Reliability Evaluation of Renewable Generation Integrated Power Grid including Adequacy and Dynamic Security Assessment

Final Project Report

S-75

Power Systems Engineering Research Center
Empowering Minds to Engineer
the Future Electric Energy System

Reliability Evaluation of Renewable Generation Integrated Power Grid including Adequacy and Dynamic Security Assessment

Final Project Report

Project Team

Chanan Singh, Project Leader

Texas A&M University

Vijay Vittal

Mojdeh Khorsand

Arizona State University

Graduate Students

Dogan Urgun

Sai Kiran Kanchari Bavajigari

Texas A&M University

Yingying Wang

Arizona State University

PSERC Publication 19-07

November 2019

For information about this project, contact:

Chanan Singh
Texas A&M University
Department of Electrical and Computer Engineering
P.O. BOX 875706
Tempe, AZ 85287-5706
Phone: 979-845-7589
Fax: 979-845-6259
Email: singh@ece.tamu.edu

Power Systems Engineering Research Center

The Power Systems Engineering Research Center (PSERC) is a multi-university Center conducting research on challenges facing the electric power industry and educating the next generation of power engineers. More information about PSERC can be found at the Center's website: http://www.pserc.org.

For additional information, contact:

Power Systems Engineering Research Center
Arizona State University
527 Engineering Research Center
Tempe, Arizona 85287-5706
Phone: 480-965-1643
Fax: 480-727-2052

Notice Concerning Copyright Material

Acknowledgements

The work described in this report was sponsored by the Power Systems Engineering Research Center (PSERC).

We express our appreciation for the support provided by PSERC's industry advisors including Orlando Ciniglio (Idaho Power); Bajarang Agrawal (APS); Evangelos Farantatos (EPRI); Anish Gaikwad (EPRI); Alan Engelmann (COMED); Feng Zhao (ISO-NE); Mark Westendorf (MISO), Lynn Hecker (MISO); Thompson Adu (MISO).

Executive Summary

The goal of this project is to investigate a reliability evaluation approach for power systems that are undergoing changes including increased penetration of renewable resources and the associated impact on system dynamic performance. These changes have brought complexities to reliability studies both for resource adequacy assessment and dynamic security assessment (DSA). Adequacy and dynamic security, the two fundamental functions of bulk power system reliability, have been analyzed separately in the system planning practice. With increasing proportion of converter-interfaced generation, the reliability evaluation methodology needs to be revisited to incorporate the evaluation of the resource adequacy as well as the adequacy of stability control capability. It is also recognized that traditional deterministic methods are not capable of addressing the increased sources of uncertainty. Alternatively, probabilistic methodology is inherently more suited to represent stochasticity and therefore is the analysis method of choice for reliability evaluation in this project.

A probabilistic reliability evaluation approach with adequacy and DSA integrated in a single framework is developed in this research. The research investigates the techniques for adequacy assessment and DSA, developing the quantification metrics for the assessment results that provide a solution to integrate adequacy assessment and DSA in an integrated framework. The sequential Monte Carlo Simulation (MCS) is selected as the analysis method as it is well suited for the analysis of large systems and is capable of considering time-vary characteristics in systems. The use of MCS for the assessment of a large number of cases can be time consuming, therefore, acceleration methods for the integrated reliability evaluation process are investigated to improve the practical application of the approach.

This report consists of three parts. Part 1 studies the assessment methods of resource adequacy and dynamic security, as well as the MCS framework to integrate the resource adequacy assessment and DSA. To improve the computational efficiency, an acceleration method for MCS and a contingency pruning process for DSA, are developed and implemented as part of the integrated reliability evaluation approach. As a proof-of-concept, the proposed approach is demonstrated on a synthetic test system and the simulation results illustrate the efficacy of an integrated reliability evaluation approach. Part 2 is focused on improving the speed of computation by using machine learning. A time-consuming operation in the adequacy assessment is the use of OPF to evaluate if in a given state the load can be satisfied. It is shown that machine learning can successfully replace OPF for this purpose. Part 3 is a study on the characteristics of the Importance Sampling. This part provides a compact reference to Cross Entropy based Importance Sampling and studies the sensitivity of computational efficiency provided by it under various situations. This can provide some guidance under which conditions it is advantageous to use this variance reduction technique.

Part I: Integrated Reliability Evaluation including Adequacy and Dynamic Security Assessment

Power system reliability assessment includes two aspects: adequacy and security. The purpose of adequacy assessment of a power system is to ensure that there are enough generation and transmission resources to supply the aggregate electric power and energy requirements of the electricity consumers while taking into consideration scheduled and reasonably expected, unscheduled outages of system components. Security assessment, on the other hand, focuses on the ability of the electric system to withstand sudden disturbances or unanticipated loss of system components. Both system adequacy and security are considered essential to provide the desired level of reliability of power systems but the two aspects are traditionally treated separately in practice.

In this part of the report, an integrated reliability evaluation combining both adequacy and dynamic security using a probabilistic approach is formulated and tested. With an integrated reliability evaluation, not only is the reliability of the post-contingency stable equilibrium point (SEP) assessed but also addresses the DSA related issue of whether a stable transition to a post-fault SEP can be achieved. The approach evaluates the system response for each selected contingency, examines transient stability in the transition from the pre-contingency to the post-contingency period, and finally evaluates the post-contingency steady-state equilibrium to ascertain that all flow limits and voltage limits are satisfied. Stochastic characteristics including renewable resources, component failures, and load variations are taken into consideration in probabilistic sampling and dynamic performance modeling. System adequacy and dynamic security are quantified in terms of MW load curtailment from the respective assessment processes and are incorporated into the calculation of the integrated reliability indices. The outcome of the integrated reliability evaluation provides the probability, frequency, and magnitude of system loss of load.

There are two important steps in the integrated reliability assessment; 1) the methodology for system state generation and 2) the technique used for state evaluation to determine the reliability indices. MCS provides a framework of probabilistic analysis based on modeling the stochastic characteristics in the system. Each system state sampled from MCS consists of generator and transmission contingency information and the system pre-contingency steady-state operating condition. The impacts of a contingency on system adequacy and dynamic security are assessed using optimal power flow (OPF) analysis of the post-contingency steady-state condition and time-domain simulation for examining the dynamic behavior, respectively. The procedure for each step is as follows:

- In each iteration of the sequential MCS, the system states are generated and incorporated with an annual 8760-hour load curve and wind speed curve. State evaluation is then conducted for each state to determine both the adequacy and the dynamic security. Dynamic security is first evaluated. When dynamic security evaluation indicates that there is a stable transition to a post-fault stable equilibrium point (SEP), the adequacy assessment is then conducted to evaluate the post-fault steady-state condition. A state is seen as a successful state only if it is steady-state stable and transiently stable. The duration time and load curtailment of all failure states are then utilized in the calculation of integrated reliability indices. Loss of load probability (LOLP), expected power not supplied (EPNS), and loss of load frequency (LOLF) are used to quantify the integrated reliability. After each iteration of the SMCS, reliability indices are updated. The coefficient of variance (COV) is calculated after each SMCS iteration to determine the convergence. The final reliability indices are obtained after the MCS converges.
- The adequacy assessment addresses the steady-state system condition after a contingency. Power flow analysis is used to assess system steady-state conditions. Considering realistic power system operating limits, AC power flow is used as the steady-state analysis technique in this work to take into consideration both bus voltage magnitude and thermal limits of transmission components. When the AC power flow results show that there is a violation of

operating limits, an AC optimal power flow (ACOPF) with remedial actions including generation rescheduling and load curtailment is deployed to optimally correct the abnormal system conditions. The remedial actions, first prioritize generation rescheduling, without curtailing any load. If the operating constraint violations still exist, remedial actions with load shedding capability are then applied. The minimum amount of load shedding that is needed to alleviate operating constraint violations is determined from the AC OPF study. In this work, the OPF package in PSS/E is utilized to solve this optimization problem.

- While adequacy studies evaluate post-contingency steady-state conditions, the DSA estimates whether the system can transition to an acceptable stable equilibrium point after a contingency. Time-domain simulation (TDS) is widely recognized as the most accurate method to describe power system transient behavior and therefore is the method of choice to perform the DSA in this work. As a result of the enhanced component modeling capabilities in TDS, critical protection systems can be modeled and their resulting actions can be simulated in the TDS. The protection systems that activate generator tripping and load shedding are represented such that following a contingency the action of these critical protections systems renders the power system to be transiently stable. Considering the requirements from the North American Electric Reliability Corporation (NERC) Protection and Control (PRC) standards, to stabilize the system, the protection systems include under-frequency load shedding, over frequency and under frequency generator tripping, over voltage and under voltage generator tripping are considered, and dynamic security is quantified by the MW load shedding due to security preserving corrective protection action.

The proposed integrated reliability evaluation procedure has been tested on a synthetic power system, which is generated to represent the realistic test system comprising of Ontario and upstate New York. The test system consists of major features of a realistic power system for transient stability and reliability studies for system planning. Compared to the traditional reliability evaluation which evaluates the adequacy and dynamic security separately, the proposed method provides the reliability indices reflecting both adequacy and dynamic security based on the quantification of the two aspects of reliability in terms of load curtailments. The quantification of the impact of dynamic security is included by the load curtailment from protection action to maintain system stability after contingencies. By including the value of load curtailment into the calculation of reliability indices, the integrated system reliability can be represented using the well-recognized reliability indices which are LOLP, EPNS, and LOLF. The results from case studies show the importance of considering the two reliability aspects together since both the steady-state and transient system performance need to be analyzed in reliability studies. Also, the two acceleration methods are verified with significant improvement of the computational efficiency. In practical applications, further methods can be explored to further reduce the computational burden.

As an outcome of this project, a research-grade software has been developed to implement and demonstrate the proposed approach.

Part II: Deep Learning and Multi-Label Learning Based MCS Methods for Composite System Reliability Evaluation

The preferred technique for composite power system reliability evaluation has emerged to be the Monte Carlo Simulation. However, this method can be quite computation intensive especially for high reliability systems. This is because for high reliability systems it takes longer time to collect enough samples with loss of load to calculate the reliability indices. This part of the report is focused on exploring the use of machine learning to improve the computational efficiency of the composite system reliability evaluation.

The reliability evaluation is performed in three phases. In the first phase samples are generated, in the second phase the samples are evaluated to determine if a problem exists and then this information is used in the third phase to find estimates of reliability indices. Most of the computation effort comes from the first two phases.

The computational effort coming from the first phase has been addressed by the use of Importance Sampling which appears to be a good tool for sampling from a distorted distributed with lower variance. In the second phase for adequacy evaluation, the time-consuming part is the optimal Power Flow whether DC or AC Power flow is used. This report is focused on exploring the use of machine learning to replace the OPF. The following results have been obtained.

- A novel method to evaluate reliability indices for composite power systems is introduced with a combination of MLL and MCS. The proposed method is implemented for MLKNN and MLRBF classifiers to identify status of buses. The case studies show that the method significantly reduces the computational burden of MCS without sacrificing accuracy. Additionally, this method advances the state of the art of using machine learning in power system reliability evaluation from the previous methods by including computation of bus indices and the transmission line failures. Moreover, the work done shows that the proposed method can be combined with the well-known variance reduction technique of IS. The outcomes for this approach show this methodology improves time efficiency of MCS even further.
- Deep learning structures are investigated to evaluate composite system reliability evaluation through MCS. A well-known deep learning topology, CNN, is implemented to characterize sampled system states for both AC and DC flow models. The results show that computational efficiency for classification using AC flow model is much higher than DC flow model since AC flow equations require nonlinear programming techniques while DC flow equations can be solved with linear techniques. The results obtained show that the proposed architecture performs state characterization with a high accuracy with COV equal or less than .01. The computation time is dramatically reduced.
- This study demonstrates that the application of the proposed machine learning methods on composite power system reliability evaluation accurately determines the system status with a substantial speed up compared with OPF based Monte Carlo Simulation methods. At the same time the accuracy is not sacrificed.
- Using machine learning to enhance the process of dynamic security assessment as well as the OPF should be investigated further. Then perhaps the Time Domain Simulation could be used to train CNN to replace this time consuming operation. This could make the inclusion of security assessment in composite reliability assessment computationally efficient and attractive to users.

Part III: Using Importance Sampling in Monte Carlo Simulation - Computation Time Sensitivity Studies

The results obtained by Monte Carlo are only estimates of true values and not the true values. Therefore the estimates have a variance. The estimates approach the true values as the variance of estimates is reduced by increasing the sample size. Importance sampling helps further by reducing the variance of the estimator and thus a smaller sample size is needed to get the same coefficient of variation. The coefficient of variatin determines the gap between the upper and lower bounds with a given level of confidence. The smaller the coefficient of variation, the tighter are the bounds around the true values. The main advantage of using variance reduction technique of Importance sampling is the reduction in computational time. This section explores the conditions under which the computation time is reduced more favorably by implementation of Importance Sampling and thus it becomes advantageous to use this variance reduction approach. It is shown that in general, the conditions which lead to higher computation time for the straight MCS tend to favor the use of Importance Sampling for relatively higher reduction of computation time by reducing the variance of estimates. The conditions which lead to higher computation time are either the ones that lead to higher reliability, i.e. , lower loss of load probability or the ones where tighter bounds on estimates are needed to have higher confidence in the estimated results.

Project Publications:

[1] Y. Wang, V. Vittal, M. Khorsand, C. Singh. (2019). Probabilistic Reliability Evaluation Including Adequacy and Dynamic Security Assessment. IEEE Transactions on Power Systems. PP. 1-1. 10.1109/TPWRS.2019.2923844.

[2] U. Dogan, C. Singh and V. Vittal "Importance Sampling using Multi Label Radial Basis Classification for Composite Power System Reliability Evaluation." IEEE Systems Journal (Under Publication).

[3] U. Dogan, C. Singh "LSTM Networks to Evaluate Composite Power System Reliability Evaluation with Injected Wind Power" 2019 IEEE Power & Energy Society (PES) General Meeting.

[4] U., Dogan, and C. Singh. "Power System Reliability Evaluation using Monte Carlo Simulation and Multi Label Classifier" National Power Systems Conference (NPSC). 2018.

[5] S.K. Bavajigari, C. Singh, "Investigation of Computational Advantage of using Importance sampling in Monte Carlo Simulation", NAPS 2019.

Student Theses:

[1] Yingying Wang. Integrated Reliability Evaluation including Adequacy and Dynamic Security Assessment. PhD Dissertation, Arizona State University, Expected: August 2021.

[2] Dogan Urgun. An Investigation on Deep Learning and Multi-Label Learning for Composite System Reliability Evaluation. PhD Dissertation, Texas A&M University, February, 2019.

Part I

Integrated Reliability Evaluation including Adequacy and Dynamic Security Assessment

Vijay Vittal
Mojdeh Khorsand
Yingying Wang, Graduate Student

Arizona State University

For information about this project, contact

Vijay Vittal
Arizona State University
School of Electrical, Computer, and Energy Engineering
P.O. BOX 875706
Tempe, AZ 85287-5706
Phone: 480-965-1879
Fax: 480-727-2052
Email: Vijay.Vittal@asu.edu

Power Systems Engineering Research Center

The Power Systems Engineering Research Center (PSERC) is a multi-university Center conducting research on challenges facing the electric power industry and educating the next generation of power engineers. More information about PSERC can be found at the Center's website: http://www.pserc.org.

For additional information, contact:

Power Systems Engineering Research Center
Arizona State University
527 Engineering Research Center
Tempe, Arizona 85287-5706
Phone: 480-965-1643
Fax: 480-727-2052

Notice Concerning Copyright Material

Table of Contents

List of Figures

List of Tables

GLOSSARY

ACOPF	Alternating current optimal power flow.
BPS	Bulk power system.
CCT	Critical clearing time.
CE	Cross-Entropy.
CE-IS	Cross-Entropy based Importance Sampling.
COV	Coefficient of variation.
DSA	Dynamic security assessment.
EENS	Expected energy not served.
EPNS	Expected power not supplied.
IS	Importance sampling.
LOLF	Loss of load frequency.
LOLP	Loss of load probability.
MCS	Monte-Carlo simulation.
MTTF	Mean time to failure in the unit of hours.
MTTR	Mean time to repair in the unit of hours.
NSMCS	Non-sequential Monte-Carlo simulation.
NERC	North American Electric Reliability Corporation.
PEBS	Potential energy boundary surface.
PDF	Probability density function.
TEF	Transient energy function.
VRT	Variance reduce technique.
OPF	Optimal power flow.
SEP	Stable equilibrium point.
SMCS	Sequential Monte-Carlo simulation.
TDS	Time domain simulation.
WTG	Wind turbine generator.

1. Introduction

1.1 Background

Power systems are designed and expected to provide reliable electrical energy to the customers with an acceptable degree of continuity and quality. However, power systems are exposed to disturbances all the time and failures in any part of the system may cause interruptions of power supply to end-users. The task of maintaining a reliable power supply is not easy. To maintain a high level of reliability in bulk power systems (BPS) is even more complicated and requires considerations of both system planning and operating time-frames. The North American Bulk Power System is an interconnected power system and its reliability needs to be maintained through both wide-area interregional planning and coordinated system operation.

According to NERC's definition, the reliability in a BPS is the degree to which the performance of the elements of that system results in power being delivered to consumers within accepted standards and in the amount desired [1]. The reliability assessment considers two basic and functional aspects: adequacy and security. The adequacy assessment of a power system addresses the ability to have the necessary generating capability and transmission line capacity. Its purpose is to ensure that there are enough generation and transmission resources to supply the aggregate electric power and energy requirements of the electricity consumers, taking into consideration scheduled and reasonably expected unscheduled outages of system components. A security assessment, on the other hand, focuses on the ability of the electric system to withstand sudden disturbances or unanticipated loss of system components, considering static security which analyzes post-disturbance system conditions and dynamic security which studies the transient period after contingencies occur. Security assessment addresses the consequences of contingencies and determines the system security margin after the contingency. In power system planning, considering both adequacy and security aspects are essential to provide the desired level of reliability for power systems. The current practice in the industry is that the two aspects are treated separately.

Both adequacy and security issues in power systems have become more challenging with the growing penetration of renewable energy resources in the system. The main sources of this challenge stem from the uncertain nature of the renewable sources like wind and solar generation and the new equipment failure characteristics and altered inertia in the system due to the necessary power electronics interface to the grid. Additionally, the deregulation of power generation and open access to transmission have driven the system to be operated in a more competitive environment and closer to the security margin. All of these new features require that the evaluation of system reliability needs to take into consideration more stochastic factors and probabilistic methods to provide more comprehensive metrics to characterize system adequacy and security.

The planning process evaluates power system reliability using different methods and objectives to assess adequacy and security, respectively. Such objectives and methods are typically known as either deterministic methods or probabilistic methods. Probabilistic methods are widely used in power system adequacy assessment. For security assessment, deterministic methods are the primary approaches used to measure system operating reliability, such as the percentage of reserve in generation capacity planning and the single-contingency principle in transmission planning. The basic weakness of deterministic methods is that they do not respond to the probabilistic nature of power system behavior including load variation, component failures, and other such factors. Probabilistic assessment methods have drawn more attention in this new environment with the increasing penetration of renewable resources and lower security margins. It is recognized that the traditional deterministic methods to incorporate the increased sources of uncertainty are not adequate and are likely to yield pessimistic results.

It is known that a power system must be both resource adequate and statically and dynamically secure to provide a high level of reliability. In current practice, system adequacy and security assessment for reliability evaluation have been treated separately. One reason for this is that assessment approaches widely used are different and the computation requirements to handle adequacy and stability at the same time can be significant.

In system adequacy assessment, following a contingency, the faulted components are assumed to be disconnected from the grid immediately and the system is assumed to return to a stable state with suitable generation rescheduling to facilitate minimum load curtailments. Although the generation rescheduling optimization problem may converge to a feasible solution representing a steady-state operating point, a stable transition to a post-fault SEP is not guaranteed. Hence, it is important to integrate dynamic security in the evaluation and provide a more comprehensive indication of the system reliability level. As a key concern of dynamic security, transient stability should, therefore, be considered in the reliability evaluation to address the system's ability to maintain stability during the transition, even though transient stability assessment is computationally burdensome and for this reason is often ignored in the overall reliability evaluation.

1.2 Objectives

Adequacy and stability analysis are important for ensuring a reliable bulk power system. With increased penetration of renewable sources and power electronics interfaces, the importance of an integrated reliability study to capture the resource adequacy and adequate control capability to maintain system stability has increased. Utilities and other stakeholders in the power grid need tools to simulate the behavior of the systems and compute risk-based measures to provide a rational basis for decision making. This project provides a tool that can be used to assess the reliability indices integrating adequacy and stability and thus provide a more realistic tool for decision making.

- Designing the Monte Carlo Simulation (MCS) framework: MCS has emerged as a preferred approach for probabilistic analysis in large power systems. MCS deals with the issue of dimensionality by sampling states based on their probabilities and drawing the inference from the sample with a convergence criterion to ensure sufficient accuracy. One of the objectives is to develop a probabilistic approach for reliability evaluation integrating system adequacy assessment and dynamic security assessment into a single framework.

- All relevant stochasticity in the system will be investigated to develop appropriate probability model representations. With proper stochastic models, those factors are thus brought into Monte Carlo simulation to sample system states.

- Integrated reliability assessment techniques: For the assessment of a state in integrated reliability evaluation, what will be examined includes the system dynamic performance after the disturbance and the post-contingency steady-state performance. The assessment techniques that can accurately evaluate the performance and can be brought into the integrated reliability measurements are investigated.
- Dynamic system performance simulation is a time-consuming process, therefore, tools that can pre-determine the stability or instability of a case can significantly reduce the computational efforts. A pruning process will be developed to determine whether stability evaluation is warranted. Another acceleration will be investigated to improve the computation efficiency of the MCS process.
- Develop quantitative metrics that integrate adequacy and dynamic security evaluation results, and develop a research-grade software to implement and demonstrate the tool.

1.3 Organization of the Report

The report is organized as follows. Chapter 2 presents a detailed literature review of the existing work in this area. Chapter 3 presents the mathematical component outage model and dynamic model for reliability evaluation, and the Monte-Carlo Simulation (MCS) method for probabilistic system analysis. Chapter 4 discusses the integrated reliability evaluation approach considering both system adequacy and dynamic security, which are assessed by steady-state analysis and transient stability analysis respectively. Two acceleration methods to speed up the integrated reliability evaluation are discussed. Chapter 5 presents the results of the application of the proposed methods along with discussions of the results. Chapter 6 summarizes the main conclusions of the work done and identifies additional research issues to be pursued as a continuation of this project.

2. Literature Review

Reliability assessment has been a continuous concern in power systems. Extensive work has been done to evaluate the adequacy and security aspects of reliability. Probabilistic analysis approaches have gained more attention because they incorporate a more complete and comprehensive representation of the system. A literature survey of existing methods of probabilistic reliability evaluation and a comprehensive consideration of adequacy and security into a single framework of reliability evaluation is presented in this Chapter.

2.1 Integrated Reliability Assessment considering Adequacy and Dynamic Security

In most power system reliability studies, the quantitative measures of the overall system reliability are based on the adequacy evaluation which only addresses the steady-state analysis. Dynamic security is an important factor to be considered in the reliability assessment since it represents whether the system is reliable before the system reaches a post-contingency SEP. Although the importance of dynamic security and its significant influence on the overall reliability is well recognized, there are limited efforts dedicated to the integrated reliability evaluation considering both adequacy and dynamic security analysis.

Felix F. Wu, et al, in [5], proposed a conceptual framework of a unified approach to probabilistic steady-state and dynamic security assessment. The authors introduced the time to insecurity as the metric of system security. The probability distribution of the time to insecurity is obtained from the solution of a linear vector differential equation whose coefficients are expressed in terms of steady-state and dynamic security regions. Reference [6] presented an integrated approach to reliability evaluation including a probabilistic assessment of transient stability. In the proposed method, system stability is determined by a sampled fault clearing time. If the fault clearing time is within the critical clearing time, the system is seen to be able to maintain transient stability after the contingency. Comprehensive reliability metrics are presented in the same paper. The probability of transient instability and the mean time to instability are used to measure the dynamic aspect of reliability. LOLP, LOLF, LOLD, and EENS are used as the composite system indices. A method of composite power system reliability including both static and dynamic processes after contingency is proposed in [7]. Sequential Monte-Carlo Simulation (SMCS) is used in this method for system states selection. In the proposed method, the reclosing time is evaluated to classify fault to be transient or permanent. The loss of load during restoration under permanent and transient faults are calculated based on a corrective OPF. Reference [8] introduced a framework for extending conventional probabilistic reliability analysis to account for system stability limits including transient and voltage stability issues. An intelligent system which is a combination of neural network and a fuzzy neural network is used to predict the security aspects on the reliability indices based on generator angle after a fault. Andrea M. Rei, et al, in [9], presented a method for integrating adequacy and security reliability evaluation using SMCS to capture stochastic features in power systems. Transient stability is evaluated by comparing the critical energy based on CCT and the PEBS calculated based on TEF. More recently, Benidris, Mitra, and Singh in [10] proposed an integrated evaluation of the reliability and stability method. A direct method is utilized for transient stability assessment based on computing the energy margin of the system under fault. Three probabilistic transient stability indices are proposed to address system instability in the reliability indices calculation.

2.2 Probabilistic Assessment of System Adequacy

Historically, reliability assessment of power systems has been conducted using deterministic approaches. Based on a list of pre-determined contingencies with some important power system components, power system planners incorporate sufficient redundancy so that system failures can be prevented [11-13]. In the complex operation of modern power systems, the need for systematic analysis of power system reliability becomes more critical. Probabilistic approaches have received considerable attention because they can assess not only the severity of an event but also the likelihood of its occurrence. Extensive work has been done for probabilistic reliability assessment but with an emphasis on system adequacy evaluation [12-15]. W. Li, in [15], addresses the fact that probabilistic reliability assessment is generally associated with four tasks: determining components outage models, the selection of system states, evaluating the consequences of the selected system states and indices calculating. The enumeration method and MCS are essentially the two different methods for system states selection. Billinton and Li in [16] compared the results using the MCS method with those obtained using the enumeration method and proposed a combined Monte-Carlo method based on a random sampling technique with an analytical approach. Both types of methods have been widely used with respective merits. However, because of the so-called curse of dimensionality, where the enumeration number increases exponentially with the number of components, the exhaustive enumeration is computationally intractable. For the assessment of a large power system, the MCS method is preferred.

The drawback of Monte-Carlo simulation is the computation time, especially when the probability of failure state occurrence is very small. Efforts have been made by many researchers to lower the variance and reduce MCS time. High-speed computing and programming paradigms have been explored to improve Monte-Carlo simulation efficiency. Gubbala and Singh in [17] described two random number generation schemes and three topologies for parallelizing the fixed interval MCS for reliability evaluation of interconnected power systems. In [18], Borges, Falcao, Mello, and Melo demonstrated two parallel methods for composite reliability evaluation using SMCS. In one of the methods, a complete simulation year is analyzed on a single processor and the many simulated years necessary for convergence are analyzed in parallel. In another method, the assessment of the system states in the simulated years is performed in parallel and the convergence is checked on one processor at the end of each simulated year. Other computational techniques such as intelligent agent technology based on which reliability evaluations are assigned to different agents [19-20], artificial neural networks [21], and object-oriented programming (OOP) [22] have also been proposed to reduce the computational cost. Another way to make the MCS method more time efficient is to use variance reduction techniques (VRTs). The objective of VRTs is to mathematically decrease the variance of the estimators of the reliability indices while not affecting their expected value. By decreasing the variance, the number of samples needed for reaching convergence can be reduced which grants a convergence speed-up. Different types of VRTs such as Antithetic Variables (AV) [23-25], Control Variables (CV) [26-27], and the combination of VRTs [28] have been tested by researchers to increase the MCS efficiency. Importance Sampling (IS) is a relatively new VRT method that has also been verified for its efficiency. But, its utilization in the study of real systems has been limited by the fact that it is difficult to find the optimized IS distribution. The application of the Cross-Entropy (CE) method provided a possible solution to that problem. In [29-31], the Cross-Entropy -based optimization process was proposed in non-sequential MCS to obtain an auxiliary sampling distribution, which helped minimize the variance of the reliability index estimators. The method has been tested in generation reliability and in composite system reliability.

2.3 Probabilistic Assessment of Dynamic Security

To evaluate the ability of the power system to withstand sudden disturbances, dynamic security studies are necessary. To better understand the existing work, the definitions of dynamic security and transient stability need to be addressed. Reference [33] described the definition of power system security and stability: Security is the ability of a power system to withstand sudden disturbances such as electric short circuits or non-anticipated loss of system components. Dynamic security analysis involves examining the different categories of system stability including rotor angle stability, frequency stability, and voltage stability. Transient stability is an integral component of dynamic security and refers to large-disturbance rotor angle stability. Transient stability analysis focuses on studying whether the power system can maintain synchronism when subjected to a severe disturbance, such as a short circuit on a transmission line.

Probabilistic dynamic security assessment approaches have been explored to consider stochastic aspects in power systems and provide the overall quantitative evaluation of system's ability to withstand contingencies. Kirschen and Jayaweera in [34-35] addressed the differences between probabilistic and deterministic security assessment methods and illustrated the benefits of probabilistic security assessment over traditional deterministic approaches. Vaahedi, W. Li, Chia, and Dommel, in [36], presented the results of the probabilistic transient stability assessment on the large-scale system of B. C. Hydro and showed that deterministic criteria produce conservative results and also that the deterministic criteria do not always correspond to the worst case scenario. The practical meaning of exploring probabilistic dynamic security assessment methods are well illustrated in those papers.

Billinton, Carvalho, and Kuruganty in [37] presented a probabilistic assessment method of power system transient stability using Lyapunov functions to determine the post-fault stable equilibrium. System risk is represented by the product of the probability of transient instability and the occurrence of the fault which caused the instability. In [38], a risk-based stability assessment method considering both transient stability and oscillatory instability is presented. The authors proposed a composite risk index which can be obtained as the summation over risks of each individual event. In [39], Acker, McCalley, Vittal, and Pecas Lopes presented a methodology to evaluate the probability and consequences of transient stability. A transient stability index based on transient energy components calculated in time domain simulations is used to quantify the transient stability of the system. EPRI and Iowa State University worked on a series of projects which tackled the risk-based security assessment issue. The identified methodology, software design and implementation of computing risk in those projects associated with line overload, transformer overload, and voltage out of limits, voltage instability, transient instability, and special protection schemes. These methods are based on the notion that risk is the product of probability and consequence. The computations were developed so that the risk result is a function of the operating condition, a specified contingency, and uncertainty related to the operating condition and the system performance following the contingency [40-42]. A probabilistic index for transient stability based on the probability distribution of the transfer admittance is described in [43] and its application in a practical system considering the probability of fault occurrence, fault location and fault clearing time is illustrated in [44]. Similar probabilistic transient stability evaluations based on the enumeration approach is described in [45] with considerations of the probabilistic characteristics of the fault. In addition, a stochastic model of a high-speed reclosing relay is proposed in this paper and system stability is based on whether the fault clearing time is less than the CCT. In [46], a bisection method is utilized to evaluate probabilistic transient stability and reduce the amount of computation time of the evaluation. An analytical model that addresses the uncertainty of the fault clearing time for probabilistic transient stability assessment of power systems is proposed in [47]. The fault clearing time is compared with the CCT to assess system stability and a corrected transient energy function-based strategy is developed for case pruning to improve the computational efficacy. In [48], the system risk is calculated using the probability distribution of the transient stability margin and the severity quantifying the impact of a contingency with a variation of stability margin.

Methods proposed in the above-mentioned papers are essentially based on an enumeration method or analytical method for probability calculation. Utilizing the conditional probability theorem, the analytical method is mainly concerned with the effect of probabilistic aspects of transient stability such as fault types, fault lines, the fault locations and fault clearing time etc.
The other approach for probabilistic evaluation is the MCS. Some work has also been done using MCS to evaluate system stability. Anderson and Bose in [49] presented a conceptual framework for an approach to the probability assessment of power system transient stability using MCS and direct method. The computer program based on this method is described in [50]. Reference [51] introduced a procedure for evaluating overall system risk due to transient instability in the framework of MCS.

2.4 Summary

The literature survey presented above shows that the merits of a comprehensive reliability evaluation with adequacy and dynamic security integrated are well recognized. The benefits of MCS as a probabilistic approach to incorporate stochastic factors are also presented in much of the literature. Although extensive work has been done on the probabilistic approaches of adequacy or dynamic security assessment, few of the works presented a practical and accurate method of an integrated reliability evaluation. The utilization of MCS as a framework for integrating the adequacy and dynamic security assessment as well as comprehensive reliability metrics need to be studied. Accurate assessment methods for adequacy and dynamic security will also be investigated in this work.

3. Power System Models for Reliability Evaluation

In this chapter, the component modeling for power system reliability assessment is discussed. For adequacy evaluation, component outage models and the concept of probabilistic sampling for selecting system states are discussed. The outage model of wind turbines, hourly wind speed curve, and hourly load curve are introduced to reflect the stochastic factors in the system. For dynamic security, the transient stability during system transition from the pre-contingency SEP to the post-contingency SEP is of concern. The dynamic performance of the components needs to be modeled to conduct the transient stability study. The dynamic model of conventional generators, wind turbines, and protection systems are also discussed in this chapter.

3.1 Outage Models of System Components

A power system consists of various components, such as generators, lines, cables, transformers, and breakers. Component outage models play an important role in reliability assessment of the power grid. The outage model for different types of components varies. Both the outage model and outage model parameters are obtained mainly from statistical studies of historical data. For bulk power system reliability evaluation, components including generators and transmission are the major concern since the contingencies that happen on these components will directly affect the continuous power supply. The outage model of conventional generators, transmission lines, and wind turbine are discussed in this section.

3.1.1 Conventional Generator Outage Model

The forced outages represent the availability of components. Most of the forced outages in a power system are repairable. An important method to model the repairable failures is the two-state Markov model using a steady up-down-up cycle process. While modeling outage cycle of conventional generators as a two-state Markov model, it is assumed that the maximum capacity of these generators are available during the up state. In the down state, the available capacity is zero. The transitions follow an exponential probability distribution. There are two ways to represent the two-stage outage model as shown in Figure 3.1 and 3.2 [52].

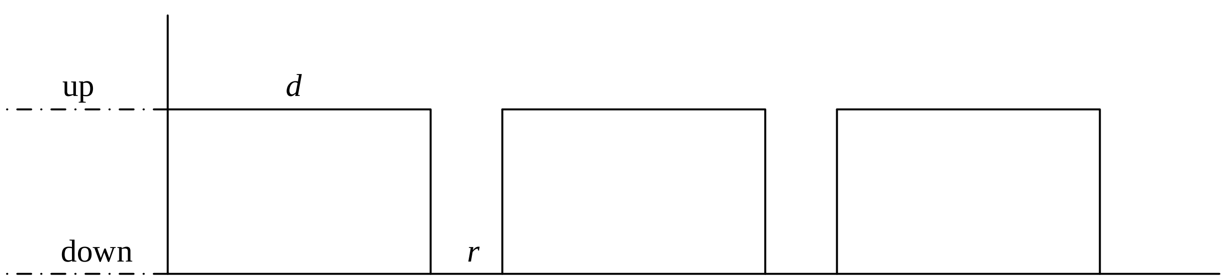

Figure 3.1 Up and down process of a repairable component

Figure 3.1 shows the cycle process between the up state and down state of a component. The parameters d and r in Figure 3.1 represent the mean time to failure and mean time to repair (in years), respectively.

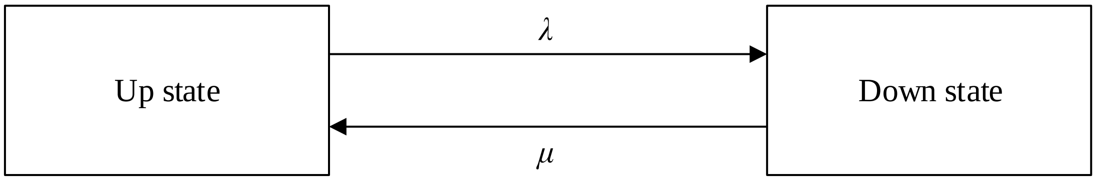

Figure 3.2 Stage space diagram of a repairable component

Figure 3.2 gives the transition diagram in the form of the transition rate between the two states. The parameter λ and μ in this figure represent the failure rate and the repair rate respectively. Mathematically, the average unavailability in the long-term process is defined by the following equation:

$$U = \frac{\lambda}{\lambda + \mu} = \frac{MTTR}{MTTR + MTTF} = \frac{f \times MTTR}{8760} \tag{3.1}$$

where U is the average unavailability, $MTTF$ the mean time to failure in the unit of hours, $MTTR$ the mean time to repair in the unit of hours, and f the average failure frequency (failures/ year). The three definitions are, in essence, the same and if only two of the parameters in are known, the others can be calculated. From the definition, we know that:

$$d = \frac{MTTF}{8760} \tag{3.2}$$

$$r = \frac{MTTR}{8760} \tag{3.3}$$

In the reliability data from different providers, the known parameters could be different. For example, the reliability data for generators that can be obtained is normally the $MTTR$ and $MTTF$, while the reliability data of transmission is normally the fault rate. Therefore, as the basis of the probabilistic fault parameter calculation, it is essential to know the relation between the parameters which are given in (3.4)- (3.9).

$$\lambda = \frac{1}{d} \tag{3.4}$$

$$\mu = \frac{1}{r} \tag{3.5}$$

$$f = \frac{1}{d+r} \tag{3.6}$$

$$U = fr \tag{3.7}$$

$$f = \frac{\lambda}{1 + \lambda r} \tag{3.8}$$

$$\lambda = \frac{f}{1 - fr} \tag{3.9}$$

3.1.2 Transmission Outage Model

For composite reliability evaluation where both generation and transmission limits are considered, the ability to supply power will be limited if the transmission fault exists even if the generation capacity is adequate. The outage cycle for transmission lines is modeled by a two-state Markov model, with transitions that follow an exponential distribution. The transition between the two states follows the same process as shown in Figure 3.1 and Figure 3.2. The parameters of transmission outage model, on the other hand, are affected by many factors, such as weather, transmission line type, voltage level, and therefore is a complicated problem. The transmission outage statistic data in available data sources is generally represented as the outage frequency (per length per year) and mean duration. For transmission lines with different length, the frequency of outage can be calculated:

$$f = f_u \times L \tag{3.10}$$

where f_u is the outage frequency per length per year, L the length of a transmission line.

Once the average failure frequency and *MTTR* are known, the other parameters can be calculated according to (3.4)- (3.9).

For transmission contingencies, the stochastic characteristics of four fault types which include three-phase faults, double line-to-ground faults, line-to-line faults, and single line-to-ground faults are also considered based on their probability of occurrence. When a fault on a transmission line is selected, the type of the fault is then determined. The probability of occurrence for the four fault types are 6.2% (three-phase), 10.0% (double line-to-ground), 8.8% (line-to-line) and 75% (single line-to-ground), as provided in [41].

3.1.3 Wind Turbine Outage Model

Because of the fact that the penetration of wind power has been increasing continuously, planning studies have to pay more attention to the impact of wind power on power system reliability. The stochastic nature of wind speed should be considered in the reliability studies as well as the outages rates of turbines.

In the assessment of the effect of wind power on power system reliability, outages of wind turbines have a significant impact on the available power. Considering the failure of each wind turbine as an independent event, the outage model of a wind turbine can be modeled as Markov components with up and down states. This outage model is considered the same with the outage model of conventional generators, yet the parameters λ and μ of wind turbines are normally quite different as compared with conventional generators. A substantial effort from both industry and research institutions has been made to develop databases and statistical analyses of wind turbine failures and reliability. For example, a large monitoring program is being pursued by Fraunhofer Institute for Wind Energy Systems (Fraunhofer IWES) Germany to establish a database that contains detailed information about reliability and availability of wind turbines. In the U.S., a CREW (continuous reliability enhancement for wind) database was developed by Sandia National Laboratories to benchmark the current U.S. wind turbine reliability performance. It collects wind farm SCADA data, downtime and reserve event records, and daily summaries of generating, unavailable, and reserve time for each turbine. From these databases, the *MTTF* and the average failure rate for the outage model can be obtained.

A three-state model has been presented in some works [53-55]. The three-state model subdivides the up state to be a rated and a de-rated state. However, since the wind turbine output is mainly based on a current wind speed value, the two-state model considering the variable wind speed manages to represent the de-rated state. In this work, an 8760-hour (365-day) wind speed data obtained from the National Renewable Energy Laboratory (NREL), National Wind Technology Center Information Portal [63] is introduced and incorporated with the power curve of wind turbines to represent the stochastic WTG output. According to the proposed model in this work, the wind turbine up or down state is determined first. If the wind turbine is in the up state, its output is determined based on the current wind speed and the power curve performance of the wind turbine. The calculation between wind speed and the output active power is based on the following equation:

$$P(\omega) = \begin{cases} 0 & ,0 \leq \omega \leq \omega_i \\ P_r \dfrac{\omega^3 - \omega_i^3}{\omega_r^3 - \omega_i^3} & ,\omega_i \leq \omega \leq \omega_r \\ P_r & ,\omega_r \leq \omega \leq \omega_o \\ 0 & ,\omega > \omega_i \end{cases} \tag{3.11}$$

where, ω is the current wind speed, $P(\omega)$ the active power under the current wind speed, ω_i the cut-in wind speed, ω_r the rated wind speed, ω_o the cut-out wind speed, and P_r the rated active power.

3.2 Probabilistic Approach for Reliability Evaluation

3.2.1 Probabilistic Analysis Approach

For reliability evaluation of power systems, two key processes are required in the evaluation. The first one is system states selection and state probability calculation. The other one is the system states assessment. By contrast with the deterministic methods, the probabilistic reliability evaluation introduces a probabilistic approach for system states selection. There are two probabilistic approaches for system state selection: analytical-based methods and simulation-based methods. The widely used methods for those two types are state enumeration and MCS, respectively.

3.2.1.1 State Enumeration

Using state enumeration method, system states are selected based on a predefined contingency list, which may comprise the whole system or a subset of the system. Combining the elements of the list, different contingency levels may be generated. The probability of a system state, therefore, can be calculated as the product of the probability of success or failure probability of each component, as given in the following equation:

$$P(s) = \prod_{i=1}^{N_f} P_f \prod_{i=1}^{N_s} P_s \tag{3.12}$$

where N_f and N_s are the numbers of failed and successful components in state s respectively. The overall system reliability can be obtained based on the probability of each selected state and its reliability condition.

