RISK-BASED MACHINE LEARNING APPROACHES FOR PROBABILISTIC TRANSIENT STABILITY

by

Umair Shahzad

A DISSERTATION

Presented to the Faculty of

The Graduate College at the University of Nebraska

In Partial Fulfillment of Requirements

For the Degree of Doctor of Philosophy

Major: Electrical Engineering

Under the Supervision of Professor Sohrab Asgarpoor

Lincoln, Nebraska

December, 2021

RISK-BASED MACHINE LEARNING APPROACHES FOR PROBABILISTIC TRANSIENT STABILITY

Umair Shahzad, Ph.D.

University of Nebraska, 2021

Advisor: Sohrab Asgarpoor

Power systems are getting more complex than ever and are consequently operating close to their limit of stability. Moreover, with the increasing demand of renewable wind generation, and the requirement to maintain a secure power system, the importance of transient stability cannot be overestimated. Considering its significance in power system security, it is important to propose a different approach for enhancing the transient stability, considering uncertainties. Current deterministic industry practices of transient stability assessment ignore the probabilistic nature of variables (fault type, fault location, fault clearing time, etc.). These approaches typically provide a conservative criterion and can result in expensive expansion plans or conservative operating limits. With the increasing system uncertainties and widespread electricity market deregulation, there is a strong inevitability to incorporate probabilistic transient stability (PTS) analysis. Moreover, the time-domain simulation approach, for transient stability evaluation, involving differential-algebraic equations, can be very computationally intensive, especially for a large-scale system, and for online dynamic security assessment (DSA).

The impact of wind penetration on transient stability is critical to investigate, as it does not possess the inherent inertia of synchronous generators. Thus, this research proposes risk-based, machine learning (ML) approaches, for PTS enhancement by

replacing circuit breakers, including the impact of wind generation. Artificial Neural Network (ANN) was used for predicting the benefit-cost ratio (BCR) to reduce the computation effort. Moreover, both ANN and support vector machine (SVM) were used and consequently, were compared, for PTS classification, for online DSA. The training of the ANN and SVM was accomplished using suitable system features as inputs, and PTS status indicator as the output. DIgSILENT PowerFactory and MATLAB was utilized for transient stability simulations (for obtaining training data for ML algorithms), and applying ML algorithms, respectively. Results obtained for the IEEE 14-bus test system demonstrated that the proposed ML methods offer a fast approach for PTS prediction with a fairly high accuracy, and thereby, signifying a strong possibility for ML application in probabilistic DSA.

Dedication

- **To my loving parents, parents-in-law, and grandmother**; whose words of encouragement and continuous support, helped me achieve this challenging feat.

- **To the love of my life, my wife, and my soulmate, Habibah**, **and my dear son, Ayaan;** without their unconditional love, patience, sacrifice, and support, the successful completion of this dissertation would have been impossible.

- **To my brothers, Asjad, and Hamza,** who have always been supportive, throughout this journey.

- **To all those who work hard to achieve their goals and never give up during the process.**

Acknowledgements

Firstly, I would like to thank ALLAH for the wisdom, determination, and perseverance that He bestowed upon me, for completing this research work.

This dissertation would not have been possible without the guidance and help of several individuals, who, in one way or another, contributed and extended their valuable assistance, in completion of this research.

I would like to express my sincere gratitude to my advisor, Dr. Sohrab Asgarpoor, for all his support, encouragement, and care about my academic progress and personal life. I consider myself privileged to have the opportunity to work under his guidance. During this arduous doctoral research journey, Dr. Asgarpoor helped me develop significant research skills, and guided me to advance in critical thinking. I would sincerely like to thank him, for his patience, motivation, and insightful feedback. His supervision helped me progress in the research, complete this dissertation, and most importantly, to become a responsible person.

I would also like to thank my degree committee members, Dr. Hudgins, and Dr. Choobineh, from the Department of Electrical and Computer Engineering at UNL, and Dr. Cohn, from the Department of Mathematics at UNL, for their immense support, guidance, and for serving as the members of the doctoral supervisory committee.

