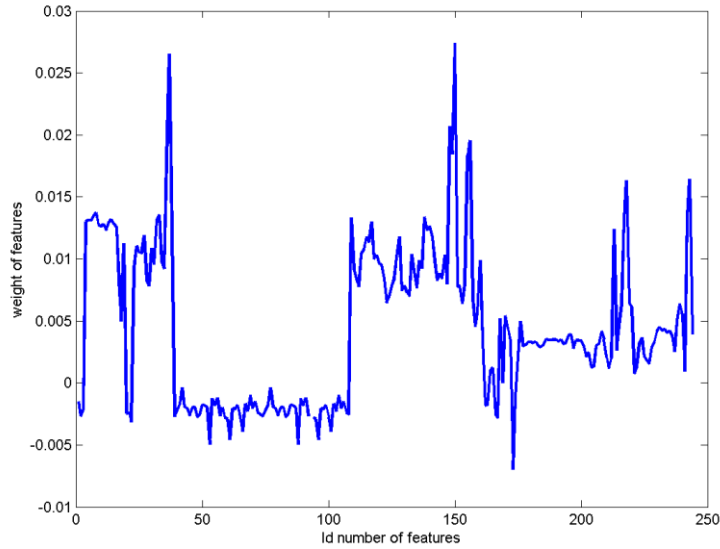


Figure 4.3: Weights of features based on small signal stability performance of the system

Relevant subset of features are selected and MLP neural network is trained. In order to observe the efficiency of relief feature selection approach, different number of features have been selected. Performance of neural network is determined for each subset of features. Table 4.10 presents the performance of MLP for different number of features.

Table 4.10: Performance of MLP neural network with different number of features

Number of features	Validation error	Training error	Test error	Time (sec)
50	2.5540e-05	9.6743e-07	7.2196e-05	14.321985
100	4.5173e-05	4.2111e-10	6.5737e-05	19.264621
150	3.3874e-05	1.0270e-06	1.0041e-04	36.292761
244	7.70804e-04	1.19142e-04	3.19039e-04	45.9559

It is obvious that, relief feature selection method contributes to increase the performance of MLP moreover it helps to decrease the length of training time.

Figure 4.4 depicts the performance of MLP which is trained by all features.

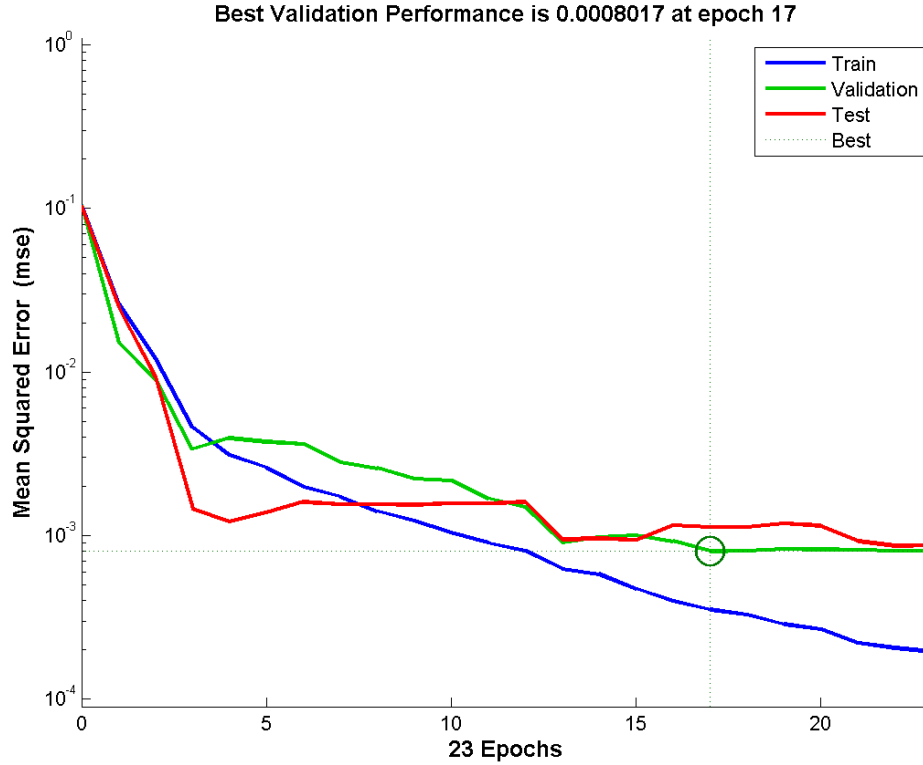


Figure 4.4: Performance of MLP for whole features corresponding to small signal stability system performance

In order to determine the number of experiments for Monte-Carlo simulation, standard deviation of the small signal stability performance of previously produced operating points is computed 0.045. Therefore required number of experiment is obtained 80 from equation (2.6) for 95% confidence level and 0.01 maximum error. Due to the reason that, these instances are used for fitting the initial distribution, maximum accepted error is considered 0.01.

80 number of operating conditions are produced by Gibbs sampling technique and OPF solution for each operating condition is obtained and small signal angle stability analysis has been performed for each of the operating points. In practice, electromechanical oscillations having damping ratio less than 3% are assumed to be critical in a power system. However, for the operating condition selected, a more conservative threshold, 5.8%, which is suitable to the distribution of the operating

points around the selected condition, is to be adopted to demonstrate the effectiveness of the proposed method towards the calculation of related risks more accurately without loss of generality. Figure 4.5 depicts the distribution of damping ratio.

Top figure of the mentioned figure illustrates the obtained distribution by using exceedance sample and the bottom one depicts the obtained distribution without using exceedance instances.