For a system with n components and each component with two states (up or down), the total number of system states is 2^n. Considering a system with 1000 components, the number of states is 2^{1000}. Sampling such a large number of states is obviously impractical. Therefore, although the enumeration method is very straightforward, this method is only suitable for small systems. For a large power system, the MCS is the preferred method.

3.2.1.2 Monte-Carlo Simulation

The MCS method samples system states based on the probability of event occurrence. Using MCS, system states transition process and random behavior are simulated. The advantage of MCS over enumeration of states is that the number of samples needed to reach a required accuracy does not depend on the size of the power system but rather on its reliability. There are two types of MCS: sequential MCS and non-sequential MCS. The NSMCS, sometimes called the state sampling approach, is based on sampling the probability of the component appearing in that state. However, similar to the state enumeration method, the NSMCS cannot simulate the chronology of time-dependent events.

Considering the load uncertainty and the wind power variability, SMCS is a more suitable approach for system states sampling since it is easy to integrate stochastic factors with chronological characteristics into states sampling.

The approach of SMCS is based on sampling a probability distribution of component state durations and includes the following steps:

Step 1: Specify initial states of all components. Normally, all components are assumed to initially be in the up state.

Step 2: Sample the duration of each component residing in its present state. The state duration distribution is assumed following an exponential distribution. Therefore, the duration of an up state or a down state is sampled according to the following equations:

$$T_{up,i}=\frac{1}{\lambda_i}lnU_i \tag{3.13}$$

$$T_{down,i}=\frac{1}{\mu_i}lnU_i \tag{3.14}$$

where U_i is a uniformly distributed random number between [0,1] corresponding to the i th component, λ_i the failure rate of the i th component, and μ_i the repair rate.

Step 3: Repeat *Step 2* in the time span considered (years) and record sampling values of each state duration for all components. The chronological state transition processes of each component in the given time span can be obtained. The illustration of the component state transition process is given in Figure 3.3.

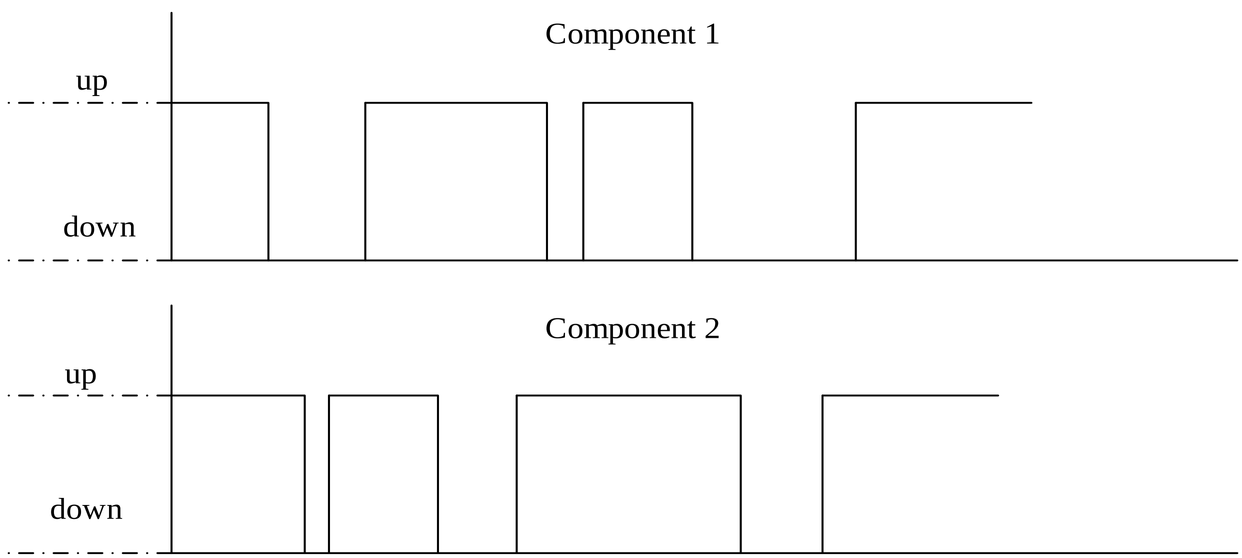

Figure 3.3 Chronological state transition processes of components

Step 4: By combining the state transition processes of all components, the chronological system state transition cycles can be obtained.

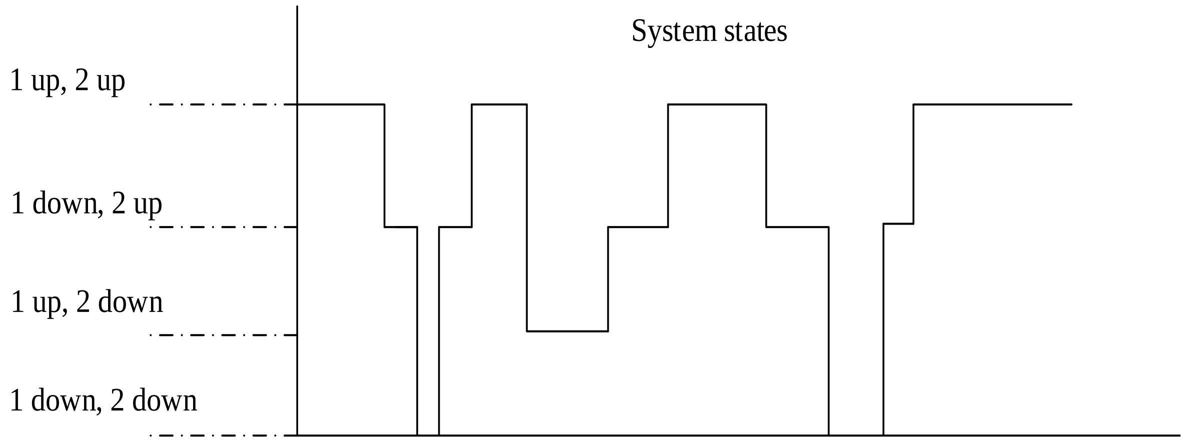

Figure 3.4 Chronological system state transition process

For example, combining the transition process of component 1 and component 2 in Figure 3.3, the system transition process is as shown in Figure 3.4. The state of each component and the duration of each system state are obtained. In power system reliability evaluation, all system states are then subject to evaluation.

3.2.2 Stochastic Load Representation

The electrical load in a power system during any time period is a stochastic process, which is difficult to describe with a simple mathematical formula. Different models are created, starting from primary load data according to the need to calculate reliability. Primary load data will provide a minimum amount of data that is needed to establish an hourly chronological load profile. Most primary load data consist of the percentage of maximum monthly load or weekly load in a year, the load in 24 hours in a typical day in each season and the maximum load in each day in a week. With the percentages of these data available and the annual peak load is known, the hourly chronological load profile can be established.

The SMCS method follows these chronological loads steps as the simulation progresses. Since the load data changes every hour, the system state transition happens at least once an hour. The load curve is then combined with generator output and the up or down state of other components, to form the system scenarios.

3.3 Dynamic models of system components

Adequacy evaluation of the power system addresses the post-contingency steady-state system condition. However, it is not guaranteed that a post-contingency SEP can always be reached. Transient stability needs to be assessed to determine the system condition during the transition.

Dynamic models of synchronous generator, wind turbine generator (WTG) and protection system are discussed in this section.

3.3.1 Synchronous Generator Dynamic Model

The detailed synchronous generator models are represented using the GENROU model in GE PSLF.

To represent synchronous generator performance in a stability study, the governor model and exciter model are essential. The IEEE (1992, 2005) ST1A model, which is a controlled-rectifier excitation system model, is used to represent excitation system for all synchronous generators.

Three governor models are used to represent the governor of a gas turbine, hydro turbine, and steam turbine respectively.

3.3.2 Wind Turbine Generator Dynamic Model

With increasing wind power penetration in power systems, the impact of these converter interfaced generators should be appropriately accounted for in the dynamic security assessment of power systems. In this project, wind turbines are detailed and aggregated in models used for system studies. With the current technology, new wind turbines are commonly equipped with capability of active power control, reactive power control and fault ride through, and these functions are modeled to represent the realistic behaviors of wind turbines during disturbances.

In this work, a GE doubly-fed asynchronous generator (DFAG) known as type-3 WTG model in PSLF is used. Physically, the rotor side of the machines is connected with the grid through a ac-dc-ac converter, hence have significantly different dynamic behavior than conventional machines. A GE WTG dynamic model includes three functional parts, as shown in Figure 3.5 [56].

- Generator/ converter model: The generator/converter model is represented using GEWTG model in PSLF, it injects real and reactive current into the network in response to control commands, and represents low and high voltage protective functions (e.g., low voltage ride through capability).
- Electrical control model: The electrical control model is represented using EXWTGE model in PSLF. It includes both closed and open loop reactive power controls based on the inputs from the turbine model (P_{ord}) and from the supervisory VAr controller (Q_{ord}), or voltage regulation with either a simplified emulator of GE's WindCONTROL system or a separate, detailed model. This model sends real and reactive commands to the generator/converter model.
- Turbine and turbine control model: The turbine and turbine control model is represented using WNDTGE model in PSLF, which represents the mechanical controls, including blade pitch control and power order (torque order in the actual equipment) to the converter; under speed trip; rotor inertia equation; wind power as a function of wind speed, blade pitch, rotor speed; and active power control.

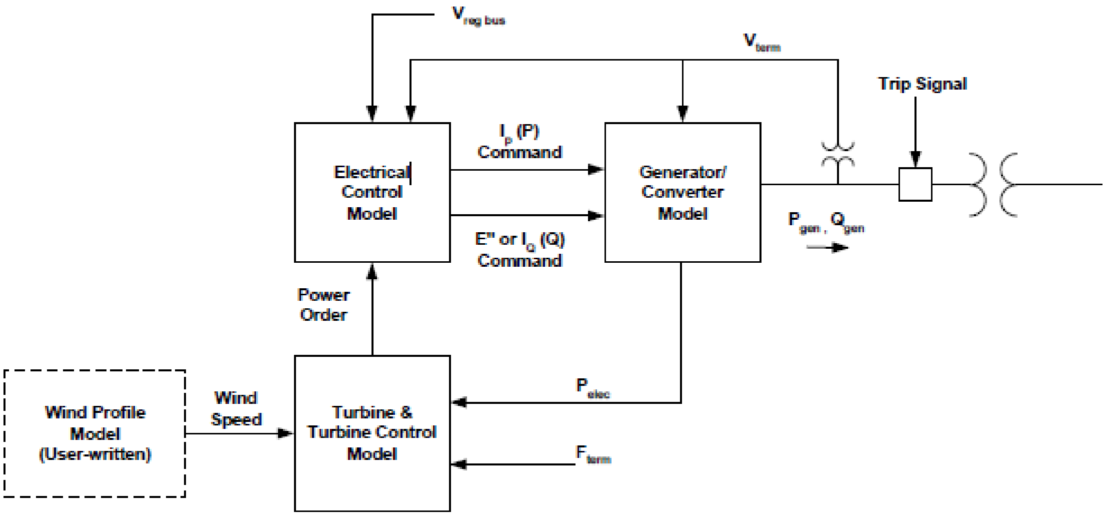

Figure 3.5 GE WTG Dynamic Model Connectivity

The active power control function can be modeled in the WindCONTROL function of the WTG model. Under normal operating conditions with near nominal system frequency, the control is either enforcing a maximum plant output or providing a specified margin by generating less power than is available from the wind (e.g., actual power generated is 95% of the available power). In response to frequency excursions, the control switches into another mode and calculates a plant power order as a function of system frequency. In the event of low system frequency, the wind plant will generate additional power in response to the loss of other generating facilities, whilst, reduce power generation in response to load loss.

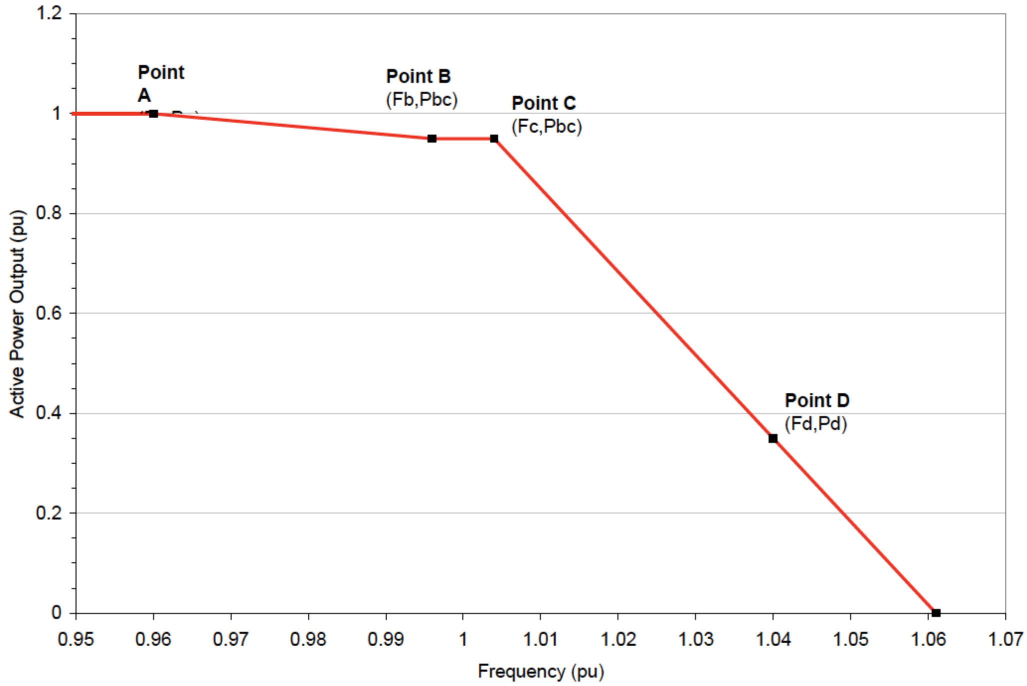

Figure 3.6 Frequency response curve

The active power control is normally set according to regional grid code requirements. An example frequency response curve is shown in Figure 5.6 [56], the APC performance is given by setting the points A through D on this response curve. The value of Pd should be greater than or equal to the minimum power, which is discussed below. The value of Fb must be less than 1, and that of Fc must be greater than 1. The value of Tpav may be increased to simulate fixed power reference.

The detailed representation of the reactive power control model is shown in Figure 3.7. Three modes of reactive power control including power factor control, voltage control, reactive power control can be used by setting proper parameters.

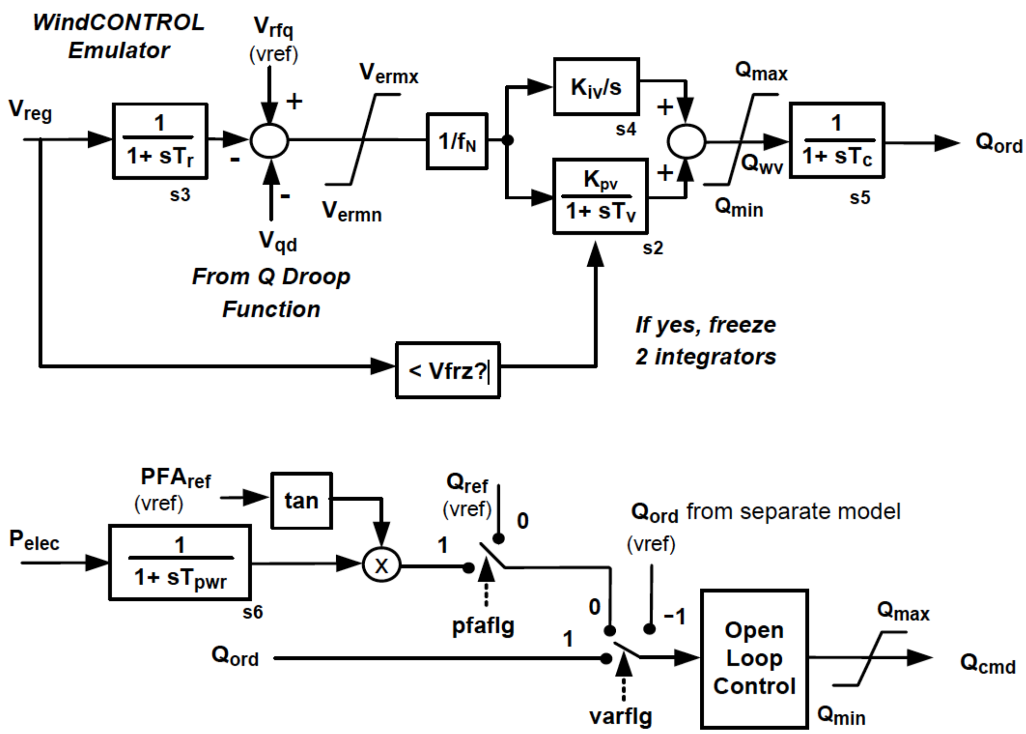

Figure 3.7 Reactive power control model

3.3.3 Modeling of System Protection

The protection systems are modeled in this work to quantify system reliability during the transition period. With the protection systems, the severity of transient instability can be measured by the amount of load that is shed or generation that is tripped after a contingency in order to keep the system stable. Three types of protection systems are modelled in this work:

- Under-frequency load shedding
- Over/ under-frequency generator tripping
- Over/ under - voltage generator tripping

3.3.3.1 Under-Frequency Load Shedding

The primary requirement of UFLS is to trip excess load to obtain generation-load balance following a disturbance which results from the tripping of lines and/or generators causing an area generation deficit [57]. Since generator turbines cannot operate at low frequencies (56-58 Hz), it is necessary to maintain frequency near the nominal frequency (60 Hz). Slow changes in load can be compensated by governor action if generators have available spinning reserve and equilibrium can be reached. However, during transient outages, the excess load is fed by the available kinetic energy of the rotating machines and frequency starts dropping. The only way to stabilize the system under such conditions is progressively shedding the load of pre-determined load centers at certain frequency thresholds.

LSDT1 model in PSLF is used to represent UFLS protection system. The setting of the load shedding relay is according to the NERC reliability standard [58]. Table 3.1 shows the UFLS criteria for the Eastern Interconnection for utilities with net peak loads greater than 100 MW. These criteria are used in the LSDT model.

Table 3.1 UFLS attributes for with net peak load greater than 100MW

Frequency Threshold (Hz)	**Total Nominal Operating Time (sec)**	**Load Shed at Stage (%)**	**Cumulative Load Shed (%)**
59.5	0.05	10	10
59.2	0.05	20	30
58.8	0.05	20	50

3.3.3.2 Over/ under-Frequency Generator Tripping

The over-frequency and under-frequency tripping of generators is required to maintain generation-load balance [57]. If any area has a load deficit, the generators start speeding up. The generator turbines are designed to operate near nominal frequency and operation at an off-nominal frequency can damage the turbine blades. To protect the costly turbine generators, the NERC reliability criteria for UFLS [58] also provides guidelines for over-frequency and under-frequency generator tripping. Figure 3.8 shows the generator over-frequency and under-frequency performance characteristics and trip modeling criteria. In this work, the generators modeled with over-frequency and under-frequency relays are tripped if the over-frequency threshold of 61.2 Hz is violated or the under-frequency threshold of 58.2 Hz for 2 s is violated.

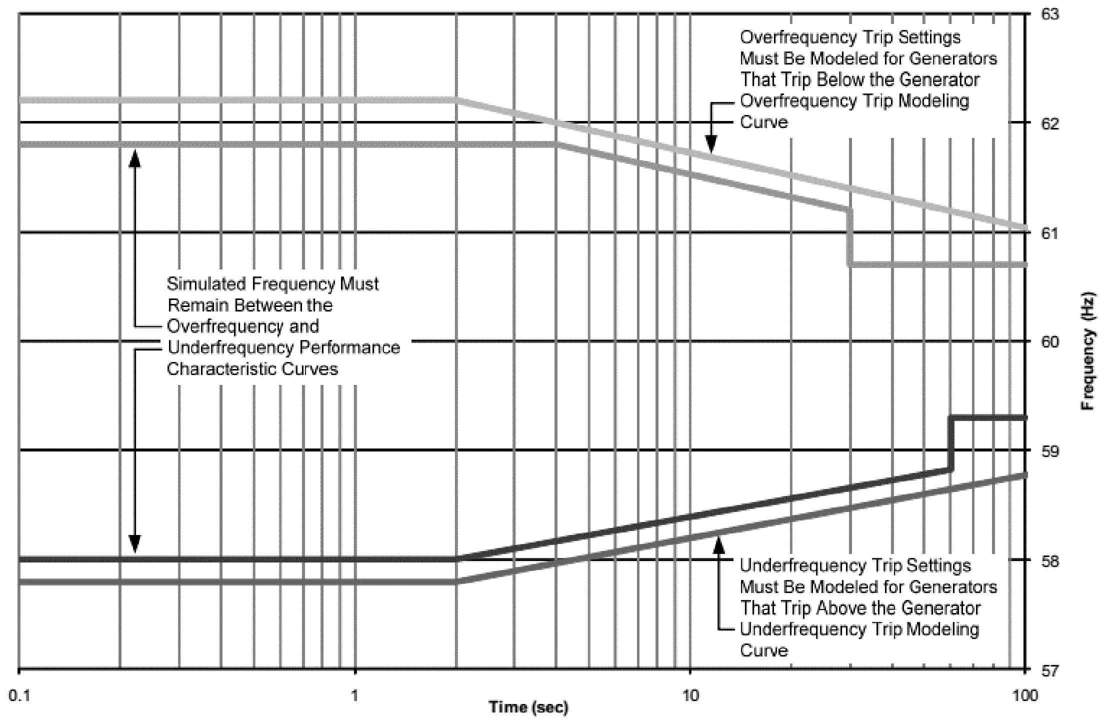

Figure 3.8 Design performance and modelling curves for over and under frequency generator trip [34]

GP1 in PSLF is used as the generator protection model in this work. This is a multifunction model and can represent protections including under/ over frequency protection, under/ over voltage protection, field over-current, stator over-current, and reverse-power protection.

3.3.3.3 Over/ under-Voltage Generator Tripping

Generators are designed to operate at a continuous minimum terminal voltage of 0.95 pu of rated voltage while delivering power at rated voltage and frequency. Under voltage can reduce the stability limit, result in excessive reactive power import and malfunctioning of voltage sensitive equipment. In the TDS, if the generator terminal voltage reduces to 0.90 pu for 1.0 s, then the generator is tripped. Generator overvoltage protection, on the other hand, is required to prevent insulation breakdown due to sustained terminal overvoltage. The generator insulation is capable of operating at a continuous overvoltage of 1.05 pu of its rated voltage. If the generator terminal voltage increases to 1.15 pu for 0.5 s, the generators are tripped.

The over/ under- voltage generator tripping protection is also modeled in GP1 in PSLF.

3.4 Summary

The first important steps in probabilistic reliability assessment is the methodology for system state generation. In this chapter, the details of using sequential MCS as the probabilistic analysis framework of the integrated reliability evaluation approach is discussed. With the modeling of the traditional generator and transmission outage occurrence, chronological wind power generation, and wind turbine outage occurrence which with increasing penetration will have important impact in system reliability, the stochasticity is addressed in system states generation. The dynamic model of generators for transient stability analysis as well as the protection systems are presented in this Chapter. With probabilistic models and dynamic models described in this chapter provides as backbone, the assessment methods and overall procedure of integrated reliability evaluation are provided in the following chapter.

4. Adequacy and Dynamic Security Integrated Reliability Evaluation

This chapter describes the integrated approach to reliability assessment addressing system adequacy and dynamic security. The assessment methods for system adequacy and dynamic security, namely power flow analysis, optimal power flow (OPF) analysis, and TDS, respectively, are discussed. The outcomes of the reliability evaluation procedure are comprehensive reliability indices that reflect power system reliability impacts from both adequacy and instability perspectives. To tackle the computational efficiency, two acceleration approaches are introduced in this work to speed up the convergence process of the MCS and the TDS for transient stability evaluation. Two widely used commercial power system analysis software packages, Siemens PSS®E OPF and GE PSLF, are used as analytical tools. The overall procedure including MCS framework is implemented in Python.

4.1 Probabilistic Approach for Integrated Reliability Evaluation

4.1.1 Procedure

The basic idea of the probabilistic approach for integrated reliability evaluation is providing an iterative process to perform reliability assessment based on probabilistic sampling using SMCS. The process includes the following four steps:

1. Selecting a system state,
2. Analyzing the system state to judge if it is a failure state,
3. Calculating risk indices for the failure state,
4. Updating cumulative indices

The flow chart of the integrated reliability evaluation procedure is provided in Figure 4.1. For each iteration of the SMCS, the system states are generated and incorporated with an annual 8760-hour load curve and wind speed curve. State evaluation is then conducted for each state. To determine both adequacy and dynamic security, each state is assessed using TDS and ACPF/ AC OPF. From TDS, if the system in the state considered is unstable during the transition after contingency, the amount of load curtailment that results from the activation of protection systems to maintain system stability is generated. The load curtailment value therefore, provides an indication of the severity of system instability. If the state is transiently stable, an AC power flow is conducted to simulate the outage stage and examine the post-disturbance stable equilibrium point. When any system operating limits are violated, the AC OPF is used for remedial action to reschedule generation and alleviate constraint violations, while avoiding load curtailment if possible or to minimize the total load curtailment if unavoidable. When load curtailment happens in either transient stability or steady-state reliability assessment, the system state is classified as a failure state.

After each iteration of the SMCS, reliability indices are updated, and SMCS convergence is determined based on the coefficient of variance (COV) of reliability indices. The final reliability indices are obtained after the MCS converges.

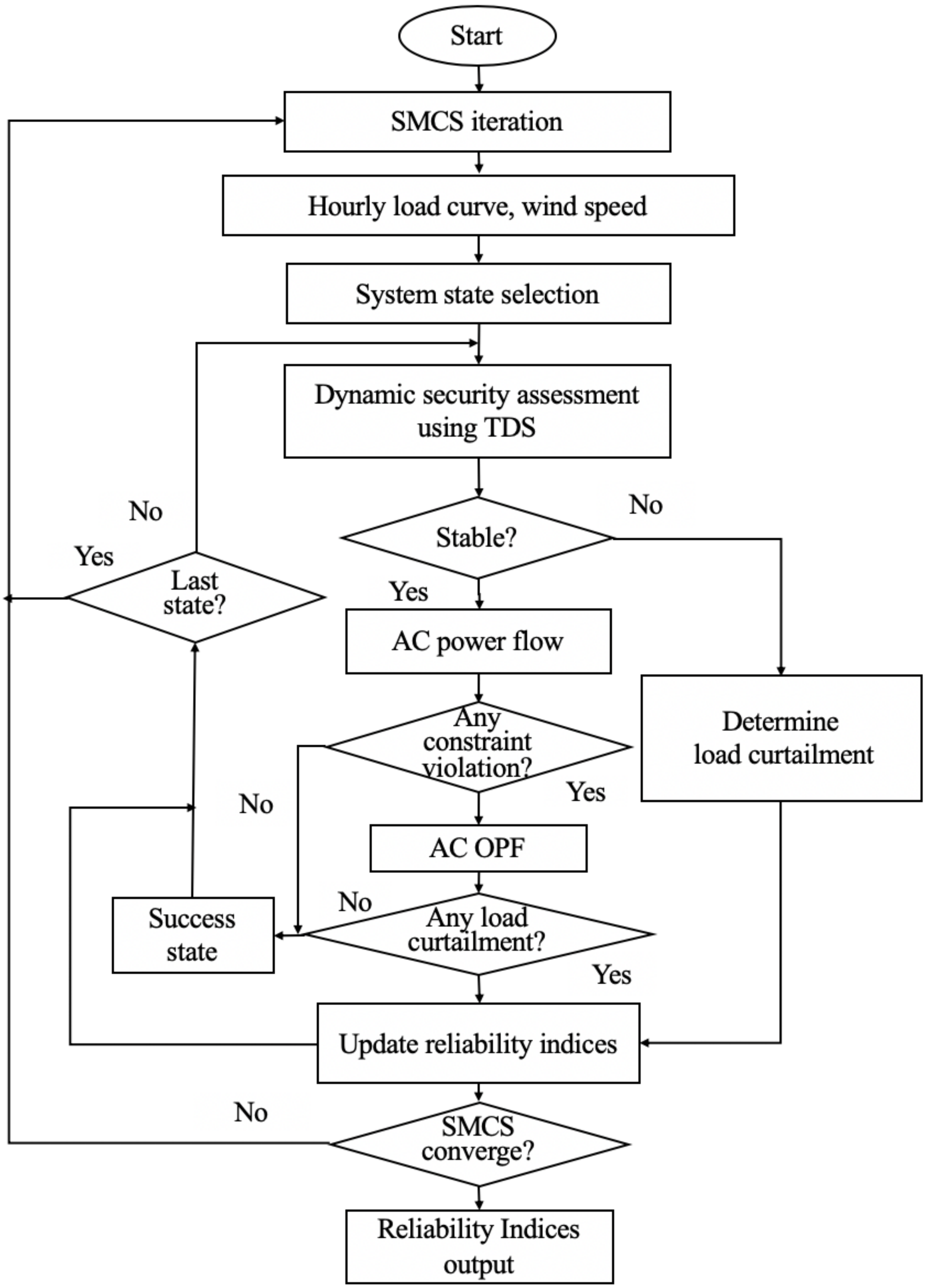

Figure 4.1 Flow chart for reliability evaluation with adequacy and transient stability integrated

4.1.2 Adequacy Assessment Methods

The objective of adequacy assessment is to determine whether the system is capable of supplying the electric load under the specified contingency without violations of the operating constraints. The assessment addresses the steady-state system condition after the contingency.

4.1.2.1 AC Power Flow Analysis

Power flow study is used to assess system steady-state conditions. Both DC power flow and AC power flow can be used as adequacy assessment methods. However, for composite system reliability study in a realistic power system, the system condition includes bus voltage and transmission limits which need to be evaluated. Therefore, the AC power flow is used as the steady-state analysis technique in this work. The equations for power flow injection at a node and power flow on a transmission line are addressed as follows:

$$P_i(V,\delta) = V_i \sum_{j=1}^{N} V_j(G_{ij}cos\delta_{ij}+B_{ij}sin\delta_{ij}) \tag{4.1}$$

$$Q_i(V,\delta) = V_i \sum_{j=1}^{N} V_j(G_{ij}sin\delta_{ij}-B_{ij}cos\delta_{ij}) \tag{4.2}$$

$$T_k(V,\delta) = max\{T_{mn}(V,\delta),T_{nm}(V,\delta)\} \tag{4.3}$$

where, N is the set of all buses in the system; G_{ij} and B_{ij} are the real and imaginary parts of the i th row and j th column element of the bus admittance matrix; δ_i and δ_j is the angle of bus i and bus j respectively, and δ_{ij} is the angle difference of bus i and bus j; $T_{mn}(V,\delta)$ and $T_{nm}(V,\delta)$ are the MVA flows at the two ends of line k. m and n are the two buses of line k. The MVA flow from bus m to n is calculated as

$$T_{mn}(V,\delta)=\sqrt{P_{mn}^2(V,\delta)+Q_{mn}^2(V,\delta)} \tag{4.4}$$

$$P_{mn}(V,\delta)=V_m^2\left(g_{m0}+g_{mn}\right)-V_mV_n(b_{mn}sin\delta_{mn}+g_{mn}cos\delta_{mn}) \tag{4.5}$$

$$Q_{mn}(V,\delta)=-V_m^2(b_{m0}+b_{mn})+V_mV_n(b_{mn}cos\delta_{mn}-g_{mn}sin\delta_{mn}) \tag{4.6}$$

where, $g_{mn} + jb_{mn}$ is the primitive admittance of line k and $g_{m0} + jb_{m0}$ is the equivalent admittance of the circuit to the ground at the end of bus m.

When the AC power flow results show that there is a violation of operating limits, then remedial actions including generation rescheduling and load curtailment are considered to correct the abnormal system conditions. The remedial actions reschedule generation first, without curtailing any load. If the operating constraint violations still exist, the remedial actions with load shedding capability are then applied to determine where and how much load shedding will be needed to alleviate operating constraint violations, which is recorded as a system failure. The results of the contingency evaluations are stored and subsequently used by the reliability calculation model to calculate the reliability indices. AC OPF is used as the technique to assess the system adequacy when the operating limits are violated.

4.1.2.2 AC OPF Model

As discussed previously, the objective of adequacy assessment is to determine whether a system is capable of supplying the electric demand considering possible generator and transmission contingencies while not violating any operating constraints. Assessment of adequacy is based on the alternating current optimal power flow (ACOPF) solution that analyzes whether load curtailment is required to maintain the system within operating limits during generator or transmission line contingencies. These contingencies are selected by the SMCS process based on the probability of fault occurrence and consideration of repair rate.

Remedial actions modeled in the ACOPF formulation, include generator active and reactive power adjustment, schedule bus voltage magnitude adjustment, transformer tap ratio, and switched shunt admittance controls. As a consequence of faulted generators or transmission outages resulting in load not being served, the remedial actions are incorporated to maintain the system within operating limits. If operating constraints violations still exist after the remedial actions, load curtailments are invoked as a last resort to bring the system back within limits. The system is considered to be adequate when the operating limits are satisfied and no load needs to be curtailed. The minimum load curtailment is the quantitative metric of system adequacy. The objective function of the ACOPF is given in (4.7), and constraints including load adjustment limits, node power balance, line flow limits, generation output limits, voltage magnitude and angle limits are shown in (4.8) - (4.16).

$$Min : \sum_{\forall L} P_{L0}(1-\psi_L) \tag{4.7}$$

Subject to

$$\psi_{L,MIN} \le \psi_L \le \psi_{L,MAX} \tag{4.8}$$

$$S_L = \psi_L \cdot S_{L0} \tag{4.9}$$

$$S_G - S_L = diag(\overline{V})\overline{Y}^*_{bus}\overline{V}^* \tag{4.10}$$

$$\left|\boldsymbol{V}_m \boldsymbol{Y}^*_{line} \boldsymbol{V}^*_n\right| \le S_{mn,MAX} \tag{4.11}$$

$$\left| \boldsymbol{V}_n \boldsymbol{Y}_{line}^* \boldsymbol{V}_m^* \right| \le S_{nm,MAX} \tag{4.12}$$

$$P_{G,MIN} \le P_G \le P_{G,MAX} \tag{4.13}$$

$$Q_{G,MIN} \le Q_G \le Q_{G,MAX} \tag{4.14}$$

$$V_{MIN} \le V \le V_{MAX} \tag{4.15}$$

$$\theta_{MIN} \le \theta \le \theta_{MAX} \tag{4.16}$$

where, ψ_L is the load adjustment factor, $\psi_{L,MIN}$ and $\psi_{L,MAX}$ are the minimum and maximum load adjustment factor respectively, P_{L0} is the initial active power of load, S_{L0} is the load initial MVA value, S_L is the MVA value of load, S_G is the MVA value of generator, $\boldsymbol{V}$ represent bus voltage, $\boldsymbol{Y_{bus}}$ is the admittance matrices, $\boldsymbol{V_m}$ and $\boldsymbol{V_n}$ are bus voltage of bus m and bus n, $\boldsymbol{Y_{line}}$ is the admittance of line connecting bus m and bus n, $S_{mn,MAX}$ is the maximum flow on line between bus m and bus n, P_G is the active power of generator, $P_{G,MIN}$ and $P_{G,MAX}$ are the minimum and maximum value of generator active power output, Q_G is the reactive power of generator, $Q_{G,MIN}$ and $Q_{G,MAX}$ are the minimum and maximum value of generator reactive power output, V is voltage magnitude, V_{MIN} and V_{MAX} is the voltage magnitude lower and upper limits, θ is voltage angle, θ_{MIN} and θ_{MAX} are the voltage angle lower and upper limits.

In this work, the PSSE OPF software tool is utilized to find the optimal solution based on a Lagrangian relaxation method. Therefore, the adequacy evaluation results are quantified by the minimum load curtailment from the optimization problem, as represented in (4.17).

$$P_{loadshed} = Min\{\sum\nolimits_{\forall L} P_{L0}(1-\psi_L)\} \tag{4.17}$$

4.1.3 Dynamic Security Assessment Methods

While adequacy studies evaluate system conditions in post-contingency steady-state conditions, the DSA estimates whether the system can maintain stability during the transition period after a contingency. If the transition is stable and a new equilibrium point is reached, the system is termed dynamically secure, otherwise it is judged to be insecure. TDS is widely recognized as the most accurate method to describe power system transient behavior and therefore is the method of choice to perform the DSA in this work. An accurate system dynamic performance with the protection systems actions can be simulated in the TDS. The protection systems including under-frequency load shedding, over frequency and under frequency generator tripping, over voltage and under voltage generator tripping as described in Chapter 3 are modeled, so that by activating these protections, following a contingency the system should be transiently stable. The dynamic security is quantified by the MW load shedding due to security preserving corrective protections, as represented in (4.18).

$$P_{loadshed} = P_{L,pre-fault} - P_{L,post-fault} \tag{4.18}$$

where $P_{loadshed}$ is the assessment result in terms of the MW load curtailment, $P_{L,pre\text{-}fault}$ is the total active power of all loads at the pre-fault SEP of the system, and $P_{L,post\text{-}fault}$ is the total active power of all loads in the post-fault SEP.

The drawback of TDS is the computational burden especially when simulating a big system with a large number of scenarios. To tackle this problem with TDS, a pruning process is introduced in this work to reduce the computational burden. The pruning process will be discussed in section 4.2.

4.1.4 Reliability Indices

The most important outcome of the probabilistic reliability assessment is the reliability indices. Different from the deterministic methods, the probabilistic reliability indices provide a quantitative indication of the overall system reliability level. The indices such as LOLP, LOLE, EPNS, EENS, LOLP, and LOLD are the widely used reliability indices in traditional system adequacy assessment. They provide an effective means to include all the system states from MCS into the reliability calculation. From SMCS, in each sampled year, the reliability can be calculated based on states duration time and load curtailment results from states assessment. LOLP, EPNS, and LOLF which are the probability index, energy index, and frequency and duration index, respectively, are three key indices used in this work. The calculation of the three indices are given as follows:

Assume that in the i th simulation year, N_i is the number of states, X_i is the set of all states, $X_{i,f}$ is the set of failure states, T_i is the sum of the durations of all states, x_j is the j th system state, τ_j is the duration time of state x_j, LC_j is the amount of load curtailment of state x_j.

1) Loss of Load Probability (LOLP) Index represents the probability of failure of the system to meet the demand. The LOLP index in the i th simulation year is calculated as:

$$LOLP_i = \frac{1}{T_i} \sum_{x_j \in X_{i,f}} \tau_j \tag{4.19}$$

2) Expected Power Not Supplied (EPNS) Index measures the expected load that will be curtailed in the cases of failure states. The EPNS index in the i th simulation year is calculated as:

$$EPNS_i = \frac{1}{T_i} \sum_{x_j \in X_{i,f}} \tau_j \cdot LC_j \tag{4.20}$$

3) Loss of Load Frequency (LOLF) Index represents the average number of load curtailment events. The LOLF index in the i th simulation year is calculated as:

$$LOLF_i=\frac{1}{T_i}\sum_{x_j\in X_{i,f}} H(x_j)\cdot 8760 \tag{4.21}$$

where, $H(x_j)$ is the indicator of success or failure of a state. In (4.21), for failure of a system state $(x_j \in X_{i,f})$, $H(x_j)=1$.

In the integrated reliability evaluation, system states are assessed by both steady-state analysis and transient stability analysis. The state is seen as a successful state only when no load curtailment is required from both steady-state and transient stability analysis.

The convergence of SMCS is typically based on the value of the coefficient of variation (COV) as it shows the extent of variability of the index in each simulation year in relation to the mean. The COV can be calculated based on any of the reliability indices. For example, we choose the LOLP as the indication index to determine the convergence. After N simulation years, the LOLP index is calculated for each simulation year.

The expected value of the index LOLP is

$$\overline{LOLP}=\frac{1}{N}\sum_{i=1}^{N} LOLP_i \tag{4.22}$$

The variance of the estimate is

$$Var(\overline{LOLP})=\frac{\sum_{i=1}^{N}(LOLP_i-\overline{LOLP})^2}{N^2} \tag{4.23}$$

Then COV is calculated as:

$$COV=\frac{\sqrt{Var(\overline{LOLP})}}{\overline{LOLP}} \tag{4.24}$$

Equation (4.24) shows that the COV is a normalized measure of the dispersion of probability distributions. Hence, the lower its value, the better is the accuracy of the estimate of $\overline{LOLP}$. To reach a demanding accuracy of SMCS, the criteria for COV is normally set to be 1% - 5%. Once the COV value attains the convergence criteria, the SMCS is completed and the reliability indices from the last simulation year represent the final reliability indices results.

4.2 Reliability Evaluation Acceleration Methodology

The reliability evaluation is basically a composite of two parts: states selection and states assessment. By using SMCS as the probabilistic method of states selection, it is easy to include chronological aspects of the power systems into the simulation. However, the computational efficiency is the main issue of SMCS. A large number of system states need to be sampled to assure accurate estimates of the reliability indices. With the objective of improving the computational efficiency, an importance sampling based on the Cross-Entropy method is introduced for SMCS.

Using TDS as the dynamic security assessment method also introduces computational burden. A pruning process is introduced to reduce the volume of cases goes through TDS. The pruning process classifies the states based on the kinetic energy gained due to the fault and the change in the magnitude of the Thévenin impedance (Z_{th}) at the point of interconnection (POI) of the generators in the post-contingency network. These two acceleration methodologies are discussed in this section.

4.2.1 CE based Importance Sampling

The number of samples required by the MCS methods can be reduced using variance reduction techniques (VRTs). Importance sampling (IS) has proved to be an effective means of improving the MCS method. The application of IS is based on the idea that certain variables have a greater impact on the estimation process of a target quantity. Thus, if these 'important' values are sampled more often based on an optimized PDF $g_{opt}(\cdot)$, the variance of the estimator should be reduced.

The selection of the new $g_{opt}(\cdot)$ is a difficult task and for this reason the application of IS has been largely limited. However, this problem has been overcome by the CE method since it provides a simple adaptive procedure to obtain the optimal probability density function. In this work, the combination of the CE method with the IS technique is used as the approach to accelerate the convergence of SMCS. A mathematical illustration is discussed below.