I cannot find words to express my appreciation and love for my best friend, my better half, and my soulmate, Habibah Ayub; thank you for being a perfect spouse, for your unconditional sacrifices, for your inspiration, and for always encouraging me, throughout this demanding journey. My heartfelt thanks go to my family, especially my parents,

parents-in-law, and grandmother, who have always supported and encouraged me. Their unwavering faith has been crucial in the successful completion of my dissertation.

My deepest thanks to my little spider-man, my beloved son, Ayaan, whose lovely smile always encouraged me to continually progress towards my goal. I love you and will always be proud to be your father. Without the emotional support of these wonderful people, this piece of work would have been impossible to achieve.

I am also grateful to United States Educational Foundation in Pakistan (USEFP) and Institute of International Education (IIE) for their financial support, through the prestigious Fulbright scholarship, due to which I could pursue, and successfully complete my Doctoral degree, in the U.S.

Special thanks to the staff of the Department of Electrical and Computer Engineering, UNL, especially, Ms. Teresa Ryans, Ms. Pamela Weise, and Ms. Patricia Worster, who have always been very kind and helpful.

Table of Contents

List of Figures

List of Tables

List of Abbreviations

Abbreviation	Meaning
AGC	Automatic Generation Control
ANN	Artificial Neural Network
AUC	Area Under the Curve
BCR	Benefit Cost Ratio
BFCL	Bridge Fault Current Limiter
BP	Back Propagation
CA	Classification Accuracy
CB	Circuit Breaker
CBA	Cost Benefit Analysis
CCT	Critical Clearing Time
CE	Classification Error
CNN	Convolutional Neural Network
DFIG	Doubly Fed Induction Generator
DL	Deep Learning
DNO	Distribution Network Operator
DPL	DIgSILENT Programming Language
DSA	Dynamic Security Assessment
DSP	Digital Signal Processor
EAC	Equal Area Criterion

EEAC	Extended Equal Area Criterion
FCT	Fault Clearing Time
FERC	Federal Energy Regulatory Commission
FN	False Negatives
FNN	Feedforward Neural Network
FOSM	First Order Second Moment
FP	False Positives
FV	Future Value
GSC	Grid Side Converter
IGBT	Insulated Gate Bipolar Transistor
ISO	Independent System Operator
LSTM	Long Short-Term Memory
MAE	Mean Absolute Error
MC	Monte Carlo
MCMC	Markov Chain Monte Carlo
ML	Machine Learning
MLPNN	Multilayer Perceptron Neural Network
MNN	Modular Neural Network
MSE	Mean Squared Error
NERC	North American Electric Reliability Corporation
OAT	One Factor at a Time

OPF	Optimal Power Flow
PCC	Point of Common Coupling
PCM	Probabilistic Collocation Method
PDF	Probability Density Function
PE	Point Estimate
PEBS	Potential Energy Boundary Surface
PI	Proportional Integral
PMF	Probability Mass Function
PTS	Probabilistic Transient Stability
PV	Photovoltaic
QMC	Quasi Monte Carlo
RBF	Radial Basis Function
RBFNN	Radial Basis Function Neural Network
RL	Reinforcement Learning
RMSE	Root Mean Squared Error
RNN	Recurrent Neural Network
ROC	Receiver Operating Characteristic
ROCOF	Rate of Change of Frequency
RPM	Revolutions Per Minute
RSC	Rotor Side Converter
RTI	Rotor Trajectory Index
RV	Random Variable