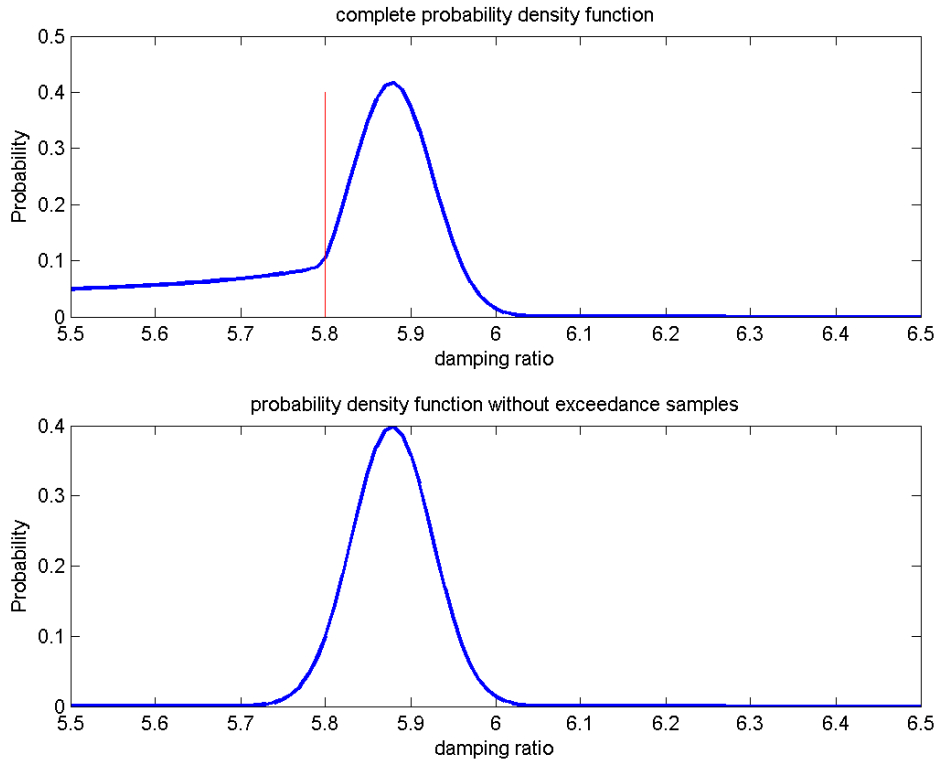


Figure 4.5: Distribution of damping ratio

Risk analysis of the forecasted operating point is done for small signal angle stability. Figure 4.6 depicts the product of probability density function and severity function total risk is the area under the obtained function.

Top figure of the above-mentioned figure is related to the function with exceedance sample and the bottom one is corresponded to the function without exceedance sample.

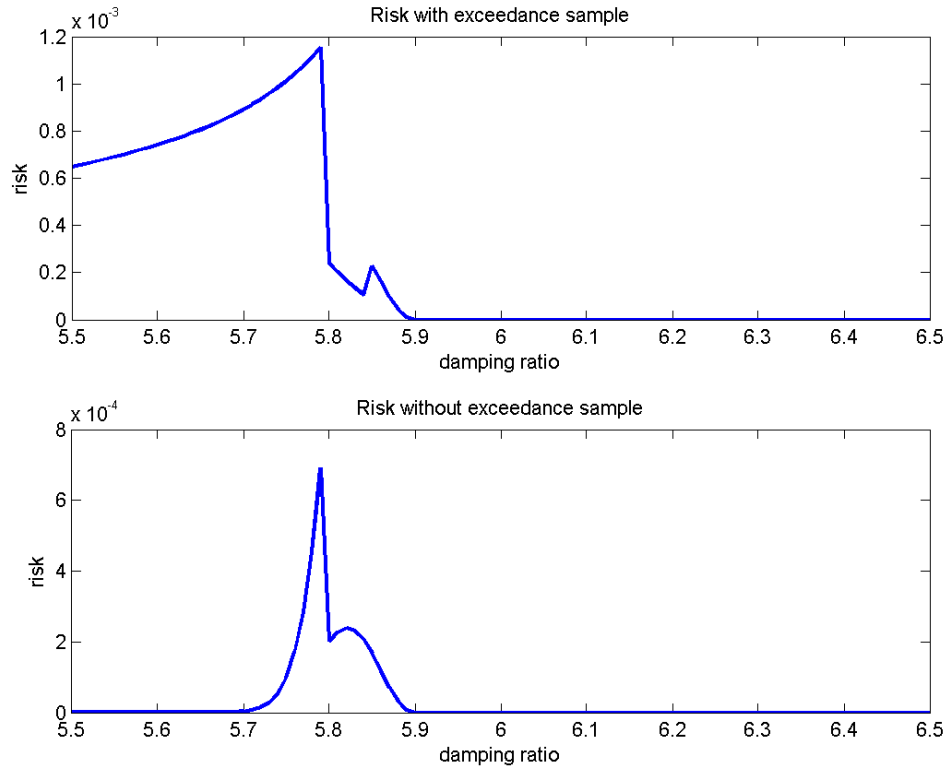


Figure 4.6: Product of probability density function and severity function for small signal angle stability

Computed total risk is presented in table 4.11.

Table 4.11: Total Risk for small signal stability of the forecasted operating condition

method	Conventional MC method	Proposed method
Risk	0.0033	0.0264

4.2.2 Results for transient stability analysis

In order to make transient stability assessment, a set of critical contingencies must be determined. Contingency scanning has been done to specify critical contingencies.

In this study, 3 critical contingencies have been specified and risk analysis has been performed for these three contingencies.

Table 4.12 presents the determined critical contingencies and their corresponded CCT values for forecasted operating point.

Table 4.12: Critical contingencies and corresponded CCT values

Contingency no	Bus	Removed Line	CCT(cycle)
78	29	29-28-1	5.2281
76	29	28-26-1	5.3892
62	22	22-21-1	5.8458

It is considered that circuit breakers act in 5 cycle therefore operating points which have CCT value less than 5 are insecure. It is obvious that, forecasted operating point can be insecure by small variation for these contingencies especially for contingency 78 therefore this forecasted operating point for the determined contingencies can be insecure and risk is not zero.

4.2.2.1 Feature selection results for each contingency

In this section feature selection results are presented for the mentioned three contingencies. Figure 4.7, 4.8 and 4.9 demonstrate weights of features for contingency 78, 76, 62 respectively which are obtained by Releif feature selection method.

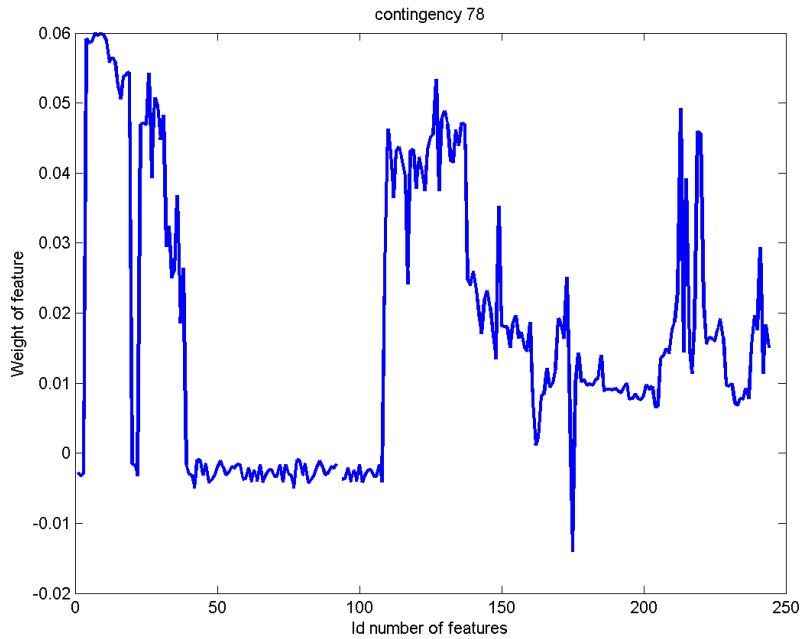
**Figure 4.7:** Weights of feature for contingency number 78