Consider the original PDF $f(\cdot)$ is based on $\mathbf{u}$ which is the original unavailability vector of each component, $\boldsymbol{X_i}=(x_{i,1},\ldots,x_{i,j},\ \ldots,x_{i,n})$ as the state of each component sampled based on $f(\cdot)$ in the ith simulation iteration, and H_i as the test function of a reliability index. Assuming that SMCS converges after N iterations, the estimation of the reliability index is

$$E_u\big(H(x)\big)=\int H(x)f(x)dx\approx\frac{1}{N}\sum_{i=1}^{N}H(\boldsymbol{X_i}) \tag{4.25}$$

If the system failures are rare, the estimation process via (4.25) is then very computational demanding. By using IS which introduces a new PDF $g(\cdot)$ in the form of a new unavailability vector $\mathbf{v}$, a system state sample $\boldsymbol{X_i}$ is drawn based on $\mathbf{v}$, and the reliability index is estimated using

$$E_v\big(H(x)\big)=\int H(x)\frac{f(x)}{g(x)}g(x)dx\approx\frac{1}{N}\sum_{i=1}^{N}H(\boldsymbol{X_i})\ \boldsymbol{W}(\boldsymbol{X_i};n,\boldsymbol{u},\boldsymbol{v}) \tag{4.26}$$

where $\boldsymbol{W}(\boldsymbol{X_i};n,\boldsymbol{u},\boldsymbol{v})$ is called the likelihood ratio to avoid any biased estimates. The likelihood ratio is given by

$$\boldsymbol{W(X_i;n,u,v)}=\frac{\prod_{j=1}^{n}(1-u_j)^{x_{i,j}}\cdot u_j^{(1-x_{i,j})}}{\prod_{j=1}^{n}(1-v_j)^{x_{i,j}}\cdot v_j^{(1-x_{i,j})}} \tag{4.27}$$

The problem now consists of finding the optimal **v** that minimizes the computational effort of the SMCS.

The CE method solves the optimal **v** issue by minimizing the *Kullback-Leibler* distance between the $g(\cdot)$ and the optimal $g_{opt}(\cdot)$ [59-60]. This distance is defined as:

$$\begin{aligned} D\left(g_{opt}(X),g(X)\right) &= E_g*\left[ln\frac{g_{opt}(X)}{g(X)}\right] \\ &= \int g_{opt}(X)lng_{opt}(X)dX-\int g_{opt}(X)lng(X)dX \end{aligned} \tag{4.28}$$

The minimization of (4.28) is equivalent to

$$max\int g_{opt}(X)lng(X)dX \tag{4.29}$$

Since $f(X)=f(X;u)$, $g(X)=f(X;v)$, $g_{opt}(X)=\frac{H(X)f(X)}{E_u(H(X))}$, then (4.29) is equal to

$$max_v\int\frac{H(x)f(X;u)}{E_u(H(x))}lnf(X;v)dx\rightarrow max_vE_u(H(X))lnf(X;v) \tag{4.30}$$

The optimal vector of parameters v_{opt} is the outcome of this optimization process. Assume now that IS can be used iteratively to solve (4.30). In the first iteration of this procedure, IS will use a new sampling function $f(X;v_0)$ with different parameters from $f(X;u)$ and $f(X;v)$. Accordingly, (4.30) is rewritten as

$$max_vE_w(H(X))\frac{f(X;u)}{f(X;v_0)}lnf(X;v) \tag{4.31}$$

The respective optimal vector of reference parameters v_{opt} is

$$v_{opt}=argmax_vE_w(H(X))W(X;v_0,u)lnf(X;v) \tag{4.32}$$

where $W(X;v_0,u)=f(X;u)/f(X;v_0)$.

One approach to solve (4.32) is to use the following stochastic program

$$\tilde{v}^*=argmax_v\frac{1}{N}\sum_{i=1}^{N}H(X_i)\,W(X_i;v_0,u)lnf(X_i;v) \tag{4.33}$$

where N is the number samples drawn from $f(X;v_0)$.

Reference [60] shows that the vector v can be calculated via

$$v_j = \frac{\sum_{i=1}^{N} H(x_i)\, W(x_i;w,u) x_{ij}}{\sum_{i=1}^{N} H(x_i)\, W(x_i;w,u)} \tag{4.34}$$

Equation (4.30) shows that it is possible to create an IS-based multi-level algorithm to iteratively improve the reference parameters v_j, $j = 1, \ldots, d$, until the optimal vector v_{opt} is obtained.

Based on the details discussed above, the process of the CE IS variance reduction algorithm for composite system reliability assessment are presented as follows. The corresponding flow chart is shown in Figure 4.2.

Step 1) Specify the CE optimization parameters: a multilevel ρ=0.1 to determine rate event percentage, a maximum sample size N_{MAX} and COV criteria β_{MAX} for IS SMCS, and the iteration counter k=1.

Step 2) Define $\mathbf{v}_0$=$\mathbf{u}$, i.e., $\mathbf{v}_0$ is equal to the vector of the original unavailability. Determine the rare event criteria γ=R_{max}. R_{max} is assumed to be 1 MW representing the max acceptable load curtailment from state evaluation.

Step 3) Generate composite system states for the current iteration $\boldsymbol{X}_1, \boldsymbol{X}_2, \ldots, \boldsymbol{X}_N$ based on $\mathbf{v}_{k-1}$. Evaluate states performance $S(\boldsymbol{X}_i)$ in the form of load curtailment needed to maintain the system within operating limit. Sort $S(\boldsymbol{X}_i)$ in the descending order so that $\mathbf{S} = [S_{[1]}, S_{[2]}, \ldots, S_{[N]}]$ and $S_{[1]} \geq S_{[2]} \geq \ldots \geq S_{[N]}$.

Step 4) Access the state performance value at the rare event multilevel $S_{[\rho N]}$. If $S_{[\rho N]} \geq \gamma$, exit the CE optimization process. If $S_{[\rho N]} < \gamma$, evaluate the test function $H(\boldsymbol{X}_i)$ for all states $\boldsymbol{X}_i$. If $S(\boldsymbol{X}_i) > \gamma$, $H(\boldsymbol{X}_i)=1$; otherwise, $H(\boldsymbol{X}_i)=0$. Calculate the likelihood ratio $\mathbf{W}(\boldsymbol{X}_i;n,\boldsymbol{u},\boldsymbol{v})$ and distorted PDF parameters $\mathbf{v}_k$ based on (4.27) and (4.34) respectively.

Step 5) k=k+1, go back to *step 3)* for the next iteration.

Step 6) after the CE optimization is completed, say after k iterations, the $\boldsymbol{v}_{opt}$=$\boldsymbol{v}_k$ gives the optimal PDF parameters.

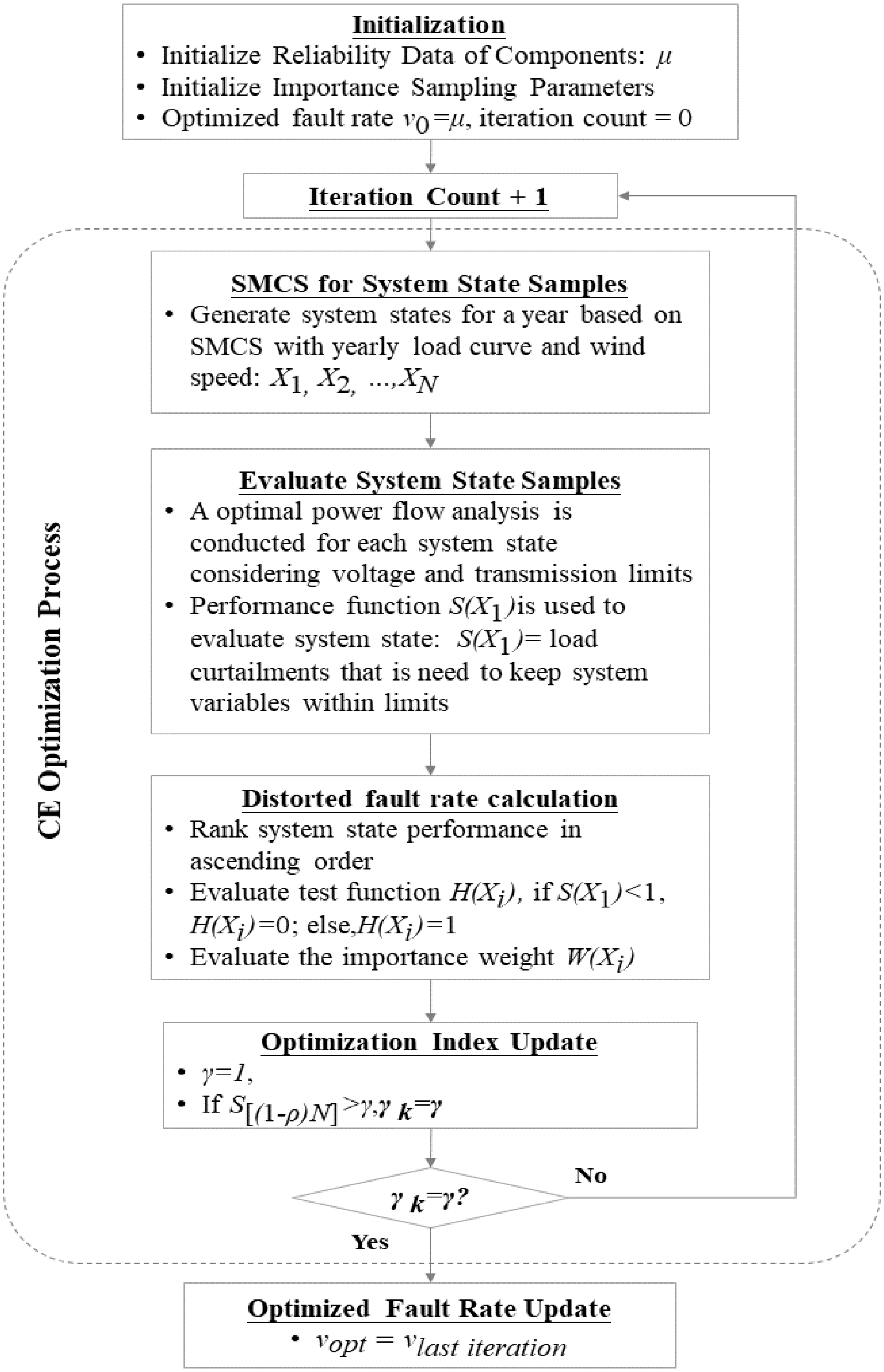

Figure 4.2 CE-IS optimization procedure

The optimal PDF parameter $\boldsymbol{v}_{opt}$ is used as the input value for the integrated reliability evaluation process. The fault rate is modified using

$$\lambda^{*}=\frac{v_{opt}}{1\text{-}v_{opt}} \tag{4.35}$$

To avoid bias, the reliability indices calculated from the SMCS based on the optimal unavailability $\boldsymbol{v}_{opt}$ are as follows:

$$T_{down}=\sum t_i\cdot\ \boldsymbol{W}(\boldsymbol{X_i};n,\boldsymbol{u},\boldsymbol{v}) \tag{4.36}$$

$$ENS=\sum t_i\cdot\Delta P(\boldsymbol{X_i})\cdot\boldsymbol{W}(\boldsymbol{X_i};n,\boldsymbol{u},\boldsymbol{v}) \tag{4.37}$$

$$OCC=\sum 1\cdot\boldsymbol{W}(\boldsymbol{X_i};n,\boldsymbol{u},\boldsymbol{v}) \tag{4.38}$$

The reliability indices are calculated as:

$$LOLP=\frac{T_{down}}{8760};\ EPNS=\frac{ENS}{8760};\ LOLF=\frac{OCC}{8760} \tag{4.39}$$

4.2.2 Transient Stability Pruning based on TEF

The dynamic security of system states selected from SMCS is assessed via TDS. It is well known that the TDS, while providing accurate results, has a significant computational burden. A screening tool for TDS is therefore investigated in this work to make the evaluation process more efficient. The proposed screening tool classifies all system states based on a two-stage approach. Firstly, an early terminated TDS is conducted for each system state to obtain system operating condition after a fault is cleared (5 cycles after the fault occurred). The system states are classified to be critical or non-critical based on the kinetic energy gained by the machines due to the fault and the maximum change in the magnitude of Thévenin impedance seen at the point of interconnection (POI) of a generator.

The reason for using these two indicators is that the stability of a power system during a fault basically depends on the kinetic energy gained by the system due to the fault and the robustness of the post-disturbance network [61]. During a fault, the ability of the network to export electrical power is severely restricted causing the machine to accelerate. Once the fault is cleared by opening the faulted line, the machine is able to export electrical power and it decelerates. The stability of the machine is dependent on its ability to decelerate in the post-disturbance condition and to reach a steady state. To estimate the ability of the system to decelerate, the change in the magnitude of Z_{th} looking into the system at the POI of the generator due to the opening of the faulted line is computed. A review of the swing equation [61] indicates that a large change in the magnitude of Z_{th} results in a substantial reduction in the peak of the post-fault swing curve as compared to the pre-faulted condition. A reduction in the peak of the post-fault swing curve limits the ability of the generator to decelerate, thereby making it prone to instability.

The kinetic energy and Thévenin impedance can be calculated from simulation results from the early terminated TDS which last to fault clear time. In this work, faults are assumed to be cleared in 5 cycles. Detailed TDSs are conducted for critical cases. Load curtailment can be calculated from TDS for each system scenario and serves as a system reliability index.

From the early terminated TDS, the angular speed of generators at the end of the fault can be obtained. The calculation of corrected kinetic energy gained during the fault is as follows:

$$\omega_{coi}=\frac{\sum_{allgens} M_i\,\omega_i}{\sum_{allgens} M_i} \tag{4.40}$$

$$M_{cr}=\sum_{critical\ gens} M_{icr} \tag{4.41}$$

$$M_{non\_cr}=\sum_{noncritical\ gens} M_{i_{non\_cr}} \tag{4.42}$$

$$\widetilde{\omega}_{cr}=\frac{\sum_{critical\ gens} M_{icr}(\omega_{icr}\text{-}\omega_{coi})}{M_{cr}} \tag{4.43}$$

$$\widetilde{\omega}_{non\_cr}=\frac{\sum_{noncritical\ gens} M_{i_{non\_cr}}\left(\omega_{i_{non\_cr}}\text{-}\omega_{coi}\right)}{M_{non\_cr}} \tag{4.44}$$

$$M_{eq}=M_{cr}*M_{non\_cr}/(M_{cr}+M_{non\_cr}) \tag{4.45}$$

$$\widetilde{\omega}_{eq}=\widetilde{\omega}_{cr}\text{-}\widetilde{\omega}_{non\_cr} \tag{4.46}$$

$$KE_{corr}=\frac{1}{2}M_{eq}(\widetilde{\omega}_{eq})^2 \tag{4.47}$$

where, ω_{coi} is the angular velocity of the center of inertia, ω_i is the angular velocity of the i th generator, M_i is the inertia constant of the$i$ th generator, M_{cr} is the inertia constants of the critical generator's inertial center, M_{non_cr} is the inertia constants of the non-critical generators inertial center, $\widetilde{\omega}_{cr}$ is the angular speed of the inertial center of the critical generator group, $\widetilde{\omega}_{non\_cr}$ is the angular speed of the inertial center of the non-critical generator group, M_{eq} and $\widetilde{\omega}_{eq}$ is the equivalent inertia constant and angular speed of the system, KE_{corr} is the corrected kinetic energy.

For the pre-fault condition, the Thévenin impedance at POI can be directly obtained by the pre-fault network. The change in the magnitude of Z_{th} can be calculated by removing faulted components based on the Z_{th} in the pre-fault network. The maximum change ΔZ_{thmax} in the magnitude of Z_{th} and the corresponding bus number of the POI where it occurs are recorded.

4.2.3 Reliability Evaluation Procedure with Accelerating Process

Previously, the overall integrated reliability assessment method with the computation accelerating methods have been discussed and the integrated procedure is shown in Figure 4.3.

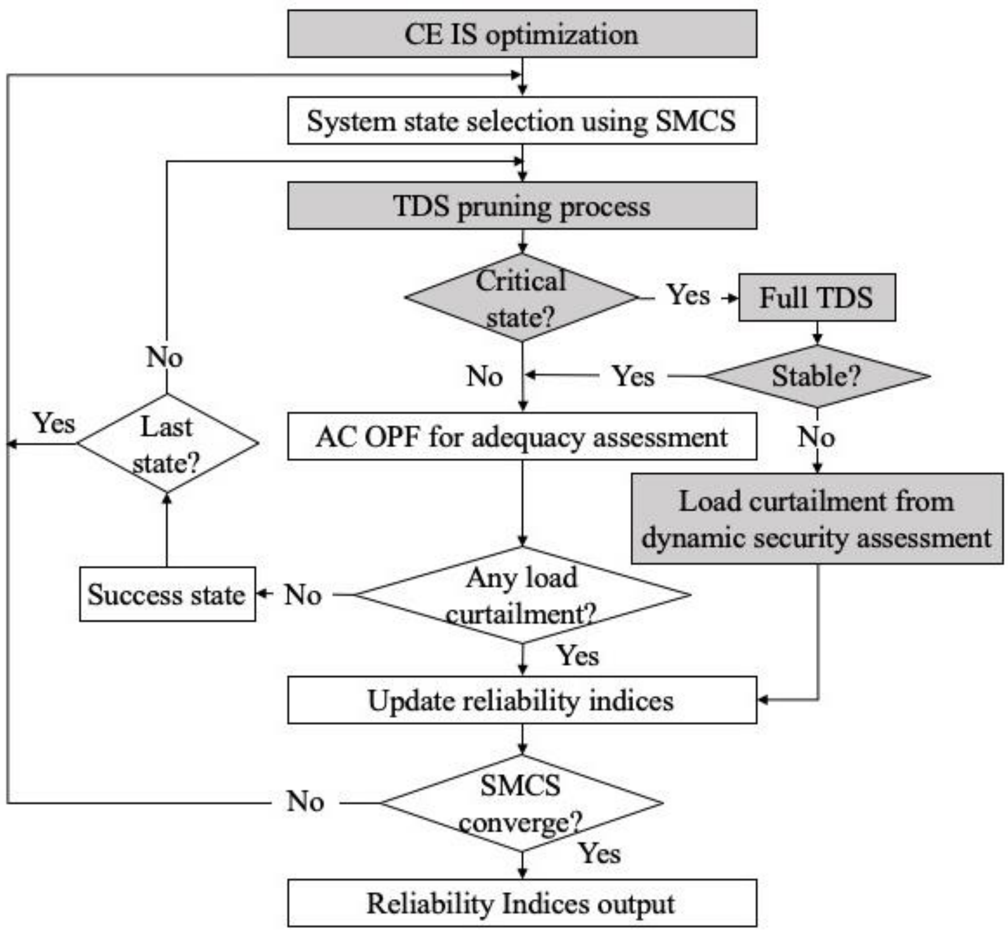

Figure 4.3 Reliability evaluation procedure with accelerating process

4.3 Summary

In this chapter, the integrated reliability evaluation procedure is presented. The assessment methods for system steady-state reliability and dynamic security are discussed. The approach evaluates the system response for each selected contingency, examines transient stability in the transition from the pre-contingency to the post-contingency period, and evaluates the post-contingency steady-state equilibrium where all flow limits and voltage limits are satisfied. Stochastic characteristics including the renewable resources, component failures, and load variations are taken into consideration in probabilistic sampling and dynamic performance modeling. System adequacy and dynamic security are quantified in terms of MW load curtailment from the respective assessment processes and are incorporated into the calculation of the integrated reliability indices. The outcome of the integrated reliability evaluation provides the probability, frequency and magnitude of system reliability represented as LOLP, LOLF, and EENS.

To address the computational efficiency problem, two accelerating techniques are presented this chapter with the objective to improve the computational efficiency of the MCS process and TDS process, respectively. The proposed approach is tested on a synthetic system, the test results are presented and discussed in the following chapter

5. Simulation, Results and Discussion

This chapter provides the detailed simulation results of the integrated reliability evaluation. The test is performed on the test system described in section 5.1. The study is conducted using analytical tools from the GE PSLF and Siemens PSS®E OPF software. The overall procedure and coordination between different software tools are done using Python.

5.1 System Description

A synthetic test system is used to perform the reliability evaluation study. The single line diagram of the test system is shown in Figure 5.1. The system consists of 11 conventional synchronous generators, 10 wind farms with type-3 WTGs, 20 transmission lines, and 6 loads at different buses. The total installed capacity of conventional generation is 17,000 MW, among that wind energy is with an installed capacity of 1,680 MW. It includes major features of a realistic power system for transient stability and reliability studies for system planning, with detailed positive sequence generator, governor and exciter models for synchronous generators, generator, and converter control models for WTGs. The power flow and dynamic data for the test system are available in [64].

Table 5.1 Synthetic test system summary

Buses	36
Generators	11 Synchronous + 10 Wind farms
Lines	30
Total Synchronous Generation	17,000 MW installed capacity
Wind Generation	1,680 MW installed capacity

5.2 Parameters used in Reliability Evaluation

5.2.1 Generator Reliability Data

Generator reliability data for the test system is assembled base on the data from IEEE Reliability Test System-1996. Unit availability data in the form of MTTF and MTTR are given in Table 5.2.

Table 5.2 Generation reliability data of the test system

No.	Bus(es)	Base (MVA)	Maximum output (MW)	MTTF (hour)	MTTR (hour)
1	1	2200	2000	1150	100
2	14, 22	2400	2000	1150	100
3	24	4500	4000	1150	100
4	2, 4, 5, 7, 26	1200	1000	1100	150

5	11	600	500	960	40
6	20	1800	1500	1100	150
7	8	33.4	30	1150	100
8	33, 35	23.4	21	1150	100
9	28, 29, 30,31	33.4	30	1100	150
10	34	23.4	21	1100	150
11	32	33.4	30	960	40
12	36	23.4	21	1500	100

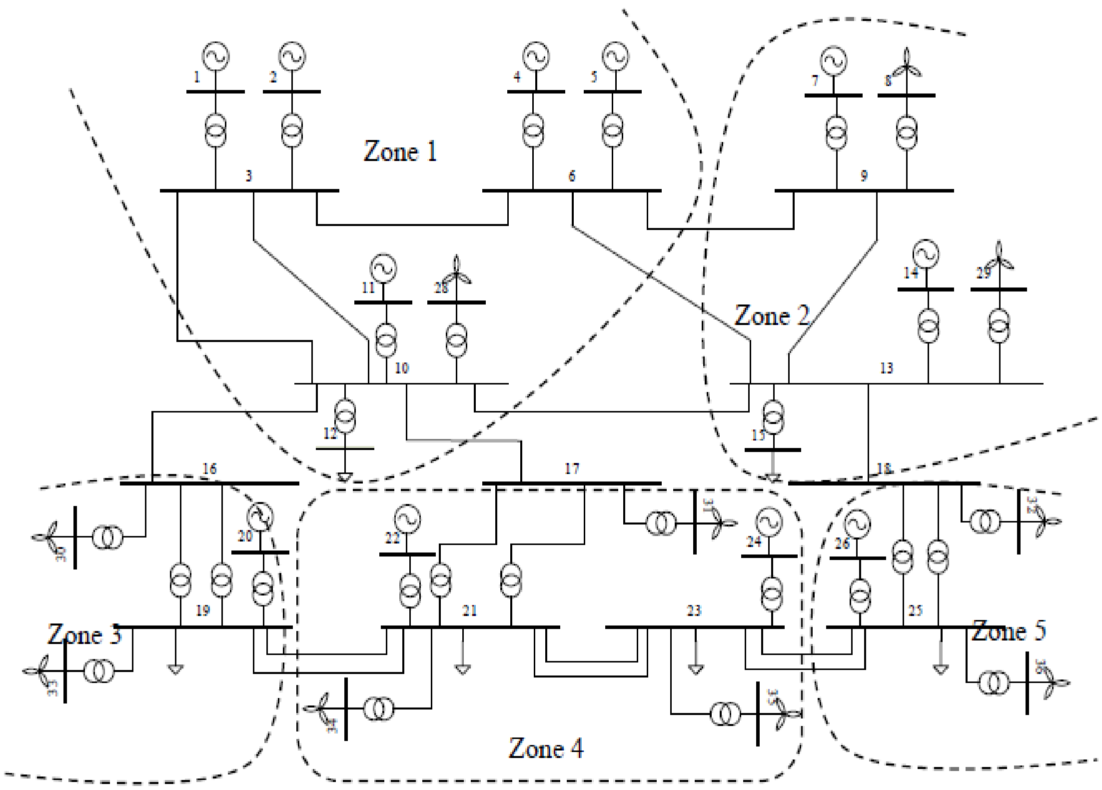

Figure 5.1 Synthetic test system

5.2.2 Transmission Reliability Data

The fault rates for transmission lines are assembled based on the data in the Canadian Electricity Authority (CEA) 2012 annual report [62]. The CEA report provides the transmission line statistics for line-related transient forced outages data in the form of the frequency of outage of transmission lines for different voltage levels in number per 100 mile-annum.

Table 5.3 Transmission reliability data of the test system

No.	**Voltage level (kV)**	**Mean fault duration (hour)**	**Unavailability (%)**
1	110	20.7	0.216
2	500	27.2	0.077

5.2.3 Load Curve Data

Hourly load data in a year is based on the hourly load data from IEEE Reliability Test System-1996. The data includes weekly peak loads in percent of the annual peak, the daily peak load in percent of the weekly peak, and the hourly load in percent of the daily peak. Once the system annual peak load is assigned, the 8760-hour load curve in the year can be calculated.

Table 5.4 Hourly Peak Load in Percent of Daily Peak

	winter weeks		**weeks**		**spring/fall weeks**	
	1 -8 & 44 - 52		**18 -30**		**9-17 & 31 - 43**	
Hour	**Wkdy**	**Wknd**	**Wkdy**	**Wknd**	**Wkdy**	**Wknd**
12-1 am	67	78	64	74	63	75
1-2	63	72	60	70	62	73
2-3	60	68	58	66	60	69
3-4	59	66	56	65	58	66
4-5	59	64	56	64	59	65
5-6	60	65	58	62	65	65
6-7	74	66	64	62	72	68
7-8	86	70	76	66	85	74
8-9	95	80	87	81	95	83
9-10	96	88	95	86	99	89
10-11	96	90	99	91	100	92
11-noon	95	91	100	93	99	94
Noon-1pm	95	90	99	93	93	91
1-2	95	88	100	92	92	90
2-3	93	87	100	91	90	90
3-4	94	87	97	91	88	86
4-5	99	91	96	92	90	85
5-6	100	100	96	94	92	88
6-7	100	99	93	95	96	92
7-8	96	97	92	95	98	100
8-9	91	94	92	100	96	97
9-10	83	92	93	93	90	95
10-11	73	87	87	88	80	90
11-12	63	81	72	80	70	85

5.3 System Adequacy Evaluation Results

In this section, system reliability is evaluated using the traditional reliability evaluation approach which considers only system adequacy. As described in section 5.1, the total generation capacity is 17,000 MW, with an additional 1,680 MW from wind power. The system peak load is 7612 MW and 2108 MVAr. The hourly load data is generated based on system peak load and the hourly load percent listed in Table 5.4. SMCS sample system states based on the components fault data listed in Table 5.2 and Table 5.3 are obtained. According to the peak load and annual load curves, the system load in 8760 hours are obtained and assigned to the selected system states according to the time stamps of the states. The same is done for the annual wind speed data in each of the 8760 hours. Since the system load level and wind speed vary in each state, the generators are dispatched accordingly. AC OPF is conducted for each system state to assess system adequacy.

The convergence criterion, coefficient of variance (COV), is set to be 5%, and the simulation converges after 746 iterations with a total simulation time of 8.95×10^5s, approximately 248 h (3600 seconds in an hour). Reliability evaluation results are given in Table 5.5. Figure 5.2 - Figure 5.5 show the convergence trajectory of the COV and reliability indices of LOLP, EPNS, and LOLF respectively. After the convergence of the SMCS, the methods found an LOLP index around 0.0015, an EPNS index around 0.0087 MW, and LOLF index around 2.7663 occ/ year. This simulation provides reference values of the reliability indices for the test system.

Table 5.5 Reliability indices from traditional SMCS method

# Iteration	**COV criteria**	**COV**	**LOLP**	**EPNS (MW)**	**LOLF (occ./y)**
746	5%	4.98%	0.0015	0.0087	2.7663

5.4 Impact of Accelerating Techniques

As illustrated in section 4.2.1, the CE IS method is introduced in the SMCS to improve the simulation efficiency. In this section the reliability results from CE IS based SMCS is presented and a comparison with the traditional SMCS is discussed. To illustrate the accuracy of the CE IS based SMCS algorithm, the simulation is conducted on the same input system peak load, annual load data, and annual wind speed data. The parameter settings of the CE-IS SMCS are as follows: $\rho = 0.1$, $N_{MAX} = 10000$, $\beta_{MAX} = 5\%$. Table 5.6 shows the adequacy evaluation results for the test system using the crude SMCS with and without the CE-IS acceleration method. Figure 5.6 - Figure 5.9 show the comparison of the convergence trajectory of the COV, LOLP, EPNS, and LOLF indices using the two methods.

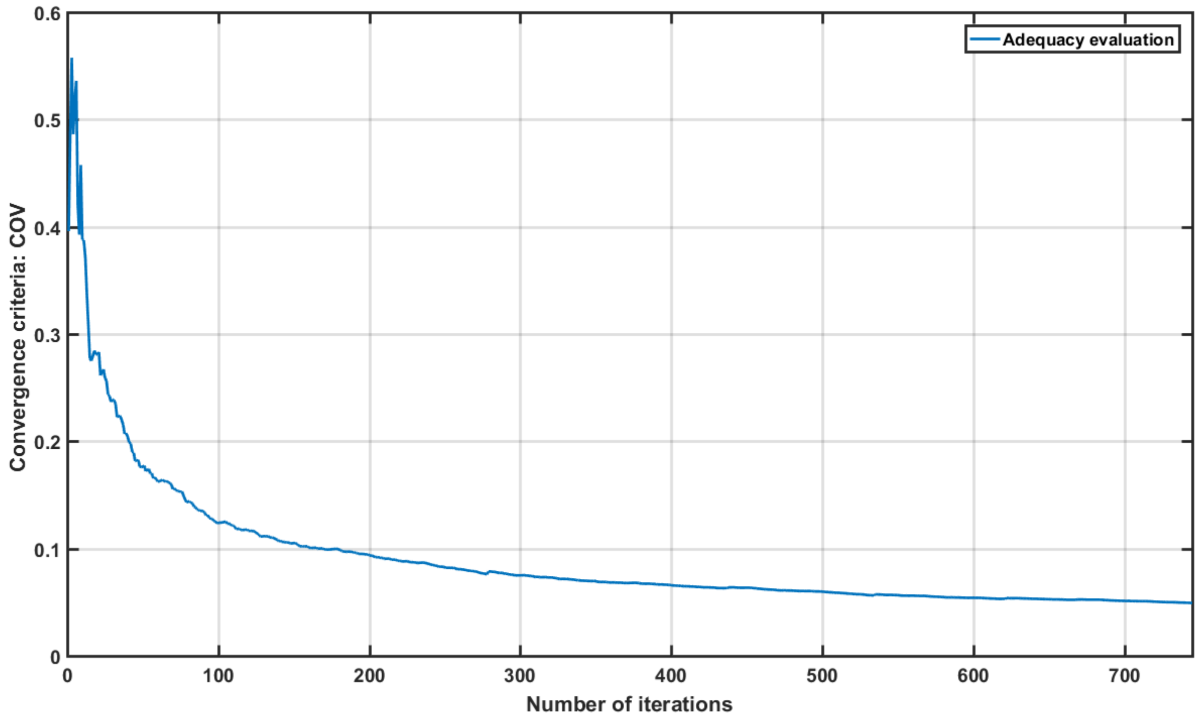

Figure 5.2 Convergence trajectory of COV from adequacy assessment based on SMCS

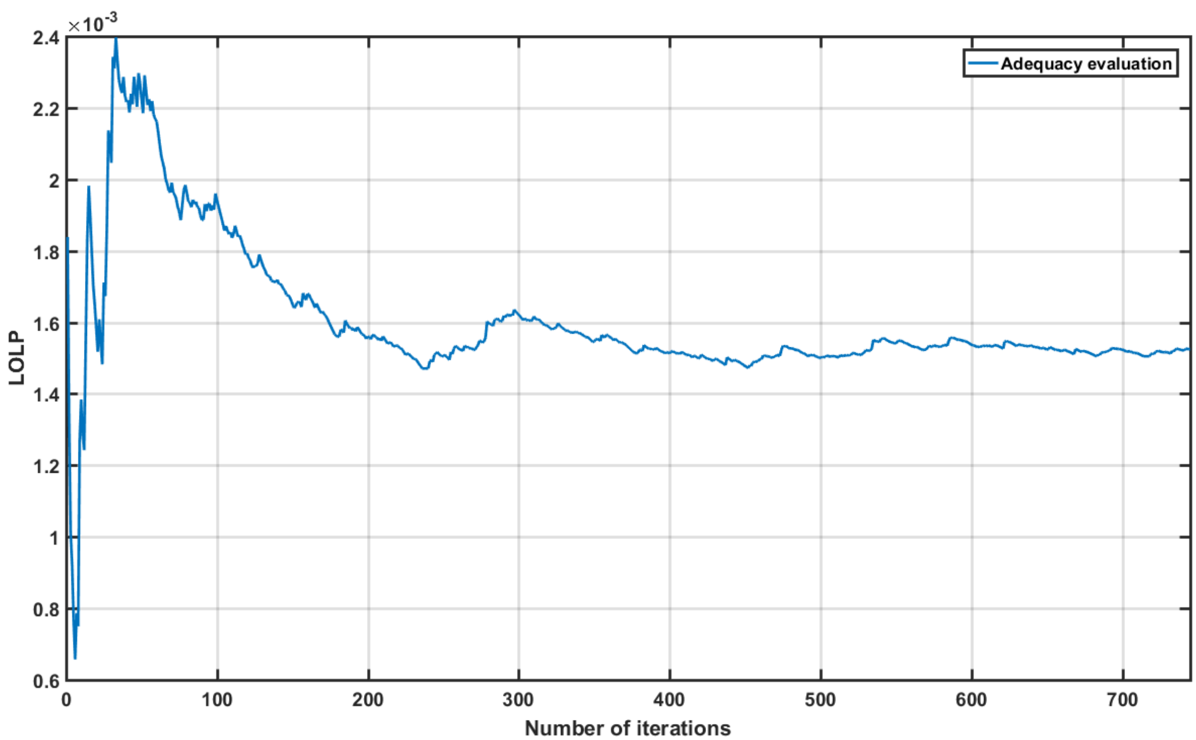

Figure 5.3 Convergence trajectory of LOLP from adequacy assessment based on SMCS

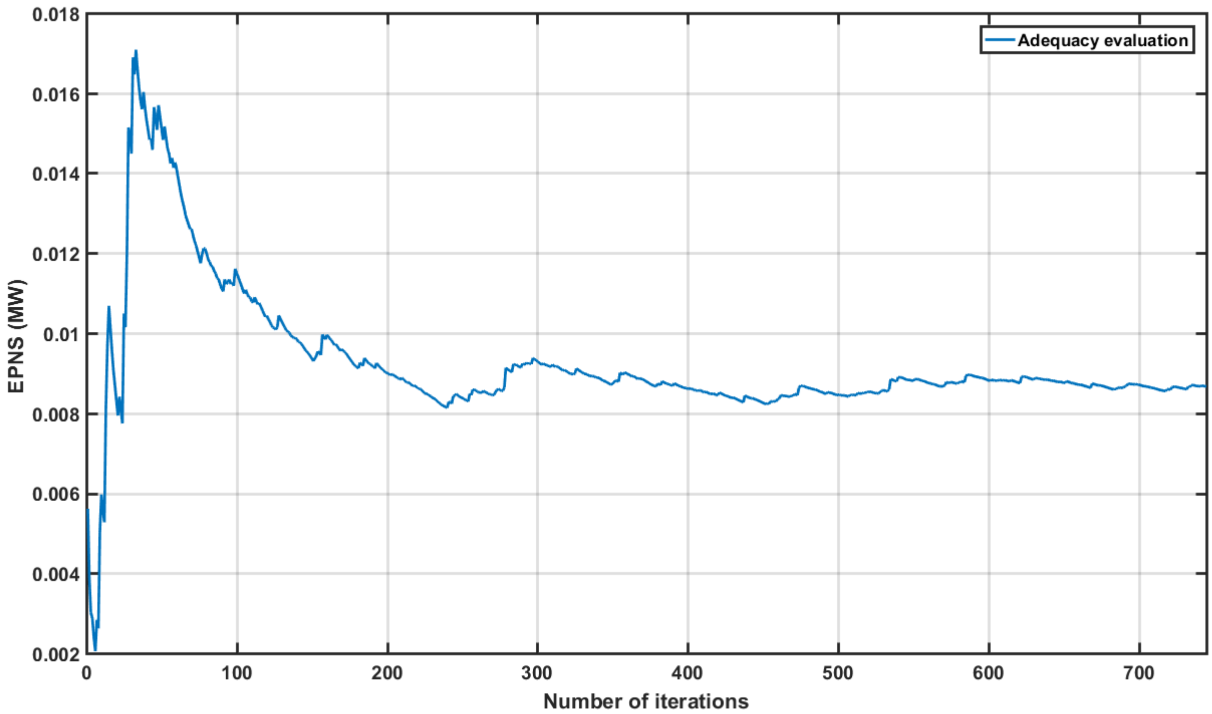

Figure 5.4 Convergence trajectory of EPNS from adequacy assessment based on SMCS

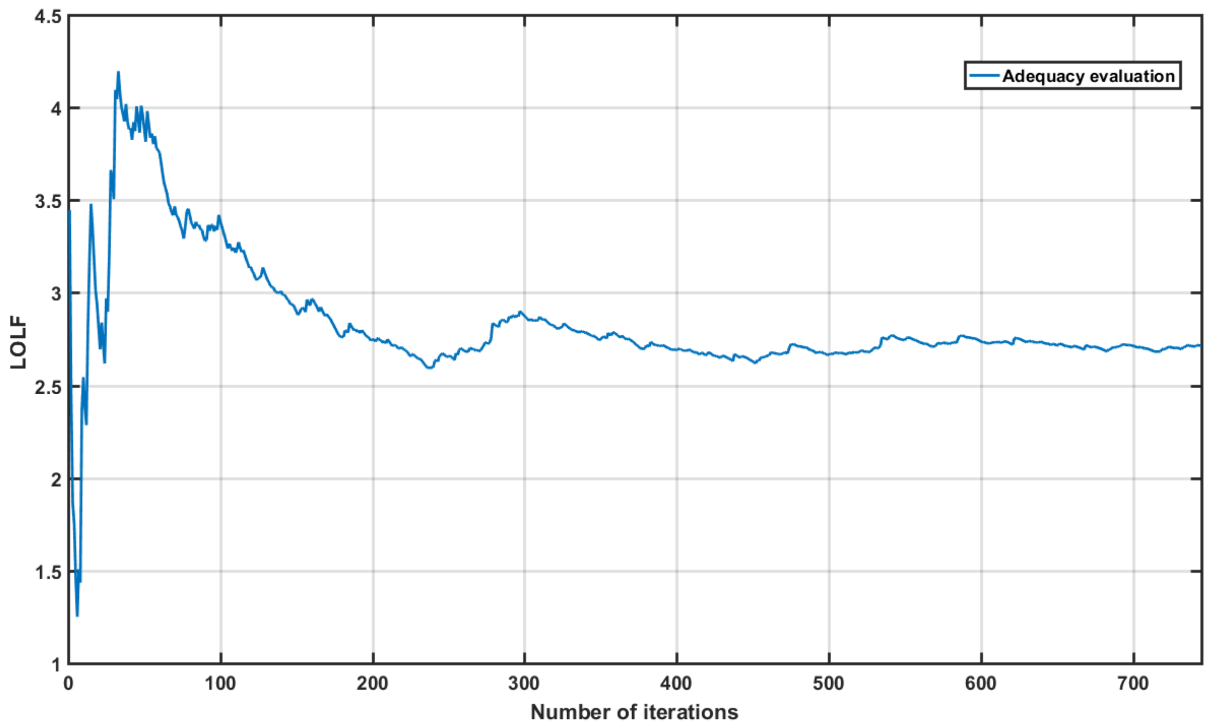

Figure 5.5 Convergence trajectory of LOLF from adequacy assessment based on SMCS

The evaluation based on SMCS with CE-IS obtains an LOLP value of 0.0014, which is within a 0.01% deviation from the crude method result. Given that both values are only estimates of the true value, the difference is negligible. However, the CE-IS SMCS method reaches convergence much faster. The simulation converges after 81 iterations with a computation time of 0.98×10^5 seconds while the evaluation without CE-IS takes 8.95×10^5 seconds to complete. Fig. 3 shows the convergence trajectory of the COV and LOLP values from the two methods. The results show that by applying the CE-IS SMCS acceleration method, a speed up factor of 9.13 in elapsed time is achieved, while maintaining approximately equal evaluation results. Hence, it can be concluded that the effect of CE-IS method is verified, it is applied for reliability evaluation studies in rest of the paper.