SFCL	Superconducting Fault Current Limiter
SG	Synchronous Generator
SL	Supervised Learning
SLT	Statistical Learning Theory
SMES	Superconducting Magnetic Energy Storage
SMIB	Single Machine Infinite Bus
SML	Supervised Machine Learning
SRM	Structural Risk Minimization
SSSC	Static Synchronous Series Compensator
STATCOM	Static Synchronous Compensator
SVC	Static VAR Compensator
SVM	Support Vector Machine
TCSC	Thyristor Controlled Series Capacitor
TDS	Time Domain Simulation
TN	True Negatives
TNSO	Transmission Network System Operator
TP	True Positives
TSA	Transient Stability Assessment
TSI	Transient Stability Index
UEP	Unstable Equilibrium Point
UL	Unsupervised Learning
UPFC	Unified Power Flow Controller

VC	Vapnik–Chervonenkis
VSC	Voltage Source Converter
WECC	Western Electricity Coordinating Council

Nomenclature

J	Moment of inertia for synchronous generator
θ	Angular displacement of the rotor, with respect to a stationary axis, on generator stator
T_m	Input mechanical torque
T_e	Output electrical torque
P_m	Input mechanical power
P_e	Output electrical power
M	Moment of inertia of the rotor mass
μ	Mean of the Normal (Gaussian) PDF
σ	Standard deviation of the Normal (Gaussian) PDF
δ_{max}	Maximum rotor angle difference
TSI_i	Transient stability index for the i^{th} MC sample
E_i	i^{th} contingency (event)
$P_r(E_i)$	Probability of i^{th} contingency (event)
$Sev(E_i)$	Severity (impact) of E_i
n_i	Number of synchronous machines which are transiently unstable for the i^{th} MC sample
S_i	Transient stability status for the i^{th} MC sample
$\hat{y}$	Predicted value of y
y_i	Actual value of y

$\bar{y}$	Mean value of y
$\text{cov}(y_i, \hat{y})$	Covariance between y_i and $\hat{y}$
σ_{yi}	Standard deviation of y_i
$\sigma_{\hat{y}}$	Standard deviation of $\hat{y}$
R	Coefficient of Regression
R^2	Coefficient of Determination
$S1$	Sensitivity (or recall)
$S2$	Specificity
F_1	F_1 score
Φ	Activation function
γ, C	SVM hyperparameters
$K(x_i, x_j)$	SVM kernel function
R_i	Transient instability risk for i^{th} MC sample
R_A	Average risk index for transient instability
$f(X_i)$	PDF for load at i^{th} bus
μ_i	Mean of the forecasted peak load for i^{th} bus
σ_i	Standard deviation of the forecasted peak load for i^{th} bus
N	Number of MC samples
N_m	Number of unstable MC samples for m^{th} line
F_i	i^{th} fault event
$\Pr(F_i)$	Probability of the i^{th} fault event

$\Pr(F_o)$	Probability of fault occurrence
$\Pr(F_L)$	Probability of fault location
$\Pr(F_T)$	Probability of fault type
$\Pr(U_i \mid F_i)$	Probability of transient instability given i^{th} fault event has occurred
$Sev(F_i)$	Severity of F_i
P_{LG}	Probability of occurrence of single line to ground fault
P_{LLG}	Probability of occurrence of double line to ground fault
P_{LL}	Probability of occurrence of line to line fault
P_{LLL}	Probability of occurrence of three phase fault
P_{LGI}	Probability of instability of single line to ground fault
P_{LLGI}	Probability of instability of double line to ground fault
P_{LLI}	Probability of instability of line to line fault
P_{LLLI}	Probability of instability of three phase fault
P_{SYS}	Probability of system instability
C_I	Installation cost for unit CB
C_F	Foundation cost for unit CB
PV	Present value
FV	Future value
r	Interest rate
BCR	Benefit cost ratio
C_{REP^m}	Replacement cost of m^{th} synchronous generator

C_i	Cost associated with an i^{th} transiently unstable sample (future monetary benefits)
C_i^p	Present value of C_i (future monetary benefits)
$EBCR_m$	Expected value of BCR for m^{th} line
DV_m	Decision variable for m^{th} line

CHAPTER 1

INTRODUCTION

This chapter presents the overview, motivation, and background of the research. Additionally, the chapter describes the problem statement, research questions, research objectives, major contributions, research impact and limitations. A brief outline of the dissertation is also provided.