Table 4.13: Weights for active power of wind generation units contingency 78

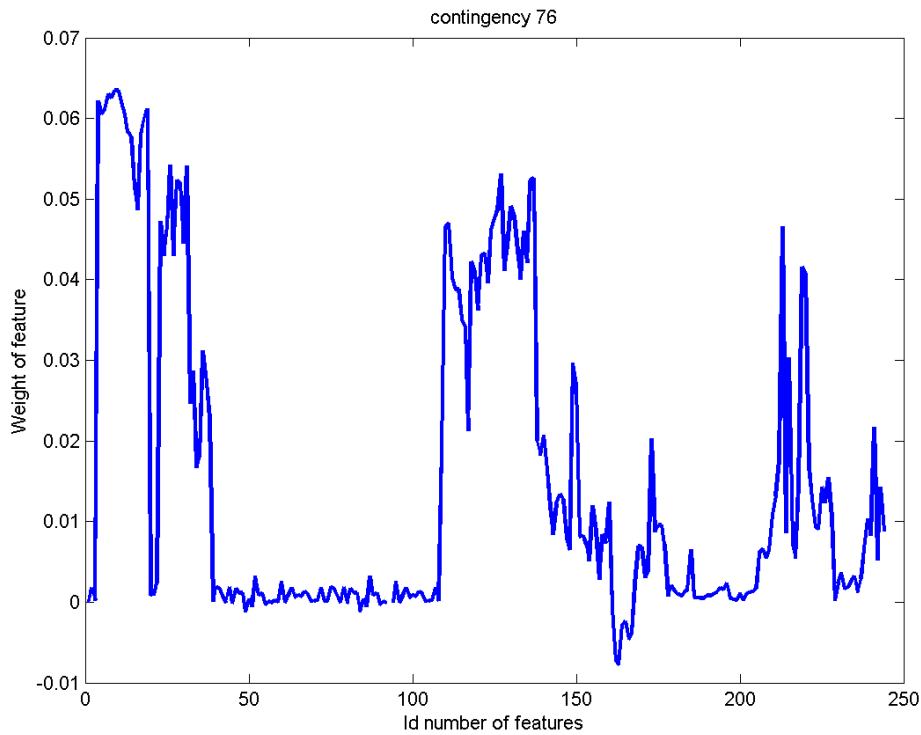
The bus number of wind generation unit	Weight of the feature
7	-0.00280
30	-0.00327
37	-0.00286

Table 4.14: Weights for active power demand of loads contingency 78

Load no	Weight	Load no	Weight	Load no	Weight	Load no	Weight	Load no	Weight
L1	-0.00162	L8	-0.00149	L15	-0.00312	L22	-0.00367	L29	- 0.00315
L2	-0.00301	L9	-0.00421	L16	-0.00276	L23	-0.00209	L30	- 0.00245
L3	-0.00320	L10	-0.00381	L17	-0.00189	L24	-0.00407	L31	- 0.00235
L4	-0.004999	L11	-0.00326	L18	-0.00218	L25	-0.00241	L32	- 0.00326
L5	-0.00096	L12	-0.00196	L19	-0.00156	L26	-0.00400	L33	- 0.00331
L6	-0.00108	L13	-0.00107	L20	-0.00166	L27	-0.00161	L34	- 0.00165
L7	-0.00162	L14	-0.00199	L21	-0.00371	L28	-0.00416	L35	- 0.00417

Table 4.15: Weights for active power of wind generation units contingency 76

The bus number of wind generation unit	Weight of the feature
7	7.98499372445943e-06
30	0.00174478703204050
37	0.000253548478796964

**Figure 4.8:** Weights of feature for contingency number 76

Negative weights of above-mentioned features in table 4.9 and 4.8 reveal that, output of wind generation units and load demands have less relevance to the system performance. It can be concluded that, output of wind generation units and power demand of loads affect the optimal power flow solution, and the output of synchronous generators, voltage of buses and voltage angle of buses which are the solution for OPF, are more relevant.

Table 4.16: Weights for active power demand of loads contingency 76

Load no	Weight	Load no	Weight	Load no	Weight	Load no	Weight	Load no	Weight
L1	0.00014	L8	0.00167	L15	0.00081	L22	0.00263	L29	0.00130
L2	0.00190	L9	0.00116	L16	0.00117	L23	- 0.00003	L30	0.00105
L3	0.00163	L10	0.00114	L17	- 0.00033	L24	0.00080	L31	0.00009
L4	0.00080	L11	-0.00119	L18	0.00006	L25	0.00177	L32	0.00018
L5	- 0.00003	L12	0.00039	L19	- 0.00008	L26	0.00058	L33	0.00022
L6	0.00190	L13	-0.00054	L20	0.00023	L27	0.00085	L34	0.00183
L7	- 0.00016	L14	0.00325	L21	0.00001	L28	0.00070	L35	0.00018

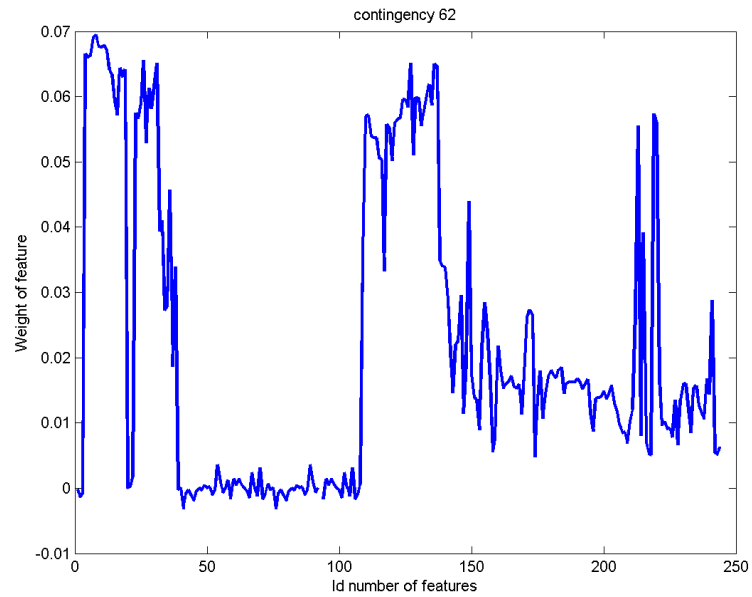
**Figure 4.9:** Weights of feature for contingency number 62

Table 4.17: Weights for active power of wind generation units contingency 62

The bus number of wind generation unit	Weight of the feature
7	-0.00019
30	-0.00138
37	-0.00079

Table 4.18: Weights for active power demand of loads contingency 62

Load no	Weight	Load no	Weight	Load no	Weight	Load no	Weight	Load no	Weight
L1	-0.00028	L8	-0.00042	L15	-0.00031	L22	0.00144	L29	0.00243
L2	0.00005	L9	-0.000007	L16	0.00362	L23	0.00048	L30	-
									0.00001
L3	-0.00318	L10	-0.00031	L17	0.00093	L24	0.00142	L31	-
									0.00119
L4	-0.00082	L11	0.00042	L18	-0.00067	L25	0.00054	L32	0.00310
L5	-0.00024	L12	0.000037	L19	0.0000007	L26	0.00003	L33	-
									0.00170
L6	-0.00117	L13	0.00020	L20	0.00119	L27	-0.00037	L34	-
									0.00094
L7	-0.00194	L14	-0.00105	L21	-0.00163	L28	-0.00148	L35	0.00066

4.2.2.2 Neural Network training results

With relevant subset of features, MLP is trained based on transient stability index targets. In order to choose best number of features, MLP is trained with different number of relevant features and performance of MLP is recorded.