Table 5.6 Reliability indices comparison between traditional SMCS and CE-IS SMCS

Method	**LOLP**	**EPNS**	**LOLF occ./yr**	**Number of iterations**	**Computation time**
Crude SMCS	0.0015	0.0087	2.7663	746	8.95×10^5s (248 h)
CE-IS SMCS	0.0014	0.0079	2.7650	81	0.98×10^5s (27 h)

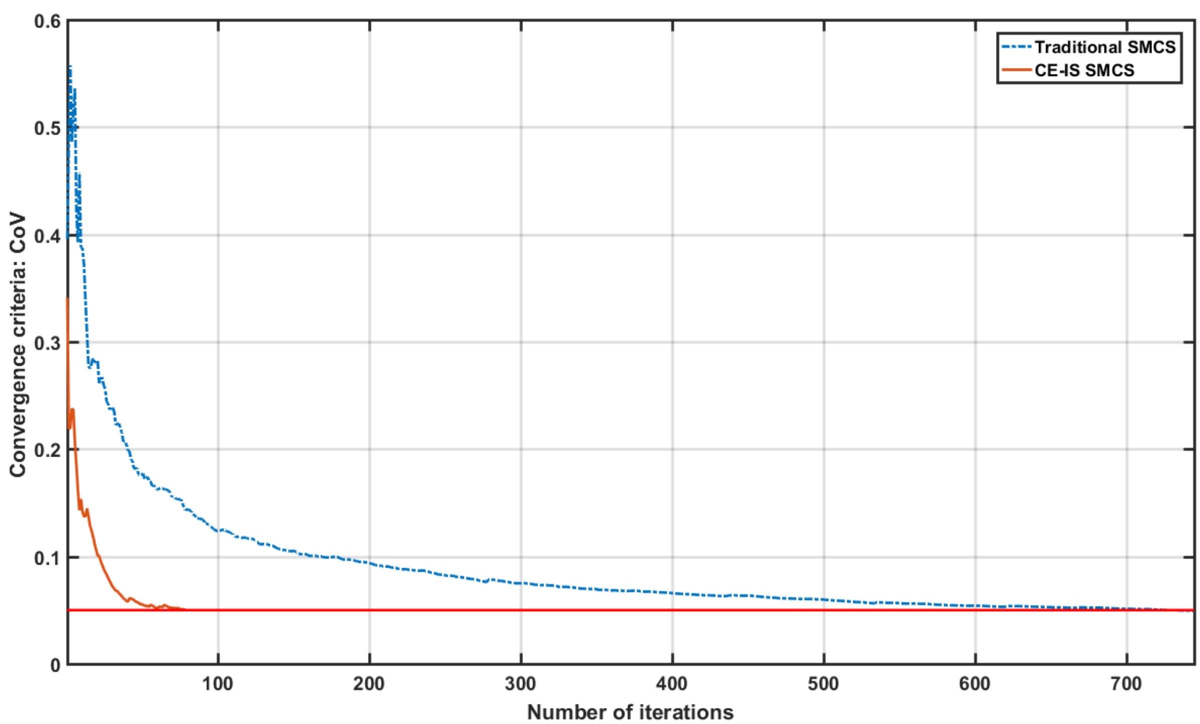

Figure 5.6 Convergence trajectory of COV from adequacy assessment based on CE-IS SMCS

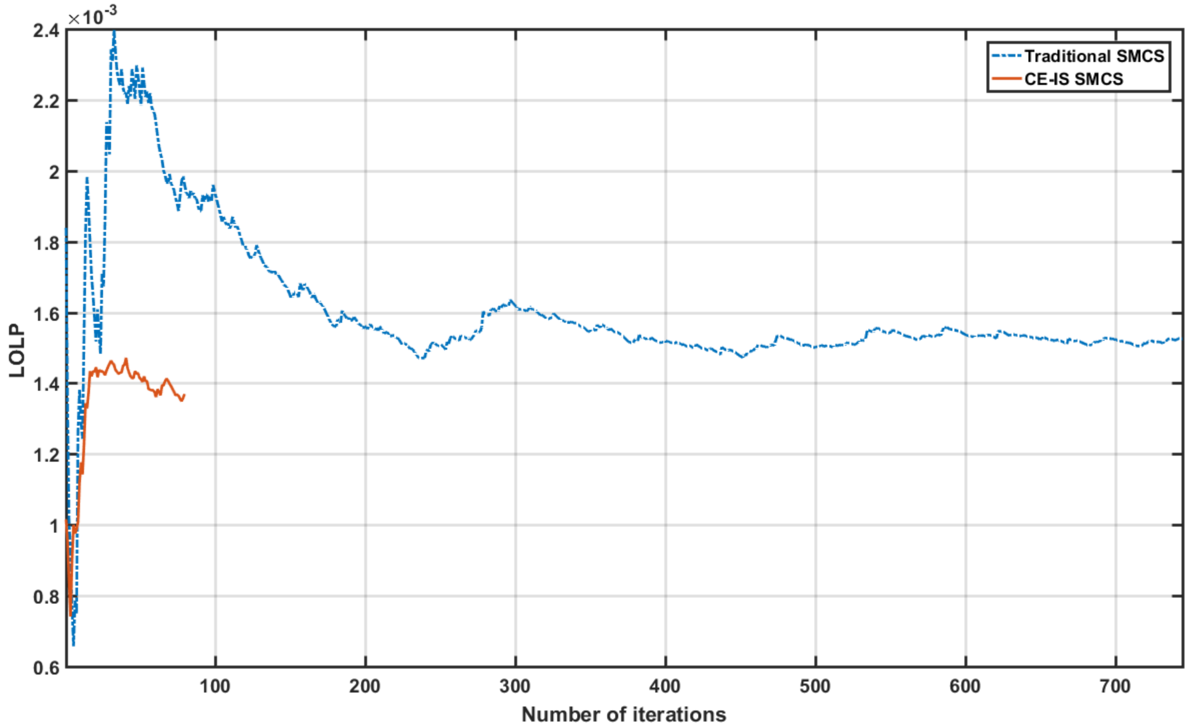

Figure 5.7 Convergence trajectory of LOLP from adequacy assessment based on CE-IS SMCS

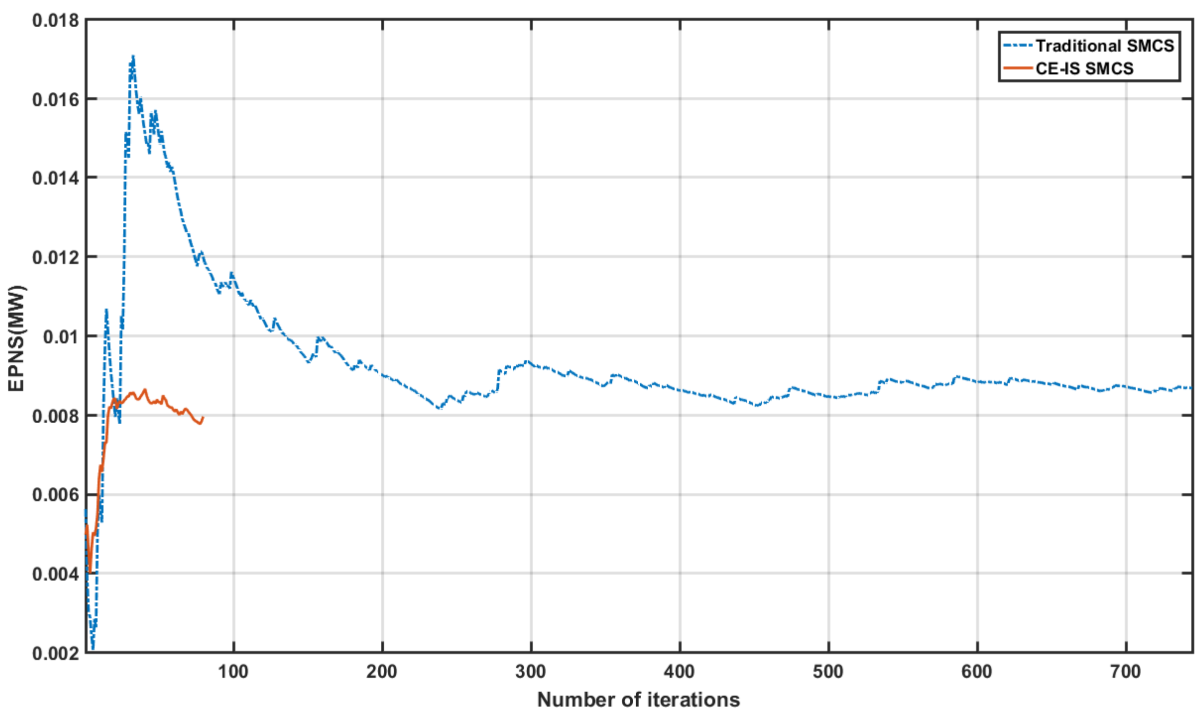

Figure 5.8 Convergence trajectory of EPNS from adequacy assessment based on CE-IS SMCS

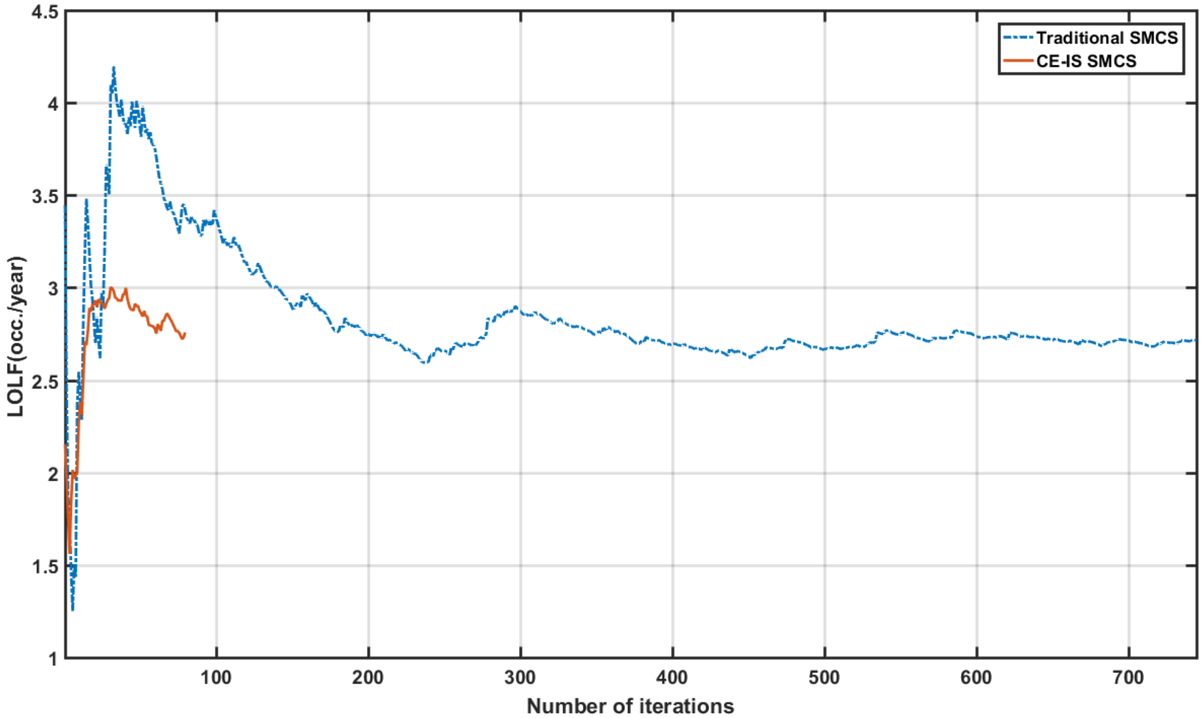

Figure 5.9 Convergence trajectory of LOLF from adequacy assessment based on CE-IS SMCS

5.5 Integrated Reliability Evaluation Results

Previous simulation results show the reliability indices from the traditional SMCS method and the CE IS based SMCS. Both methods reached a very similar reliability index associated with LOLP, while CE IS based SMCS demonstrated a computational speed-up of 9.13 times. Based on the proposed speed-up of the SMCS, this section will discuss the integrated reliability results considering both adequacy and transient stability. The integrated results provide a measure for both steady-state and dynamic evaluation of the ability of the system to meet electrical demand. The reliability results are compared in Table 5.7. With system dynamic security being quantified by load shedding due to the stability corrective control, the proposed method provides the quantitative integrated reliability evaluation results considering both system adequacy and dynamic security. The comparison between the results from the two approaches reflects the influence of including impact of dynamic security in the reliability evaluation.

From Table 5.7, it can be observed that for this test system, accounting for dynamic security has a significant impact on all three reliability indices resulting in an increase in the LOLP, EPNS, and LOLF indexes. The LOLP index and LOLF index increase from 0.0014 to 0.0939 and from 2.7650 to 35.6944, respectively. The increase of these two indices indicates that among all sampled states in a year, there is large portion of cases and longer duration for which the system cannot provide reliable power supply because of dynamic security problems. The increase of the EPNS index is due to the fact that the load shedding value from DSA is included in the EPNS index calculation. Theoretically, the increase of all three reliability indices was to be expected because a stable transition to post-fault SEP is not always guaranteed after contingencies. However, these results provide a quantitative analysis of this impact. Results in Table 5.7 also show that the number of iterations for the SMCS to converge is reduced in the integrated approach. The reason for this is that with DSA considered, the values of indices increase and the number of iterations is inversely proportional to the index being calculated. However, despite the fewer iterations, the computation time is still much higher than the adequacy evaluation computation time because of the computational burden introduced by TDS for DSA. A statistical summary of states evaluation in one iteration of SMCS shows that: among all 11992 system states that are sampled, there are 408 cases labeled as steady-state unreliable from adequacy assessment, 5002 cases labeled as dynamically insecure, and 4603 cases labeled as steady-state reliable yet dynamically insecure. Among the 4603 cases, 8 cases are N-1 contingencies and 4595 cases are N-k contingencies with k>1. The statistical data confirms that the reliability study will give optimistic results if the DSA is not considered.

Table 5.7 Reliability indices comparison: Adequacy Vs. Adequacy and transient stability

Method	**LOLP**	**EPNS**	**LOLF occ./yr**	**Number of iterations**	**Computation time**
Integrated reliability	0.0939	72.80	35.6944	20	3.89×10^5s (108 h)
Adequacy	0.0014	0.0079	2.7650	81	1.16×10^5s (27 h)

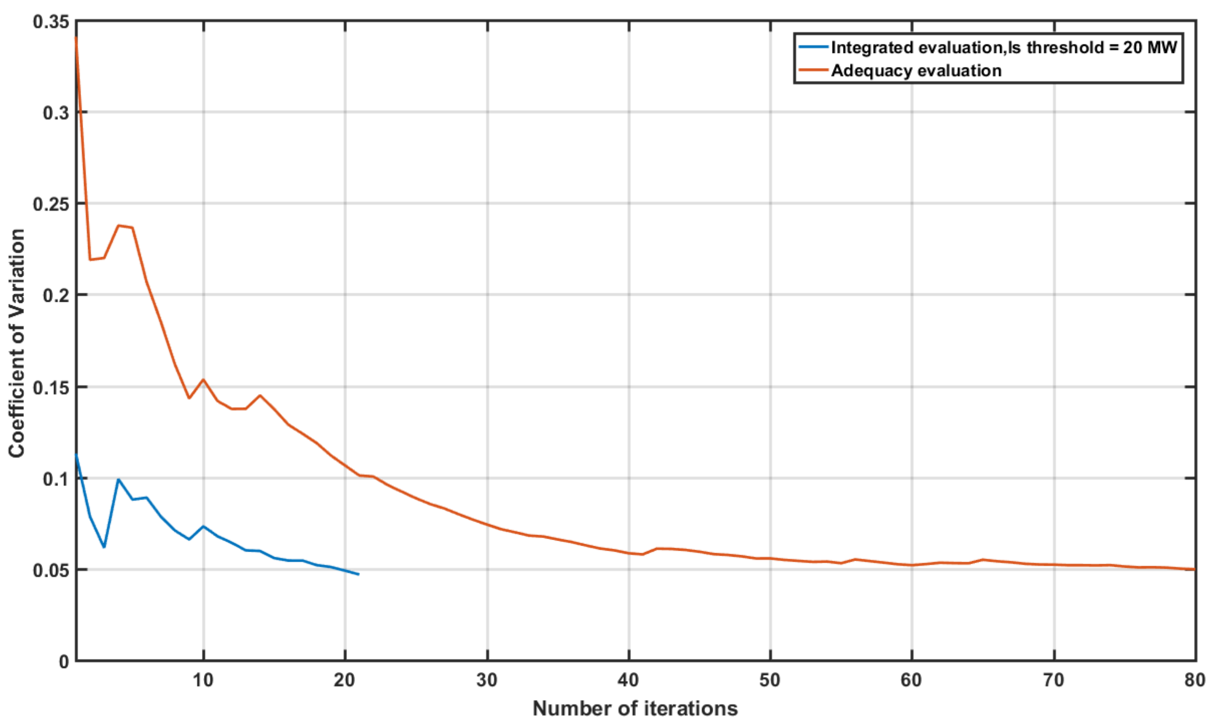

Figure 5.10 Convergence trajectory of COV comparison: adequacy Vs. adequacy and transient stability

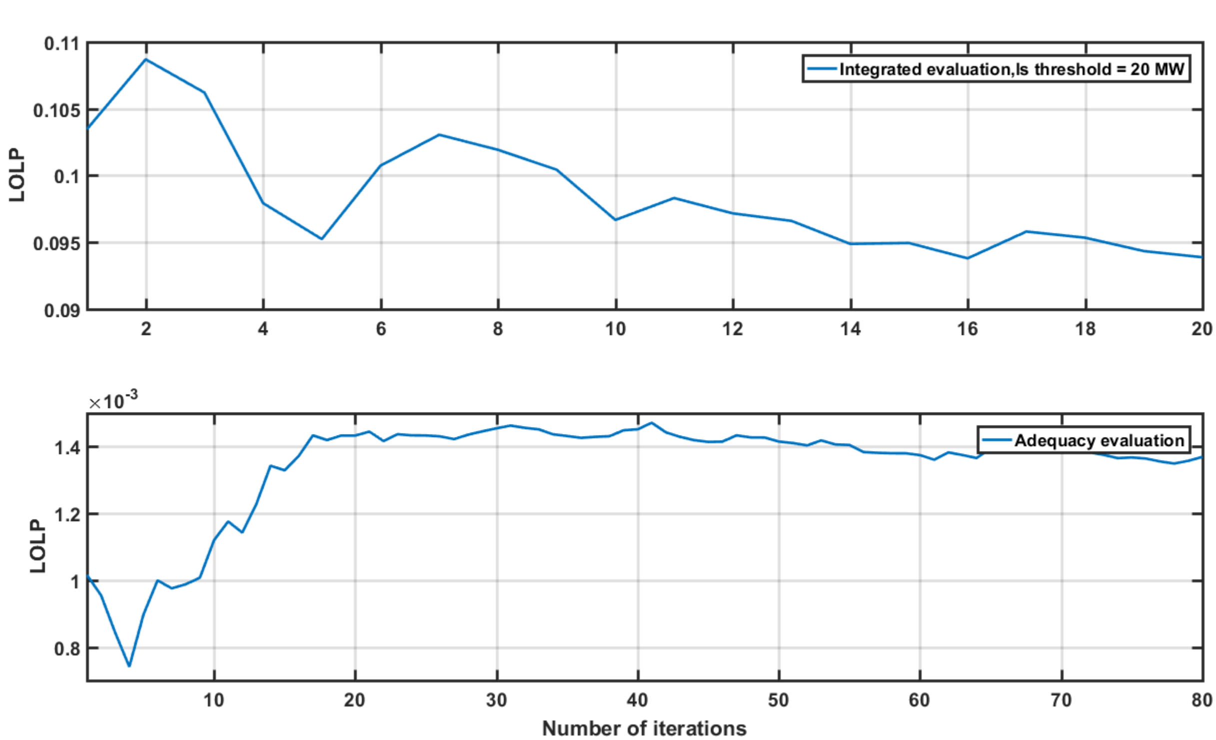

Figure 5.11 Convergence trajectory of LOLP comparison: adequacy Vs. adequacy and transient stability

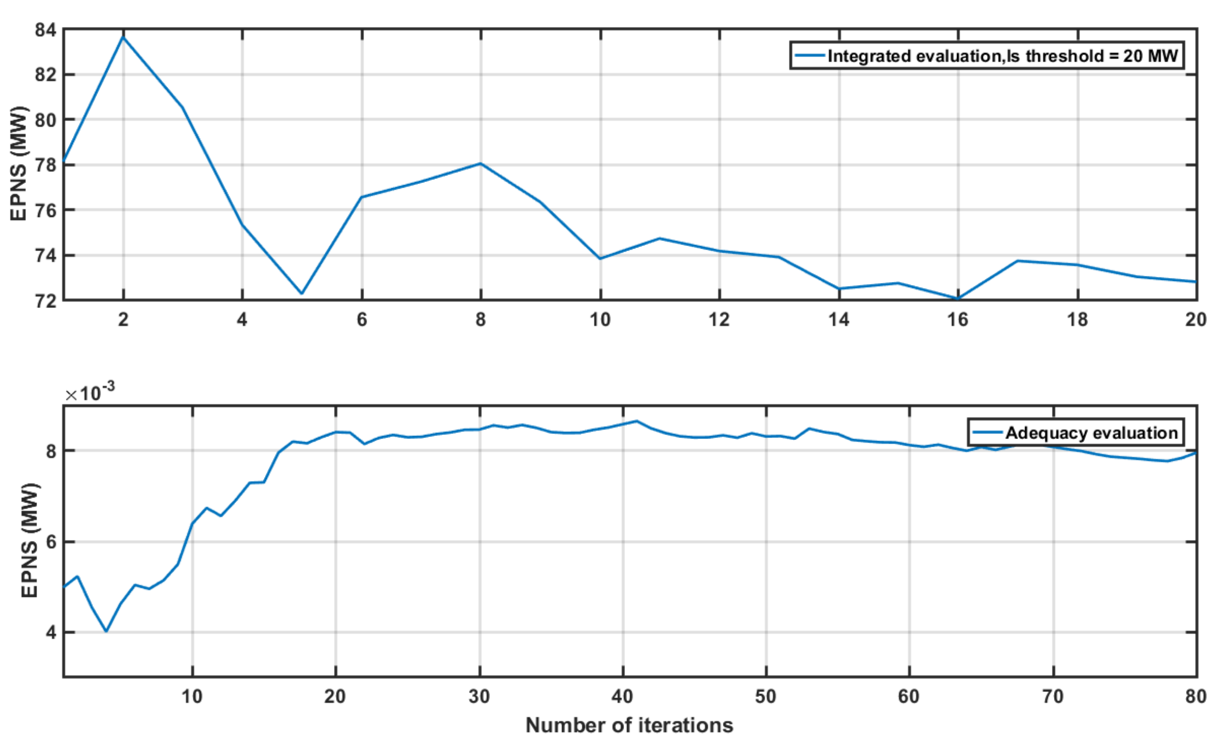

Figure 5.12 Convergence trajectory of EPNS comparison: adequacy Vs. adequacy and transient stability

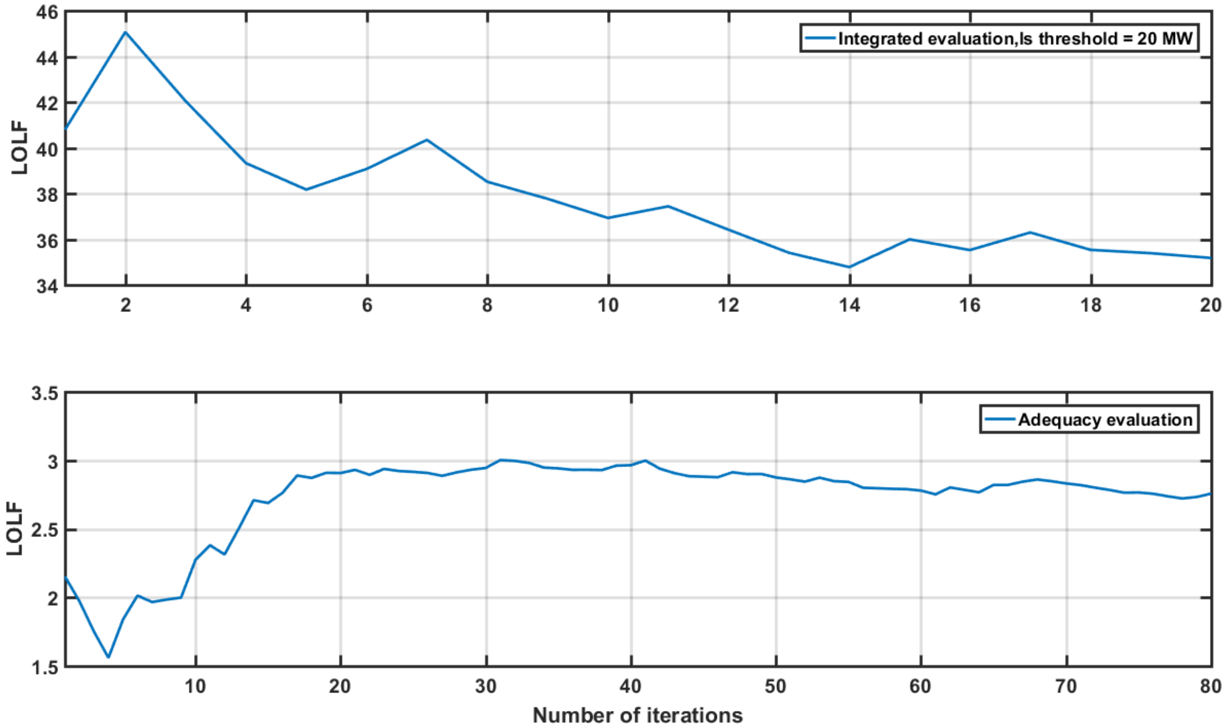

Figure 5.13 Convergence trajectory of LOLF comparison: adequacy Vs. adequacy and transient stability

The trajectories of COV and the values of the three reliability indices are compared in Figure 5.10 – Figure 5.13 to show the impact on reliability assessment when dynamic security is considered. As shown in Figure 5.10, the simulation of the integrated reliability evaluation converges faster than that without dynamic security. The reason for this is that with dynamic security considered, more cases are detected to be unreliable and are therefore considered in the reliability calculation. With a larger number of unreliable cases being viewed as important, the variance is reduced which provides the same effect as importance sampling.

The analysis of a specific case is essential to understand the states that are steady-state reliable but dynamically insecure. A state with 6375 MW load and 6581 MW generation in the pre-fault condition is chosen to conduct this analysis. The set of contingencies in this selected state are listed in Table 5.8.

For this case, the adequacy analysis using AC OPF gives a load curtailment of 0.185 MW. This assessment result indicates that the system state is steady-state reliable. By contrast, the results from TDS for DSA show that 1554.4 MW load needs to be shed to maintain stability. The load shedding protection actions are listed in Table 5.9. From Table 5.9, it can be observed that two stages of under-frequency load shedding protection are activated after the fault. The first stage protection was activated at around 2.5 s when the load bus frequency dropped to 59.5 Hz, and the second stage of protection action was initiated between 4.2 s-4.3 s when the load bus frequency dropped to 59.2 Hz, as shown in Figure 5.14. The two stages of protection action brought the system back to a stable operating condition. Figure 5.15 shows the active power output of the 11 synchronous generators in the system. Since the generators at bus 24 and bus 26 are the contingency components, the loads at bus 23 and bus 25 which were primarily supplied by generators at bus 24 and bus 26, suddenly lost their power supply. In addition, the transmission outage from bus 13 to bus 18 limited the power supply from the generator at bus 14 to the heavy load area in zone 4 and zone 5. The resulting system frequency decline, therefore, quickly triggered the load shedding corrective actions. Noticeably, this unreliability which is significant cannot be captured by the steady-state assessment and therefore it is essential to incorporate DSA in system reliability evaluation.

Table 5.8 Pre-fault and post-fault condition of outage components

No.	Outage component	Rating (MVA)	Pre-fault condition	Fault at time
1	Gen6 on bus 24	4500	1849.3 MW generation	1 s
2	Gen8 on bus 26	1200	405.9 MW generation	1 s
3	Wind farm on bus 808	33.4	9.80 MW generation	1 s
4	Wind farm on bus 3404	23.4	6.90 MW generation	1 s
5	Wind farm on bus 3405	23.4	6.90 MW generation	1 s
6	Line from bus 13 to bus 18	1500	488.6 MW flow	1 s

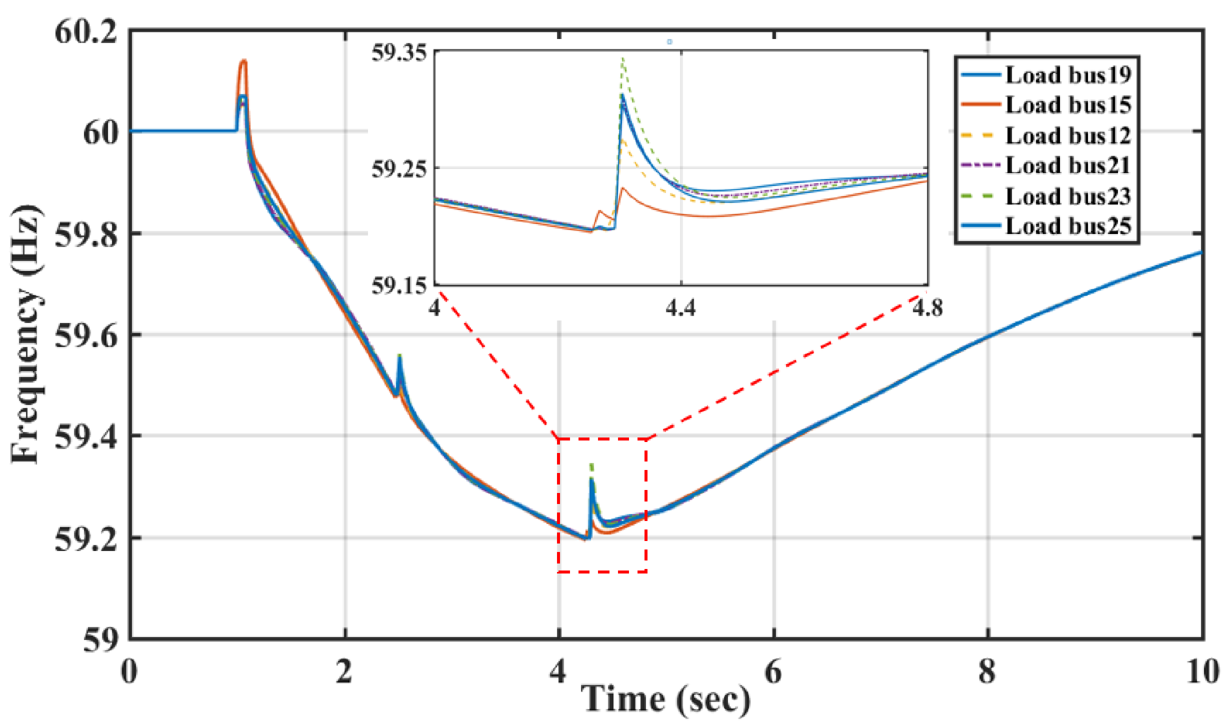

Figure 5.14 Load bus frequency under contingency

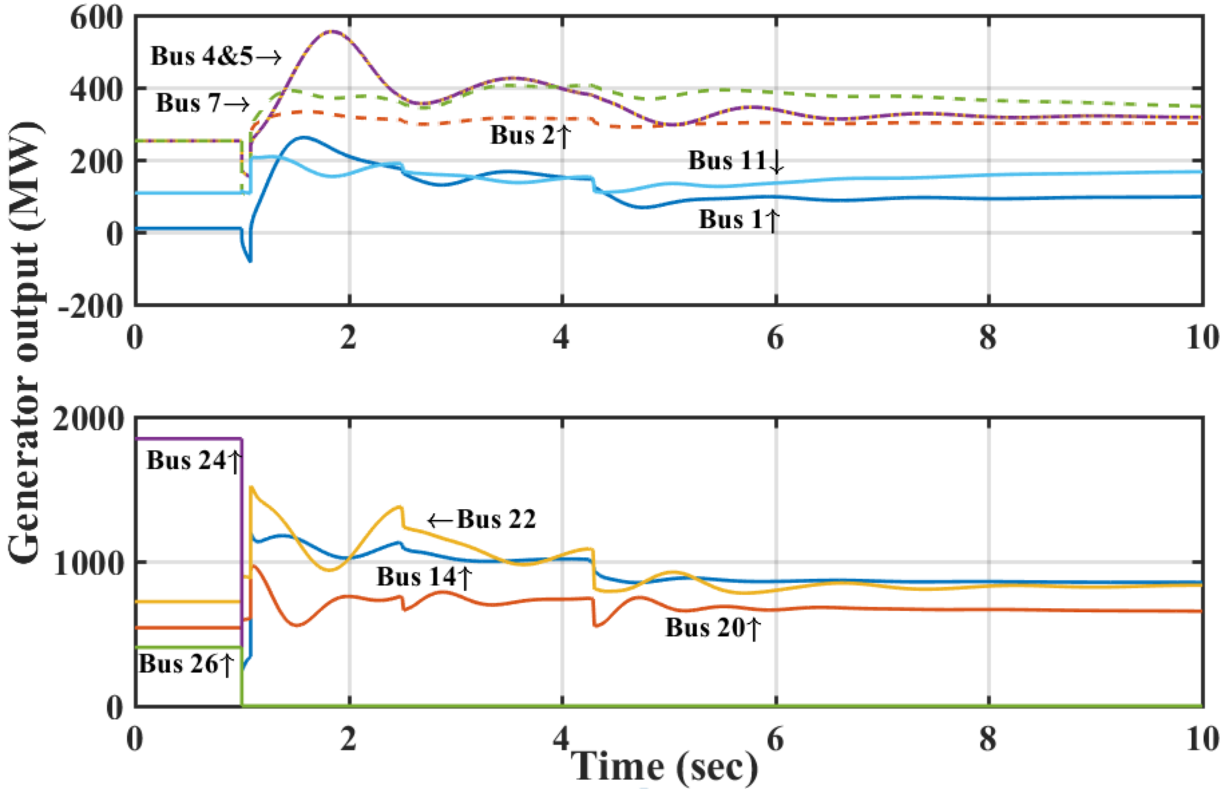

Figure 5.15 Generator output under contingency

Table 5.9 Post-contingency protection action report

Time (s)	Switching Action	Details
2.4835	Stage 1 tripped	Load at bus 15
2.4960	Stage 1 tripped	Load at bus 12
2.5044	Stage 1 tripped	Load at bus 25
2.5127	Stage 1 tripped	Load at bus 23
2.5169	Stage 1 tripped	Load at bus 19
2.5169	Stage 1 tripped	Load at bus 21
4.2629	Stage 2 tripped	Load at bus 15
4.2875	Stage 2 tripped	Load at bus 12
4.3004	Stage 2 tripped	Load at bus 23
4.3045	Stage 2 tripped	Load at bus 19
4.3045	Stage 2 tripped	Load at bus 21
Total Load shedding:		1554.4 MW

5.6 Transient Stability Pruning Effects

Previous simulation results show the reliability indices obtained from the traditional SMCS method and the CE IS based SMCS. Both methods reached a similar result. This section shows the simulation results with the accelerating process applied. This section shows the evaluation results of two cases to illustrate the effect of the pruning process. One case is a reliability evaluation with the pruning process and the other is with no pruning process applied. Different load shed values *ls_TDS* are used to determine whether a system is transiently unstable and should be considered in the reliability indices calculation. If the load curtailment of a state from the TDS is higher than *ls_TDS*, then the state is a transiently unstable state and the amount of load curtailment, as well as state duration are introduced in the index's calculation.

For a chosen iteration, the pruning process eliminated 3842 states among 11992 states from the detailed TDS analysis, with a 32.04% speed up of the DSA. The criteria for the two stability estimation metrics are KE_{cr}=0.5×10^{-5}pu and ΔZ_{th}=0.005 pu obtained by conducting a sensitivity study. Table 5.10 gives the comparison of LOLP results from the evaluation process with and without the TDS pruning process. The LS_{TDS} in Table 5.10 represents the criteria for determining dynamically insecure cases. When the load shedding results from TDS are larger than LS_{TDS}, the state is considered to be dynamically insecure. Simulation results show that the reliability evaluation with the pruning process gives similar results compared with no pruning process be applied. The deviations vary from 2.4494%-4.9645%. Additionally, from the sensitivity study of the different load shed threshold value, it can be seen that when LS_{TDS} is equal to 20 MW, 100 MW, and 200 MW, the LOLP results are very close to each other. Thus, we can choose any of the three values to be the load shed threshold to determine whether a state is transiently unstable from TDS.

Table 5.10 Comparison of LOLP indices results: with and without pruning process

LS_{TDS} criteria (MW)	**Without Pruning Process**	**With Pruning Process**	**Deviation (%)**
20	0.0939	0.0916	2.4494
100	0.0939	0.0916	2.4494
200	0.0936	0.0913	2.4573
400	0.0753	0.0721	4.2497
500	0.0564	0.0536	4.9645
600	0.0388	0.0375	3.3505
700	0.0269	0.0259	3.3457
800	0.0264	0.0257	2.2727
Average deviation:			3.1924

6. Conclusions and Future Work

6.1 Conclusions

This research proposes a probabilistic reliability evaluation method with resource adequacy and DSA integrated in a single framework. Sequential MCS was used because it provides a flexible approach for considering time-variant stochastic characteristics in the system. Stochastic characteristics have been considered in this work and include components outages, different transmission fault types, chronological load variance, and stochastic wind power output. Compared to the traditional reliability evaluation which evaluates adequacy and dynamic security separately, the proposed method provides the reliability indices reflecting both adequacy and dynamic security based on the quantification of the two aspects of reliability in terms of load curtailments.

The quantification of the impact of dynamic security is included by the load curtailment from protection action to maintain system stability after contingencies. By including this value of load curtailment into the calculation of reliability indices, the integrated system reliability can be represented using the well-recognized reliability indices which are LOLP, EPNS, and LOLF. The proposed method is tested on a synthetic test system and the results show the importance of considering the two reliability aspects together since both the steady-state and transient system performance need to be analyzed in reliability studies. Also, the computational effort in the evaluation process is significant because of the sequential MCS and the TDS for dynamic security assessment. Two acceleration methods are introduced to lighten the computational burden. In practical applications, the computational time could be substantially reduced further by using parallel or distributed computing as the SMCS is amenable to such implementations.

The work done in this project presents an approach for integrating adequacy and dynamic security assessment into a single framework. Dynamic security is of significant importance for power system reliability evaluation, the case studies show that when dynamic security is not considered in the reliability evaluation, many unstable system conditions are overlooked in the planning phase. The integrated reliability evaluation results can provide as a decision-making support to identify system conditions with inadequate resource or inadequate stabilization capability. An obstacle to integrating adequacy and security in reliability evaluation is the overwhelming number of cases that need to be assessed. The effect of the proposed Cross-Entropy based Importance Sampling method developed to speed up the convergence process of the MCS and the pruning process developed to reduce the computational burden of the transient stability assessment have been verified in the case studies. The main conclusions of the study are as follows:

1. The proposed reliability evaluation approach provides an effective method of integrating adequacy and dynamic security into a single framework. Stochastic and time-variant characteristics in the system can easily be considered in the evaluation using SMCS.

2. The dynamic security of a system state is quantified by the amount of load shed that is needed to keep the system stable during the transition. By introducing this value of load shed into the calculation of reliability indices, the overall system reliability can be represented using the well-recognized reliability indices which are LOLP, EPNS, and LOLF.

3. The proposed CE-IS method greatly speeds up the convergence process of the MCS and the pruning process considerably reduces the number of cases needing to be evaluated by TDS. The two acceleration methods were found to be accurate in the sense that they do not introduce bias into the calculation of the reliability indices.

6.2 Future work

- Computational efficiency: The large computational burden of the reliability evaluation is due to a large number of system states that are needed to reach an expected value of reliability indices. Although the approach proposed in this work is not targeted on real-time evaluation, the computational efficiency is expected to be improved. Since each iteration of the MCS is independent of each other, parallel computing techniques and multiple CPU cores to execute reliability evaluation can be applied in this work.
- Stochastic relay performance: In this work, it is assumed that the faults on transmission lines are cleared 5 cycles after the relay is tripped. However, in a practical scenario there is some uncertainty associated with the fault clearing time due to relay mis-operation which can be incorporated into the simulations. Also, if historical data on relay failures are available, the probability of relay failure can be incorporated into the dynamic security assessment.

Appendix A. Generator Dynamic Model Data

A.1 Synchronous Generator Dynamic Data

Generator, governor, and exciter are modeled for a synchronous generator in this project. Steam turbine, gas turbine, and hydro turbine are modeled with different governors that are represented by TGOV1, GGOV1 and HYGOV respectively in PSLF. The parameters of governor model are given in Table A.1. The parameters of synchronous generator model (GENROU) and exciter model are given in Table A.2.

Table A.1 Synchronous generator governor dynamic data

TGOV1									
R	**T1**	**Vmax**	**Vmin**	**T2**	**T3**	**Dt**	**/**		
0.05	0.5	1.0	0.0	3.0	10.0	0.0	/		
GGOV1									
r	**rselect**	**Tpelec**	**Maxerr**	**Miner**	**Kpgov**	**Kigov**	**Kdgov**	**Tdgov**	**vmax**
0.04	1.0	1.0	0.05	-0.05	10.0	2.0	0.0	1.0	1.0
vmin	**Tact**	**Kturb**	**wfnl**	**Tb**	**Tc**	**Flag**	**Teng**	**Tfload**	**Kpload**
0.15	0.5	1.5	0.2	0.1	0.0	1.0	0.0	3.0	2.0
Kiload	**Ldref**	**Dm**	**ropen**	**rclose**	**Kimw**	**Pmwset**	**/**		
0.67	1.0	0.0	0.10	-0.1	0.002	80.0	/		
HYGOV									
Rperm	**rtemp**	**Tr**	**Tf**	**Tg**	**Velm**	**Gmax**	**Gmin**	**Tw**	**At**
0.04	0.3	5.0	0.05	0.5	0.2	1.0	0.0	1.0	1.2
Dturb	**qn1**	**ttrip**	**tn**	**tnp**	**db1**	**eps**	**db2**	**GV0**	**Pgv0**
0.5	0.08	0	0	0	0	0	0	0	0
GV1	**Pgv1**	**GV2**	**Pgv2**	**GV3**	**Pgv3**	**GV4**	**Pgv4**	**GV5**	**Pgv5**
0	0	0	0	0	0	0	0	0	0
hdam	**Bgv0**	**Bgv1**	**Bgv2**	**Bgv3**	**Bgv4**	**Bgv5**	**bmax**	**tblade**	**/**
1.0	0	0	0	0	0	0	0	100	/

Table A.2 Synchronous generator and exciter dynamic data

GENROU								
Tpdo	**Tppdo**	**Tpqo**	**Tppqo**	**H**[1]	**D**	**Ld**	**Lq**	**Lpd**
7.0	0.025	0.75	0.05	6.0	0.0	2.2	2.1	0.22
Lpq	**Lppd**	**L1**	**S1**	**S12**	**Ra**	**Rcomp**	**Xcomp**	**/**
0.416	0.2	0.147	0.109	0.3	0.0	0.0	0.0	/
EXST1								
Tr	**Vimax**	**Vimin**	**Tc**	**Tb**	**Ka**	**Ta**	**Vrmax**	**Vrmin**
0.0	0.1	-0.1	1.0	10.0	100.0	0.02	5.0	-5.0
Kc	**Kf**	**Tf**	**Tc1**	**Tb1**	**Vamax**	**Vamin**	**Xe**	**Ilr**
0.05	0.0	1.0	1.0	1.0	5.0	-5.0	0.04	2.8
Klr	**/**							
5.0	/							

A.2 Wind Turbine Dynamic Data

The DFAG WTG are modeled using GEWTG, WNDTGE, EXWTGE in GE PSLF. The dynamic data used in this paper are provided in Table A.3 - Table A.5.