1.1 Overview

Driven by various techno-economic and environmental factors, the electric energy industry is anticipated to undergo a paradigm shift, with a significantly augmented level of renewables, especially, wind and solar power sources, gradually replacing conventional power production sources (coal, diesel, natural gas, etc.). This increasing demand of large-scale wind integration in the conventional power system, along with the inherent and external uncertainties of the system, brings a lot of challenges [1-2]. One of them is the power system transient stability. Power systems are regularly exposed to unanticipated faults. Such faults can cause transient instability and can consequently lead to prevalent outages [3]. To preserve system security, system operators and planners perform analysis to make critical operating and planning decisions that will ensure safe operation of the power system after the occurrence such faults. The current general practice, within the power industry, is to use the deterministic approaches, with significant safety margins, to cover all possible uncertainties. With the adoption of a deterministic criterion for system security, power systems generally operate with a large security margin. Usually, these

deterministic criteria provide safe, but conservative limits for system operating conditions. The most crucial security criterion is the (*N-1*) security criterion that guarantees safe operation of the power system, after the failure of a single element of the system, where *N* is the total number of system components [3].

In the past several years, there has been a significant increase in connections of intermittent and stochastic, power electronics interfaced renewable energy generation sources. These uncertainties, coupled with load uncertainties, are becoming one of the vital characteristics of modern power systems. The transient stability assessment of such systems, using traditional deterministic methodology, is swiftly becoming inappropriate and thus, innovative probabilistic assessment approaches are desirable [4].

The conventional approaches for transient stability are computationally intensive [5-6]. Various soft computing approaches (these approaches are useful in solving complex problems, which cannot be realized by classical numerical methods), based on machine learning (ML) [7], may provide a better solution in this regard, especially, for online transient stability classification, incorporating various uncertainties.

1.2 Motivation

The current industry practices use the deterministic approach for transient stability assessment [8-9]. Although, the deterministic approaches result in highly secured power systems, but they do not consider the probability of operating conditions. Apart from the high cost due to conservative designs, the key drawback with the deterministic assessment techniques is that they consider all security problems to have equal risk [10]. Various literature (including research papers, technical reports, dissertations, and white papers) [9-

30] mention that probabilistic risk-based transient stability, and incorporating risk in power planning procedures, is a future research area, and consequently, work needs to be done in this domain. Moreover, planning guides/manuals of various utilities [31-40] recommend using risk-based probabilistic approaches in the near future. Moreover, the integration of renewable generation, introduces more stochasticity in the system, making the application of probabilistic practices essential in the transient stability assessment process. Probabilistic assessment methods [41-42] can be helpful when uncertain parameters are included into system stability assessment. It is, thus, of great implication to propose a risk-based approach, for overcoming the inadequacies of the deterministic approach. The integration of wind farms in the power system decreases the inherent inertia and hence, impacts transient stability. The time-domain simulation method for transient stability assessment involves numerical integration and differential-algebraic equations. This can pose a huge computation burden, especially for large-scale systems, and for online dynamic security assessment (DSA). A different and unique approach is required to reduce these computation efforts. These knowledge gaps have paved the motivation for this research, which is divided into parts: (1) proposing a risk-based probabilistic ML approach, using artificial neural network (ANN) for transient stability enhancement, by replacement of circuit breakers, incorporating wind generation, and (2) using both ANN and support vector machine (SVM) to classify transient stability status, for online application.

1.3 Problem Statement

Presently, majority of the power utilities resort to deterministic transient stability approaches, but to cater for future power networks (incorporating inherent multiple

uncertainties), it is essential to include risk-based stability criterion. The proposed approach provides a better and realistic quantification of system uncertainties, as it uses risk-based stability criterion for decision-making, as opposed to the currently used deterministic procedures. Moreover, it is important to propose faster methods for transient stability prediction, considering various uncertainties. This is, particularly, important for online DSA. The proposed approach, thus, consists of two main parts: (1) proposing a risk-based probabilistic ML approach ANN for transient stability enhancement, by replacement of circuit breakers, and consequently analyzing the impact of wind generation, and (2) using both ANN and SVM to classify transient stability status, for online application, in the presence of uncertainties.