Table 4.19, 4.20, 4.21 presents the performance of MLP with different number of features for contingency number 78, 76 and 62 respectively.

Table 4.19: MLP performance of different number of features for contingency 78

Number of features	Validation error	Training error	Test error	Time (sec)
50	2.5616e-04	3.1533e-05	4.3215e-04	5.067287
100	8.8812e-04	6.9510e-06	0.0017	14.172820
150	0.0051	0.0028	0.0074	30.310461
244	0.022139	1.50845e-3	0.014574	38.7898

Table 4.20: MLP performance of different number of features for contingency 76

Number of features	Validation error	Training error	Test error	Time (sec)
50	2.6393e-04	2.2867e-04	2.5238e-04	4.339165
100	0.0019	4.9872e-04	9.5878e-04	12.183364
150	0.0064	1.5791e-04	0.0048	29.973048
244	0.01006	2.05345e-4	0.00538	38.7898

Training results clear that, best performance of MLP is obtained for 50 number of features. In addition, training 50 number of feature takes less time and reduces the complexity of the model therefore 50 number of relevant features is selected to estimate the system performance for an operating point by MLP neural network. Figure 4.10, 4.11 and 4.12 depicts the best performance of MLP with 50 number of features for contingency number 78, 76 and 62 respectively.

Table 4.21: MLP performance of different number of features for contingency 62

Number of features	Validation error	Training error	Test error	Time (sec)
50	5.5270e-04	1.2710e-04	0.0010	4.313037
100	0.0024	1.1285e-05	0.0048	12.209338
150	0.0049	3.5265e-07	0.0060	30.444468
244	0.05067	1.17070e-2	0.03601	38.7898

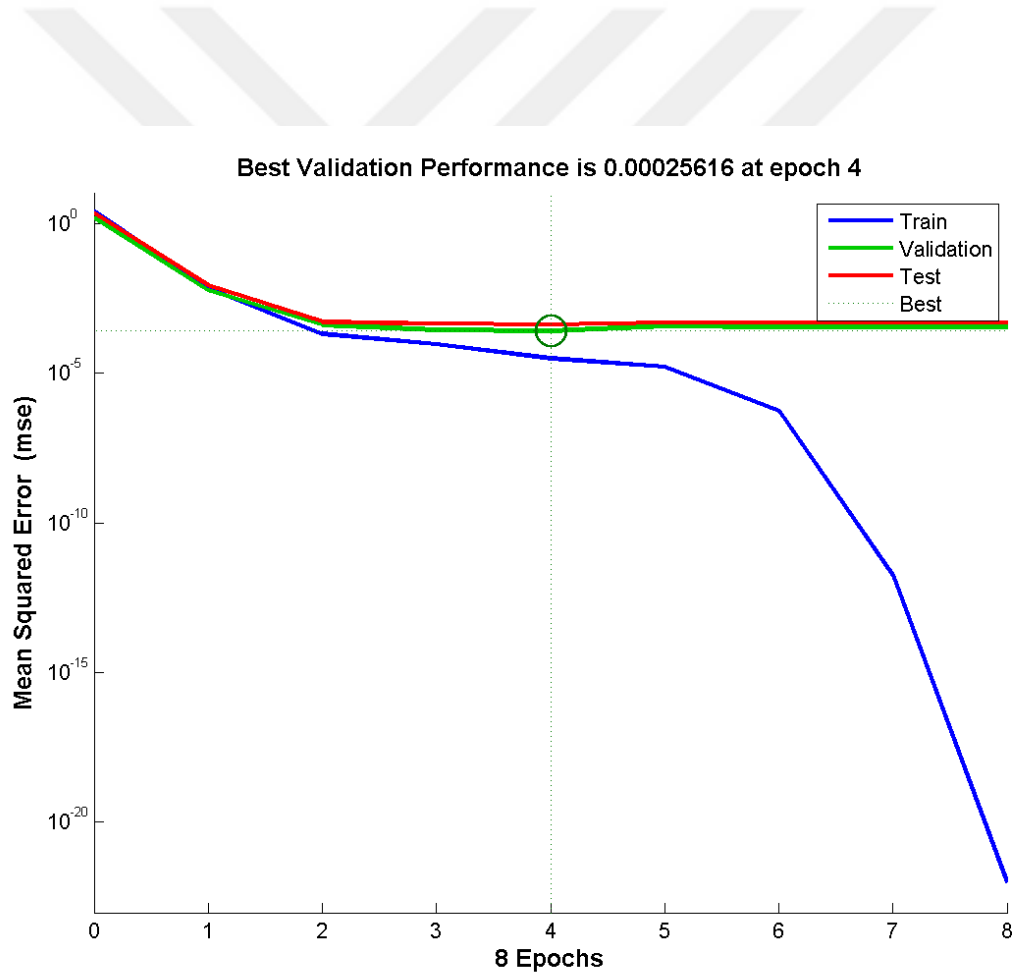


Figure 4.10: Best MLP performance with 50 number of features for contingency number 78

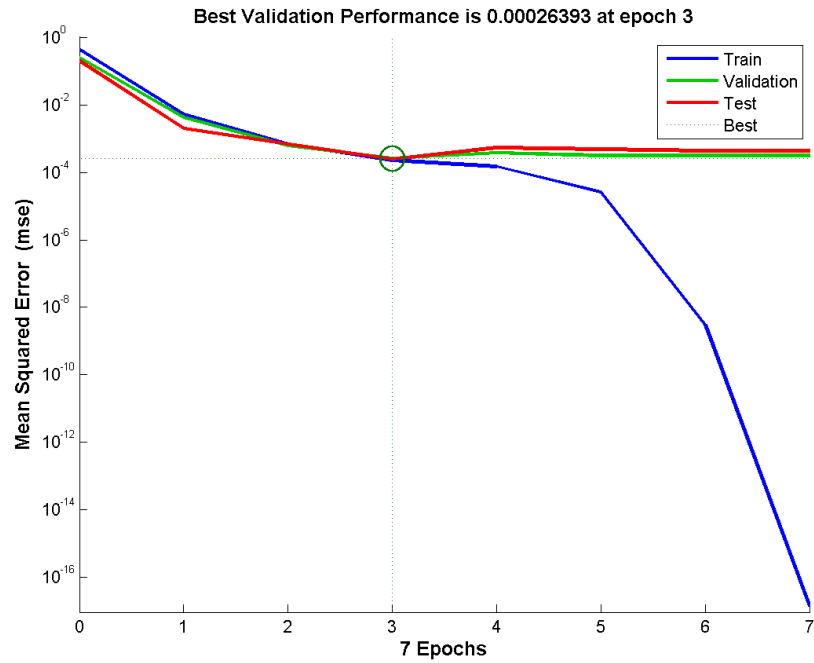


Figure 4.11: Best MLP performance with 50 number of features for contingency number 76