Table A.3 Wind turbine generator/converter model

lpp	**dVtrp1**	**dVtrp2**	**dVtrp3**	**dVtrp4**	**dVtrp5**	**dVtrp6**	**dTtrp1**	**dTtrp2**
0.8	-0.25	-0.5	-0.7	-0.85	0.1	0.15	1.9	1.2
dTtrp3	**dTtrp4**	**dTtrp5**	**dTtrp6**	**fcflg**	**rrpwr**	**brkpt**	**zerox**	**/**
0.7	0.2	1	0.1	0	10	0.9	0.5	

[1] For hydro turbine governor, H is 5.91 s; for steam turbine governor and gas turbine governor, H is 6, 9.7 or 14 depends on the MVA rating of a unit.

Table A.4 Wind turbine electrical control model

spdw1	**tp**	**tpc**	**kpp**	**kip**	**kptrq**	**kitrq**	**kpc**	**kic**	**Pimax**	**Pimin**
0	0.3	0.05	150	25	3	0.6	3	30	27	0
Pwmax	**Pwmin**	**Pwrat**	**H**	**nmass**	**Hg**	**Ktg**	**Dtg**	**Wbase**	**Tw**	**Apcflg**
1.12	0.04	0.45	4.94	1	0	0	0	0	1	1
Pa	**Pbc**	**Pd**	**Fa**	**Fb**	**Fc**	**Fd**	**Pmax**	**Pmin**	**Kwi**	**dbwi**
1	0.95	0.4	0.96	0.996	1.004	1.04	1	0.2	10	0.0025
Twowi	**urlwi**	**drlwi**	**Pmxwi**	**Pmnwi**	**wfflg**	**Td1**	**Tpset**	**Pirat**	**Tpav**	**Tlpwi**
5.5	0.1	-1	0.1	0	0	0.15	5	10	0.15	1

Table A.5 Wind turbine converter control model

varflg	**Kqi**	**Kvi**	**Vmax**	**Vmin**	**qmax**	**qmin**	**xiqmax**	**xiqmin**	**tr**	**tc**	**kpv**
-1	0.1	40	1.1	0.9	0.436	-0.436	1.45	0.5	0.02	0.15	18
kiv	**vl1**	**vh1**	**tl1**	**tl2**	**th1**	**th2**	**ql1**	**ql2**	**ql3**	**qh1**	**qh2**
5	-9999	9999	0	0	0	0	0	0	0	0	0
qh3	**pfaflg**	**fn**	**tv**	**tpwr**	**ipmax**	**xc**	**kqd**	**tlpqd**	**xqd**	**vermn**	**vfrz**
0	0	1	0.05	0.05	1.22	0	0	5	0	0.1	0.7

References

[1] NERC Transmission Planning Standards, TPL-002-0b, 2014.

[2] NERC Glossary of Terms Used in NERC Reliability Standards, 2018.

[3] Reliability Assessment Guidebook, North American Electric Reliability Corporation, ver. 3.1, Aug. 2012.

[4] P. Kundur, Power System Stability and Control. New York: McGraw Hill Inc., 1994.

[5] F. F. Wu, Yu-Kun Tsai and Y. X. Yu, "Probabilistic steady-state and dynamic security assessment," IEEE Transactions on Power Systems, vol. 3, no. 1, pp. 1-9, Feb. 1988.

[6] A.M. Leite da Silva, J. Endrenyi and L. Wang, "Integrated treatment of adequacy and security in bulk power system reliability evaluations," IEEE Transactions on Applied Superconductivity, vol. 3, no. 1, pp. 275-285, Mar. 1993.

[7] G.M. Huang, Yishan Li, "Power system reliability indices to measure impacts caused by transient stability crises," IEEE Power Engineering Society Winter Meeting, New York, Jan. 27, 2002 – Jan. 31, 2002.

[8] N. Amjady, "A framework of reliability assessment with consideration effect of transient and voltage stabilities," IEEE Transactions on Power Systems, vol. 19, no. 2, pp. 1005-1014, May. 2004.

[9] A. M. Rei, A. M. Leite da Silva, J. L. Jardim, J. Carlos de Oliverira Mello, "Static and dynamic aspects in bulk power system reliability evaluations," IEEE Transactions on Power Systems, vol. 15, no. 1, pp. 189-195, Feb. 2000.

[10] M. Benidris, J. Mitra, C. Singh, "Integrated evaluation of reliability and stability of power systems," IEEE Transactions on Power Systems, vol. 32, no. 5, pp. 4131-4139, Sep. 2017.

[11] O. Gerald Ibe, I. Kelechi, "Adequacy analysis and security reliability evaluation of bulk power system," IOSR Journal of Computer Engineering, vol. 11, no. 2, pp. 26-35, May, 2013.

[12] R. Allan, "Power system reliability assessment—a conceptual and historical review," Reliability Engineering and System Safety, vol. 46, no. 1, pp. 3-13, 2013.

[13] R. Billinton, R. N. Allan, *Reliability Evaluation of Power Systems*. New York: Springer., 1996.

[14] R. Billinton, W. Li, *Reliability Assessment of Electric Power Systems Using Monte Carlo Methods*. New York: Springer., 1994.

[15] W. Li, *Risk Assessment of Power Systems: Models, Methods and Applications*. Wiley, 2004.

[16] R. Billinton, W. Li, "Hybrid approach for reliability evaluation of composite generation and transmission systems using Monte-Carlo simulation and enumeration technique", IEE Proceedings C – Generation, Transmission and Distribution, vol. 138, no. 3, pp. 233-241, May, 1991.

[17] N. Gubbala, C. Singh, "Models and considerations for parallel implementation of Monte Carlo simulation methods for power system reliability evaluation," IEEE Transactions on Power Systems, vol. 10, no. 2, pp. 779-787, May. 1995.

[18] C. L. T. Borges, D. M. Falcao, J. Carlos O. Mello, A. C. G. Melo, "Composite reliability evaluation by sequential monte carlo simulation on parallel and distributed processing environments," IEEE Transactions on Power Systems, vol. 16, no. 2, pp. 203-209, May. 2001.

[19] M. A. da Rosa, V. Miranda, L.Carvalho, A. M. Leite da Silva, "Modern computing environment for power system reliability assessment," IEEE 11th International Conference on Probabilistic Methods Applied to Power Systems, Singapore, Jun. 14, 2010 – Jun. 17, 2010.
[20] M. A. da Rosa, A. M. Leite daSilva, V. Miranda, "Multi-agent systems applied to reliability assessment of power systems ," Electrical Power and Energy Systems, vol. 42, no. 1, pp. 367-374, 2012.
[21] A. M. Leite daSilva, L. Chaves de Resende, L. A. da Fonseca Manso, V. Miranda, " Composite reliability assessment based on monte carlo simulation and artificial neural networks," IEEE Transactions on Power Systems, vol. 2, no. 3, pp. 1202-1209, Aug. 2007.
[22] J. A. Dias, C. L. Tancredo Borges, "Object oriented model for composite reliability evaluation including time varying load and wind generation," IEEE 11th International Conference on Probabilistic Methods Applied to Power Systems, Singapore, Jun. 14, 2010 – Jun. 17, 2010.
[23] A. Sankarakrishnan, R. Billinton, "Sequential monte carlo simulation for composite power system reliability analysis with time varying load," IEEE Transactions on Power Systems, vol. 10, no. 3, pp. 1540-1545, Aug. 1995.
[24] S. R. Huang, S. L. Chen, "Evaluation and improvement of variance reduction in monte carlo production simulation," IEEE Transactions on Energy Conversion, vol. 8, no. 4, pp. 610-620, Dec. 1993.
[25] A. Jonnavithula, R. Billinton, "Composite system adequacy assessment using sequential monte carlo simulation with variance reduction techniques," IEE Proceedings - Generation, Transmission and Distribution, vol. 144, no. 1, pp. 1-6, Jan. 1997.
[26] G. C. Oliveira, M. V. F. Pereira, S. H. F. Cunha, "A technique for reducing computational effort in monte carlo based composite reliability evaluation," IEEE Transactions on Power Systems, vol. 4, no. 4, pp. 1390-1315, Nov. 1989.
[27] M.V.F. Pereira , M.E.P. Maceira , G.C. Oliveira ,L.M.V.G. Pinto, "Combining analytical models and Monte-Carlo techniques in probabilistic power system analysis," IEEE Transactions on Power Systems, vol. 7, no. 1, pp. 265-272, Feb. 1992.
[28] S. H. F. Cunha ; M. V. F. Pereira ; L. M. V. G. Pinto ; G. C. Oliveira, "Composite generation and transmission reliability evaluation in large hydroelectric systems," IEEE Transactions on Power Apparatus and Systems, vol. PAS-104, no. 10, pp. 2657-2663, Oct. 1985.
[29] A. M. Leite da Silva, R. A. G. Fernandez, C. Singh, "Generating capacity reliability evaluation based on monte carlo simulation and cross entropy methods," IEEE Transactions on Power Systems, vol. 25, no. 1, pp. 129-137, Feb. 2010.
[30] R. A. Gonzalez-Fernandez, A. M. Leite da Silva, "Reliability assessment of time-dependent systems via sequential cross entropy monte carlo simulation," IEEE Transactions on Power Systems, vol. 26, no. 4, pp. 2381-2389, Nov. 2011.
[31] R. A. Gonzalez-Fernandez, A. M. Leite da Silva, L. C. Resende, M. T. Schilling, "Composite systems reliability evaluation based on monte carlo simulation and cross entropy methods," IEEE Transactions on Power Systems, vol. 28, no. 4, pp. 4598-4606, Nov. 2013.
[32] G. L. LUCARELLI, C. YAN, Z. BIE, C. Wang, T. Ding, F. Xiong, "Three approaches to use cross entropy in reliability evaluation of composite system with high penetration

wind energy," IEEE ASIA-Pacific Power and Energy Conference, Xi'an, Oct. 25- Oct. 28, 2016.
[33] P. Kundur, J. Paserba, V. Ajjarapu, G. Andersson, A. Bose, C. Canizares, N. Hatziargyriou, D. Hill, A. Stankovic, C. Taylor, T. Van Cutsem, V. Vittal, "Definition and Classification of Power System Stability," IEEE Transactions on Power Systems, vol. 19, no. 2, pp. 1387-1401, May. 2004.
[34] N. Maruejouls, V. Sermanson, S.T. Lee, P. Zhang, "A practical probabilistic reliability assessment using contingency simulation," IEEE PES Power Systems Conference and Expostition, Oct. 10 – Oct. 13, 2004.
[35] D.S. Kirschen, D. Jayaweera, D.P. Nedic, R.N. Allan, " A probabilistic indicator of system stress," IEEE Transactions on Power Systems, vol. 19, no. 3, pp. 1650-1657, Aug. 2004.
[36] E. Vaahedi, W. Li, T. Chia, H. Dommel, "Large scale probabilistic transient stability assessment using BC Hydro's on-line tool," IEEE Transactions on Power Systems, vol. 15, no. 2, pp. 661-667, May. 2000.
[37] R. Billinton, P.R.S, Kuruganty, M.F. Carvalho, "An Approximate Method for Probabilistic Assessment of Transient Stability," IEEE Transactions on Reliability, vol. R-28, no. 3, pp. 255-258, Aug. 1979.
[38] J.D. McCalley, A.A. Fouad, V. Vittal, A.A. Irizarry-Rivera, B.L. Agrawal, R.G. Farmer, "A risk-based security index for determining operating limits in stability-limited electric power systems," IEEE Transactions on Power Systems, vol. 12, no. 3, pp. 1210-1219, Aug. 1997.
[39] V. Van Acker, J.D. McCalley, V. Vittal, J.A. Pecas Lopes, "Risk-based transient stability assessment," PowerTech Budapest 99, Aug. 29 – Sep. 2, 1999.
[40] M. Ni, J.D. McCalley, V. Vittal, T. Tayyib, "Online risk-based security assessment," IEEE Transactions on Power Systems, vol. 18, no. 1, pp. 258-265, Feb. 2003.
[41] J. D. McCalley, V. Vittal, M. Ni, "On-line risk-based security assessment," EPRI Final Report TR-113276, July 1999.
[42] J. D. McCalley, V. Vittal, Y. Dai, W. Fu, A. Irizarry-Rivera, V. Van Acker, H. Wan and S. Zhao, "Risk based security assessment," EPRI Final Report TR-113276, July 1999.
[43] R. Billinton, P. R. S. Kuruganty, "A Probabilistic Index for Transient Stability," IEEE Transactions on Power Apparatus and Systems, vol. PAS-99, no. 1, pp. 195-206, Jan. 1980.
[44] R. Billinton, P. R. S. Kuruganty, "Probabilistic assessment of transient stability in a practical multimachine system," IEEE Transactions on Power Apparatus and Systems, vol. PAS-100, no. 7, pp. 3634-3641, Jul. 1981.
[45] S. Aboreshaid, R. Billinton, M. Fotuhi-Firuzabad, "Probabilistic transient stability studies using the method of bisection," IEEE Transactions on Power Systems, vol. 11, no. 4, pp. 1990-1995, Jul. 1981.
[46] R. Billinton, S. Aboreshaid, "Stochastic modelling of high-speed reclosing in probabilistic transient stability studies," IEE Proceedings – Generation, Transmission and Distribution, vol. 142, no. 4, pp. 350-354, Jul. 1995.
[47] D.Z. Fang, L. Jing, T.S. Chung, "Corrected transient energy function-based strategy for stability probability assessment of power systems," IET Generation, Transmission and Distribution, vol. 2, no. 3, pp. 424-432, May. 2008.

[48] A. Dissanayaka, U. D. Annakkage, B. Jayasekara, B. Bagen, "Risk-based dynamic security assessment," IEEE Transactions on Power Systems, vol. 26, no. 3, pp. 1302-1308, Aug. 2011.
[49] P. M. Anderson, A. Bose, "A probabilistic approach to power system stability analysis," IEEE Transactions on Power Apparatus and Systems, vol. PAS-102, no. 8, pp. 2430-2439, Aug. 1983.
[50] K. J. Timko, A. Bose, P. M. Anderson, "Monte carlo simulation of power system stability," IEEE Transactions on Power Apparatus and Systems, vol. PAS-102, no. 10, pp. 3453-3459, Oct. 1983.
[51] R. Billinton, P.R.S. Kuruganty, M.F. Carvalho, "An Approximate Method for Probabilistic Assessment of Transient Stability," IEEE Transactions on reliability, vol. R-28, no. 3, pp. 255-258, Aug. 1979.
[52] R. Billinton, P.R.S. Kuruganty, M.F. Carvalho, "Risk assessment of power system models method and application," IEEE Transactions on reliability, vol. R-28, no. 3, pp. 255-258, Aug. 1979.
[53] K. Clark, N. Miller, and J. Sanchez-Gasca, "Modeling of GE wind turbine-generators for grid studies," GE Energy, Version 4.5.
[54] "Generic type-3 wind turbine-generator model for grid studies," Version 1.1, WECC Wind Generator Modeling Group, Sept. 2006.
[55] I. A. Hiskens, "Dynamics of type-3 wind turbine generator models," IEEE Transactions on Power Systems, vol. 27, no. 1, pp. 465-474, Feb. 2012.
[56] General Electric International, Inc., GE PSLF User's Manual, v19.0_02, Schenectady,USA, 2016.
[57] The Department of Energy (DOE), "The Quadrennial Energy Review: Energy Transmission, Storage, and Distribution Infrastructure," April 2015.
[58] NERC Protection and Control Standards, PRC-006-1, Atlanta, USA, 2013.
[59] R. Y. Rubinstein and D. P. Kroese, The Cross-Entropy Method. A Unified Approach to Combinatorial Optimization, Monte-Carlo Simulation, and Machine Learning. New York, NY, USA: Springer, 2004.
[60] R. Y. Rubinstein and D. P. Kroese, Simulation and the Monte Carlo Methods, 2nd ed. New York, NY, USA: Wiley, 2007.
[61] A.A. Fouad, V. Vittal, Power System Transient Stability Analysis Using the Transient Energy Function Method, New Jersey, Prentice-Hall inc., 1992.
[62] Canadian Electricity Association, "Forced outage performance of transmission equipment," Annual Report on Equipment Reliability Information System, 2012.
[63] Jager, D.; Andreas, A.; (1996). NREL National Wind Technology Center (NWTC): M2 Tower; Boulder, Colorado (Data); NREL Report No. DA-5500-56489. http://dx.doi.org/10.5439/1052222
[64] S. Datta, V.Vittal, "Operational Risk Metric for Dynamic Security Assessment of Renewable Generation", IEEE Transactions on Power Systems, vol. 32, no. 2, pp. 1389-1399, Mar. 2017.
[65] P. Mitra, V. Vittal, B. Keel, J. Mistry, "A systematic approach to n-1-1 analysis for power system security assessment," IEEE Power and Energy Technology Systems Journal, vol. 3, no. 2, pp. 71-80, June. 2016.

Part II

Deep Learning and Multi-Label Learning Based MCS Methods for Composite System Reliability Evaluation

Chanan Singh
Dogan Urgun, Graduate Student

Texas A&M University

For information about this project, contact

Chanan Singh, Regents Professor and Irma Runyon Chair Professor
Department of Electrical and Computer Engineering
Texas A&M University
3128 TAMU
College Station, Texas 77843-3128
Phone: (979) 845-7589
Email: singh@ece.tamu.edu

Power Systems Engineering Research Center

The Power Systems Engineering Research Center (PSERC) is a multi-university Center conducting research on challenges facing the electric power industry and educating the next generation of power engineers. More information about PSERC can be found at the Center's website: http://www.pserc.org.

For additional information, contact:

Power Systems Engineering Research Center
Arizona State University
527 Engineering Research Center
Tempe, Arizona 85287-5706
Phone: 480-965-1643
Fax: 480-727-2052

Notice Concerning Copyright Material

Table of Contents

List of Figures

List of Tables

Glossary

AI	Artificial Intelligence
ANN	Artificial Neural Networks
CE	Cross Entropy
CMCS	Crude Monte Carlo Simulation
CNN	Convolutional Neural Network
COV	Coefficient of Variation
IS	Importance Sampling
MAP	Maximum a Posteriori
MCS	Monte Carlo Simulation
MLL	Multi Label Learning
MLKNN	Multi Label K-Nearest Neighbor
MLRBF	Multi Label Radial Basis Function
LHS	Latin Hypercube Sampling
LOL	Loss of Load
LOLP	Loss of Load Probability
KNN	K- Nearest Neighbor
OPF	Optimal Power Flow
PDF	Probability density function
RELU	Rectified Linear Unit
RBF	Radial Basis Function
RTS	Reliability Test System

1. Introduction

Tremendous progress has been made on developing probabilistic methods for power system reliability evaluation over the past several decades. In most cases, these methods can be grouped into two categories, analytical solution methods and simulation-based methods. In analytical methods the system is represented by mathematical models and reliability indices computed using mathematical solutions. Even though these methods give exact solutions within assumptions made, deriving these models can become a challenging problem especially for large power systems.

Among the many simulation methods developed (including Importance Sampling or Latin Hypercube Sampling), MCS based techniques are currently the most widely used methods to estimate the reliability indices of composite power systems. MCS methods sample system states with the basic concept that their occurrence is proportional to their probabilities. For most cases, MCS is more suitable for composite system analysis because of its simplicity and flexibility in estimating complex system parameters in various conditions [1, 2]. Despite the advantages of MCS, it requires solving optimization equations to perform optimal power flow (OPF) analysis for characterization of each sampled state and repetitive states most of the time. Therefore, MCS suffers from long computation time to produce statistically converged reliability indices. This indicates a need of research for efficient simulation methods in reliability analysis of large power systems.

Considerable amount of research has been done on increasing computational efficiency of these simulations in the past few decades. Some of these approaches use variance reduction techniques [3], state space pruning [4], fuzzy optimal power flow [5] or more efficient sampling techniques like LHS [1] or IS [6]. Some of these researches also use population-based intelligent search (PIS) methods as an alternative to search for meaningful states to decrease the computational burden of these simulation methods. Some of classical examples of these methods are genetic algorithms (GA) [7,8], particle swarm optimization (PSO) [9] or ant colony optimization (ACO) [10]. Some of the researchers also implement pattern classification techniques to reduce the number of states to be evaluated in power system reliability assessment. Some examples of these methods are Artificial Neural Network (ANN) based classifiers [11], Artificial Immune Recognition System (AIR) [12] or Least Squares Support Vector Machine based classifiers [13]. Pattern classification-based techniques have shown significant performance to reduce the computational burden required for reliability analysis, however, still more research is needed for increasing classification accuracy and model flexibility.

The remainder of this part of the report is organized as follows: Section 2 provides background information about composite system adequacy analysis and MCS techniques. Section 3 investigates different types of deep learning structures and demonstrates performance of those algorithms on case studies. Section 4 first gives a background information about multilabel learning then performance of this type of learning for composite system reliability evaluation is explored through case studies. Section 5 focuses on combination of multilabel learning and importance sampling combination within MCS. After briefly explaining the theory behind importance sampling, performance results for proposed method is presented in this section. The conclusions and outlook are given in Section 6. References are listed at the end.

2. Concept of Composite Power System Reliability Evaluation and Monte Carlo Simulation

2.1 Introduction

There are two main categories of power system reliability evaluation techniques, analytical and simulation. In analytical modeling method, a model is built that reasonably approximates the physical system and is also amenable to calculations. Monte Carlo Simulation methods, on the other hand, are based on sampling and estimating the indices from the samples. MCS based techniques are able to handle any type of probability distribution associated with component state durations, capture systematic and temporal dependencies, and evaluate probability distributions of resultant indices. In general, they provide more flexibility to incorporate complex operating conditions in assessing especially large and complex power systems compared to analytical solutions [14,15]. In this section first MCS is introduced later power system models, test systems and metrics used to measure performance of AI based power system reliability evaluation methods are described.

2.2 Monte Carlo Simulation

Monte Carlo simulation is a representative simulation method that is usually adopted to deal with reliability evaluation of large-scale or complex power systems. MCS methods are classified into two categories, non-sequential simulation and sequential simulation. Non-sequential simulation is based on a random sampling algorithm, where a component state is selected according to its probability distribution without considering chronological connection. By using this approach, reliability indices such as loss of load probability (LOLP) and expected unserved energy (EUE) can be directly estimated. Assuming reliability coherence, the indices of frequency and duration can be estimated through a conditional probability approach [16,17], or calculated directly from sampled failure states using the frequency balance property [18].
The main handicap of non-sequential approach emerges when an event chronology is required to reflect the inherent variability of reliability estimations or to incorporate time-varying characteristics. Sequential Monte Carlo Simulation approach becomes more suitable [19] in these circumstances. As an example, if aging factor is considered as a practical issue in reliability, then component failure rates that increase with time are naturally incorporated by sequential simulation [20]. Sequential simulation steps through system states in time domain, where a state of each component is chronologically connected to its preceding and succeeding states. A realistic history is created by combining sequences of component state durations and system load over a given time horizon. In this manner, using sequential simulation LOLP, EUE, LOLF or LOLD indices are calculated more simply and accurately. Also, economic indices such as loss of load cost (LOLC) can be estimated more accurately [21]. Compared to non-sequential approach, sequential simulation provides simplicity of accurately incorporating time-dependent variables and their

correlations, however, this approach requires considerably more computing time to converge than non-sequential simulation. In non-sequential simulation, any two sampled system states can be completely independent, on the other hand, in sequential simulation, any two consecutive system states differ by a realization of one random variable. As a result, the overall state space is less represented by sequential simulation than by non-sequential simulation considering the same number of sampled states. Therefore, sequential simulation would require a larger number of states to reach the same convergence criterion. This problem is especially critical for composite systems where their state evaluation involves analysis of power flow and optimization-based remedial action.

2.2.1 Non-Sequential Monte Carlo Simulation

In non-sequential MCS approach system states are randomly sampled from the state space. In the following, non-sequential MCS is described in three main steps.

1.) Select a state of the power system by random sampling of the states of all components and the load levels.
2.) Characterize the selected state as success or failure through a test function, by performing an adequacy analysis, which usually involves optimal power flow (OPF).
3.) Update the estimate $E(F)$, the expected value of the system reliability indices using the results obtained in step 2. $E(F)$ is described in (eq 2.1).

$$E(F) = \frac{1}{N}\sum_{i=1}^{N} F(x_i) \tag{2.1}$$

where N is the number of simulated states.

4.) If the stopping criterion is satisfied then stop the simulation, otherwise, return to step 1.

The estimate of uncertainty is usually represented by the coefficient of variation β. An acceptable value of the estimate of uncertainty is used as stopping criteria for the simulation. Besides variance, a determined specified number of samples can also be used as stopping criteria. Calculation of β is described in (2.2).

$$\beta = \frac{\sqrt{V\big(E(F)\big)}}{\big(E(F)\big)} \tag{2.2}$$

where $V\big(E(F)\big)$ is the variance of the estimate $E(F)$.

2.2.2 Sequential Monte Carlo Simulation

Sequential MCS is a type of MCS in which each system state is related to the previous set of system states. By doing this a sequential time evaluation of system behavior is created which enables evaluation of a wider range of reliability indices [22]. Sequential simulation can be generally implemented with two methods, fixed time interval method and next event method. Both methods are described in the following subsections.

2.2.2.1 Fixed Time Interval Method

In this method, sequence of time intervals is stepped through, where component states are selected according to their transition probabilities. Its steps are described as follows.

1.) Initialize component states with random sampling from their probabilities of being up or down.
2.) Sample for component states in the next transition using each component transition probability matrix in (eq 2.3), where $\Delta\tau$ is a chosen as small time step.

$$\begin{array}{cc} & \begin{array}{cc} Up & Down \end{array} \\ \begin{array}{c} Up \\ Down \end{array} & \begin{bmatrix} 1-\lambda\Delta\tau & \lambda\Delta\tau \\ \mu\Delta\tau & 1-\mu\Delta\tau \end{bmatrix} \end{array} \tag{2.3}$$

3.) Generate a load level for step $\Delta\tau$ from historical chronology.
4.) Evaluate current system state with contingency analysis. If no bus has loss of load then load curtailment is zero otherwise remedial action is called to find a load curtailment.
5.) Repeat Steps 2–4 while updating reliability indices. If convergence criterion is satisfied, stop the program.

The length of time step $\Delta\tau$ will affect simulation accuracy. A smaller step results in higher accuracy, but will require a larger number of states to be evaluated and thus result in higher computational cost. This issue imposes a computational limitation for fixed time interval method to be used in practice even though it is theoretically feasible.

2.2.2.2 Next Event Method

In this method, simulation proceeds by keeping a record of the time when the next event occurs, where the residence time of each component state is determined by the value of a random variable from its continuous distribution. Its steps are given as follows:

1.) Initialize component states with random sampling from their probabilities of being up or down.
2.) Generate the state (up or down) duration τ for each component$i$. Draw a pseudo-random number $z \sim U(0,1)$ and substitute it into the inverse transform of distribution function F_t in (eq 2.4).

$$\tau_i = F_{t_i}^{-1}(z) \tag{2.4}$$

3.) Update the associated load sequence in correspondence to component sequence.

4.) Evaluate each state of system sequence obtained in Steps 2–3 in the similar way as seen in Step 4 of Fixed time interval method.
5.) Repeat steps 2–4 while updating reliability indices. If convergence criterion is satisfied, stop the procedure.

2.3 Composite System Adequacy Analysis

MCS techniques are currently the most widely used methods to assess the adequacy analysis of a composite system [23]. The MCS is based on a combination of state sampling with direct approach for system analysis and the utilization of a minimization model for load curtailment. This method is especially well suited for large power systems and allows multi state representation of components as well. Minimization model for load curtailment is usually required to solve an optimization problem based on power flow equations. The power flow equations for analysis use either DC or AC power flow model. Following subsections presents the mathematical information required to perform load curtailment analysis based on DC or AC power follow model respectively.

2.3.1 DC Power Model

In composite system reliability studies, power flow analyses are usually carried out in solving optimization problems for minimum of load curtailment. There are usually two types of power flow analysis used to characterize a system state, DC and AC power flow analysis. The DC power flow model is described by the nodal equation

$$\beta\delta + G = D \tag{2.5}$$

and the line flow equation

$$bA\delta = F \tag{2.6}$$

where

$N_b = Number\ of\ Buses$

$N_t = Number\ of\ Transmission\ Lines$

$b = N_t x N_t\ primitive\ matrix\ of\ transmission\ linesusceptances$

$A = N_t x N_b\ \ element\ node\ incidence\ matrix$

$\beta = N_b x N_b\ augmented\ node\ susceptance\ matrix$

$\delta = N_b\ vector\ of\ bus\ voltage\ angles$

$G = N_b\ vector\ of\ bus\ generation\ levels$

$D = N_b\ vector\ of\ bus\ loads$

$F = N_t\ vector\ of\ transmission\ line\ flows$

Load curtailment can be found by solving following linear programming model

$$\text{Loss of Load} = \min\left(\sum_{i=1}^{N_b} C_i\right) \tag{2.7}$$

Subject to

$$\begin{gathered} \beta\delta + G + C = D \\ G \leq G^{max} \\ C \leq D \\ bA\delta \leq F^{max} \\ -bA\delta \leq F^{max} \\ G, C \geq 0 \\ \delta, unrestricted \end{gathered} \tag{2.8}$$

2.3.2 AC Power Flow Model

Following set of equations describes the formulation and incorporation of the objective function of minimum load curtailment in the non-linear programming problem. This objective function is subject to equality and inequality constraints of the power system operation limits. The equality constraints include the power balance at each bus and the inequality constraints are the capacity limits of generating units, power carrying capabilities of transmission lines, voltage limits at the nodes and reactive power capability limits. The minimization problem is formulated as follows [24]

$$\text{Loss of Load} = \min\left(\sum_{i=1}^{N_b} C_i\right) \tag{2.9}$$

Subject to

$$\begin{gathered} P(V,\delta) - P_D + C = 0 \\ Q(V,\delta) - Q_d + C_q = 0 \\ P_G^{min} \leq P(V,\delta) \leq P_G^{max} \\ Q_G^{min} \leq Q(V,\delta) \leq Q_G^{max} \\ V^{min} \leq V \leq V^{max} \\ S(V,\delta) \leq S^{max} \\ 0 \leq C \leq P_D \\ \delta, unrestricted \end{gathered} \tag{2.10}$$

In (2.9) and (2.10), C_i is the load curtailment at bus i, C is the vector of load curtailments, C_q is the vector of reactive load curtailments, V is the vector of bus voltage magnitudes, δ is the vector of bus voltage angles, P_D and Q_D are the vectors of real and reactive power loads, P_G^{min}, P_G^{max}, Q_G^{min} andQ_G^{max} are the vectors of real and reactive power limits of the generators, V^{min} and V^{max} are the vectors of minimum and maximum allowed voltage magnitudes, S (V, δ) is the vector of power flows in the lines, S^{max} is the vector of power rating limits of the transmission lines and P(V, δ) and Q(V, δ) are the vectors of real and reactive power injections. Moreover, N_b is the number of buses, $N_d$is the number of load buses, N_t is the number of transmission lines and N_g is the number of generators. In the standard minimization problem given by (2.9) and (2.10), all generation and network constraints have been taken into consideration. It has been assumed that one of the bus angles is zero in the constraints (2.10) to work as a reference bus.

2.4 Reliability Test Systems

IEEE 30 bus test system or IEEE RTS is used to demonstrate performance of proposed methods in this study. In this subsection, test systems are described.

2.4.1 IEEE 30 Bus Test System

There are 41 transmission lines in this system, with 435 MW maximum generation and 255 MW maximum load. There are 9 generation units for 6 generation buses in this case study. Since there is no reliability data associated with this system, corresponding parameters are chosen from RTS, for simplicity, all generators share the same failure rates and repair time and all transmission lines share the same failure rates and repair time. A detailed schematic of IEEE 30 Bus Test System is given in figure 1. Data of the original IEEE 30 Bus Test System can be accessed through the [25].

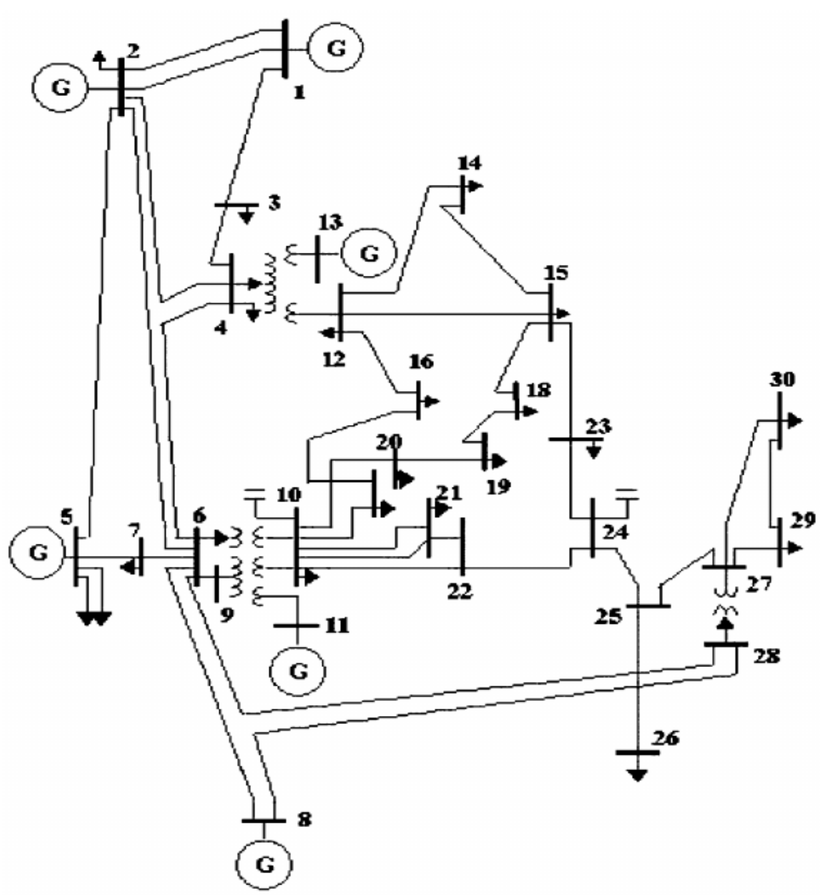

Figure 2.1: Schematic for IEEE 30 Bus Test System

2.4.2 IEEE RTS 79 Test System

This system has 24 buses (10 of them are generation buses), 38 transmission lines and 32 generation units. The total installed capacity is 3405 MW and the system has 2850 MW at its annual peak. A detailed schematic of RTS is given in figure 2. Data of the original IEEE RTS can be accessed through the [26].

For some of case studies Modified RTS (MRTS) is preferred. MRTS is designed for the studies on transmission line reliability for composite systems. In this system, generation capacities are doubled and all the loads are multiplied by 1.8 while rest of the system parameters remain unchanged. In this way effect of transmission lines on overall system reliability is increased and becomes more observable. The total installed capacity is 6810 MW and the system has 5130 MW at annual peak.

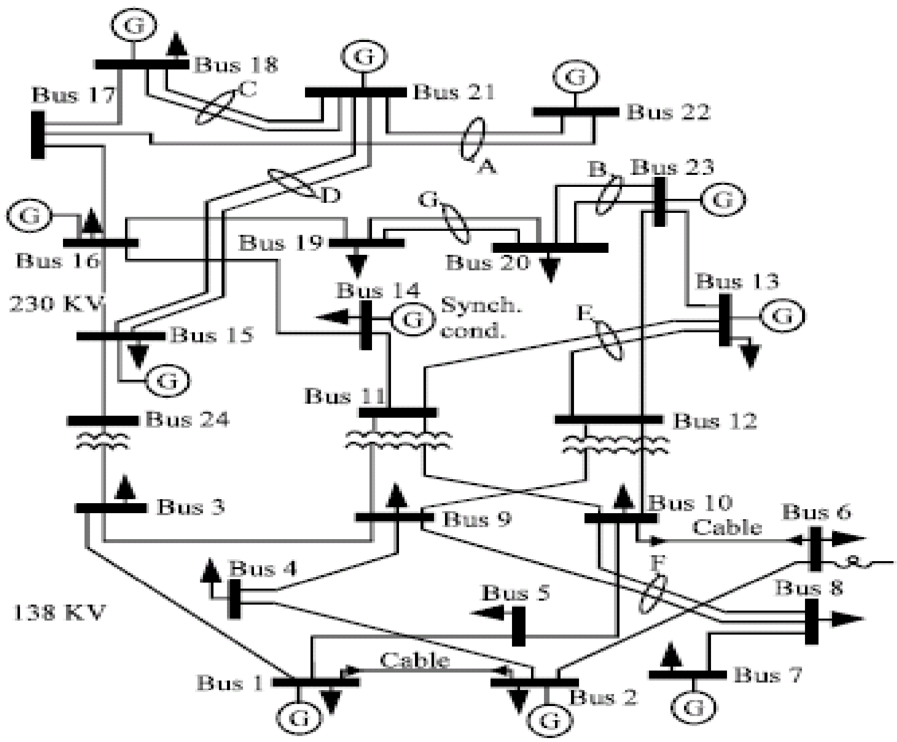

Figure 2.2: Schematic for IEEE RTS

2.5 Performance Evaluation Metrics

Performance of an AI based composite system reliability evaluation techniques is commonly considered as binary classification. To test performance of these systems usually statistical measures of Sensitivity (also termed as recall) and Specificity are utilized. These measurements are based on the terminologies of True Positive (TP), False Positive (FP), True Negative (TN), and False Negative (FN). Here, True and False refers the assigned classification being correct or incorrect, while positive or negative refers to assignment to the positive or the negative category.

In a binary classification, sensitivity expresses proportion of correctly identified positives to all predictions classified as positive. Calculation of sensitivity is described in (eq 2.11).

$$Sensitivity = \frac{TP}{TP+FN} \tag{2.11}$$

Sensitivity measures the performance of classifier in reducing computational time of MCS since each False Negative requires an analysis by power flow equation. Specificity on the other hand, measures the proportion of correctly identified negative samples of all samples that are classified as negative. Specificity is described in (2.12).

$$Specificity = \frac{TN}{TN+FP} \tag{2.12}$$

In terms of reliability evaluation, specificity measures the accuracy of classifier in estimating reliability parameters since incorrect classifications of negative cases tend to change calculated reliability parameters. In addition to the metrics described above, performance of the proposed method in estimating composite power system reliability indices is evaluated based on Loss of Load Probability (LOLP). Definition of LOLP is given in (2.13).

$$LOLP = \frac{Total\ number\ of\ Failures}{TotalNumberof\ Samples} \tag{2.13}$$

3. Multi label classification for composite system reliability evaluation

3.1 Introduction

In this section a new approach for reliability evaluation of composite power systems by combining Monte Carlo simulation (MCS) and Multi Label Learning (MLL) is described. Multi Label K-Nearest Neighbor (MLKNN) algorithm is used as a classifier to show effectiveness of the proposed method. MLL is a classification technique in which the target vector of each instance is assigned into multiple classes. In this research MLL method is used to classify states (failure or success at bus level) of a power system without requiring optimal power flow (OPF) analysis, except in the training phase. As a result, the computational burden to perform OPF is reduced dramatically. For illustration, the proposed method is applied to the IEEE 30 BUS Test System and IEEE Reliability Test System (IEEE RTS). The results from various case studies demonstrate that MLKNN based reliability evaluation provides promising results in both classification accuracy and computation time in evaluating the composite power system reliability. Details of the proposed method are presented in following subsections.

3.2 Multi Label Learning for Power System Reliability Evaluation

Multi label classification is a type of learning where each sample is associated with multiple labels, making it suitable for calculation of bus indices. The multi-label learning methods can be explored in two main groups which are algorithm adaptation and problem transformation methods. Algorithm adaptation methods mainly target to extend some specific single class learning algorithms to handle multi label classification problems directly. Some examples of this group include MLKNN, neural networks based Multi Label classification or decision trees. The transformation methods, on the other hand, aim to transform a multi label classification problem into a single label classification problem. Binary reverse method or pair-wise method can be given as examples for this method. In this part of research, a combination of MCS and MLKNN classifier is used to evaluate reliability indices of composite power systems.

The main contribution of the proposed method is a technique that minimizes computational burden of classification of sampled states with MCS and reduces the need for OPF for the reliability evaluation, except in the training stage and to extend the capability to bus level classification. MLKNN has one of the most time efficient structures among many MLL methods. This feature allows increasing computational efficiency of MCS. Moreover, experiments show that performance of MLKNN is superior to those of some well-established multi-label learning methods [27, 28, 29, 30, and 31].

3.3 Multi Label K-Nearest Neighbor Algorithm

MLKNN approach is an MLL algorithm which is derived from the traditional KNN. In this method, KNN for each element in the training set is identified. Then statistical information is gained from the label sets for each instance. Lastly maximum a posteriori (MAP) principle is applied to determine the label for the test instance. In this study MLKNN algorithm is chosen for evaluation of reliability indices of composite power system because of its classification performance and time efficient structure.

Before explaining the algorithm, several notations are introduced. Let there be an instance m and its associated label set Y ⊆ ý. Let $\overrightarrow{y_m}$ be the category vector for m, where its q_{th} component $\overrightarrow{y_m}(q)$ ($q \in Y$) takes the value of 1 if $q \in Y$ and 0 otherwise. In addition, let N (m) denote the set of KNNs of m identified in the training set. Thus, based on the label sets of these neighbors, a membership counting vector can be defined as:

$$\overrightarrow{C_m}(q) = \sum_{a \in N(m)} \overrightarrow{y_a}\,(q),\ \ q \in y \qquad (3.1)$$

Where $\overrightarrow{C_m}(q)$ counts the number of neighbors of m belonging to the qth class. This vector is used to determine how many samples in number of neighbors N of sample m has labeled for each class described. In terms of composite system evaluation, the equation describes how many load failures occurred for the sample m in K number of neighbors. These numbers can be obtained by counting the occurrence of failures in training target matrix.