1.4 Research Questions

This research attempts to answer the following research questions:

1. How to enhance the current deterministic industry practices to cater for the future needs of the power systems?
2. How to incorporate transient instability risk, for decision-making regarding CBs, considering system uncertainties?
3. How does wind generation impact transient instability risk?
4. How can cost-benefit analysis be applied to a power system, in the presence of uncertainties?
5. How can artificial intelligence (e.g., ANN) be used to enhance computational efficiency of the traditional transient stability assessment methods?

6. How can artificial intelligence methods (ANN and SVM) be used for online probabilistic transient stability prediction (which is a significant component of online DSA)?

1.5 Research Scope and Objectives

The overall aim of this research is two-fold: (1) to apply ANN for probabilistic transient stability (PTS) enhancement, by replacing circuit breakers, and (2) to apply ANN and SVM for online PTS prediction. These aims are fulfilled through the following objectives (O1-O9).

O1. Literature review to summarize the state-of-the-art of research in PTS and risk-based transient stability.

O2. Literature review to summarize the state-of-the-art of research in application of ANN and SVM for transient stability classification.

O3. Concept of marginal transient stability using time-domain approach.

O4. Replacement of circuit breakers based on transient instability risk.

O5. Analyzing the impact of wind generation on transient instability risk.

O6. Ranking of circuit breakers for line and bus faults (based on transient instability risk).

O7. Assess the impact of network topology on system instability probability and transient instability risk.

O8. Application of ANN to reduce computation time in a probabilistic, time-domain simulation, transient stability problem.

O9. Application of ANN and SVM and consequently, comparing their performance for online PTS classification.

1.6 Main Contributions of the Research

The work within this dissertation contributes to the area of power system stability and is exclusively focused on transient stability, in the presence of uncertain parameters. The main outcome of this research is the proposed probabilistic cost benefit analysis approach for replacement of circuit breakers, based on transient instability risk. Consequently, ML algorithms (ANN and SVM) are applied for improving the computational efficiency, and for the application of online DSA. The following are the major contributions of the research.

- Replacement of circuit breakers based on transient instability risk.
- Concept of marginal transient stability using time-domain approach.
- Application of ANN to reduce computation time in a probabilistic, time-domain simulation, transient stability problem.
- Application of ANN and SVM and consequently, comparing their performance for online PTS classification.

1.7 Impact of the Proposed Research

- Presently, majority of the power utilities resort to deterministic approaches for transient stability assessment, but to cater for future power networks (incorporating inherent multiple uncertainties), it is essential to include risk-based stability criterion for decision-making.

- The proposed approach provides a better and realistic quantification of system uncertainties.
- The proposed approach will help power system planners to replace circuit breakers based on transient instability risk.
- The risk-based approach is suitable for justifying the economic investments rather than the presently used worst-case deterministic approach.
- The approach will be useful for future power systems (which will include abundant renewable generation), as it will assess the impact of high wind penetration on transient instability risk.
- The ML approaches (ANN and SVM) used will help in reducing computation efforts, especially for online application.

1.8 Limitations of the Research

The following are the limitations of the research.

1. Only type 3 wind generator, i.e., doubly fed induction generator (DFIG), was used to assess the impact of wind generation of transient instability risk.
2. The external uncertainties (weather, cyber attacks, human error, etc.) were not considered in this research.
3. The research is directly associated with only transient stability. Other kinds of stability (frequency, voltage) were not considered.
4. *(N-1)* contingency criterion was used in this research.
5. Only ANN and SVM ML algorithms were utilized in this research.

1.9 Dissertation Outline

The dissertation is organized in five chapters. A brief description of each chapter is presented below.