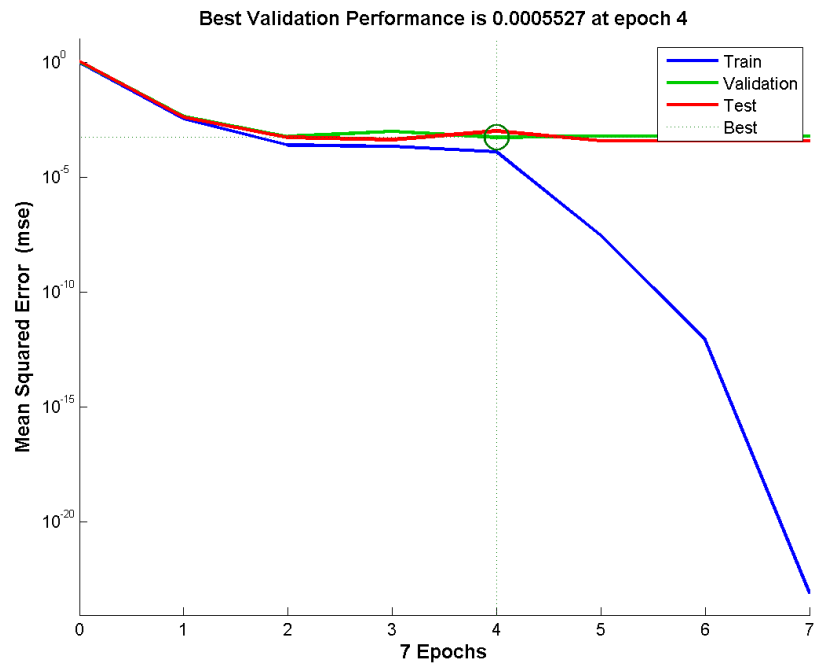


Figure 4.12: Best MLP performance with 50 number of features for contingency number 62

4.2.2.3 Results for fitting the distribution

To obtain an initial approximation of the desired distribution, system performances of MC instances which are obtained by Gibbs method is sufficient. Standard deviation for CCT values which corresponded to the operating points generated by LHS approach is obtained and Required number of MC instances for specified confidence level and accepted error and computed standard deviation is presented in table 4.22.

Table 4.22: Required number of MC instances for three critical contingencies

Contingency number	error	Confidence level	n
78	0.05	95	106
76	0.05	95	100
62	0.05	95	241

In the process of exceedance sample generation, the total number of instances including discarded instances are determined. Table 4.23 presents the total number of instances which are produced and their CCT is estimated by trained ANN in order to obtain exceedance sample.

Table 4.23: Total number of produced instances to obtain exceedance sample

Contingency number	Maximum number of iterations for exceedance sample
78	238
76	772
62	4638

The MC instances are used to fit the normal distribution which is an initial approximation of the output distribution moreover exceedance instances which are distinguished by the contribution of machine learning tools, are utilized to fit the left tailed region of the distribution. Top part of figures 4.13, 4.14 and 4.15 demonstrate the probability distribution for system performance (CCT) which is obtained with MC instances and exceedance sample for contingency number 78, 76 and 62 respectively and bottom part of the mentioned figures depicts the distribution of CCT without exceedance sample. Red vertical line in the figures separate the GP distribution and

normal distribution. Left distribution of the red line is GP distribution and right part is Normal distribution.