For each test instance t, MLKNN firstly identifies its KNNs $N(t)$ where N is the training set. Let H_1^q be the event that t has label q, while H_0^q be the event that t does not have label l. Furthermore, let E_j^q ($j \in \{0,1, \ldots, K\}$) denote the event that, among the KNNs of t, there are exactly j instances which have label q. Therefore, based on the membership counting vector $\overrightarrow{C_t}$ the category vector $\overrightarrow{y_t}$ is determined using the following maximum a posteriori principle:

$$\overrightarrow{y_t}\,(q) = \operatorname{argmax}_{b \in \{0,1\}} P\left(H_b^q | E^q_{\overrightarrow{C_t}(q)}\right),\ \ l \in Y \qquad (3.2)$$

Using the Bayesian rule, Eq. (2) can be rewritten as:

$$\overrightarrow{y_t}\,(q) = \operatorname{argmax}_{b \in \{0,1\}} \frac{P(H_b^q) P\left(E^q_{\overrightarrow{C_t}(q)} | H_b^q\right)}{P\left(E^q_{\overrightarrow{C_t}(q)}\right)} \qquad (3.3)$$

$$= \operatorname{argmax}_{b \in \{0,1\}} P(H_b^q) P\left(E^q_{\overrightarrow{C_t}(q)} | H_b^q\right) \qquad (3.4)$$

$$P(q) = P\left(H_1^q | E^q_{\overrightarrow{C_t}(q)}\right) = \frac{P(H_1^q)P\left(E^q_{\overrightarrow{C_t}(q)} | H_1^q\right)}{\sum_{b\in\{0,1\}} P(H_b^q)P\left(E^q_{\overrightarrow{C_t}(q)} | H_b^q\right)} \quad (3.5)$$

Eq. (3.2 – 3.4) explain how to calculate prior and conditional probabilities. Prior probability term is used to describe the loss of load probability for each bus in overall training dataset. The output of this process is a Qx1 vector where Q is number of buses in a system. Conditional probability terms represent the loss of load probabilities for each bus of sample m in all K neighbors. This probability is also calculated by counting the occurrences of failures in all K neighbors for each bus Q. The output of this process is a $Qx(K+1)$ matrix where Q is number of buses in a system and K is the number of neighbors specified for classifier.

As shown in Eq. (3.4), in order to determine the category vector $\overrightarrow{y_t}$, all the information that is needed is the prior probabilities $P(H_b^q)$ $(j \in \{0,1, \dots, K\})$. Actually, these prior and posterior probabilities can all be directly estimated from the training set based on frequency counting.

Correspondingly, $C'[j]$ counts the number of training instances without label q whose k nearest neighbors contain exactly j instances with label q. Finally, using the Bayesian rule, steps from (3.5) the algorithm's outputs based on the estimated probabilities can be computed.

3.4 MLKNN for Power System Reliability Evaluation

In this section, first formulation of general composite system reliability evaluation parameters is made, later, implementation of MLKNN algorithm is fully explained by steps.

3.4.1 General Definition of MLKNN Parameters

In this study, total generation capacities for buses of composite system are taken as input parameter for MLKNN classifier. So, each bus which is capable of generation in the system is considered as an element of input matrix G for every sample (instance) M as described in (3.6).

$$G_{input} = \begin{bmatrix} G_{11} & G_{12} & G_{1N} \\ G_{21} & G_{22} & G_{2N} \\ G_{M1} & G_{M2} & G_{MN} \end{bmatrix} \quad (3.6)$$

Where N is the number of the generation buses and M is the total number of samples in the input matrix.

A target matrix T is also created for training of the MLKNN classifier which includes state information for each bus of the system for M different samples, described in (3.7).

$$T = \begin{bmatrix} T_{11} & T_{12} & T_{1Q} \\ T_{21} & T_{22} & T_{2Q} \\ T_{M1} & T_{M2} & T_{MQ} \end{bmatrix} \quad (3.7)$$

Where Q is the number of the load buses in the system and S is the status information of bus q. While defining status of buses '-1' is taken to describe 'success state' and '1' for 'failure state'. Desired output for this classifier P_{out}, contains failure probabilities for each bus of composite reliability system for each sample M, described in (3.8).

$$P_{out} = \begin{bmatrix} P_{11} & P_{12} & P_{1Q} \\ P_{21} & P_{22} & P_{2Q} \\ P_{M1} & P_{M2} & P_{MQ} \end{bmatrix} \tag{3.8}$$

3.4.2 Explanation of MLKNN Procedure

After giving definitions of general parameters for MLKNN classifier, the training and testing procedure is now explained in steps. Before starting explanation, a few parameters are described:

m: defines index of current sample of the total M samples.
q: defines the bus index of total Q buses of system.
T_m defines the state of bus *q* at sample *m* so;

$$T_m(q) = \begin{cases} -1 \; where \; bus \; q = success \\ \;\; 1 \; where \; bus \; q = failure \end{cases}$$

K indicates the determined nearest neighbor index used in classification.

3.4.2.1 Training Procedure

1. MLKNN is a classification technique which uses the k-nearest neighbor algorithm for finding the closest relationship between training samples. So, the first step of training procedure is creating a distance matrix. In this study, Euclidian distance method is used to create this matrix described in (3.9).

$$\sum_{i=1}^{M} \sqrt{(a_i - b_i)^2} \tag{3.9}$$

For further explanation, a vector is described to represent sum of squares for input vectors for each sample in (3.10).

$$G_{ss} = [G_1^2 + G_2^2 \dots G_N^2] \tag{3.10}$$

Where G_{ss} sum of squares for each generation bus and N is the number of total generation buses. Based on equation (3.10) a concurrent generation matrix can be described in (3.11).

$$G_{concur} = \begin{bmatrix} G_{ss1} & G_{ss1} & G_{ss1} \\ G_{ss2} & G_{ss2} & G_{ss2} \\ G_{ssM} & G_{ssM} & G_{ssM} \end{bmatrix} \tag{3.11}$$

Where G_{concur} is the concurrent matrix created to calculate distance used in k-mean algorithm and M is number of total samples. At this point G_{concur} is applied to equation (9) described at (3.12).

$$Dist = \sqrt{G_{concur} + G_{concur}^{T} - 2\left(G_{input} x G_{input}^{T}\right)} \tag{3.12}$$

Where $Dist$ is the matrix including the data of distances between samples. Finally, the distance matrix is described in (3.13) below.

$$Dist = \begin{bmatrix} Dist_{11} & Dist_{12} & Dist_{1M} \\ Dist_{21} & Dist_{22} & Dist_{2M} \\ Dist_{M1} & Dist_{M2} & Dist_{MM} \end{bmatrix} \tag{3.13}$$

Where each element of the MxM matrix describes the distances between samples.

1- In the second step prior probabilities of failure for each bus are calculated based on counting instances as shown in (3.1). Calculation of prior probabilities is described in (3.14).

$$P^{1}(q) = \frac{\sum_{i=1}^{M} Y_i(q)}{M} \tag{3.14}$$

$$P^{0}(q) = 1 - P^{1}(q) \tag{3.15}$$

At the end of process, a Qx1 Prior and a Qx1 Compliment Prior probability matrix are obtained which gives prior probabilities of each bus.

2- In the third step conditional and conditional compliment probabilities for buses are calculated for K nearest neighbors based on counting. In this step the algorithm first determines how many of K nearest neighbors for sample m have failure at bus q. Later the process is repeated for all M samples to determine probability of failure for bus q conditional to the occurrence in nearest numbers. This process is formulated in (3.16).

$$Cond(k, q_b) = P\left(C_{(k)\in\{0,K\}]}^{q} | H_{b\in\{0,1\}}^{q}\right) \tag{3.16}$$

where $C_{(k)\in\{0,K\}]}^{q}$ denote the number of instances which have failure on bus q. Also H_1^q describes the event bus q has failure at the sample x as likewise H_0^q describes the event bus q has not failure at the sample x.

At the end of these steps four required probability matrices for system are obtained.

-Prior (Q) is a Q x 1 matrix that defines prior probability of failure for all buses in system.
-Prior Negative (Q) is a Q x 1 matrix that defines prior probability of failure for all buses in system.
-Cond ($K|Q$) is a Q x (K+1) matrix that shows conditional probabilities of failure for all buses in system according to kth closest neighbor.
-Cond Negative ($K|Q$) is a Q x (K+1) matrix that shows negative of conditional probabilities of failure for all buses in system according to kth closest neighbor.

After training parameters are obtained, testing process is used to calculate probability of failure for each bus for given sample m based on using Bayesian rule showed in (3.5).

3.4.2.2 Testing Procedure

After training of MLKNN classifier is completed, testing process can be used to identify bus statuses of a composite power system. Testing process has 3 main steps:

1- As in training, the first step of testing is also calculating distance matrix between test data and training data. The same process as described in (3.9) is used in this step. At the end of this step $M_{test}xM_{train}$ distance matrix is obtained.

2- In this step, the number of failures in K nearest neighbors is determined based on the counting process (3.1). Results of the counting indicates the required indices for Cond ($k|q$) where q indicates the bus number and $k \in \{0, \dots, K\}$.

3- In the last step Bayesian rule described in (3.5) used for determining failure probabilities of busses in test database.

$$P(q) = \frac{P(q)P(k|q)}{\sum_{b\in\{0,1\}} P\left(C_{(k)}^{q} | H_{b}^{q}\right)} \tag{3.17}$$

Afterwards the probability matrix obtained previously specified threshold can be used to determine if a bus is in failure state or not. Overall flow diagram for the proposed MLKNN classifier for IEEE RTS 79 test system which has 24 buses (10 of them are generation buses), 38 transmission lines and 32 generation units is given in figure 3.1.

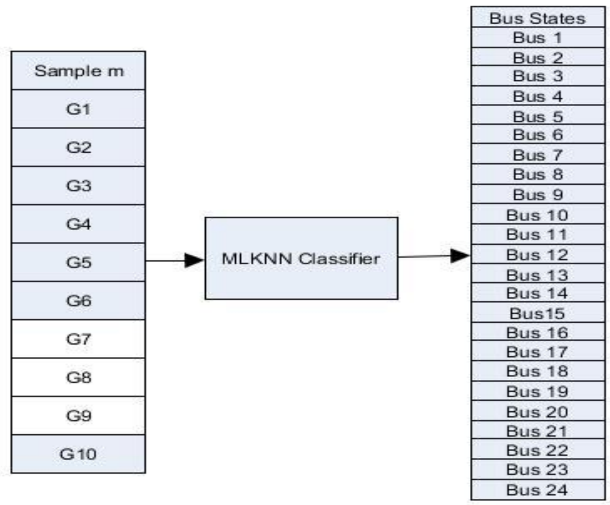

Fig 3.1: Overall diagram for MLKNN Classifier on RTS 79

3.5 Proposed Topology

This section explores the MLKNN classifier in conjunction with the MCS to reduce the computational requirements while evaluating the composite power system reliability. In this section firstly, the basic steps of the MCS are described and then detailed computational procedure of the proposed MLKNN classifier for the composite power system reliability evaluation is explained.

In this study, the non-sequential MCS is used as a benchmark for testing the performance of the proposed method. Non-sequential MCS is generally preferred for composite test system because of simplicity of method and its computational efficiency. The basic steps for the composite reliability evaluation by the non-sequential MCS are explained as follows;

1- Select a random state for all components of the power system as $x = (x_1, x_2 \ldots x_m)$ where m is the component number in power system.

2- Classify each of the selected state x (as success or failure) through KLMNN classifier (classifier is trained with a proper training database created in section A).

3- Update the estimate E (F), the expected value of the system reliability indices using the results gathered from step 2 described in (3.18).

$$E(F) = \left(\frac{1}{N} \sum_{i}^{N} x_i\right) \qquad (3.18)$$

Where N is the number of simulation steps.

4- If the determined criteria for variance is reached stop the simulation otherwise return to the first step.

The estimate of uncertainty is usually represented by the coefficient of variation β. An acceptable value of the estimate of uncertainty is used as stopping criteria for the simulation. Besides variance, determined specified number of samples can also be used as stopping criteria. Composite power systems are generally highly reliable systems, so the probability that failed state occurs is much less than success states. Therefore, the steps described in this section repeat many times. As a result, reducing the computational burden of power flow analysis by using KLMNN classifier provides significant time and computational efficiency.

The first step of implementing MLKNN classifier in composite system evaluation with combination of MCS is generating a training database. A proper dataset is created using a set of sampled states and the corresponding state classification labels for each bus (success or failure), which are obtained from the MCS. Once the appropriate training patterns are obtained, then the MLKNN classifier is trained, which would then be used for the state space classification of the testing database to evaluate the reliability. In this database, input vector is created by using generation buses while output of the classifier is defined as state (success or failure) of each bus of the selected test system. Proposed method is applied on two different test systems; IEEE 30 Bus Test System (case 1) and IEEE RTS 79 (case 2). Performance of proposed system is also tested on varying load levels (case 3) and under circumstance of failing transmission lines (case 4). In these test systems input data sampled from 6, 10, 24 and 47 input buses respectively while target matrices sampled from 24 and 30 buses for RTS and IEEE 30 Test System. General algorithm of the proposed method for composite reliability evaluation is shown in Fig. 3.2 and its detailed implementation procedure is outlined below.

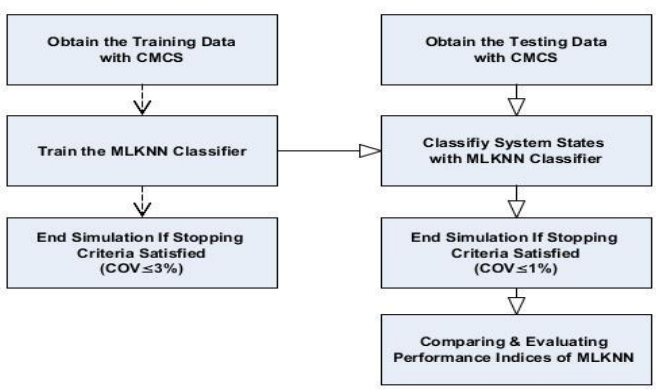

Fig 3.2: Flowchart of overall process of MCS-MLKNN method

Training data samples for MLKNN classifier are obtained by the MCS. Size of these data samples can be determined by either using pre-specified number of samples or the convergence of the coefficient of variation method. In this study, size of dataset is determined by specific number of samples proportional to the LOLP of the system to get the sufficient number of attributes in the input training patterns. In this training dataset, input vectors are created based on information of the total generation for each bus, and target matrices are created by using corresponding system state characteristics (success or failure) for each bus of the testing system. In the target matrix 1 and -1 are used to represent the desired output of the success and failure states, respectively. In this study, the repetitive states of MCS are not included in the training database to reduce the number of samples included in the training process. It should be also noted that the number of obtained success states is much higher than failed states, some of the success states are discarded to prevent overemphasis of classification. Consequently, the number of training patterns is decreased to speed up the MLKNN training and a balanced training patterns is created to increase overall performance of classifier.

After training patterns have been generated, the next step is training the MLKNN for selected set of input/output patterns. Once the MLKNN classifier is trained, the MCS follows the same steps as described early in this section with exception that the state characterization is now performed by the trained MLKNN instead of running the DC-OPF calculations. With this procedure, the composite reliability indices can be calculated without requiring power flow calculations. In this way, computational time necessary for evaluating the composite power system reliability is reduced dramatically. In this study, the simulations are run until coefficient of variation reaches (COV) 1% for testing stage in all case studies. COV indices are calculated based on overall failures of specified test system. It is because overall the failure rate of a system is much higher than bus level failures. The overall performance of the MLKNN classifier is measured by using the parameters: overall accuracy, sensitivity, specificity, g-mean and simulation time.

3.6 Case Studies and Results

In this section four case studies were conducted on IEEE 30 Bus System and IEEE RTS respectively to analyze the performance of MLKNN classifier.

In case study one, performance of the proposed method is demonstrated on IEEE 30 Bus System. Following that, case studies are performed on the IEEE RTS since it is used by most of the developers of new algorithms for composite system reliability evaluation studies. Performance of the proposed method on RTS shown in case study two and case study three is for peak load level and hourly varying load level respectively. For the first three case studies the capacity and admittance constraints of transmission lines are considered. However, the transmission line failures are not considered as these have much smaller probabilities than the generator failures.
In case study four, transmission line failures are also considered. For this case study, Modified RTS (MRTS) is preferred. MRTS is designed for transmission line reliability studies for composite systems. In this system, generation capacities are doubled and all the loads are multiplied by 1.8 while rest of the system parameters remain unchanged. In this way effect of transmission lines on overall system reliability is increased and become more observable.

All the simulations are performed using MATLAB (2017b) platform on a PC with Intel Core i7-4510 CPU (~2.6GHz), 16 GB Memory. Simulation results for case studies are discussed in following subsections.

3.6.1 Case Study 1

In case study 1, to test performance of the proposed classifier, the system is tested on IEEE 30 Bus Test System for single load level of 255 MW (annual peak). A total of 34867 samples are characterized through MCS to create a training dataset with 33744 success and 1123 failure states. In these samples a total of 2500 states are selected to train the classifier. After the training, the system is tested until COV reaches the limit of ≤ 1%. During this phase, 301583 samples are classified with 291243 success and 10340 failure states. The overall performance of MLKNN classifier is stated in table 3.1 as well as simulation times for each model (MLKNN Classifier and OPF). The classification performance is compared to results obtained from DC-OPF analyses in table 3.2. The performance indices to present obtained results are calculated according to metrics described before.

According to the results stated in table 3.1, MLKNN Classifier can successfully identify overall LOLP of IEEE 30 Bus Test system with a small error rate. Table 3.1 also shows that MLKNN classification reduces the computation time for calculating reliability indices dramatically comparing to OPF based MCS methods. Table 3.2 shows that MLKNN classification method provides reasonably accurate classification of bus states of RTS composite power system (success or failure).

Table 3.1: Comparison of Overall System Performance for MLKNN & CMCS for IEEE 30 Bus Test System

	CMCS	**MLKNN**
Success States	291792	291502
Failure States	9791	10081
Loss of Load	0.032	0.033
Sensitivity	N/A	0.99812
Specificity	N/A	1
G-Mean	N/A	0.999
Analysis Time (Sec)	28659	1639

Table 3.2: Comparison of MLKNN & CMCS in Bus Level for IEEE 30 Bus Test System

	CMCS		MLKNN		Accuracy		
	FN	LOL (%)	FN	LOL (%)	Sensitivity	Specificity	G-Mean
Bus1	0	0.00	0	0.00	N/A	N/A	N/A
Bus2	1712	0.57	1831	0.61	0.998	0.988	0.9929
Bus3	1519	0.50	1511	0.50	0.999	0.984	0.9914
Bus4	3379	1.12	3363	1.12	1	0.993	0.9964
Bus5	3558	1.18	3562	1.18	0.999	0.996	0.9974
Bus6	0	0.00	0	0.00	N/A	N/A	N/A
Bus7	2187	0.73	2194	0.73	0.999	1	0.9995
Bus8	2030	0.67	2022	0.67	0.999	0.991	0.9949
Bus9	0	0.00	0	0.00	N/A	N/A	N/A
Bus10	1049	0.35	1038	0.34	1	0.981	0.9904
Bus11	0	0.00	0	0.00	N/A	N/A	N/A
Bus12	2366	0.78	2348	0.78	0.999	0.986	0.9924
Bus13	0	0.00	0	0.00	N/A	N/A	N/A
Bus14	2747	0.91	2729	0.90	1	0.989	0.9944
Bus15	1686	0.56	1671	0.55	0.999	0.985	0.9919
Bus16	1068	0.35	1045	0.35	0.999	0.974	0.9864
Bus17	784	0.26	771	0.26	1	0.971	0.9853
Bus18	1412	0.47	1394	0.46	0.999	0.982	0.9904
Bus19	1500	0.50	1486	0.49	0.999	0.982	0.9904
Bus20	1436	0.48	1423	0.47	0.999	0.981	0.9899
Bus21	1755	0.58	1868	0.62	0.999	0.985	0.9919
Bus22	0	0.00	0	0.00	N/A	N/A	N/A
Bus23	1313	0.44	1291	0.43	1	0.977	0.9884
Bus24	1370	0.45	1358	0.45	1	0.983	0.9914
Bus25	0	0.00	0	0.00	N/A	N/A	N/A
Bus26	1194	0.40	1182	0.39	1	0.976	0.9879
Bus27	0	0.00	0	0.00	N/A	N/A	N/A
Bus28	0	0.00	0	0.00	N/A	N/A	N/A
Bus29	469	0.16	451	0.15	1	0.958	0.9787
Bus30	1464	0.49	1447	0.48	1	0.981	0.9904

3.6.2 Case Study 2

In case study two, to test performance of the proposed classifier, the system is tested on single area IEEE RTS for single load level of 2850 MW (annual peak). There are 10 generation buses in RTS which are considered as input vector. To train the classifier, 14682 samples are obtained through MCS with 13381 successes and 1301 failures in this process. After obtaining adequate number of samples, the training patterns are recombined to generate a balanced training dataset (some of the success states are discarded to prevent overtraining). A total of 3000 samples are selected with 1301 failure and 2699 success states for this dataset. It should be noted that most of the success states are ignored during this process to emphasize classification of failure states (which is reasonable in order to calculate reliability indices). After MLKNN classifier is successfully trained, the proposed system is tested until COV reaches the limit of $\leq$ 1%. After testing is completed, 109743 samples are classified with 100470 successe and 9273 failure states.

The overall classification performance of MLKNN classifier and time comparison between the proposed method and CMCS is shown in table 3.3. The classification performance at bus level is stated with comparison of results obtained from DC-OPF analyses in table 3.4.

According to the results in table 3.3, MLKNN Classifier can successfully identify overall LOLP of RTS system with a small error rate. Similar to Case study one, MLKNN classification reduces computation time of reliability indices significantly comparing to OPF based CMCS.

Table 3.3: Comparison of Overall System Performance for MLKNN & CMCS

	CMCS	MLKNN
Success States	100470	100402
Failure States	9273	9341
Loss of Load Probability	0.0845	0.085
Sensitivity	N/A	0.997
Specificity	N/A	0.993
G-Mean	N/A	0.994
Analysis Time (Sec)	6271	448

Table 3.4: Comparison of MLKNN & CMCS in Bus Level

	MLKNN		CMCS		Accuracy		
	FN	**LOL (%)**	**FN**	**LOL (%)**	**Sensitivity**	**Specificity**	**G-Mean**
Bus1	2973	2.71	2934	2.67	0.998	0.974	0.9859
Bus2	2937	2.68	2953	2.69	0.999	0.981	0.9899
Bus3	0	0.00	0	0.00	N/A	N/A	N/A
Bus4	0	0.00	0	0.00	N/A	N/A	N/A
Bus5	936	0.85	952	0.87	0.999	0.972	0.9854
Bus6	33	0.03	29	0.03	0.999	0.896	0.9460
Bus7	5032	4.59	5063	4.61	0.998	0.981	0.9894
Bus8	648	0.59	642	0.59	0.999	0.986	0.9924
Bus9	3	0.0	1	0.0	0.999	0.33	0.5741
Bus10	20	0.02	21	0.02	0.999	0.952	0.9752
Bus11	0	0.00	0	0.00	N/A	N/A	N/A
Bus12	0	0.00	0	0.00	N/A	N/A	N/A
Bus13	456	0.42	471	0.43	0.999	0.946	0.9721
Bus14	208	0.19	224	0.20	0.999	0.902	0.9492
Bus15	0	0.00	0	0.00	N/A	N/A	N/A
Bus16	56	0.05	51	0.05	1	0.882	0.9391
Bus17	0	0.00	0	0.00	N/A	N/A	N/A
Bus18	1018	0.93	1034	0.94	0.999	0.907	0.9518
Bus19	48	0.04	51	0.05	0.999	0.921	0.9592
Bus20	3802	3.46	3879	3.53	0.997	0.946	0.971
Bus21	0	0.00	0	0.00	N/A	N/A	N/A
Bus22	0	0.00	0	0.00	N/A	N/A	N/A
Bus23	0	0.00	0	0.00	N/A	N/A	N/A
Bus24	0	0.00	0	0.00	N/A	N/A	N/A

As shown in Table 3.4, MLKNN classification method provides reasonably accurate classification of bus states of RTS composite power system (success or failure). It should be noted that if failure rate at a bus is very small (like bus-9) classifier may show lower performance than average. The reason is that, there are not enough samples generated to adjudicate properly for those buses neither in training nor testing stages.

3.6.3 Case Study 3

In case study three, performance of the proposed method is tested using varying load levels based on information provided in hourly load chart of RTS. To be able to classify varying load levels, load information should be added to the input of the classifier. For this purpose, the input equation (3.1) is modified and the new equation is described below in (3.23).

$$I_{input} = \begin{bmatrix} G_{11} - D_{11} & G_{12} - D_{12} & G_{1N} - D_{1N} \\ G_{21} - D_{21} & G_{22} - D_{22} & G_{2N} - D_{2N} \\ G_{M1} - D_{M1} & G_{M2} - D_{M2} & G_{MN} - D_{MN} \end{bmatrix} \tag{3.23}$$

In this equation G represents generation and D represents load at bus N for a total of M number of samples.

However, size of classifier can become too large while classifying system states in varying load levels which reduces computational efficiency of the classifier. In this study, instead of using one classifier, multiple classifiers are trained for different load levels to handle this problem efficiently. Total state space sampled is divided to five main levels based on available generation data. For each of those levels a unique classifier is trained. In the testing stage, a decision tree is used to determine which classifier to be used for classification for every random sample. The overall diagram of the algorithm used in this case study is given in figure 3.3.

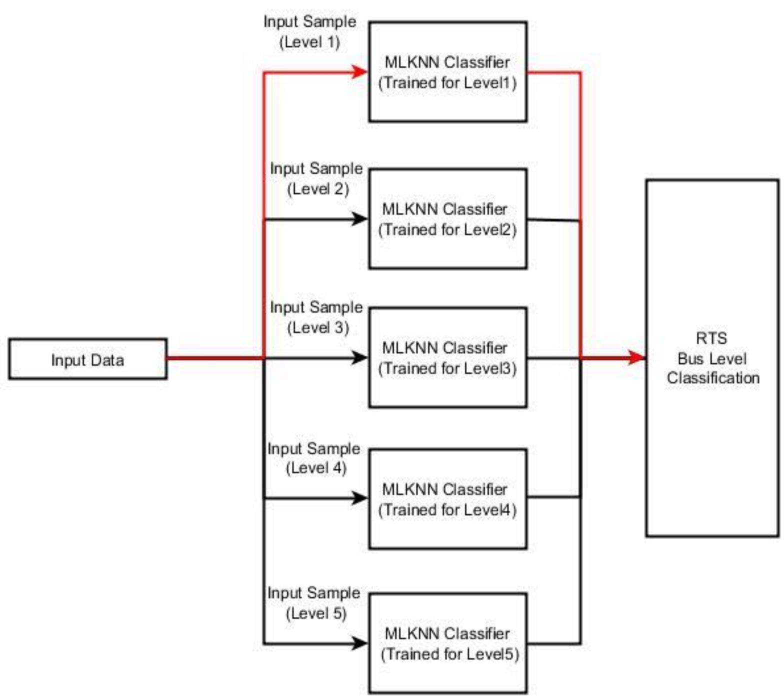

Figure 3.3: Overall Diagram of Proposed Classifier for Variable Load Levels

In the training phase of classifier, MCS is run for each load level until a total of 5000 samples are obtained with 3000 success and 2000 failure states. Most of the success states are discarded to prevent overtraining as in previous case studies. After MLKNN classifier is successfully trained, the proposed system is tested until COV reaches the limit of ≤ 1%. After testing is completed, 7540967 samples are classified with 7540967 successes and 9052 failure states.

The overall classification performance of proposed method for characterizing random samples is given in table 3.5. The classification performance at bus level is presented with comparison of results obtained from DC-OPF analyses in table 3.6.

Results obtained in table 3.5 show that the proposed classifier can successfully identify overall LOLP of RTS system with an acceptable error rate. The results also show that computation time of reliability indices is reduced significantly comparing to CMCS. As presented in Table 3.6 MLKNN classifier can characterize system buses very accurately in varying load levels.

Table 3.5: Comparison of System Performance for Varying Hourly Load

	CMCS	**MLKNN**
Success States	7540967	7540881
Failure States	9052	9138
Loss of Load	0.0012	0.0012
Sensitivity	N/A	0.99
Specificity	N/A	0.96
G-Mean	N/A	0.97
Analysis Time (Sec)	375249	33492

Table 3.6: Comparison of System Performance for Varying Hourly Load in Bus Level

Location	CMCS	MLKNN	Sensitivity	Specificity	G-Mean
Bus1	2115	2178	0.99	0.97	0.98
Bus2	1157	1204	0.99	0.98	0.98
Bus3	N/A	N/A	N/A	N/A	N/A
Bus4	N/A	N/A	N/A	N/A	N/A
Bus5	606	612	0.99	0.97	0.98
Bus6	36	51	0.99	0.92	0.95
Bus7	4143	4051	0.99	0.91	0.95
Bus8	325	371	0.99	0.97	0.98
Bus9	N/A	N/A	N/A	N/A	N/A
Bus10	N/A	N/A	N/A	N/A	N/A
Bus11	N/A	N/A	N/A	N/A	N/A
Bus12	N/A	N/A	N/A	N/A	N/A
Bus13	269	291	0.99	0.88	0.94
Bus14	247	278	0.99	0.89	0.93
Bus15	0	0	N/A	N/A	N/A
Bus16	43	49	1	0.91	0.95
Bus17	0	0	N/A	N/A	N/A
Bus18	1797	1658	1	0.86	0.93
Bus19	239	291	1	0.92	0.96
Bus20	4484	4303	1	0.91	0.95
Bus21	N/A	N/A	N/A	N/A	N/A
Bus22	N/A	N/A	N/A	N/A	N/A
Bus23	N/A	N/A	N/A	N/A	N/A
Bus24	N/A	N/A	N/A	N/A	N/A

3.6.4 Case Study 4

In case study four, the proposed method is applied to include transmission line failures in the classification scheme. To create proper patterns to classify transmission line failures, states of transmission lines should be combined with the generation data as an input of classifier. Many of the previous studies that have a focus on machine learning techniques for state classification of composite power systems ignore system failures sourced from transmission lines by making some assumptions. The main difficulty lies here is combining information of available generation and transmission line capacity at the input of the classifier to create proper patterns. In this study, a new approach is proposed to achieve classification of failures reasoned from transmission line failures. In this approach, equation (3.1) is modified by applying discrete time convolution to the information obtained from generation and transmission line states. New input equation is described below in (3.24) below.

$$I_{input} = \begin{bmatrix} G_{11} & G_{12} & G_{1N} \\ G_{21} & G_{22} & G_{2N} \\ G_{M1} & G_{M2} & G_{MN} \end{bmatrix} * \begin{bmatrix} Tr_{11} & Tr_{12} & Tr_{1L} \\ Tr_{21} & Tr_{22} & Tr_{2L} \\ Tr_{M1} & Tr_{M2} & Tr_{ML} \end{bmatrix} \quad (3.24)$$

Where G represents generation at bus N and Tr represents available transmission line capacity for transmission line L at total M number of samples. The description of discrete time convolution which symbolized as ‘*’ is presented in equation (3.25).

$$(G * Tr) = \sum_m G(i-m)xTr(m) \quad (3.25)$$

In the training phase, MCS is run for each load level until a total of 10000 samples are obtained with 5000 failure and 5000 success states. As in the previous case studies, the remaining success states are discarded to prevent overtraining. After MLKNN classifier is successfully trained, the proposed system is tested until COV reaches the limit of $\leq 1\%$. After testing is completed, 200387 samples are classified with 189659 successes and 10728 failure states. It should be noted that load level is considered as constant at its annual peak level.

The overall classification performance of the proposed method for characterizing random samples is given in table 3.7. The classification performance at bus level is presented with comparison of results obtained from DC-OPF analyses in table 3.8. Finally, table 3.9 demonstrates the classification performance of proposed approach only for system failures sourced from transmission line failures.

Results obtained in table 3.7 show that proposed classifier can successfully identify overall LOLP of RTS system with an acceptable error rate. The results also show that time of computing reliability indices is reduced significantly comparing to CMCS. As presented in Table 3.8 MLKNN classifier can characterize system buses very accurately when transmission line failures considered.

Table 3.7: Performance Comparison when Transmission Line Failures Considered

	CMCS	**MLKNN**
Success States	190659	190583
Failure States	9728	9804
Loss of Load	0.0485	0.0489
Sensitivity	N/A	0.99
Specificity	N/A	0.98
G-Mean	N/A	0.99
Analysis Time (Sec)	12390	637

Table 3.8: Comparison of Bus Level Classification Performance for Case 4

	CMCS	MLKNN	Sensitivity	Specificity	G-Mean
Bus1	961	986	0.98	0.97	0.97
Bus2	304	321	0.99	0.98	0.98
Bus3	42	51	N/A	N/A	N/A
Bus4	194	206	N/A	N/A	N/A
Bus5	360	372	0.99	0.97	0.98
Bus6	2003	2028	0.99	0.92	0.95
Bus7	1282	1307	0.99	0.97	0.98
Bus8	440	452	0.99	0.99	0.99
Bus9	38	47	0.99	0.98	0.98
Bus10	258	271	0.99	0.97	0.98
Bus11	N/A	N/A	N/A	N/A	N/A
Bus12	N/A	N/A	N/A	N/A	N/A
Bus13	1623	1702	0.99	0.93	0.96
Bus14	4929	5013	0.99	0.91	0.95
Bus15	1	3	N/A	N/A	N/A
Bus16	821	833	1	0.91	0.95
Bus17	N/A	N/A	N/A	N/A	N/A
Bus18	721	739	0.99	0.89	0.94
Bus19	320	332	0.99	0.93	0.96
Bus20	400	421	0.99	0.97	0.98
Bus21	N/A	N/A	N/A	N/A	N/A
Bus22	N/A	N/A	N/A	N/A	N/A
Bus23	N/A	N/A	N/A	N/A	N/A
Bus24	N/A	N/A	N/A	N/A	N/A

Table 3.9: Classification Performance for System Failures only Sourced from Transmission Line Failures

	CMCS	MLKNN	Sensitivity	Specificity	G-Mean
Bus1	10	13	0.98	1	0.99
Bus2	N/A	N/A	N/A	N/A	N/A
Bus3	22	22	1	1	1
Bus4	34	36	0.99	0.97	0.98
Bus5	26	23	0.97	0.85	0.91
Bus6	127	123	0.98	0.95	0.96
Bus7	7	8	0.99	0.88	0.93
Bus8	44	46	0.98	0.96	0.97
Bus9	17	15	1	0.88	0.94
Bus10	58	57	1	0.98	0.99
Bus11	N/A	N/A	N/A	N/A	N/A
Bus12	N/A	N/A	N/A	N/A	N/A
Bus13	3	3	1	1	1
Bus14	42	45	0.99	0.98	0.99
Bus15	24	24	1	1	1
Bus16	14	15	1	0.91	0.9391
Bus17	N/A	N/A	N/A	N/A	N/A
Bus18	N/A	N/A	N/A	N/A	N/A
Bus19	18	18	1	1	1
Bus20	16	16	1	1	1
Bus21	N/A	N/A	N/A	N/A	N/A
Bus22	N/A	N/A	N/A	N/A	N/A
Bus23	N/A	N/A	N/A	N/A	N/A
Bus24	N/A	N/A	N/A	N/A	N/A

It is shown at table 3.9 that overall system failures sourced from transmission line failures are characterized with a high accuracy. The same table also shows that proposed method has a good classification accuracy for bus level characterization.

3.7 Summary

This section has presented a method to evaluate reliability indices for composite power systems. The proposed method uses a MLKNN classifier to identify status of buses that significantly reduces the computational burden of OPF analysis. The importance of reducing the computational time can be understood by two examples.

In Monte Carlo Simulation, the accuracy of convergence is very important. Convergence is measured by the COV, smaller COV means better convergence and more confidence in the estimates. Now the sample size and consequently computational time is inversely proportional to COV or directly proportional to accuracy of convergence. The proposed method reduces the required time for reliability analysis considerably for the same level of accuracy defined by the coefficient of variation. Alternatively, for the same time this allows convergence to a higher level of accuracy. Another example is that for optimal planning of resources, reliability studies may need to be done many times. So, the reduction of computational time helps in optimal planning by being able to conduct more studies in the same time.

The effectiveness of the proposed method is demonstrated on the IEEE 30 Bus Test System and IEEE RTS respectively in four different case studies. In the first two case studies, the load level of system is considered constant at its peak value, in the third case study performance of proposed method is tested on varying hourly load data of RTS and finally, the proposed system is used on MRTS with considering transmission line failures. The accuracy of classification is evaluated by considering the parameter of sensitivity, specificity and g-mean. The training samples are chosen using CMCS. After classifier is trained, testing is completed with CMCS until COV reaches 1% limit. All classifications in this step are made by MLKNN instead of DC-OPF analysis.

The results show that the proposed method shows good performance for classifying success and failure states at constant load level for both IEEE 30 bus system and RTS. In some buses with lower failure rate, however, classification accuracy performs slightly below the average in some cases. This could be amended by adding more samples to the training dataset. It should also be mentioned that, classification of success states showed better performance than classification of failure states.

In the third case study, demand information of sampled states is included in the input of the classifier for characterizing system state includes varying load levels. In this stage, a decision tree is applied to choose one of the five different classifiers which are trained with a focus of load level of sampled states. The simulation results presented demonstrate that the proposed method can execute the classification with a good accuracy.

In the last case study, a new approach is introduced to handle classification of failures sourced from inadequate transmission line capacities. Results presented for this case study proves that proposed approach can characterize those states with a high accuracy.

It is also shown in all case studies that, the time required for calculating composite system reliability indices with MLKNN classifier is much less than OPF based reliability evaluation methods. These case studies demonstrate that the application of the proposed method to composite power system reliability evaluation accurately determines the states status with a huge speed up compared with OPF based Monte Carlo Simulation methods. This method advances the state of the art of using machine learning in power system reliability evaluation from the previous methods by including computation of bus indices and the transmission line failures.

4. Multi Label Classification & Importance Sampling Combination for Composite System Reliability Evaluation

4.1 Introduction

In this section a new approach for evaluation of power system reliability indices with Monte Carlo Simulation is explained with a combination of Multilabel Radial Basis Function (MLRBF) classifier and Importance sampling (IS). Multilabel classification algorithms are different from single label approaches, in which each instance can be assigned into multiple classes. This characteristic gives MLRBF capability to be used to classify composite power system states (success or failure) without requiring optimal power flow (OPF) analysis, with the exception of training and cross-entropy optimization phases. The proposed method is applied to the IEEE RTS for different load levels. The outcomes of case studies show that MLRBF algorithm together with importance sampling provides good classification accuracy in reliability evaluation while reducing computation time substantially. The details of proposed method are explained in the following sections.

4.2 MLRBF Classification for Power System Reliability Evaluation

RBF is one of the most popular approaches among neural network classification methods. RBF Neural Networks are generally comprised of two layers of neurons. In RBF, each hidden neuron (basis function) in the first layer is associated with a prototype vector while each output neuron corresponds to a possible class. Usually training an RBF neural network is handled in a two-stage procedure. In the first layer, the basis functions are learned by performing clustering analysis on training instances while weights are optimized by solving a linear problem in the second layer. Comprehensive descriptions of RBF neural networks are available in [32]. In this section, first, a general formulation of composite system reliability evaluation parameters for MLRBF classification is explained, later, application of the proposed method is described in steps. Finally, a flow chart of MLRBF is provided for clearer understanding in Figure 4.1.

4.2.1 General Definitions for MLRBF in Power System Reliability Evaluation

In this study, total generation capacities and total demand for each bus of composite system are taken as input parameters for MLRBF classifier. So, generation and demand information for each bus in the system is considered as an element of input matrix I for every sample (instance) m as described in (4.1).

$$I_{input} = \begin{bmatrix} G_{11} - D_{11} & G_{12} - D_{12} & G_{1N} - D_{1N} \\ G_{21} - D_{21} & G_{22} - D_{22} & G_{2N} - D_{2N} \\ G_{M1} - D_{M1} & G_{M2} - D_{M2} & G_{MN} - D_{MN} \end{bmatrix} \tag{4.1}$$

where N is the number of the buses and M is the total number of samples in the input matrix. Status of state information for each bus of the system for M different samples is stored in a target matrix T for the purpose of training the MLRBF classifier which is described in (4.2).

$$T = \begin{bmatrix} S_{11} & S_{12} & S_{1Q} \\ S_{21} & S_{22} & S_{2Q} \\ S_{M1} & S_{M2} & S_{MQ} \end{bmatrix} \tag{4.2}$$

where Q is the number of the load buses in the system and S is the status information of bus q. While defining status of buses '-1' is used for 'success states' and '1' for 'failure states' for the corresponding bus. P_{out}, contains failure probability for each bus of composite system for each sample M as the output for this classifier which described in (4.3).

$$P_{out} = \begin{bmatrix} P_{11} & P_{12} & P_{1Q} \\ P_{21} & P_{22} & P_{2Q} \\ P_{M1} & P_{M2} & P_{MQ} \end{bmatrix} \tag{4.3}$$

Now, training and testing procedure is explained in steps in the following subsection.

4.2.2 Explanation of MLRBF Classification Procedure

It is necessary to describe some related parameters before starting explanation;

m: defines index of current sample of total M samples.
i_m: defines the input vector for sample m.
q: defines the bus index of total Q buses of system.
Y_m defines the state of bus *q* in sample *m* so;

$$Y_m(q) = \begin{cases} 1\ (failure)\ where\ \ T_{iq} = 1 \\ 0\ (success)\ where\ \ T_{iq} = -1 \end{cases}$$

Let I=R^d be the input space and Q= {1, 2…, Q} be the finite set of Q possible classes. Given a multilabel training dataset DSet= $\{(i_m,Y_m)| \le m \ge M\}$, where $i_m \in I$ is a single instance and $Y_m \subseteq Q$ is label set associated with i_m.
In this study, K-Means Clustering is applied for each class $q \in Q$ on the set of instances U_q with label q which described in (4.4).

$$U_q= \{i_m\ |(i_m Y_{m)} \in DSet,\ q \in Y_m\} \tag{4.4}$$

In the next step, k_q number of clustered groups are formed for class q and the jth centroid ($1 \le j \le k_q$) is regarded as a prototype vector c^q_j of basis function $\alpha^l_j(.)$. It should be noted that, k_q is taken as a fraction of the total number of instances in U_q which is described as α.