- Chapter 1 presents the overview and motivation of the research. Additionally, the chapter describes the problem statement, research questions, research objectives, major contributions, impact and limitations of the research.
- Chapter 2 presents the background and overview related to power system stability, transient stability, probabilistic risk-based transient stability, ML, ANN, and SVM.
- Chapter 3 describes the problem formulation and proposed approaches. Moreover, various assumptions are listed.
- Chapter 4 describes the test system (IEEE 14-bus), case studies, results, and associated discussion. A validation approach, using sensitivity analysis, is also presented.
- Chapter 5 concludes the dissertation and provides numerous recommendations for relevant future research work. Moreover, generic recommendations for promoting risk-based decision making in power systems are outlined.

1.10 References

[1] M. Ma, Y. H. Liu, and D. M. Zhao, “Research on the impact of large-scale integrated wind farms on the security and stability of regional power system,” *International Conference on Power System Technology*, 2010, pp. 1-6.

[2] Z. Liu, C. Liu, Y. Ding, and G. Li, "Transient stability studies of power system with shared transmission of wind power and thermal power," *2nd IET Renewable Power Generation Conference*, 2013, pp. 1-4.

[3] M. Ilic, Toward standards for dynamics in electric energy systems, Power Systems Engineering Research Center (PSERC), Jun. 2012. [Online]. Available: https://pserc.wisc.edu/documents/publications/papers/fgwhitepapers/Ilic_PSERC_Future_Grid_White_Paper_Dynamics_June_2012.pdf

[4] P. N. Papadopoulos and J. V. Milanović, "Probabilistic framework for transient stability assessment of power systems with high penetration of renewable generation," *IEEE Transactions on Power Systems*, vol. 32, no. 4, pp. 3078-3088, Jul. 2017.

[5] S. Jin, Z. Huang, R. Diao, D. Wu, and Y. Chen, "Parallel implementation of power system dynamic simulation," *IEEE Power & Energy Society General Meeting*, 2013, pp. 1-5.

[6] A. Sabo and N. I. A. Wahab, "Rotor angle transient stability methodologies of power systems: a comparison," *IEEE Student Conference on Research and Development*, 2019, pp. 1-6.

[7] H. Sawhney and B. Jeyasurya, "On-line transient stability assessment using artificial neural network," *Large Engineering Systems Conference on Power Engineering (IEEE Cat. No.04EX819)*, 2004, pp. 76-80.

[8] P. Zhang, K. Meng, and Z. Dong, "Probabilistic vs deterministic power system stability and reliability assessment," in *Emerging Techniques in Power System Analysis*. Springer, Berlin, Heidelberg, 2010.

[9] K. N. Hasan, R. Preece, and J. V. Milanović, "Existing approaches and trends in uncertainty modelling and probabilistic stability analysis of power systems with renewable generation," *Renewable and Sustainable Energy Reviews*, vol. 101, pp. 168-180, Mar. 2019.

[10] A. Dissanayaka, U. D. Annakkage, B. Jayasekara, and B. Bagen, "Risk-based dynamic security assessment," *IEEE Transactions on Power Systems*, vol. 26, no. 3, pp. 1302-1308, Aug. 2011.

[11] W. Li and J. Zhou, "Probabilistic reliability assessment of power system operations," *Electric Power Components and Systems*, vol. 36, no. 10, pp. 1102-1114, Sep. 2008.

[12] V. Eric Van, "Transient stability assessment and decision-making using risk," *PhD Dissertation*, Iowa State University, USA, 2000.

[13] J. V. Milanovic, "Probabilistic stability analysis: the way forward for stability analysis of sustainable power systems," *Philosophical Transactions*, vol. 375, no. 2100, pp. 1-22, Jul. 2017.