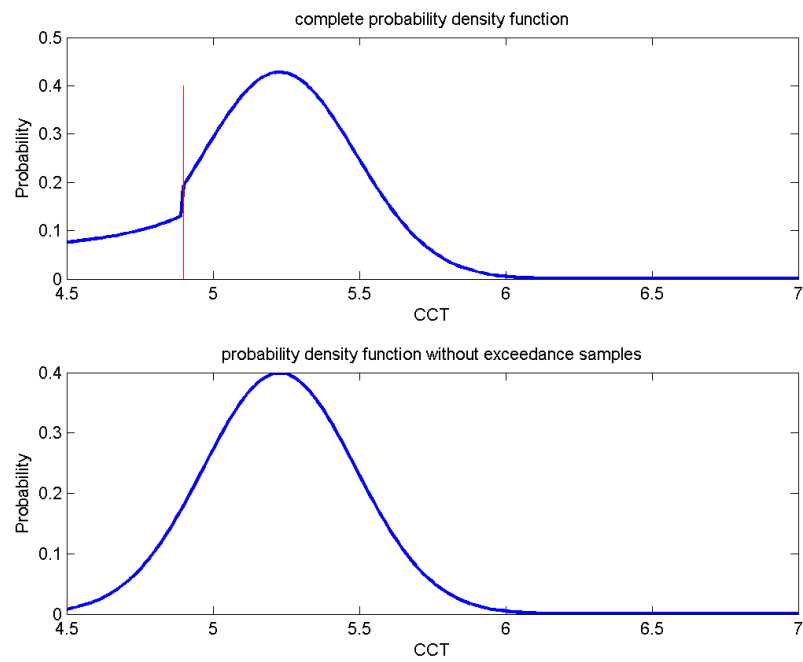


Figure 4.13: Distribution of CCT for contingency number 78

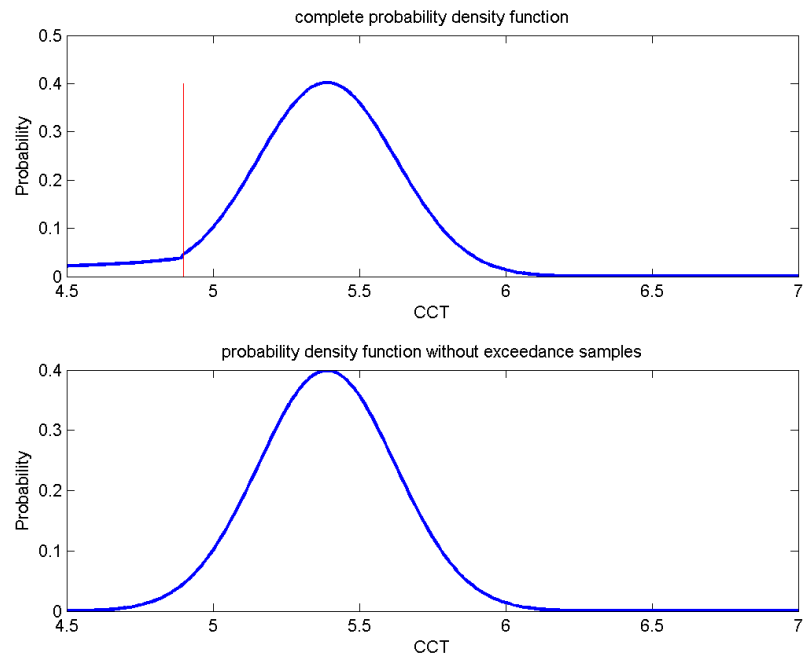


Figure 4.14: Distribution of CCT for contingency number 76

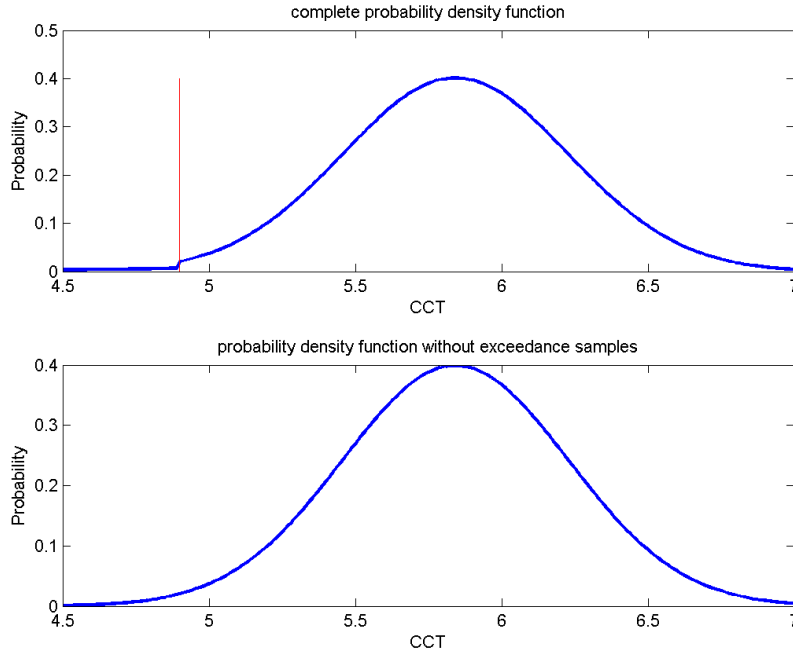


Figure 4.15: Distribution of CCT for contingency number 62

4.2.2.4 Risk analysis

In this section, risk analysis of the forecasted operating point is done for each of three contingencies. Figure 4.16, 4.17 and 4.18 depicts product of probability density function and severity function and total risk is the area under the obtained function. Top figures of the above-mentioned figures are related to the distribution with exceedance sample and the bottom ones are corresponded to the distribution without exceedance sample.

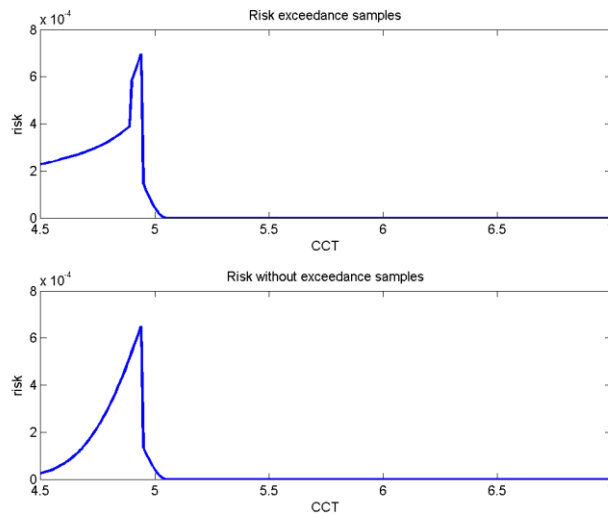


Figure 4.16: Product of probability density function and severity function for contingency number 78

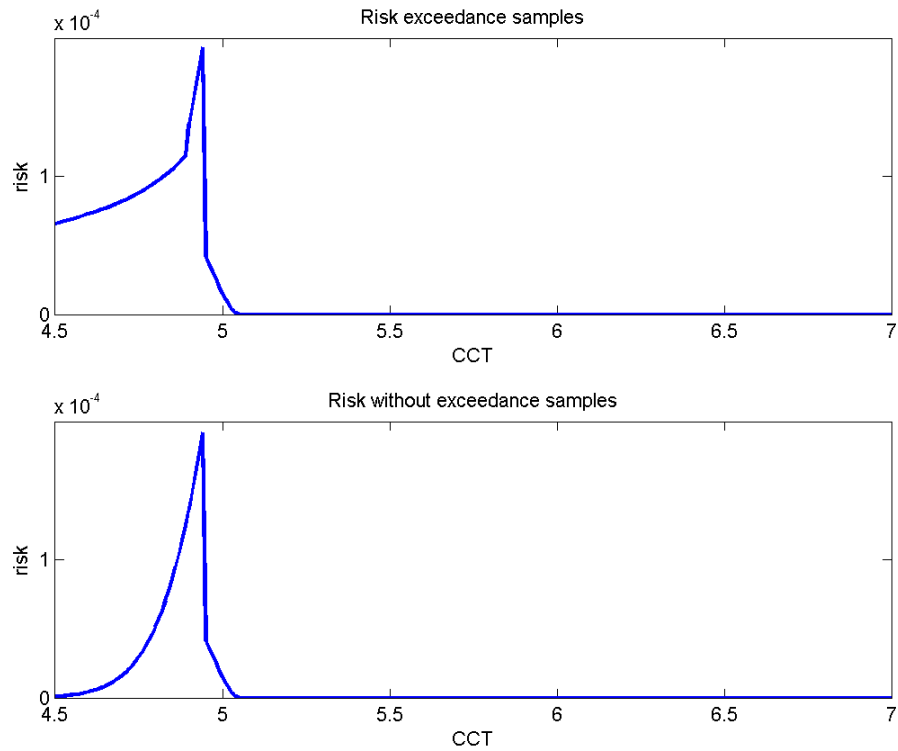


Figure 4.17: Product of probability density function and severity function for contingency number 76