As each output neuron of the MLRBF neural network is related to a possible class, weights between hidden and output layer can be shown as (4.5).

$$W=[w_{jq}]_{(K+1)\ XQ} \tag{4.5}$$

Here, $K = \sum_{q=1}^{Q} k_q$ shows the total number of prototype vectors retained in the hidden layer. The weight matrix W can be learned by minimizing the following sum-of-squares error function as described below (4.6).

$$E = \frac{1}{2}\sum_{m=1}^{M}\sum_{q=1}^{Q}\left(Y_q(i_m) - T_q^m\right)^2 \tag{4.6}$$

Where T_q^m represents the output of i_m on the q-th class, which takes the values of +1 if $q \in Y_i$ and -1 otherwise. So, the output of i_m for the *q*-th class can be calculated as presented below (4.7).

$$y_q(i_m) = \sum_{j=0}^{Q} w_{jq}\phi_j(i_m) \tag{4.7}$$

In this study, the basis function a_j isrepresented with the following widely-used Gaussian style activation (4.8).

$$\phi_j(i_m) = \exp(-\frac{dist(i_m,c_j)^2}{2\sigma_j^2}) \tag{4.8}$$

Here dist (i_m,c_j) calculates the distance between i_m and the *j*-th prototype vector c_j with the usual Euclidean distance algorithm. The smoothing parameter σ is shown with the equation below (4.9).

$$\sigma = \left(\frac{\sum_{p=1}^{K-1}\sum_{r=p+1}^{K} dist(c_p,c_r)}{\frac{K(K-1)}{2}}\right) \tag{4.9}$$

Differentiating the error function (4.6) with respect to w_{jq} and setting the derivative to zero will result in the equation given below (4.10).

$$(\phi^T\phi)W = \phi^T T \tag{4.10}$$

In equation 4.10, ϕ = $[\phi_{mj}]_{m\ x\ [K+1]}$ with elements, $\phi_{mj} = \phi_j(i_m)$, W= $[w_{jq}]_{Q\ x\ [K+1]}$ and T=$[t_{mq}]_{m\ x\ Q}$ with elements $t_{mq} = t_q^m$.

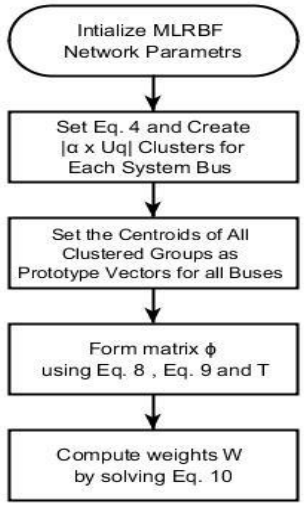

Figure 4.1: Flowchart Describing Training Phase of MLRBF

4.3 Importance Sampling

Importance sampling is one of the most successful variance reduction techniques used in reliability evaluation of composite power systems [34]. Importance sampling changes probability density function of occurrences by emphasizing certain values of a random variable which have greater impact, when compared with others, on the estimation process of a target quantity.

Consequently, values which have more importance are sampled more often and the variance of the estimator is reduced faster. IS aims to select a probability density function different from the original to minimize variance of samples [35]. To be able to obtain the maximum performance from importance sampling, selected probability density function should be equal or close to optimum $f_{optimum}(.)$ which is initially unknown.

At this stage the CE method can be utilized for estimating the optimal, or at least close to optimal, reference parameters by minimizing the distance between the original sampling density and the optimal sampling density $f_{optimum}(.)$ iteratively.

Detailed technical information can be obtained from [36]. In following subsections general definitions regarding IS are presented and later implementation of CE method for power system reliability evaluation is described.

4.3.1 General Definition of Importance Sampling

Consider a power system with G_N generating stations. Also, assume that the system has J identical and independent units, each one with a capacity G_{nj} for each of N stations. Let u_{nj} be a vector which contains original probability of unavailability of all generation units in the system. Under this assumption, the analytical problem of evaluating the LOLP index can be described by the following equation:

$$LOLP = \frac{1}{M}\sum_{i=1}^{M} H(X_i) \quad (4.11)$$

Where X_i represents ith sample of M total samples and H represents the test function which takes value of 1 if sample X_i has loss of load and 0 otherwise. Under these assumptions IS can be applied to the system by using the new unavailability vector v_{nj} to calculate H_{IS}.

$$LOLP_{IS} = \frac{1}{M_{IS}}\sum_{i=1}^{M_{IS}} H_{IS}(X_i)W(X_i; m_{IS}; u; v) \quad (4.12)$$

The expression $W(X_i; m_{IS}; u; v)$ is called likelihood ratio. This value represents a necessary correction in the sampling process because of the changed unavailability vector v. In this study, $W(X_i; m_{IS}; u; v)$ is calculated using (4.13).

$$W(X_i; m_{IS}; u; v) = \frac{\prod_{j=1}^{N_G}(1-u_j)^{x_j^G}(u_j)^{1-x_j^G}}{\prod_{j=1}^{N_G}(1-v_j)^{x_j^G}(v_j)^{1-x_j^G}} \quad (4.13)$$

where x_j^G represents the availability of generation unit j. The main problem in this process is defining optimal v values to minimize the computation time. In this study CE algorithm is utilized for this purpose which is explained in following subsection.

4.3.2 CE Algorithm

In this subsection, CE algorithm to determine optimal unavailability (v) values for each generation unit is described. Detailed information about CE can be found in [37].

The algorithm used in this study converges to optimal v parameters using an iterative procedure. During each iteration, v parameters are updated by using predefined number of system state samples. CE algorithm contains 6 main steps. While optimal v parameters are estimated in steps (1-4), loss of load indices are calculated with IS-MCS in steps (5-6).

1- Define the initialization parameters as sample size used for each iteration N (e.g. 50,000), multilevel parameter p (e.g. between 0.01 and 0.1). Define v_0=u, t=1 and ϕ= L_{LMAX} where v represents updated unavailability vector, t iteration number, ϕ stopping criteria for performance function and L_{LMAX} maximum peak load of the system.

2- Generate a set of random samples of states X_1, $X_2 \ldots X_N$ from the densities f (., v_{t-1}). Evaluate the performance of selected states S= [S_1, S_2 …, S(X_N)] according to the selected performance function. Sort the performances of the states in an increasing order so that S

$_{[1]} \leq S_{[2]} \leq \ldots \leq S_{[XN]}$ and then compute the performance of state (p) quantile of the performances, $S_{[(p)XN]}$.

3- Set the $\phi_t = S_{[(1-p)\ XN]}$ provided that ϕ_t is less than ϕ otherwise set $\phi_t = \phi$. Evaluate the indicator function $H(X_i)$ such that $H(X_i) = 1$ if $S(X_i) > \phi_t$ otherwise $H(X_i)=0$.

4- Use the sample from step 2 to update the new unavailability vector

$$v_{tj} = 1 - \frac{\sum_{i=1}^{X_N} H_t(X_i) W_{i,t-1} X_{ij}}{\sum_{i=1}^{X_N} H_t(X_i) W_{i,t-1}} \tag{4.14}$$

Where

$$W_{i,t-1} = W(X_i; u; v_{t-1}) \tag{4.15}$$

5- If $\phi_t = \phi$ criteria has been satisfied, optimal parameters has been found otherwise increase iteration number as t=t+1 and go back to the step 2.

6- Calculate loss of load indices with the equation (12) by using the optimal parameters defined in step 5.

A flowchart is provided for a clear understating in explaining cross-entropy method at figure 4.2.

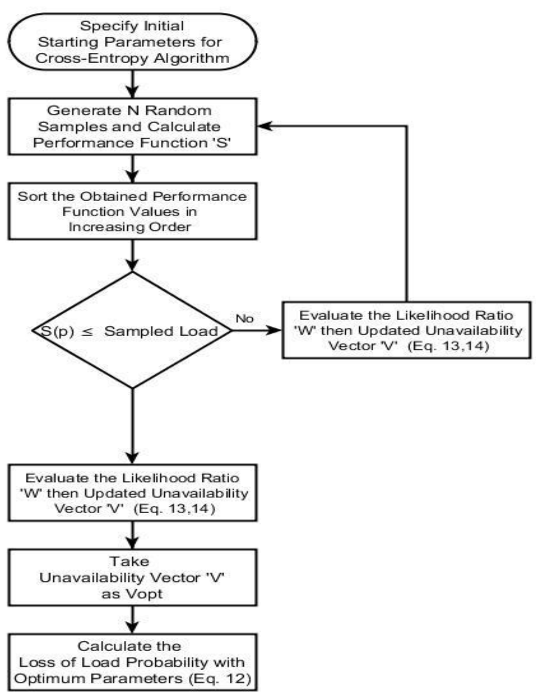

Figure 4.2: Flow Chart of Importance Sampling

4.4 Application Procedure of the Proposed Method

In this section application procedure of the proposed method is explained. This novel approach calculates reliability indices of composite power systems by combining multi label classifier and importance sampling technique within the framework of Monte Carlo Simulation (MCS). Generally, the most time-consuming part of composite system reliability evaluation is the optimal power flow analysis (OPF). This approach proposes use of a faster Multilabel classifier instead of OPF analysis after proper training. The proposed method can be applied while using either non-sequential or sequential MCS, however, non-sequential approach is chosen to illustrate the performance in this study because of simpler architecture. A benchmark created with Crude Monte Carlo Simulation (CMCS) analysis is also provided for comparison purposes.

The first step of applying combination of IS and ML classification is determining optimum unavailability vector via CE method as described in section B. After optimum unavailability vector is created then multi label classifier is trained and tested to use state space classification for evaluating the reliability indices of composite power system. Detailed implementation of multi label classifier is described below in two subsections defined as training and testing process.

4.4.1 Training Process

Training data samples for this study are created through MCS. The generation and demand information for each bus of selected sample states is used to create input vector while status of each bus for those states are recorded as target vector as shown in (1-2). To increase the training performance, input vector variables are normalized between -1 and 1. In this step, failure and non-failure status observed for each bus are labeled as 1 and -1 respectively. It should also be noted that an equal number of success and failure states is chosen to create training dataset to prevent overtraining.

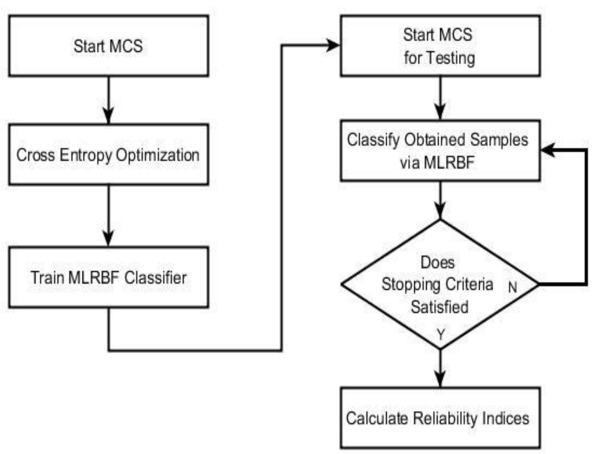

Figure 4.3: Overall Diagram of Proposed Method

4.4.2 Testing Process

In the testing process, MCS with importance sampling is used for generating random samples and these states are classified by multi label classifier until simulation reaches a previously determined stopping criterion. Stopping criteria for this study is defined as coefficient of variation (COV) to represent the estimated uncertainty.

Reliability indices are calculated and a comparison is made with the results obtained from Crude MCS and IS benchmarks. Reliability indices are evaluated based on Loss of Load Probability (LOLP). Complete flowchart of the proposed methodology is presented in figure 4.3.

4.5 Case Studies and Results

IEEE Reliability Test System (RTS) is chosen for demonstration of the proposed method. Two case studies are implemented to demonstrate the performance of the proposed method. In the first case study, load level of RTS is considered constant at its annual peak. In the second case study hourly load data of RTS is divided to 5 different load levels by considering their occurrence probabilities similar to [38]. After the application procedure described in previous section is completed, performance comparison of the proposed multi label classifier is made with the results obtained from CMCS benchmark and standard importance sampling process after calculation of reliability indices for all system buses. Since system losses caused from transmission line failures are much lower than the ones occurring from generation unit failures, states of transmission lines are considered as available at all the time. The capacities of transmission lines are, however, considered. Initial parameters of cross entropy optimization are determined as sample size N=50000, multilevel parameter p=0.05.

All simulations of this study are performed using MATLAB (2012) planform on a PC with Intel Core i7-4510 CPU (~2.6GHz), 16 GB Memory. It should be noted that the results presented below are the average of the 10 simulations.

4.5.1 Case Study 1

In this case study, load level of RTS is considered to be constant and at its peak value of 2,850W. To train MLRBF classifier 6000 samples are selected of which 3000 are success states and the rest are failure states. Clustering rate α is chosen as 0.25 for this case study which means number of clusters created is equal to one fourth of times of total failures for each bus as described in (6). After training of MLRBF classifier is completed, MCS is simulated until COV reaches 1% for all simulation types as it specified for stopping criteria of testing phase. In this process a total of 109,743 samples were obtained with 100459 characterized as successes and 9284 as failures. The comparison of obtained results is made in Table 4.1 and 4.24. Simulation results for overall system classification performance and time comparison of MLRBF Classifier on RTS are presented in

Table 4.1. The simulation results for bus level classification performance are stated in Table 4.2. In this study, performance comparison in this study is made based on Loss of Load Probability (LOLP). Table 4.1 shows that MLRBF classification method can compute overall LOLP of RTS with a small fraction of error and computational time required to evaluate the reliability indices can be significantly reduced by proposed MLRBF – IS combination when compared to standard MCS methods.

Table 4.1: Comparison on Overall Performance Analysis and CPU Time Spend

Algorithm	Success States	Failure States	LOLP $x\ 10^{-2}$	CPU Time (Sec $x\ 10^{3}$)
CMCS	100459	9284	8.46	5.49
CEIS	18786	1748	8.51	1.03
ML-CEIS	21305	1988	8.53	0.069

Table 4.2 also shows that proposed method shows reasonably accurate classification on characterizing the failed bus states of RTS.

Table 4.2: Comparison of Classification Performance at Bus Level Based on LOLP

Location	CMCS (LOLP) $x\ 10^{-2}$	CEIS (LOLP) $x\ 10^{-2}$	ML-CEIS (LOLP) $x\ 10^{-2}$
Bus 1	2.67	2.71	2.82
Bus 2	2.69	2.74	2.84
Bus 3	0.00	0.00	0.00
Bus 4	0.00	0.00	0.00
Bus 5	0.87	0.84	0.89
Bus 6	0.03	0.05	0.06
Bus 7	4.61	4.72	4.75
Bus 8	0.59	0.6	0.68
Bus 9	0.00	0.00	0.00
Bus 10	0.02	0.01	0.04
Bus 11	N/A	N/A	N/A
Bus 12	N/A	N/A	N/A
Bus 13	0.43	0.47	0.52
Bus 14	0.20	0. 24	0.29
Bus 15	0.00	0.00	0.00
Bus 16	0.05	0.1	0.15
Bus 17	N/A	N/A	N/A
Bus 18	0.94	1.1	0.81
Bus 19	0.05	0.12	0.02
Bus 20	3.53	3,59	3.76
Bus 21	N/A	N/A	N/A
Bus 22	N/A	N/A	N/A
Bus 23	N/A	N/A	N/A
Bus 24	N/A	N/A	N/A

4.5.2 Case Study 2

In this case study load level of the system is chosen randomly from the original load data of RTS. There are 8736 different load levels specified in annual hourly load values of RTS. As in the first case study, samples used for training MLRBF classifier are obtained through MCS. Since size of classifier is one of the most determining factor in classification time, multiple classifiers are trained for different load levels to handle this problem efficiently instead of training one large network.

For this purpose, five different thresholds are defined based on available power. For each level a unique classifier is trained. In testing stage, a decision tree is used to determine which classifier to be used for classification for every random sample. The overall diagram of the algorithm used in this case study is given in figure 4.4.

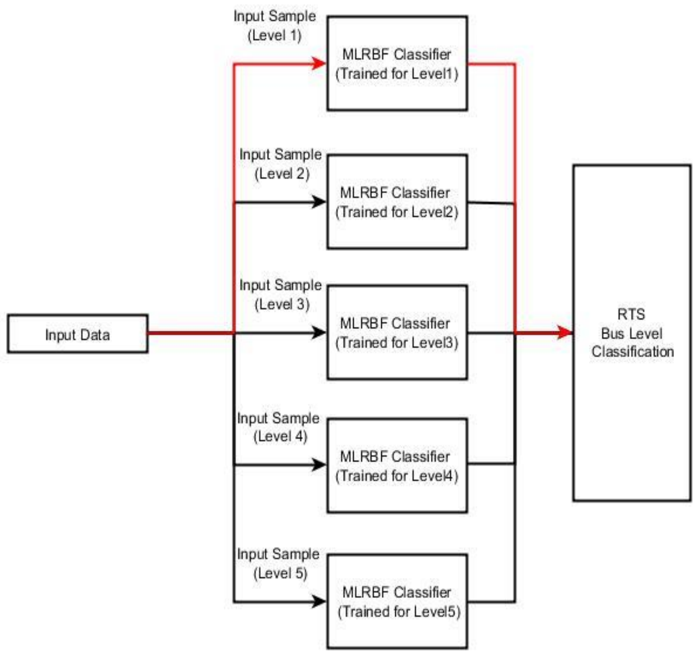

Figure 4.4: Overall Diagram of Proposed Classifier for Variable Load Level

For this case study each level is trained by 10000 samples which are obtained through MCS sampled with optimal unavailability vector obtained through the CE process. Similar to the first case study, clustering rate α is selected as 0.25 in this process.

After training is completed, MCS is run until COV reaches 1% as in first case study. In this process a total of 7,442,879 samples were obtained by CMCS with 7,433,754 characterized as successes and 9,125 as failures. The comparison on classification performance and simulation time for obtained results is presented in Table 4.3 and 4.4.

Comparison on overall classification performance and simulation times for the proposed method is given in Table 4.3.

Table 4.3: Comparison on Overall Performance Analysis and CPU Time Spend for varying Load Levels

Algorithm	Success States	Failure States	LOLP x 10^{-3}	CPU Time (Sec x 10^{3})
CMCS	7433754	9125	1.20	376.690
CEIS	473232	1740	1.23	23.710
ML-CEIS	496277	1817	1.23	1.110

It is clear from Table 4.3 that the proposed method can classify failure states of RTS in multi load level with a close performance. It is also observed in Table 4.3 that the proposed method provides a huge boost in terms of calculation time. Table 4.4 shows performance of the proposed method for bus level classification. It can be observed from the results that the proposed method can compute the LOLP accurately.

Table 4.4: Comparison of Classification Performance at Bus Level Based on LOLP for varying Load Levels

Location	**CMCS (LOLP) x 10^{-5}**	**CEIS (LOLP) x 10^{-5}**	**ML-CEIS (LOLP) x 10^{-5}**
Bus 1	28.06	28.7	33.4
Bus 2	15.35	16.43	17.51
Bus 3	0	0.0	0
Bus 4	0	0	0
Bus 5	8.04	8.42	9.14
Bus 6	0.49	0.43	0.21
Bus 7	54.95	54.02	57.42
Bus 8	2.31	1.90	3.28
Bus 9	0	0	0
Bus 10	0	0	0
Bus 11	N/A	N/A	N/A
Bus 12	N/A	N/A	N/A
Bus 13	3.58	3.72	5.97
Bus 14	3.28	3.10	4.51
Bus 15	0	0	0
Bus 16	0.58	0.99	1.2
Bus 17	N/A	N/A	N/A
Bus 18	10.81	12.26	19.58
Bus 19	2.18	2.21	3.84
Bus 20	60.47	61.28	64.91
Bus 21	N/A	N/A	N/A
Bus 22	N/A	N/A	N/A
Bus 23	N/A	N/A	N/A
Bus 24	N/A	N/A	N/A

4.6 Summary

In this study, a new method is presented to evaluate reliability indices for composite power systems. The proposed method uses an MLRBF classifier to identify status of buses in a way that does not require OPF analysis during Monte Carlo Simulation. The effectiveness of the proposed method is demonstrated on the IEEE RTS.

As can be observed from the results, MLRBF classifier can classify loss of load states with good accuracy most of the times. It should also be noted that rate of classification error increases in states with low frequency of failures. The main reason of this performance loss for those buses is lack of adequate samples in the training dataset. Performance of the proposed method can be increased by adding more samples to the training dataset as a natural outcome.

The main advantage of the proposed method is the ability of reducing the time for reliability analysis considerably which is shown in two different case studies.

5. Deep Learning for Composite Power System Reliability Evaluation

5.1 Introduction

In reliability analysis there is focus on searching the system state space for states that represent events of interest, like failure of the system to meet the required demand for a set of specific nodes. This indicates a need for methods that efficiently determine states to be examined and then evaluated. Artificial Intelligence (AI) based methods have been studied for this purpose either in themselves or in conjunction with widely used methods like Monte Carlo Simulation (MCS). In recent years, deep learning techniques have received considerable attention and showed significant promise in many fields when compared to other AI techniques. In this section a novel methodology based on combination of deep Convolutional Neural Networks (CNN) and MCS is presented for evaluation of composite power system reliability. This approach is applied to the IEEE Reliability Test System (RTS) by using both AC and DC power flow models for different load levels. The case studies show that the proposed method has a superior performance in both classification accuracy and reducing computational burden of reliability evaluation compared to previous AI based studies.

5.2 Convolutional Neural Networks

Recently, deep learning algorithms have drawn significant attraction in the area of artificial intelligence. This terminology is basically an extension of traditional artificial neural networks (ANN). These algorithms have dramatically improved the state of the art in areas like speech recognition, visual object recognition, object detection and many other domains such as drug discovery and genomics. Deep learning allows computational models that are composed of multiple processing layers to learn representations of data with multiple levels of abstraction. Their immense capability of learning optimal features from raw input data allows avoiding feature engineering. Through these algorithms, pattern classification performance of machines increased even more than humans in some applications [39, 40]. This section explores a new method for composite system reliability evaluation with combination of CNNs and MCS by considering both DC and AC approaches.

CNN is a deep feed forward artificial neural network algorithm which is one of the most used architectures among deep learning methods. It can simply be described as neural networks that use convolution in place of general matrix multiplication in at least one of their layers. CNNs are inspired by research done on the visual cortex of mammals and how they perceive the world using a layered architecture of neurons in the brain, and the overall architecture is reminiscent of the LGN–V1–V2–V4–IT hierarchy in the visual cortex ventral pathway [41,42]. The CNNs can encode certain properties into the architecture which results in less feature engineering requirements compared to other algorithms. Also, CNNs are easier to train and have much fewer parameters than fully connected networks with the same number of hidden units for this reason. In CNNs multi-level neural networks are trained with less neuron requirements. The ability to characterize system features in its own system makes CNNs more suitable for many pattern recognition problems [43].

A typical CNN architecture consists of three stages including convolutional layers, pooling layers and fully connected layers. Input data is sampled into smaller sized feature maps by filters in convolution layer. This process is done by computing the dot product between the entries of the filter and the input. Then pooling layers are applied to reduce the size of the data obtained in convolutional layer. This is followed by connected layer. In this layer activation function is applied to the features gathered in the previous layers, as seen in regular neural networks. At the end, predictions for trained classes can be obtained by applying a SoftMax function. Remaining of this section describes basic concepts of CNNs in subsections. Rigorous theoretical explanation on CNNs can be found in [44,45].

5.2.1 Convolutional Layer

Typically, in convolutional layers input data is applied to a convolutional operation to transfer the results to the next layer. In convolutional network terminology, the first argument to the convolution is often referred to as the input, and the second argument as the kernel or feature map. In usual convolution process, the kernels have flipped relative to the input. This process is not necessary in neural network implementations. Instead, many neural network algorithms implement a related function called as cross-correlation, which has a similar process with convolution but without flipping the kernels. Cross-Correlation operation has been described in eq. (5.1) for a 1-dimensional input.

$$(Input * Kernel)(i) = \sum_{m} Input(i+m)Kernel(m) \qquad (5.1)$$

5.2.2 Rectified Linear Unit (ReLU)

In the standard way of modeling, neuron's output in a neural network can be described with either tangent hyperbolic function (tanh) or sigmoid function (sigm). In terms of training time with gradient descend these saturating nonlinearity functions consume much more time when compared to non-saturating nonlinearity function, ReLU. This function can be replaced with previous functions used for increasing the nonlinear properties of the decision function without affecting the receptive fields of the convolution layer significantly. Usage of ReLU is also helpful to alleviate the vanishing gradient problem, which is the issue where the lower layers of the network train very slowly because the gradient decreases exponentially through the layers [46]. ReLU function is described in eq (5.2).

5.2.3 Pooling Layer

Pooling is an important concept used in CNNs. Although classification can be achieved without implementing any pooling, this process is commonly used in CNNs. The pooling layer is useful in reducing the number of parameters and amount of computation in the network. Briefly a pooling layer is summarizing the outputs of neighboring groups of neurons in the same kernel map. This process is usually done by using one of the several non-linear functions. Max pooling function which is the most common function used in pooling, is chosen in this study.

5.2.4 Dropout Technique

Deep neural networks are strong classification tools though those architectures, especially the ones consisting of large number of parameters, suffer from a serious problem called overfitting. Overfitting describes an incorrect optimization problem for an artificial intelligence model, where the weights are too closely trained for a set of data, and this may result in false positive characterization. Combining the predictions of many different neural nets is a very successful way to handle this problem but this solution could become very expensive in terms of computational effort. The technique called as “dropout” is proposed to deal with this problem by combining different models in a very efficient way which only costs about a factor of two during training.
The main idea in this technique is randomly dropping neurons from the neural network during training stage with a probability of 0.5. Dropped neurons do not participate in the forward and back-propagation stages. So, every time an input is processed, the neural network basically samples a different architecture, but all architectures share weights. In the test stage, all neurons are used but their outputs multiplied by 0.5, which is a reasonable approximation of geometric mean of the predictive distributions produced by dropout technique. In this way, expected value of an output neuron can be in same range as in the training stages [47].

5.2.5 Fully Connected Layer

Neurons in a fully connected layer have connections to all activations in the previous layer, as seen in regular neural networks. Their activations can hence be computed with a matrix multiplication followed by a bias offset.

5.3 Implementation of CNN in Composite System Reliability Evaluation

In this section, a new approach based on CNNs and MCS is presented for evaluating composite power system reliability indices. The main motivation in this study is designing a new methodology for reducing computation time of traditional MCS for evaluation of composite power system reliability indices using either AC or DC power flow model and including transmission line failures and capacity constraints.

Typically, in MCS, reliability indices of a power system are computed in three main stages. First, states of power system components are sampled based on their probability distributions and then the resulting system states are characterized as system success or failure using a power flow model. Finally, calculation of desired reliability indices is done based on previously characterized sample states. Characterization of power system states is usually done by a power flow analysis. This process can be considered as a linear or nonlinear programming problem depending on power flow model chosen for an application. Performing a power flow analysis for every sample obtained by MCS can create a significant computational burden especially in large composite power systems, especially with high reliability. Systems with high reliability will need a significant sample size for converging with acceptable accuracy.

The proposed method aims to reduce amount of power flow analysis by using CNN as a pre-classifier. To achieve this a CNN is designed to characterize overall system status as success or failure by using the information gathered from sampled states of system components as an input. Then afterwards, power flow analysis is applied only for samples classified as failure states.

In this study, CNN architecture is implemented to classify system status obtained from both DC and AC power flow analysis. To increase efficiency of classification and reduce to time spent for obtaining training samples a heuristic algorithm is used for generating required training samples. This method can create much more detailed datasets in a considerably shorter time.

Non-sequential MCS approach is used to analyze performance of the proposed method because of simplicity of model. However, the characterization process is the same in the sequential MC and therefore there is no additional difficulty in using it for sequential MC. The proposed method is explained in detail in three subsections.

5.3.1 Creating Proper Training Datasets

Creating a proper training dataset is one of the most important aspects that affects the classification performance of an AI based classifier.

The first step in creating a training dataset is to decide its size. Optimal size of a training dataset for an ANN commonly depends on parameters like input data, number of classes or number of neurons in network and varies a lot based on the application. In power system reliability evaluation, AI based classifiers are generally used to increase the time efficiency of MCS. For this reason, the size of training dataset is determined proportional to the sample size used in MCS.

One of the most common challenges in creating training dataset is called the class imbalance problem. This problem usually occurs when one or more target classes in a dataset are underrepresented (minor class) in comparison with the other classes (major class). Previous studies on this problem show that the negative effect of this problem in classification accuracy can be significantly reduced by applying down-sampling to the major class as well as applying an over-sampling to the minor class [48].

In a typical power system, occurrences of failure events (minor class) are much lower when compared to success events (major class). This feature of power systems creates a class imbalance problem and prevents effective training of AI based classifiers. The majority of previous studies using AI based classifiers for characterization of power system states were using unproportioned training datasets by simply reducing the amount of the success states.

In this study, a three-step heuristic algorithm is used to deal with class imbalance problem in power system state classification. This algorithm creates a training dataset which includes failure and success states in an equal proportion by applying down-sampling to major class (success) and over-sampling to the minor class (failure). It can be applied to any power system and AI based classifier without adding considerable computational cost. The proposed algorithm consists of three main steps which are described in the following.

-First, size of the training dataset is determined. The most common parameter to determine number of samples used in MCS is coefficient of variation (COV). Typically, an acceptable value of the estimate of uncertainty is determined as stopping criteria before the simulation and then MCS is run until the stopping criteria is satisfied. Calculation of COV, presented as β is formulated below.

$$\beta = \frac{\sqrt{V\big(E(F)\big)}}{\big(E(F)\big)} \tag{5.2}$$

In this equation $V\big(E(F)\big)$ represents the variance of estimated value E(F) and its square root is the standard deviation.

A looser stopping criterion is used in terms of COV for this purpose (e.g. 10%). Normally, power flow analysis is required in system characterization of each sample obtained. In this step, the system is failed if total active power generation is less than total active load to avoid additional computational burden of power flow analysis. When COV is reached to the previously described stopping criteria the total number of obtained samples is taken as the size of training dataset.

-In the next step, power flow analysis is performed for a small portion of determined dataset size (e.g. 5%). The results obtained in this step are used for calculating the proportion of classes. If the number of success states obtained in this process is higher than the number of failure states unavailability rate of all components is increased by multiplying a constant step size Δw. The process is repeated until an equal amount of success and failure states obtained by considering a tolerance.

-Finally, a training dataset is created in size calculated in the first step by using unavailability values obtained in previous step.

5.3.2 Designed CNN Architecture

In this study a CNN architecture is presented for state characterization within the MCS process for composite power system reliability calculations. The proposed architecture is designed to consider all aspects of a composite power system to characterize a sampled system state as success or failure based on the presence of active load curtailment. Due to the different characteristics of the generation and transmission line a multi input CNN is preferred to extract system features more efficiently. Considering the correlation between active and reactive power, generation information is included in the network in the same branch but in different channels. Then transmission line capacities are added as a second input branch.

Most of the current methods use DC power flow, which is often called OPF, to evaluate the reliability of composite power systems. OPF consists of a series of approximations in the usual power flow equations which reduces the problem down to a set of linear equations that would be normally represented by nonlinear equations. OPF can find the optimum solution for a power system state significantly faster when compared to AC flow model and this feature makes it very suitable for power system reliability analysis. However, OPF ignores the effects of the voltage and reactive power constraints on the reliability indices. For this reason, reliability analysis performed by using OPF can be considered optimistic and can be different when compared to AC flow analysis. In other words, some states recognized as success states by using the OPF, may be characterized as failure by AC model since the failures of these states is usually because of voltage and reactive power limit violations. For this reason, AC flow model is also considered in state characterization of samples in this study. Extensive theoretical explanations and comparison for both models can be found in [49-50]. Details of designed CNN architecture are described in the following subsections.

5.3.2.1 Proposed CNN Model

In this subsection the designed CNN architecture is described. It is essential to form the inputs of neural network by considering the required information to classify patterns. Available generation capacity and available transmission line capacities are used to create input vectors. Since DC flow model considers only active power, one channel input is created with generation and demand information of a sampled state. For the AC flow model, reactive power information is included to input of classifier as an additional channel. Furthermore, transmission line information is added to the network in a second branch to be able to classify the system failures caused by transmission lines.

The first input branch of the network is created by using the information of available generation and demand at each bus of composite power system. Input data representing the active generation is described by equation (5.4).

$$Input_{i,1} = G_i - D_i \tag{5.3}$$

Where, G represents maximum available generation and D for demand at bus i.

The second channel used for classifying the states of AC flow model contains the information of maximum and minimum reactive power generation that can be produced from each bus of power system. The reactive power demands for that bus input data of second channel is described in equation (5.4).

$$Input_{i,2} = Qmax_i - Qmin_i - D_{qi} \quad (5.4)$$

where Q_{max} and Q_{min} represent maximum and minimum limits of reactive generation, and D is demand at bus i. Finally, the transmission capacities are included in the network as raw data in the second branch as a vector. After the information is prepared a min-max normalization is applied to all input vectors.

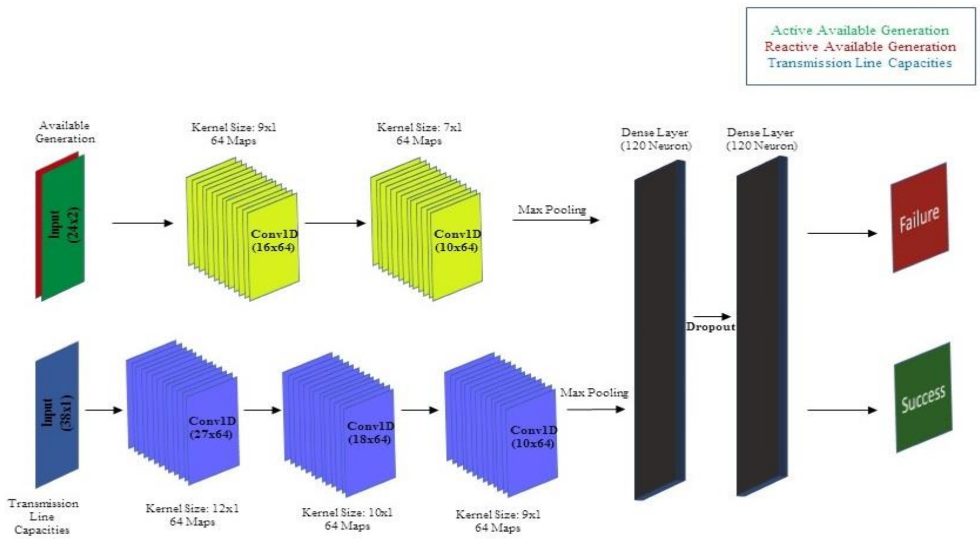

Figure 5.1: Overall Diagram of Proposed CNN Structure

The designed architecture consists of two main input branches for generation and transmission line information.

Two convolutional layers following a one max-pooling layer are applied for the first branch. In the first convolution layer, input is extracted to low level feature maps by using 64 kernels with size of 9x1. Then second layers of convolution are applied to gather more detailed features by using 64 kernels 7x1. Following convolutional layers, a max-pooling layer is applied to the extracted features.

As for the second branch, transmission line features are extracted by three convolutional layers following a one max-pooling layer. In the convolution layers input is extracted to low level feature maps by using 64 kernels with size of 12x1, 10x1 and 9x1 respectively. Following convolutional layers, a max-pooling layer is applied for reducing the sample size.

The obtained features are applied to two fully connected layers. Each of these layers includes 120 neurons. One layer of dropout is used with proportion of 0.5 between fully connected layers. At the end a SoftMax function is used to accomplish binary classification. A diagram of designed CNN architectures is presented in Fig 5.1 for a clearer understanding. Moreover, application procedure of the proposed deep CNN and MCS combination demonstrated in a flowchart in Fig 5.2.

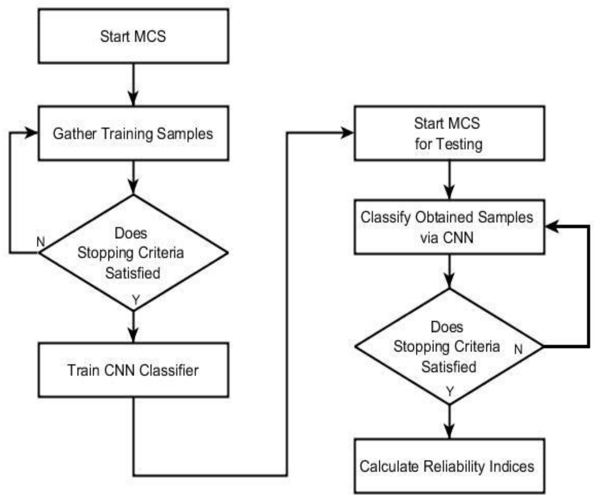

Figure 5.2: General Structure of Proposed Method

5.4 Case Studies

Performance of the proposed method for both AC and DC flow models is illustrated using three different case studies for constant peak-load level, varying load level and considering failures caused by insufficient transmission line capacities respectively. Case studies are described in detail in the following.

5.4.1 Case 1

In this case, the proposed method is tested on single area IEEE RTS for single load level of 2850 MW (annual peak). There are 10 generation buses in RTS which are considered as input vector. To train the classifier, initial COV for training dataset creator algorithm is chosen as 10%. After obtaining adequate number of samples, training of the proposed CNN is completed within 100 iterations. After the classifier is successfully trained, the proposed model is run for computing indices until COV reaches the stopping criterion threshold $\leq 1\%$ for both AC and DC flow models.

After the calculation phase is completed, 110826 samples are classified by DC-OPF with 101470 successes and 9356 failure states. Similarly, 53194 samples are classified by AC-OPF with 44643 successes and 8551 failure states.

Table 5.1 presents the comparative results between CMCS (Crude Monte Carlo Simulation) and the proposed approach for DC flow model while Table 5.2 provides results of similar analysis for AC flow model. Performance indices are calculated as described in eq (8-10).

Table 5.1: Performance Comparison of CNN-MCS & CMCS for Constant Load (DC Flow Model)

	CMCS	CNN
Success States	101470	101460
Failure States	9356	9342
LOLP	0.084	0.084
Sensitivity	N/A	0.99
Specificity	N/A	0.99
Analysis Time (Hr)	1.47	0.05

It is clear from Table 5.1 that the proposed method shows an outstanding performance for both classification accuracy and reliability evaluation with significantly reduced computational effort for DC flow model.

Table 5.2 shows the classification performance of the proposed method for AC power flow. The proposed CNN classifier can characterize the system states with a small error rate in terms of both classification accuracy and reliability evaluation. It should be noted that comparative computational time for AC power flow model is significantly reduced compared to DC flow model since this model requires nonlinear programming to solve AC flow equations. It should be noted that the time for AC analysis is approximately 13 hours in CMCS where as in the CNN approach it is less than two minutes.

Table 5.2: Performance Comparison of CNN-MCS & CMCS for Constant Load (AC Flow Model)

	CMCS	CNN
Success States	44643	44648
Failure States	8551	8523
LOLP	0.160	0.160
Sensitivity	N/A	0.99
Specificity	N/A	0.99
Analysis Time (Hr)	12.92	0.031

5.4.2 Case 2

In this case, the system is tested on single area IEEE RTS for the original annual hourly load data. All 24 buses of RTS are considered as input vector as described in eq (6-7). To train the classifier, initial COV for training dataset creator algorithm is chosen as 10%. After obtaining adequate number of samples, training of the proposed CNN is completed in 200 iterations. After the classifier is successfully trained, the proposed system is run for calculation phase until COV reaches the limit of $\leq 1\%$ as stopping criteria for both AC and DC flow models. After calculation phase is completed, 7466186 samples are classified by DC-OPF with 7457892 successes and 8294 failure states. Similarly, 5701891 samples are classified by AC-OPF with 5692084 successes and 9807 failure states.

Table 5.3 presents the comparative results between CMCS and the proposed approach for DC flow model while Table 5.4 provides results of similar analysis for AC flow model. Performance indices are calculated as described in equations (8-10).

It is clear from Table 5.3 that the proposed method shows very good accuracy in classification of system states and predict reliability indices with a small error rate. Table 5.3 also shows that proposed method can significantly reduce the computational effort for DC flow model.

Table 5.3: Performance Comparison of CNN-MCS & CMCS in Varying Load (DC Flow Model)

	CMCS	**CNN**
Success States	7457892	7457973
Failure States	8294	8212
LOLP	0.0011	0.0011
Sensitivity	N/A	0.99
Specificity	N/A	0.99
Analysis Time (Hr)	98.5	2.92

Table 5.4 shows that the proposed method has an outstanding performance for both classification accuracy and reliability evaluation with significantly reduced computational effort in AC flow model.

Table 5.4: Performance Comparison of CNN-MCS & CMCS in Varying Load (AC Flow Model)

	CMCS	**CNN**
Success States	5692084	5692198
Failure States	9807	9693
LOLP	0.0017	0.0017
Sensitivity	N/A	0.99
Specificity	N/A	0.99
Analysis Time (Hr)	1599.3	2.54

5.4.3 Case 3

In this case study, performance of the proposed system is tested while transmission line failures are considered. The system load level is considered constant at 2850 MW (annual peak). After the classifier is successfully trained with the same method as described in previous sections, the proposed system is used to compute indices until COV reaches the stopping criteria of $\leq 1\%$ for both AC and DC flow models. After testing is completed, 110452 samples are classified by DC-OPF with 101065 successes and 9387 failure states. Similarly, 53640 samples are classified by AC-OPF with 44532 successes and 9108 failure states.

Table 5.5 presents the comparative results between CMCS (Crude Monte Carlo Simulation) and the proposed approach for DC flow model while Table 5.6 provides results of similar analysis for AC flow model. Performance indices are calculated as described in equations (8-10).
It is clear in Table 5.5 that the proposed method shows very good accuracy in classification of system states and estimate reliability indices with a small error rate and significantly reduced time.

Table 5.5: Performance Comparison of CNN-MCS & CMCS for Line Failures (DC Flow Equation Model)

	CMCS	**CNN**
Success States	101065	101079
Failure States	9387	9354
LOLP	0.084	0.084
Sensitivity	N/A	0.99
Specificity	N/A	0.99
Analysis Time (Hr)	1.46	0.05

Table 5.6 shows that the proposed method shows high performance in classification accuracy and dramatically reduces computational effort for AC flow model compared to CMCS. Both tables clearly state that proposed multi input CNN structure shows outstanding performance in reliability evaluation of a complete composite system.

Table 5.6 : Performance Comparison of CNN-MCS & CMCS for Line Failures (AC Flow Equation Model)

	CMCS	**CNN**
Success States	44532	44546
Failure States	9108	9064
LOLP	0.170	0.169
Sensitivity	N/A	0.99
Specificity	N/A	0.99
Analysis Time (Hr)	12.71	0.029

5.5 Summary

A multi input CNN structure is proposed to increase the computational efficiency of MCS for evaluating composite power system reliability.