[14] A white paper on the incorporation of risk analysis into planning processes, EPRI Report, 2015. [Online]. Available: https://pubs.naruc.org/pub.cfm?id=536DCF19-2354-D714-5117-47F9BA06F062

[15] W. Li and P. Choudhury, "Probabilistic planning of transmission systems: Why, how and an actual example," *IEEE Power and Energy Society General Meeting - Conversion and Delivery of Electrical Energy in the 21st Century*, 2008, pp. 1-8.

[16] W. Li, "Framework of probabilistic power system planning," *CSEE Journal of Power and Energy Systems*, vol. 1, no. 1, pp. 1-8, Mar. 2015.

[17] J. Choi, T. Tran, A. A. El-Keib, R. Thomas, H. Oh, and R. Billinton, "A method for transmission system expansion planning considering probabilistic reliability criteria," *IEEE Transactions on Power Systems*, vol. 20, no. 3, pp. 1606-1615, Aug. 2005.

[18] Probabilistic assessment improvement plan, NERC Report, Dec. 2015. [Online]. Available: https://www.nerc.com/comm/PC/PAWG%20DL/ProbA%20%20Summary%20and%20Recommendations%20final%20Dec%2017.pdf

[19] K. Bell, Methods and tools for planning the future power system: issues and priorities, *IET Report*, 2015. [Online]. Available: https://strathprints.strath.ac.uk/54341/1/Bell_IET_2015_Methods_and_tools_for_planning_the_future_power_system.pdf

[20] A study on probabilistic risk assessment for transmission and other resource planning, EPRI Report, 2015. [Online]. Available: https://pubs.naruc.org/pub.cfm?id=536DCE1C-2354-D714-5175-E568355752DD

[21] Developing a framework for integrated energy network planning (IEN-P), EPRI Report, 2018. [Online]. Available: https://www.epri.com/#/pages/product/3002010821/?lang=en-US

[22] Stochastic operations and planning, PNNL Technical Report, PNNL-23680, 2015. [Online]. Available: https://pdfs.semanticscholar.org/53e4/3719e74f6f9f6d39f942c79fcebca53ab9bb.pdf

[23] Transmission planning white paper, 2014. [Online]. Available: https://pubs.naruc.org/pub.cfm?id=53A151F2-2354-D714-519F-53E0785A966A

[24] Incorporating risk-based methods in planning, 2014. [Online]. Available:

https://www.ieee-pes.org/presentations/gm2014/2014-PES-GM_RiskBasePlan_2014Jul28_IEEE.pdf

[25] Transmission planning, 2020. [Online]. Available: https://www.epri.com/portfolio/programs/027570

[26] S. W. Kang, E. Meier, N. Kadel, J. Jin, and D. Aryal, "An approach for probabilistic composite power system transmission planning - ERCOT," *2020 IEEE Power & Energy Society General Meeting (PESGM)*, 2020, pp. 1-5.

[27] New approaches to power system planning. [Online]. Available: https://esmap.org/node/4626

[28] NERC and regional probabilistic studies, 2019. [Online]. Available: https://www.nerc.com/comm/PC/PAWG%20DL/Proabilistic_Analysis_Forum_Presentations_December_11_13_2019.pdf

[29] C. Lin, "A review of transmission development planning based on risk evaluation," *MATEC Web of Conferences*, 2015, pp. 1-4.

[30] P. Zhang, "Utility application experience of probabilistic risk assessment method," *IEEE/PES Power Systems Conference and Exposition*, 2009, pp. 1-7.

[31] PJM region transmission planning process, Oct. 2020. [Online]. Available: https://www.pjm.com/~/media/documents/manuals/m14b.ashx

[32] Transmission Planning, 2020. [Online]. Available: https://www.michigan.gov/documents/mpsc/MPG_Adv_Planning_1.19.21_Presentation_713269_7.pdf

[33] 2020 MRO regional risk assessment. [Online]. Available:

https://www.mro.net/MRODocuments/2020%20Regional%20Risk%20Assessment%20Final.pdf

[34] Southwest Power Pool, 2020. [Online]. Available: https://spp.org/documents/62876/spc%20workshop%20materials%202020%2008%2031%20v3.pdf