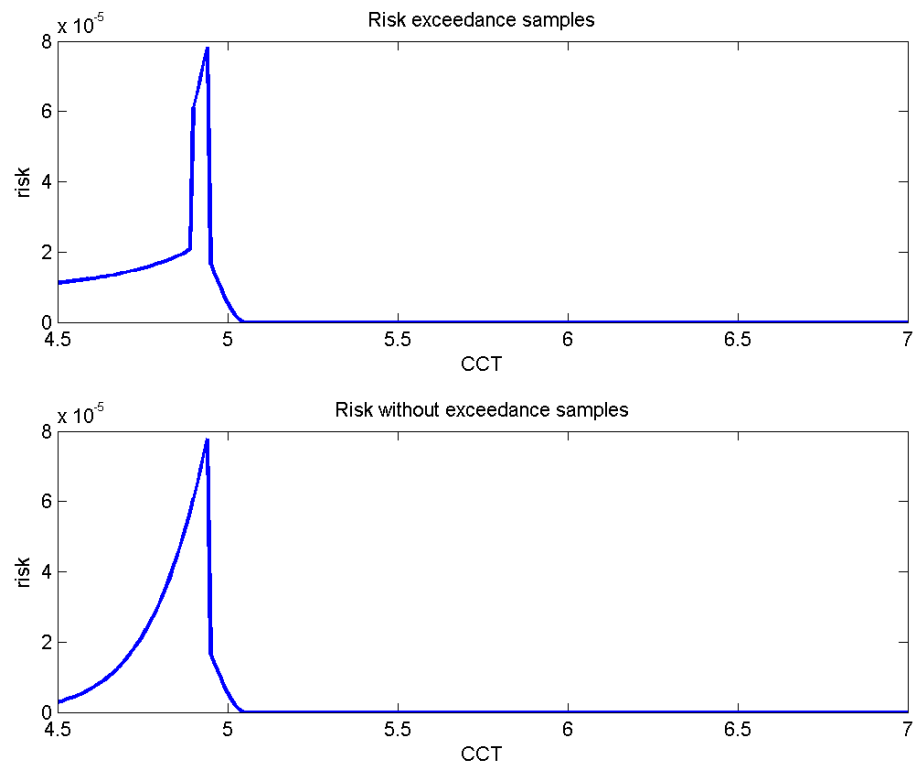


Figure 4.18: Product of probability density function and severity function for contingency number 62

4.2.2.5 Comparison of the proposed methodology with other methods

In order to show the efficiency of the proposed methodology, it is compared with deterministic security assessment and probabilistic security assessment without exceedance sample.

In addition, the proposed methodology is much faster with utilization of machine learning tools therefore its required time is compared with a probabilistic security assessment with exceedance sample generation without machine learning tools. At the end, effect of the accuracy of the estimator is studied by simple estimator model.

Table 4.24 presents the comparison of the proposed methodology with probabilistic security assessment without GP distribution and deterministic security assessment.

Table 4.24: Total risk of forecasted operating point for each contingency with different methods

Contingency number	Risk for deterministic assessment	Risk with conventional MC probabilistic assessment	Risk with proposed probabilistic assessment
78	0	0.0112	0.0154
76	0	0.0022	0.0044
62	0	0.0012	0.0010

The time duration which is required for each step of the proposed methodology is presented in tables 4.25, 4.26 and 4.27 for contingency number 78, 76 and 62 respectively.

Table 4.25: Length of time duration taken for each step of the methodology for contingency 78

Operation	Time length in seconds	Time length in minutes
Creating database for training NN	1283.012840	21.3835
Selecting features	0.524463	0.0087
Training MLP	4.060896	0.06768
Sampling with Gibbs technique	681.050741	11.3508
Fitting and risk analysis	0.545658	0.00909
Generating exceedance sample for GP	395.250705	6.5875
Total time	2364.4	39.4067

The results points out that the proposed method requires much less time than the time required for the method which uses the dynamic simulation.

In the proposed methodology, the usage of machine learning approaches for the generation of an exceedance sample decreases the computation time drastically and makes the overall method more efficient

Table 4.26: Length of time duration taken for each step of the methodology for contingency 76

Operation	Time length in seconds	Time length in minutes
Creating database for training NN	1298.364955	21.6394
Selecting features	0.597270	0.0100
Training MLP	4.231993	0.0705
Sampling with Gibbs technique	655.757746	10.9293
Fitting and risk analysis	0.545658	0.00909
Generating exceedance sample for GP	452.237899	7.5373
Total time	1573.3	26.2210

Exceedance sample generation has been done without utilization of machine learning tools and the length of time taken is recorded. Table 4.28 reveals the fact that, machine learning tools speed up the process drastically.

Table 4.27: Length of time duration taken for each step of the methodology for contingency 62

Operation	Time length in seconds	Time length in minutes
Creating database for training NN	1282.165039	21.3694
Selecting features	0.511183	0.0085
Training MLP	4.060896	0.06768
Sampling with Gibbs technique	1549.677922	25.8280
Fitting and risk analysis	0.545658	0.00909
Generating exceedance sample for GP	850.421544	14.1737
Total time	1573.3	26.2210

Table 4.28: Required time for exceedance sample generation with machine learning tools and without machine learning tools

method	Time for co 78	Time for co 76	Time for co 62
With machine learning tools	395.250705	452.237899	850.421544
Without machine learning tools	1554	5041.3	30287

5. CONCLUSION

Conventional dynamic security assessment methods using deterministic approaches are not reliable for modern power systems integrated with distributed generations introducing a considerable amount of uncertainties in systems' operating conditions. In this thesis, a novel methodology is proposed for probabilistic dynamic security assessment of a large power system.

Monte-Carlo simulation methods for probabilistic security assessment is not appropriate due to their computational burden and time consuming procedures. The Monte-Carlo approach for performing dynamic security assessment requires a large number of operating points, therefore a large computation time, to obtain an acceptable accuracy. Instead of this approach, it is more reasonable to just generate larger number of instances in the tailed region of the distribution in order to get accurate models which has done in the proposed method for the probabilistic security assessment. In the proposed methodology, exceedance instances are used to increase the accuracy of the model for left tailed region of the system performance' distribution. The risk which is computed by both proposed methodology and conventional Monte-Carlo confirms that, proposed methodology is more accurate in comparison with the conventional Monte-Carlo.

Exceedance sample generation can be time consuming due to the low probability of its instances being sampled; therefore a large number of simulations are needed to achieve a required number of exceedance instances. In this step, the proposed neural networks trained by the dataset, which is obtained by a feature selection method, contribute to speed up the process.

In order to demonstrate the efficiency of the proposed method, a test system with 68 bus and 3 wind generation units is used. The results show that the proposed probabilistic security assessment method gives a more realistic level of the risk when compared to the deterministic studies. Moreover, the method is fast enough to be applied for online probabilistic assessment. The proposed methodology is also

compared with the one using the conventional Monte-Carlo approach. The risk for one forecasted operating condition is computed with both Monte-Carlo method and the proposed methodology. Since the proposed methodology uses more information to approximate the probability distribution of the system performance, the risk obtained by the proposed method is more accurate than the risk which is computed by conventional Monte-Carlo approach. In addition, the time which is required for exceedance sample generation by the proposed methodology is compared with the time needed by using directly the dynamic simulation to distinguish the exceedance instances. The results points out that the proposed method requires much less time than the time required for the method which uses the dynamic simulation. In the proposed methodology, the usage of machine learning approaches for the generation of an exceedance sample decreases the computation time drastically and makes the overall method more efficient. The performance of MLP neural network is obtained by training the dataset with all feature and the dataset with relevant features. The results illustrate that not only the error of the MLP neural network is increased but also the training time of the neural network is reduced drastically.