It is critical for any binary classification problem to create a balanced dataset. In this study, a new heuristic approach is proposed to create a proper dataset by applying down-sampling to success states as well as applying over-sampling to failure states. The proposed algorithm can be applied to any artificial intelligence-based method without any additional computational cost.

The proposed method is tested for both DC and AC flow models by using one and two channel CNN structure respectively. Generation and demand information of a sampled state is utilized to create input vector of the classifier.

First two case studies are used to demonstrate the performance of the proposed MCS-CNN combination based on constant load and hourly load models respectively. Third case study shows the performance of the proposed approach on system states with transmission line failures. Sensitivity and Specificity parameters are used to show classification performance of CNN classifier while LOLP is chosen as metric to demonstrate reliability evaluation of the proposed method.

Results show that the proposed method can accurately characterize system states for both DC and AC flow model for fixed peak load and varying hourly load and therefore reliability indices can be evaluated without significant error and with a significantly reduced simulation time.

It can be seen from simulations that AC and DC models can give quite different results on the same test system. Computational power required for classification using AC flow model is much higher than DC flow model since AC flow equations requires nonlinear programming techniques while DC flow equations can be solved with linear techniques. It is shown that proposed approach reduces the simulation time to complete a reliability analysis based on AC flow model to the same level as OPF based analysis using DC flow.

Finally, simulations show that designed multi input CNN architecture can successfully characterize system states with transmission line failures in order to increase computational efficiency even further.

It should also be pointed out that deep learning here is used to replace the optimal power flow based on DC or AC equations. The objective is to reduce the time of characterization of a state since state characterization is the major part of computational burden. Although we have used the basic MC, this proposed method can be also used with MC and variance reduction techniques like Importance Sampling. The combination would obviously reduce the computational time even further.

6. Conclusions and Future Work

6.1 Conclusions

In this research, a novel method to evaluate reliability indices for composite power systems is introduced with combination of MLL and MCS. The proposed method is implemented for MLKNN and MLRBF classifiers to identify status of buses. The case studies show that the method significantly reduces the computational burden of MCS.

Additionally, this method advances the state of the art of using machine learning in power system reliability evaluation from the previous methods by including computation of bus indices and the transmission line failures.

Moreover, the work done shows that the proposed method can be combined with the well-known variance reduction technique of IS. The outcomes for this approach show this methodology improves time efficiency of MCS even further.

Finally, deep learning structures are investigated to evaluate composite system reliability evaluation through MCS. CNN, a well-known deep learning topology, is implemented to characterize sampled system states for both AC and DC flow models. The results show that computational efficiency for classifying using AC flow model is much higher than DC flow model since AC flow equations require nonlinear programming techniques while DC flow equations can be solved with linear techniques. The results obtained show that the proposed architecture performs state characterization with a high accuracy with COV equal or less than .01.

The importance of reducing the computational time can be understood by two examples. In Monte Carlo Simulation, the accuracy of convergence is very important. Convergence is measured by the COV, smaller COV means better convergence. Now the sample size (or computational time) is inversely proportional to COV or directly proportional to accuracy of convergence. The proposed method reduces the required time for reliability analysis considerably for the same level of accuracy defined by the coefficient of variation. Alternatively, for the same time this allows convergence to a higher level of accuracy. Another example is that for optimal planning of resources, reliability studies may need to be done many times. So, the reduction of computational time helps in optimal planning by being able to perform more studies in the same time.

This study demonstrates that the application of the proposed methods on composite power system reliability evaluation accurately determines the states status with a substantial speed up compared with OPF based Monte Carlo Simulation methods.

6.2 Future Work

This effort has focused on improving the computational efficiency of composite power system reliability evaluation by replacing the OPF by machine learning. The future efforts could be directed to:

1. Using the machine learning based OPF for other applications such as operational planning.
2. Using machine learning to enhance the process of dynamic security assessment as well as the OPF. Then perhaps the Time Domain Simulation could be used to train CNN to replace this time consuming operation. This could make the inclusion of security assessment in composite reliability assessment computationally efficient.

References

[1] Singh, Chanan, and Roy Billinton. *System reliability, modelling and evaluation*. Vol. 769. London: Hutchinson, 1977.

[2] Singh, C., and P. Jirutitijaroen. "Monte Carlo simulation techniques for transmission systems reliability analysis." *A Tutorial Paper presented at IEEE Power Engineering Society General meeting, Tampa, Florida*. 2007.

[3] Zhaohong, B., & Xifan, W. (2002). Studies on variance reduction technique of Monte Carlo simulation in composite system reliability evaluation. *Electric Power Systems Research*, *63*(1), 59-64.

[4] Singh, C., & Mitra, J. (1997). Composite system reliability evaluation using state space pruning. *IEEE Transactions on Power Systems*, *12*(1), 471-479.

[5] Saraiva, J. T., Miranda, V., & Pinto, L. M. V. G. (1995, May). Generation/transmission power system reliability evaluation by Monte Carlo simulation assuming a fuzzy load description. In *Power Industry Computer Application Conference, 1995. Conference Proceedings. 1995 IEEE* (pp. 554-559). IEEE.

[6] Miranda, V., de Magalhães Carvalho, L., Da Rosa, M. A., Da Silva, A. M. L., & Singh, C. (2009). Improving power system reliability calculation efficiency with EPSO variants. *IEEE Transactions on Power Systems*, *24*(4), 1772-1779.

[7] Samaan, N., & Singh, C. (2002). Adequacy assessment of power system generation using a modified simple genetic algorithm. *IEEE Transactions on Power Systems*, *17*(4), 974-981.

[8] Samaan, N., & Singh, C. (2003, July). Assessment of the annual frequency and duration indices in composite system reliability using genetic algorithms. In *Power Engineering Society General Meeting, 2003, IEEE* (Vol. 2, pp. 692-697). IEEE.

[9] Earla, R., Mitra, J., & Patra, S. B. (2004, August). A particle swarm based method for composite system reliability analysis. In *North American Power Symposium*.

[10] Wang, L., & Singh, C. (2008). Population-based intelligent search in reliability evaluation of generation systems with wind power penetration. *IEEE transactions on power systems*, *23*(3), 1336-1345.

[11] Singh, C., & Wang, L. (2008). Role of artificial intelligence in the reliability evaluation of electric power systems. *Turkish Journal of Electrical Engineering & Computer Sciences*, *16*(3), 189-200.

[12] da Silva, A. M. L., de Resende, L. C., da Fonseca Manso, L. A., & Miranda, V. (2007). Composite reliability assessment based on Monte Carlo simulation and artificial neural networks. *IEEE Transactions on Power Systems*, *22*(3), 1202-1209.

[13] Pindoriya, N. M., Jirutitijaroen, P., Srinivasan, D., & Singh, C. (2011). Composite reliability evaluation using Monte Carlo simulation and least squares support vector classifier. *IEEE Transactions on Power Systems*, *26*(4), 2483-2490.

[14] M. V. F. Pereira and N. J. Balu, "Composite generation/transmission reliability evaluation," Proc. IEEE, vol. 80, no. 4, pp. 470–491, Apr. 1992.

[15] W. Li and R. Billinton, "Effect of bus load uncertainty and correlation in composite system adequacy evaluation," IEEE Trans. Power Syst., vol. 6, no. 4, pp. 1522–1529, Nov. 1991

[16] A. C. G. Melo, M. V. F. Pereira, and A. M. L. da Silva, "A conditional probability approach to the calculation of frequency and duration indices in composite reliability evaluation," IEEE Trans. Power Syst., vol. 8, no. 3, pp. 1118–1125, Aug. 1993.

[17] F. F. C. Véliz, C. L. T. Borges, and A. M. Rei, “A comparison of load models for composite reliability evaluation by nonsequential Monte Carlo simulation,” IEEE Trans. Power Syst., vol. 25, no. 2, pp. 649–656, May 2010
[18] J. Mitra and C. Singh, “Pruning and simulation for determination of frequency and duration indices of composite power systems,” IEEE Trans. Power Syst., vol. 14, no. 3, pp. 899–905, Aug. 1999.
[19] R. Billinton and W. R. Li, Assessment of Electric Power Systems Using Monte Carlo Methods. New York, NY, USA: Plenum, 1994
[20] H. Kim and C. Singh, “Reliability modeling and simulation in power systems with aging characteristics,” IEEE Trans. Power Syst., vol. 25, no. 1, pp. 21–28, Feb. 2010.
[21] P. Wang and R. Billinton, “Time sequential distribution system reliability worth analysis considering time varying load and cost models,” IEEE Trans. Power Del., vol. 14, no. 3, pp. 1046–1051, Jul. 1999.
[22] Sankarakrishnan, A., and Roy Billinton. "Sequential Monte Carlo simulation for composite power system reliability analysis with time varying loads." *IEEE Transactions on power Systems* 10.3 (1995): 1540-1545.
[23] Billinton, R., et al. "Reliability assessment of composite generation and transmission systems." *IEEE Tutorial*, *IEEE Winter Power meeting*. 1990
[24] Benidris, Mohammed, and Joydeep Mitra. "Reliability and sensitivity analysis of composite power systems considering voltage and reactive power constraints." *IET Generation, Transmission & Distribution* 9.12 (2015): 1245-1253.
[25] Luo, X., A. D. Patton, and C. Singh. "Real power transfer capability calculations using multi-layer feed-forward neural networks." *IEEE Transactions on Power Systems* 15.2 (2000): 903-908.
[26] Force, RTS Task. "IEEE reliability test system." *IEEE Trans. on PAS* 98.6 (1979): 2047-2054.
[27] Tsoumakas, G., & Katakis, I. (2006). Multi-label classification: An overview. *International Journal of Data Warehousing and Mining*, *3*.
[28] Zhang, M. L., & Zhou, Z. H. (2007). ML-KNN: A lazy learning approach to multi-label learning. *Pattern recognition*, *40*(7), 2038-2048.
[29] Zhang, M. L., & Zhou, Z. H. (2006). Multilabel neural networks with applications to functional genomics and text categorization. *IEEE transactions on Knowledge and Data Engineering*, *18*(10), 1338-1351.
[30] Thabtah, F. A., Cowling, P., & Peng, Y. (2004, November). MMAC: A new multi-class, multi-label associative classification approach. In *Data Mining, 2004. ICDM'04. Fourth IEEE International Conference on* (pp. 217-224). IEEE.
[31] Wu, T. F., Lin, C. J., & Weng, R. C. (2004). Probability estimates for multi-class classification by pairwise coupling. *Journal of Machine Learning Research*, *5*(Aug), 975-1005.
[32] Bishop, Christopher M. Neural networks for pattern recognition. Oxford university press, 1995.
[33] Zhang, Min-Ling. "MLRBF Neural Networks for Multi-Label Learning." Neural Processing Letters 29.2 (2009): 61-74.
[34] da Silva, Armando M. Leite, Reinaldo AG Fernandez, and Chanan Singh. "Generating capacity reliability evaluation based on Monte Carlo simulation and cross-entropy methods." *IEEE Transactions on Power Systems* 25.1 (2010): 129-137.

[35] González-Fernández, Reinaldo A., et al. "Composite systems reliability evaluation based on Monte Carlo simulation and cross-entropy methods." *IEEE Transactions on Power Systems* 28.4 (2013): 4598-4606.
[36] Rubinstein, Reuven Y., and Dirk P. Kroese. *Simulation and the Monte Carlo method*. Vol. 10. John Wiley & Sons, 2016.
[37] T. Homem-deMello and R.Y Rubinstein, Estimation of Rare Event Probabilities Using Cross-Entropy, in Proc. Winter Simulation Conf., San Diego, CA, USA, Dec 2002, vol. 1, pp. 310-319.
[38] Luo, Xiaochuan, Chanan Singh, and Alton D. Patton. "Power system reliability evaluation using learning vector quantization and Monte Carlo simulation." *Electric Power Systems Research*66.2 (2003): 163-169.
[39] LeCun, Yann, Yoshua Bengio, and Geoffrey Hinton. "Deep learning." *nature* 521.7553 (2015): 436.
[40] Deng, Li, and Dong Yu. "Deep learning: methods and applications." *Foundations and Trends® in Signal Processing*7.3–4 (2014): 197-387.
[41] Hubel, David H., and Torsten N. Wiesel. "Receptive fields, binocular interaction and functional architecture in the cat's visual cortex." *The Journal of physiology* 160.1 (1962): 106-154.
[42] Felleman, Daniel J., and DC Essen Van. "Distributed hierarchical processing in the primate cerebral cortex." *Cerebral cortex (New York, NY: 1991)* 1.1 (1991): 1-47.
[43] Taigman, Y., Yang, M., Ranzato, M. & Wolf, L. Deepface: closing the gap to human-level performance in face verification. In *Proc. Conference on Computer Vision and Pattern Recognition* 1701–1708 (2014).
[44] Goodfellow, Ian, et al. *Deep learning*. Vol. 1. Cambridge: MIT press, 2016.
[45] Hu, Wei, et al. "Deep convolutional neural networks for hyperspectral image classification." *Journal of Sensors* 2015 (2015).
[46] Krizhevsky, Alex, Ilya Sutskever, and Geoffrey E. Hinton. "Imagenet classification with deep convolutional neural networks." *Advances in neural information processing systems*. 2012.
[47] Srivastava, Nitish, et al. "Dropout: a simple way to prevent neural networks from overfitting." *The Journal of Machine Learning Research* 15.1 (2014): 1929-1958.
[48] Chawla, Nitesh V., et al. "SMOTE: synthetic minority over-sampling technique." *Journal of artificial intelligence research* 16 (2002): 321-357.
[49] Wang, Peng, et al. "Reliability assessment of power systems considering reactive power sources." *Power & Energy Society General Meeting, 2009. PES'09. IEEE*. IEEE, 2009.
[50] Overbye, Thomas J., Xu Cheng, and Yan Sun. "A comparison of the AC and DC power flow models for LMP calculations." *System Sciences, 2004. Proceedings of the 37th Annual Hawaii International Conference on*. IEEE, 2004.

Part III

Using Importance Sampling in Monte Carlo Simulation – Computation Time Sensitivity Studies

Chanan Singh
Sai Kiran Kanchari Bavajigari, Graduate Student

Texas A&M University

For information about this project, contact

Chanan Singh, Regents Professor and Irma Runyon Chair Professor
Department of Electrical and Computer Engineering
Texas A&M University
3128 TAMU
College Station, Texas 77843-3128
Phone: (979) 845-7589
Email: singh@ece.tamu.edu

Power Systems Engineering Research Center

The Power Systems Engineering Research Center (PSERC) is a multi-university Center conducting research on challenges facing the electric power industry and educating the next generation of power engineers. More information about PSERC can be found at the Center's website: http://www.pserc.org.

For additional information, contact:

Power Systems Engineering Research Center
Arizona State University
527 Engineering Research Center
Tempe, Arizona 85287-5706
Phone: 480-965-1643
Fax: 480-727-2052

Notice Concerning Copyright Material

Table of Contents

List of Tables

1. Introduction

The Monte Carlo Simulation is one of the frequently used probabilistic methods for composite system reliability evaluation. One of the factors affecting the simulation time is the variance of the estimator. In an attempt to improve computational efficiency of MCS, several approaches have been used to reduce the variance of the estimators. These techniques include stratified sampling [8, 9], dagger sampling [10, 11], Latin Hypercube sampling [12, 13] and Importance sampling [14]. Of these approaches, importance sampling appears to be gaining more popularity recently. In this paper a Cross Entropy (CE) based Importance sampling method is investigated for its sensitivity range of computational efficiency. Reference [1] describes various approaches where CE method can be used.

In the cross entropy Monte Carlo simulation a secondary pdf is generated using Importance Sampling. Generating an optimal secondary pdf is critical as the variance may not be reduced if a proper pdf is not generated. The secondary pdf is used to calculate importance weights. The CE method provides an iterative procedure to generate the secondary pdf through which we calculate the Importance weights which help in reducing the variance and consequently computation time.

The CE method using non-sequential MCS has been implemented in Generation Capacity reliability (GCR) evaluation, where the system transmission constraints are ignored [2] and the method is tested using a fixed load model and a multilevel load model. A CE based Sequential Monte Carlo simulation method for GCR evaluation is implemented in [3], where time dependent systems are considered and a comparison is provided between different CE based and non-CE based Monte Carlo simulation algorithms. They are tested on an IEEE RTS 96 and a modified RTS 96 system. These papers show that CE method is a computationally improved method compared to simple Monte Carlo methods, as it reduces the computation time as well as the sample size.

Reference [4] implements the CE method in a composite power system model using non sequential Monte Carlo Simulation, where the indices are calculated for both single area and multi area power systems. Reference [5] implements the CE method using quasi sequential Monte Carlo methods, where renewable energy sources are integrated in the test system. The CE method has been improved in [6] by assuming the load to follow a Gaussian distribution and using a truncated Gaussian distribution for the load in the training phase instead of having a fixed load. Here a different mathematical model is used for DC OPF where instead of calculating the load curtailment, the excess load served is calculated. These are implemented on a single and multi-area reliability test systems.

A three stage CE IS method is implemented in [7] for degenerate cases. Here a third stage is employed before the normal CE algorithm to detect the degenerate parameters. A parallel cross entropy optimization method has been implemented in [15].

In all the references using the CE method to calculate the reliability indices of power systems, there is no sensitivity analysis which explores the limits of CE method under different conditions. This report provides a study of how changing the system parameters affects the reliability indices using a new parallel computing approach to sequential simulation.

The remaining of this part of report is organized as follows; Section 2 describes the Cross Entropy method, section 3 describes the CE Algorithm, section 4 discusses the parallel computing methodology, section 5 discusses the results, section 6 conclusion and finally appendices are provided at the end.

2. Cross Entropy Method

2.1 Introduction

Importance Sampling is a variance reduction technique, where using a secondary probability density $g^*(X)$, the rare events or failure events are sampled more frequently. As there is no direct procedure to generate the secondary probability density, an iteratively updated secondary density $f(X_i,v)$ is generated by distorting the original parameters. The Cross Entropy method [1] gives an adaptive iterative procedure to find the distorted parameters. In Importance Sampling the secondary density is chosen such that the distance between the optimal secondary density and iteratively updated secondary density is minimum. A particular measure of distance between the two distributions is the Kullback-Leibler distance (Appendix A), which is termed as the Cross Entropy between the optimal $g^*(X)$ and $f(X_i,v)$. The Cross Entropy based approach is an accelerated Monte Carlo approach which improves the computation efficiency.

The system State X_i is generated as $[X_G,X_L,X_{load}]$, which is a vector containing generator states , transmission states and load level. The X_G and X_L are calculated using the component unavailability vector $u = [u_G,u_L]$ where u_G and u_L are sub vectors for generation and transmission. The up/down states of generator and transmission lines are determined after generating random numbers for each component and comparing with its unavailability vector. The load is randomly generated from the load curve. The Reliability indices such as Loss of Load Probability (LOLP), Expected Energy not Supplied (EENS) are used for reliability assessment. For a random sample $X_1, X_2\ldots X_N$ generated from $[u_G,u_L]$ the reliability index calculated from a Monte Carlo simulation is given by

$$E(H) = \frac{1}{N}\sum_{i=1}^{N} H(X_i) \tag{2.1}$$

Where $H(X_i)$ is the test function for computing the reliability adequacy index.

For a system using Importance Sampling where rare events are sampled more often, the reliability index is calculated from the samples $X_1,X_2,\ldots.X_N$ generated from distorted $[v_G, v_L]$, with Likelihood ratio W given by

$$\boldsymbol{E(H) = \frac{1}{N}\sum_{i=1}^{N} H(X_i)W(X_i;u,v)} \tag{2.2}$$

2.2 Likelihood Ratio

In section 2.1 the expression W (X_i;u,v) is the likelihood ratio between the two probability density functions and is a correction factor introduced to avoid any biased estimates.

$$W(X_i; u, v) = \frac{f(X_i;u)}{f(X_i;v)} = \frac{\prod_{j=1}^{Nc}(1-u_j)^{x_j}(u_j)^{(1-x_j)}}{\prod_{j=1}^{Nc}(1-v_j)^{x_j}(v_j)^{(1-x_j)}} \qquad (2.3)$$

X_i=X_1, X2… X_N are random samples of generating states, u_j represents unit unavailability, v_j represents distorted unavailability. x_j represents availability of a component, with a value 1 if the component is available and 0 if not. Nc is the total number of components.
Here

$$W = W_G{*}W_L \qquad (2.4)$$

where,

$$W_G = \frac{\prod_{G=1}^{N_G}(1-u_G)^{x_G}(u_G)^{(1-x_G)}}{\prod_{G=1}^{N_G}(1-v_j)^{x_G}(v_G)^{(1-x_G)}} \qquad (2.5)$$

$$W_L = \frac{\prod_{L=1}^{N_L}(1-u_L)^{x_L}(u_L)^{(1-x_L)}}{\prod_{L=1}^{N_L}(1-v_L)^{x_L}(v_L)^{(1-x_L)}} \qquad (2.6)$$

W_G and W_L are the likelihood ratios of generators and transmission lines respectively. N_G, N_L are total number of generators and transmission lines. u_G, v_G, u_L, v_L are generator and transmission line original and distorted unavailability. x_G, x_L are the Generator and Transmission Lines states represented by 1 if up and 0 if down.

2.2.1 Distorted Parameters for the Sequential Simulation

The initial undistorted unavailabilities of the power system components are given by $u = \frac{\lambda}{(\lambda+\mu)}$. Where λ and μ are the component failure and repair rates respectively. During the Cross Entropy procedure, a distortion is applied to the unavailability and new distorted unavailability parameters are generated. So, during the sequential simulation to calculate the residence time of each state the new failure and repair rates, λ* and μ* rates generated from distorted parameters are used. The new distorted parameters [3] are given by,

$$\mu^* = \mu \qquad (2.7)$$

$$\lambda^* = \frac{v\mu^*}{(1-v)} \qquad (2.8)$$

To maximize the number of failure events in a time period the distortion is applied only to the failure rate.

3. Cross Entropy Algorithm

3.1 Training Phase

Step1: Initialize all the parameters such as N (number of samples), ρ (multi-level parameter), α (smoothing parameter), N_{max} (maximum sample size). Limiting Load (L_d)

Step 2: Define V_0 = u that is the initial undistorted unavailabilities of Generators and Transmission lines. Set t = 1 (iteration counter).

Step 3: Generate system states X_1, X_2... X_N from the initial unavailabilities according to the Bernoulli mass function.

Step 4: Evaluate the System Performance function P (X_i) for all X_i. A DC power flow analysis is performed and load curtailment is calculated. If any power flow violations occur then an optimization algorithm based on linear programming, described in Appendix B, is solved. P (X_i) is the sum of capacity of all the generators. If a load curtailment occurs then P (X_i) is recalculated as

$$P\ (Xi) = L_{max} - \text{load Curtailment} \tag{3.1}$$

Step 5: Sort the Calculated Performance functions P (X_i) in a descending order such as P = [P_1, P_2… P_N], P_1>P_2>…>P_N. Then calculate the (1-ρ) th quantile of performance function P [(1-ρ)*N].

Step 6: If P [(1-ρ)*N] <L_d, set L = L_d, otherwise set L = P [(1-ρ)*N]. Then evaluate the function H (X_i) for all Xi, such that H (X_i) = 1 if P (X_i) <L and H (X_i) = 0, otherwise.

Step 7: Calculate the Likelihood ratios W (Xi, u, v_{t-1}), where W = W_G*W_L. Update the new distorted parameters v_{Gt}, v_{Lt}.

$$v_{Gtj} = 1 - \frac{\sum_{i=1}^{N} I_{\{P(X_i)<L\}} W(X_i; u, v) X_{ij}}{\sum_{i=1}^{N} I_{\{P(X_i)<L\}} W(X_i; u, v)} \tag{3.2}$$

$$v_{Ltj} = 1 - \frac{\sum_{i=1}^{N} I_{\{P(X_i)<L\}} W(X_i; u, v) X_{ij}}{\sum_{i=1}^{N} I_{\{P(X_i)<L\}} W(X_i; u, v)} \tag{3.3}$$

Step 8: If $L=L_d$, then the training phase ends and go to Step 9 or else increase the iteration counter t and go to step 3 for next iteration.

Step 9: Start the Testing Phase.

3.2 Testing Phase (Sequential Simulation)

The testing phase is the phase during which the reliability evaluation is finally done. For the testing phase of sequential simulation, the optimal distorted parameters are derived from the initial training phase. Here the load is taken from the hourly load curve and is not distorted.

Step 1: From the distorted parameter vector v the new transition rate vectors μ* and λ* are generated for the generators and transmission lines. Initialize NY max (maximum simulated years ~5000)

Step 2: Generate the random sample X_1 from the new distorted transition rate vectors and the calculate sample residence time (T_{res} (X_i)). Initialize T_{sim} (~8736 hours), T_Down, T_WDown, T_Up, T_WUp to Zero.

Step 3: Evaluate the current sample likelihood ratios W (X_i;u,v).

Step 4: Transition to the next system states, and sample the residence time from the chronological load model and the distorted transition rate vectors. Calculate the cumulative sample times T_{res} total = ΣT_{res} (X_i). If the total residence time after the current sample is greater than Tsim, the residence time of the current sample is reduced and same sample is used as starting sample for the next year.

Step 5: Once all the sample states and likelihood ratios for a simulation year is generated all the states are evaluated to generate each sample up time and down time. Here a parallel computing technique is used to calculate the sample up and down times to reduce the computational time. The MATLAB parallel tool box is used to reduce the computational time by evaluating all the states in parallel using the multi cores of the processor. Go to step 6 if down time or go to step 7 if up time.

Step 6: Accumulate the Down time

$$T\_Down = T\_Down+t_i; \quad (3.4)$$

$$T_W Down = T_W Down+(t_i*W(X_i,u,v)); \quad (3.5)$$

Step 7: Accumulate the Up time

$$T\_Up = T\_Up+t_i; \quad (3.6)$$

$$T_W Up = T_W Up + (t_i*W(X_i,u,v)) \quad (3.7)$$

Step 8: The LOLP index for this simulation year is evaluated using a weighted mean approach.

$$\omega(NY) = \frac{(T_W UP + T_W Down)}{T_{sim}} \quad (3.8)$$

$$LOLP(NY) = \frac{(T_W Down * \omega(NY))}{(T_W Up + T_W Down)} \quad (3.9)$$

Step 9: The Coefficient of Variation (β) is calculated and compared with the β_{max}. If it falls below β_{max} or $NY>NY_{max}$ the simulation is stopped. Or else go to step 2.

Step 10: Evaluate the LOLP index

$$LOLP = \frac{\sum_{i=1}^{NY} LOLP(i)}{\sum_{i=1}^{NY} \omega\,(i)} \quad (3.10)$$

3.3 Acceleration Using Parallel Pooling

Using the parallel computing capacity of any desktop or laptop for simulations helps us in improving the computation efficiency. An average laptop/Desktop has 4 to 8 cores.

Using Matlab for parallel computing [18], we need to first assign number of cores we need for simulation, depending on the availability of cores. Once the number of cores is specified the main Matlab creates the same number of worker Matlabs. Main Matlab divides the work and sends the data and code to the workers. The workers execute the assigned iterations and send results back to the main Matlab. Then main Matlab combines results and continues executing statements after parallel computing. This causes an extra overhead time but for a large system the parallel computing benefit is far higher than the overhead time.

For example, if the main Matlab has to evaluate 100 samples with four cores it divides the work between the workers and each worker evaluates 25 samples.

4. Results and Discussions

4.1 Introduction

The Sequential MCS-Cross Entropy Methods are implemented on IEEE RTS 79 [16]. To show the sensitivity of computation time the parameters are varied and the change in computation time is recorded. The IEEE RTS 79 is a 24-bus system with 32 generating units and 38 transmission lines. The maximum generation capacity is 3405 MW with a peak load of 2850 MW. The load is a correlated load between the buses. All the simulations are performed on Matlab using an Intel 4 core, 3.4GHz processor.

4.2 Varying the multi core for Computational efficiency

Here the number of cores or workers used for computing is varied and computational time is calculated. All the computations are done at system peak load of 2850 MW until a 2% convergence is reached.

It can be observed from the Table 4.1 that as the number of cores of the computer utilized for evaluating the states increases the computational time decreases. As expected, the improvement ratios as a function of cores are about the same in Cross Entropy-Importance Sampling based MCS (CE-ISMC) and Simple MCS (SMC). However, the times taken can be substantially different.

Table 4.1. Results with varying cores

	Number of cores	LOLP (10^{-3})	Ny (years)	Time (Secs)	Improvement Ratio in Time
CE-ISMC	1	1.17	182	5,753	1
	2	1.16	177	2,893	1.98
	4	1.18	181	1,987	2.89
SMC	1	1.1	6990	46,498	1
	2	1.1	6957	24,919	1.86
	4	1.1	7061	17,130	2.71

4.3 Varying Coefficient of Variation (COV)

The COV value is varied from 5% to 1% and the change in LOLP and computation time is observed. This simulation is implemented at a system peak load of 2850 MW using 4 cores.

It can be seen that for COV of 5%, the CE IS reduces computation time by 2395 seconds whereas for 1% the time is reduced by 60,658 seconds. Therefore CE-IS becomes computationally more advantageous as value of COV is made tighter.

Table 4.2. Results with varying Coefficient of variation

	COV (β) (%)	LOLP (10^{-3})	Ny (years)	Time (Secs)	Time Saving
CE-ISMC	5	1.1	26	335	2359
	2	1.1	179	1,846	15,284
	1	1.1	775	8,000	60,658
SMC	5	1.1	1077	2,694	
	2	1.1	7061	17,130	
	1	1.1	28204	68,658	

4.4 Varying the System Peak Load

In this case the system peak load is increased and decreased by 300 MW from the base peak load of 2850 MW of the chronological Load Curve. The LOLP and the computation time is evaluated and compared with the Simple Monte Carlo Simulation. All the values are calculated for a 2% convergence using 4 cores.

Table 4.3. Results with varying system peak load

	LOAD (MW)	LOLP (10^{-3})	Ny (years)	Time (Secs)	Time Saving
CE-ISMC	3150	5.9	163	1,637	5101
	2850	1.1	185	1,978	15152
	2550	0.14	224	2,666	63098
SMC	3150	6.0	1696	6,738	
	2850	1.1	7061	17,130	
	2550	0.14	39329	65,784	

As can be seen from the table III. The system takes a greater number of samples and increased computation time before converging as the load decreases. This is because the LOLP increases with higher peak load and simulation time is inversely proportional to the LOLP being estimated [19]. Therefore, the CE-IS MC becomes computationally more advantageous with higher reliability systems.

4.5 Varying the System Outage Rates

The component outage rates are varied and the change in LOLP and computation time are observed. The Forced outage rate is the component probability of failure. The forced outage rate is changed uniformly for all the generating components. This is carried out at a system peak load of 2850 MW and 2% convergence criteria.

It can be observed from the Table IV that increasing the forced outage rates increases the loss of load and decreases the computational time. Similar to the previous case, increase reliability leads to higher savings in computational time with the CE-IS use in MC.

Table 4.4. Results with varying Forced Outage Rates

	Outage Rate (Multiplier)	**LOLP (10^{-3})**	**NY (Years)**	**Time (Secs)**	**Time Saving**
CE-ISMC	**1**	**1.1**	**194**	**2,099**	**15031**
	1.25	**2.2**	**156**	**1,808**	**8854**
	2	**9.4**	**85**	**1,058**	**3746**
SMC	**1**	**1.1**	**7061**	**17,130**	
	1.25	**2.2**	**3617**	**10,662**	
	2	**9.6**	**989**	**4,804**	

5. Conclusions

It should be noted but is often forgot that the results obtained by Monte Carlo are only estimates of true values and not the true values. Therefore the estimates have a variance. The estimates approach the true values as the variance of estimates is reduced by increasing the sample size. Importance sampling helps further by reducing the variance of the estimator and thus a smaller sample size is needed to get the same coefficient of variation. The coefficient of variance determines the gap between the upper and lower bounds with a given level of confidence. The smaller the coefficient of variation, the tighter are the bounds around the true values. The main advantage of using variance reduction technique of Importance sampling is the reduction in computational time. This paper has explored the conditions under which the computation time is reduced more favorably by implementation of IS and thus it becomes advantageous to use this variance reduction approach. In general, the conditions which lead to higher computation time for the straight MCS tend to favor the use of IS for relatively higher reduction of computation time by reducing the variance of estimates. The conditions which lead to higher computation time are either the ones that lead to higher reliability, i.e. , lower loss of load probability or the ones where tighter bounds on estimates are needed to have higher confidence in the estimated results.

Appendices

Appendix A: Derivation of v parameter using Kullback-Leibler Distance

This appendix provides a detailed derivation on calculating the v parameter for IS as described in [1]. With samples X_1, X_2… X_N generated from secondary density g*(X) the reliability index is calculated using an unbiased estimator

$$r = \frac{1}{N}\sum_{i=1}^{N} I_{\{P(X_i<L)\}} \frac{f(X_i;u)}{g(X_i)} \tag{6.1}$$

The best way to estimate r is given by

$$g(X_i) = \frac{I_{\{P(X_i<L)\}} f(X_i;u)}{r} \tag{6.2}$$

Using this g we will have a zero variance estimator for r and it requires only one sample. But this approach is unworkable because of the unknown parameter r which we want to estimate. So, the idea of cross entropy is to choose g from a family of densities f(.;v), i.e. to calculate the reference parameter v such that the distance between the densities g* and f(.;v) is minimum. This distance between the densities is represented by Kullback -Leibler distance or Cross Entropy. The Kullback - Leibler distance or Cross Entropy is defined as

$$D(g^*, f) = E(ln\left(\frac{g^*(X)}{f(X;v)}\right))$$

$$= \int g^*(X)\ln(g^*(X))\,dx - \int g^*(X)\ln\big(f(X;v)\big)\,dx \tag{6.3}$$

Minimizing Kullback-Leibler distance is equivalent to maximizing

$$\max \int g^*(X)\ln\big(f(X;v)\big)\,dx \tag{6.4}$$

This can be written as:

$$Max\ D(v) = \max E\ \left(I_{\{P(X<L)\}}\ln\big(f(X,v)\big)\right) \tag{6.5}$$

Using Importance Sampling and a change of measure f(.;v) we can rewrite it as

$$\text{Max } D(v) = \max E\,(I_{\{P(X)<L\}}\, W(X;u,v) \ln(f(X;v))) \tag{6.6}$$

For any reference parameter v, where

$$W(X;u,v) = \frac{f(X;u)}{f(X;v)} \tag{6.7}$$

The optimal solution v* can be written as

$$v^* = \operatorname{argmax} E_w\,(I_{\{P(X)<L\}}\, W(X;u,v) \ln(f(X;v))) \tag{6.8}$$

The D (v) is differential with respect to v, and the solution can be obtained by solving the following system of equations.

$$\frac{1}{N}\sum_{i=1}^{N} I_{\{P(X_i)<L\}} W(X_i;u,v) \nabla\, ln(f(X_i,v)) = 0 \tag{6.9}$$

Now

$$\frac{\partial}{\partial v_j}(\ln(f(X_i;v))) = \frac{-x_i}{v_j(1-v_j)} + \frac{1}{v_j} \tag{6.10}$$

Substituting this equation in the above equation, the j_{th} equation becomes

$$\sum_{i=1}^{N} I_{\{P(X_i)<L\}} W(X_i;u,v)\left(\frac{-X_{ij}}{v_j(1-v_j)} + \frac{1}{v_j}\right) = 0 \tag{6.11}$$

By solving the equation (29) we get

$$v_j = 1 - \frac{\sum_{i=1}^{N} I_{\{S(X_i)<L\}} W(X_i;u,v) X_{ij}}{\sum_{i=1}^{N} I_{\{S(X_i)<L\}} W(X_i;u,v)} \tag{6.12}$$

Appendix B: State Evaluation: Heuristic and Linear Program Using DC Power Flow

This appendix describes the DC Power flow model used in the simulations.

The DC power flow equation and line flow equations are

$$B\theta+G = D \tag{6.13}$$

$$b\hat{A}\theta = F \tag{6.14}$$

Where;

N_b = Number of buses
N_t = Number of transmission lines
b = N_t x N_t primitive matrix of transmission line susceptances
$\hat{A}$ = $N_t x N_b$ element node incidence matrix
B = $N_b x N_b$ augmented node susceptance matrix
θ = N_b vector bus voltage angles
G = N_b vector of bus Generation levels
D = N_b vector of bus loads
F = N_t vector of transmission line flows

A computationally efficient selective approach based on DC power flow as given in [17] is first used to find a feasible flow. This approach consists of the following steps.

Step 1: The total injection at all buses are calculated by subtracting the bus loads from available generations at buses.

Step 2: If the sum of positive injections is greater than the sum of negative injections, the positive injections are scaled down proportionately so that the sum equals that of negative injections and vice versa if net negative injections are greater than net positive injections.

Step 3: once power balance is accomplished the G vector generated from step 2 is used in DC Power flow equation (6.13) to calculate θ, then θ is used in line flow equation (6.14) to calculate the line flows.

If the line flows satisfy flow constraints a feasible flow is found and if load is curtailed then the reliability indices are updated. If the line flows do not satisfy the flow constraints a Linear Programming (LP) model is implemented to calculate the optimized line flows and load curtailment. This LP model is described as follows:

Minimize Load Curtailment = $\text{Min}\sum_{i=1}^{N} LC_i$

Subject to Constraints;

Power balance: $B\theta + G + LC = D$

Generation limit: $G \leq G^{max}$

Flow Limits: $b\hat{A}\theta \leq F^{max}$

$-b\hat{A}\theta \leq F^{max}$

Load Limits: $LC \leq D$

Boundaries: $G, LC \geq 0$

θ unrestricted

where;

LC = N_b vector of Load curtailments

G^{max} = N_b vector of maximum available bus generation levels

F^{max} = N_t vector of flow capacities of transmission levels

References

[1] Rubinstein, Reuven Y, and Dirk P Kroese. *The Cross-Entropy Method: A Unified Approach to Combinatorial Optimization, Monte-Carlo Simulation and Machine Learning*. Springer, 2004.

[2] A. M. Leite da Silva, R. A. G. Fernandez and C. Singh, "Generating Capacity Reliability Evaluation Based on Monte Carlo Simulation and Cross-Entropy Methods," in *IEEE Transactions on Power Systems*, vol. 25, no. 1, pp. 129-137, Feb. 2010.

[3] R. A. Gonzalez-Fernandez and A. M. Leite da Silva, "Reliability Assessment of Time-Dependent Systems via Sequential Cross-Entropy Monte Carlo Simulation," in *IEEE Transactions on Power Systems*, vol. 26, no. 4, pp. 2381-2389, Nov. 2011.

[4] R. A. González-Fernández, A. M. Leite da Silva, L. C. Resende and M. T. Schilling, "Composite Systems Reliability Evaluation Based on Monte Carlo Simulation and Cross-Entropy Methods," in *IEEE Transactions on Power Systems*, vol. 28, no. 4, pp. 4598-4606, Nov. 2013.

[5] A. M. L. da Silva, R. A. González-Fernández, S. A. Flávio and L. A. F. Manso, "Composite reliability evaluation with renewable sources based on quasi-sequential Monte Carlo and cross entropy methods," *2014 International Conference on Probabilistic Methods Applied to Power Systems (PMAPS)*, Durham, 2014, pp. 1-6.

[6] E. Tómasson and L. Söder, "Improved Importance Sampling for Reliability Evaluation of Composite Power Systems," in *IEEE Transactions on Power Systems*, vol. 32, no. 3, pp. 2426-2434, May 2017.

[7] C. Yan, L. G. Luca, Z. Bie, T. Ding and G. Li, "A three-stage CE-IS Monte Carlo algorithm for highly reliable composite system reliability evaluation based on screening method," *2016 International Conference on Probabilistic Methods Applied to Power Systems (PMAPS)*, Beijing, 2016, pp. 1-6.

[8] A. C. G. Melo, G. C. Oliveira, M. Morozowski Fo and M. V. F. Pereira, "A hybrid algorithm for Monte Carlo/enumeration based composite reliability evaluation (power systems)," *1991 Third International Conference on Probabilistic Methods Applied to Electric Power Systems*, London, UK, 1991, pp.70-74.

[9] D. Li, L. Dong, H. Shen, B. Li and Y. Liao, "Reliability evaluation of composite generation and transmission systems based on stratified and gradual importance sampling algorithm," *2011 International Conference on Advanced Power System Automation and Protection*, Beijing, 2011, pp. 2082-2087.

[10] H. Kumamoto, K. Tanaka, K. Inoue and E. J. Henley, "Dagger-Sampling Monte Carlo for System Unavailability Evaluation," in *IEEE Transactions on Reliability*, vol. R-29, no. 2, pp. 122-125, June 1980.

[11] Sun, R.; Singh, C.; Cheng, L.; Sun, Y. Short-term reliability evaluation using control variable based dagger sampling method. Electr. Power Syst. Res. 2010, 80, 682–689

[12] Z. Shu and P. Jirutitijaroen, "Latin Hypercube Sampling Techniques for Power Systems Reliability Analysis with Renewable Energy Sources," in *IEEE Transactions on Power Systems*, vol. 26, no. 4, pp. 2066-2073, Nov. 2011.

[13] Z. Shu, P. Jirutitijaroen, A. M. Leite da Silva and C. Singh, "Accelerated State Evaluation and Latin Hypercube Sequential Sampling for Composite System Reliability Assessment," in *IEEE Transactions on Power Systems*, vol. 29, no.4, pp.1692-1700, July2014.

[14] D. Lieber, A. Nemirovskii and R. Y. Rubinstein, "A fast Monte Carlo method for evaluating reliability indexes," in *IEEE Transactions on Reliability*, vol. 48, no. 3, pp. 256-261, Sept. 1999.

[15] G. E. Evans, J. M. Keith and D. P. Kroese, "Parallel cross-entropy optimization," *2007 Winter Simulation Conference*, Washington, DC, 2007, pp. 2196-2202.

[16] P. M. Subcommittee, "IEEE Reliability Test System," in *IEEE Transactions on Power Apparatus and Systems*, vol. PAS-98, no. 6, pp. 2047-2054, Nov. 1979.

[17] C. Singh and J. Mitra, "Composite system reliability evaluation using state space pruning," in *IEEE Transactions on Power Systems*, vol. 12, no. 1, pp. 471-479, Feb. 1997.

[18] MATLAB and Parallel Computing Toolbox Release 2017a, The MathWorks, Inc., Natick, Massachusetts, United States.

[19] C. Singh., P. Jirutitijaroen and J. Mitra, 2018. *Electric Power Grid Reliability Evaluation: Models and Methods*. Wiley-IEEE Press, 2019.