As a future work, probabilistic preventive control actions will be studied to improve the risk of the forecasted operating condition.



REFERENCES

- [1] Wang, D., Yu, Y., Fu, C., Zhang, J., Wu, J., Jia, H., et al. (2005). Security Region Based Probabilistic Security Assessment of Power Transmission System a Basic Concepts. *IEEE*.
- [2] Rueda, J. L., Colome, D. G., & Erlich, I. (2009). assessment and enhancement of small signal stability considering uncertainties. *IEEE Trans*, 198-207.
- [3] Rueda, J., & Erlich, I. (2012). probabilistic frame work for risk analysis of power system small signal stability. *Proc. Inst. Mech. Eng*, 118-133.
- [4] Wu, F. F., & Tsai, Y.-K. (1983). Probabilistic Dynamic Security Assessment of Power Systems: part 1- Basic Model. *IEEE Trans*.
- [5] Wu, F. F., Tsai, Y.-K., & Yixin, Y. (1987). probabilistic steady state and dynamic security assessment. *IEEE Trans*.
- [6] Billinton, R., & Kuruganty, P. R. (1981). Probabilistic assessment of transient stability in a practical multi-machine system. *IEEE Trans*.
- [7] Ming, D., Renching, D., & Yacheng, L. (1999). Monte-Carlo simulation approach to probability stability. *Journal of Tsinghua University*.
- [8] Yongji, G. (1989). The principle application of power system reliability. *Tsinghua University*.
- [9] Verbic, G., & Canizares, C. A. (2006). probabilistic optimal power flow in electricity markets based on two point estimate method,. *IEEE Trans*, 1883-1893.
- [10] Preece, R., Wooley, N. C., & Millanovic, J. V. (2013). The probabilistic collocation method for power system damping and voltage collapse studies in the presence of uncertainties. *IEEE Trans*, 2253-2262.
- [11] Steinmetz, C. P. (1920). Power control and stability of electric generating stations. *AIEE Trans*, 1215-1287.
- [12] Vassell, G. S. (1991). Northeast blackout of 1965. *IEEE Power Engineering Review*, 4-8.
- [13] Wang, X., Song, Y., & Irving, M. (2008). *Modern Power Systems Analysis*. Springer.
- [14] Kundur, P. (1944). *Power system stability and control*. New York: Mc.Graw-Hill.
- [15] Grigsby, L. L. (2001). *the electric power engineering handbook*. IEEE Press.
- [16] Bettiol, A. L., Wehenkel, L., & Pavella, M. (1999). Transient stability-constrained maximum allowable transfer. *IEEE Transactions on Power Systems*.
- [17] Preece, R., & Milanovic, J. V. (2015). Risk-based Small-disturbance Security Assessment of Power Systems. *IEEE PES Transactions*.
- [18] Ruhle, O., & Lerch, E. (2010). *Ranking of system contingencies in DSA systems – First experiences*. IEEE Power Energy Soc. Gen. Meeting.
- [19] Ni, M., McCalley, J. D., Vittal, V., & Taysib, T. (2003). Online Risk-Based Security Assessment. *IEEE Trans*.

- [20] **Carpentier, J.** (1979). Optimal power flows. *International Journal of Electrical Power and Energy Systems*, 3-15.
- [21] **Cain, M. B., O'Neill, R. P., & Castillo, A.** (2012). *History of Optimal Power Flow and Formulations*.
- [22] **Dunn, W., & Shultis, J. K.** (2011). *Exploring Monte Carlo Methods*.
- [23] **van Ravenzwaaij, D., Cassey, P., & Brown, S. D.** (2016). A simple introduction to Markov chain Monte-Carlo sampling. *Psychonomic Bulletin and Review*.
- [24] **Smith, A. F., & Roberts, G. O.** (1993). Bayesian computation via the Gibbs sampler and related Markov chain Monte Carlo methods. *Journal of the Royal Statistical Society*.
- [25] **Stein, M.** (1987). *Large sample properties of simulations using latin hypercube sampling*. Technometrics.
- [27] **Kucuktezcan, C. F., & Genc, V. M.** (2010). Dynamic security assessment of a power system based on Probabilistic Neural Networks. *IEEE Pes Innovative Smart Grid Technologies Conference Europe*.
- [28] **Bettiol, A. L., Souza, A., Todesco, J. L., & Tesch, J. R.** (2003). Estimation of critical clearing times using neural networks. *IEEE Power-Tech Conference*. Bologna.
- [29] **Wang, S. C.** (2012). *Interdisciplinary computing in Java programming* (Vol. 743). Springer Science & Business Media.
- [30] **Rosenblatt, & Frank, X.** (1961). *Principles of neuro-dynamics: perceptrons and the Theory of Brain Mechanisms*. Washington DC: Spartan Books.
- [31] **Rumelhart, David, E., Hinton, G. E., & Williams, R. J.** (1986). *Learning Internal Representations by Error Propagation*. MIT Press.
- [32] **Alpaydin, E.** (2010). *Introduction to machine learning- 2nd edition*. Massachusetts Institute of Technology.
- [33] **Levenberg, & Kenneth .** (1944). *A Method for the Solution of Certain Non-Linear Problems in Least Squares*. Quarterly of Applied Mathematics.
- [34] **Marquardt, & Donald.** (1963). An Algorithm for Least-Squares Estimation of Nonlinear Parameters. *SIAM Journal on Applied Mathematics*.
- [35] **BenBassat, M.** (1982). *Pattern recognition and reduction of dimensionality*. North Holland: Hand books of Statistics.
- [36] **John, G. H., Kohavi, R., & Pfleger, K.** (1994). irrelevant features and the subset selection problem. *Eleventh International Conference on machine learning*, (pp. 121-129).
- [37] **Guyon, I., & Elisseeff, A.** (2003). An Introduction to Variable and Feature Selection. *Journal of Machine Learning Research* 3, 1157-1182.
- [38] **Kira, K., & Rendell, L. A.** (1992). The feature selection problem: traditional methods and new algorithm. *Ninth National Conference on Artificial Intelligence*, (pp. 129-134).
- [39] **Pickands, J.** (1975). *Statistical Inference Using Extreme Order Statistics*. The Annals of Statistics.
- [40] **Rogers, G.** (2000). *Power system oscillations*. Norwell: MA: Kluwer.
- [41] (1994). Retrieved from mathworks.
- [42] (2015). *Transient security assessment tool model manual*. Powertech Labs Inc.

- [43] **Zimmerman, R. D., & Murillo-Sánchez, C. E.** (2016). *Matpower 6.0 User's Manual*. Power Systems Engineering Research Center .





CURRICULUM VITAE



Candidate's full name: Sevda Jafarzadeh

Place and date of birth: Tabriz, Iran – 26/07/1993

E-mail Address: jafarzadeh.sevda@gmail.com

Education:

B.Sc.: Electrical Engineering – University of Tabriz