[35] Black Hills/Colorado Electric, LLC, 2010. [Online]. Available: http://www.oasis.oati.com/BHCT/BHCTdocs/BHCE_Method_Criteria_Process_Business_Practice_04082010.pdf

[36] BC Hydro transmission system studies guide, 2014. [Online]. Available: https://www.bchydro.com/content/dam/BCHydro/customer-portal/documents/corporate/regulatory-planning-documents/transmission-planning/transmission-system-studies-guide.pdf

[37] WECC, 2013. [Online]. Available: https://www.wecc.org/Reliability/131009-2013Plan-TheGuide-Final-v5.pdf

[38] Western planning regions and transmission planning coordination, 2015. [Online]. Available: https://westernenergyboard.org/wp-content/uploads/2015/04/04-15-15-ICF-WIEB-Western-Planning-Regions.pdf

[39] The future of electricity resource planning, Sep. 2016. [Online]. Available: https://eta-publications.lbl.gov/sites/default/files/lbnl-1006269.pdf

[40] 2020 SERC reliability risk report. [Online]. Available: https://www.serc1.org/docs/default-source/committee/ec-reliability-risk-working-group/2020-reliability-risk-report.pdf?sfvrsn=e80ea39_2

[41] S. Aboreshaid, “Probabilistic evaluation of transient stability of a power system incorporating a UPFC,” *Canadian Conference on Electrical and Computer Engineering*, 2003, pp. 441-446.

[42] M. Al-Sarray, H. Mhiesan, M. Saadeh, and R. McCann, “A probabilistic approach for transient stability analysis of power systems with solar photovoltaic energy sources,” *IEEE Green Technologies Conference*, 2016, pp. 159-163.

CHAPTER 2

BACKGROUND AND OVERVIEW

This chapter presents the background and overview related to power system stability, transient stability, probabilistic risk-based transient stability, machine learning, artificial neural network (ANN), and support vector machine (SVM).

2.1 Power System Reliability, Security, and Stability

Power systems usually consist of three different stages: generation, transmission, and distribution, as illustrated by Figure 2.1. In the first stage (generation), the electric power is generated, generally by using synchronous generators. Then, the transformers increase the voltage level, before the power is transmitted, to reduce the line currents, which consequently reduces the power transmission losses. After the transmission, the voltage is stepped down using transformers to be distributed to consumers [1].

The reliability of an electric power system is generally defined as "the probability that the power system will perform the function of delivering electric energy to consumers on a continuous basis and with acceptable service quality" [2]. Power system reliability evaluation is generally divided into two categories: adequacy and security [3-4], as illustrated by Figure 2.2. Adequacy involves the assessment that there are sufficient generation facilities in the system to meet the customer load demands, considering scheduled and rationally anticipated unscheduled outages of system components [5]. Adequacy is a steady-state issue and deals with both generation and transmission capacity. Security deals with the response of the system to sudden disturbances, such as line outages

or faults. Security is further divided into two categories: static and dynamic. Static security deals with steady-state analysis of post-disturbance system conditions to verify that there is no line overload or bus voltage violations. Dynamic security assessment (DSA) refers to the investigation required to determine whether a power system can satisfy the specified reliability criteria, under small or large disturbances [6]. Power system stability is "the ability of an electric power system, for a given initial operating condition, to regain a state of operating equilibrium, after being subjected to a physical disturbance, with most system variables bounded so that practically the entire system remains intact [7]." Stability is an important part of power system reliability and security.

The usual practice for DSA has been to use a deterministic approach. Generally, the power system is designed to endure a set of contingencies, chosen on the basis that they have a substantial possibility of occurrence. In practice, these contingencies are defined as "the loss of any single element in a power system, either spontaneously or preceded by a single, double, or a three-phase fault." This method is generally known as the (*N*-1) criterion, as it examines the behavior of an *N*-component network, following the loss of any one of its components [7-8].

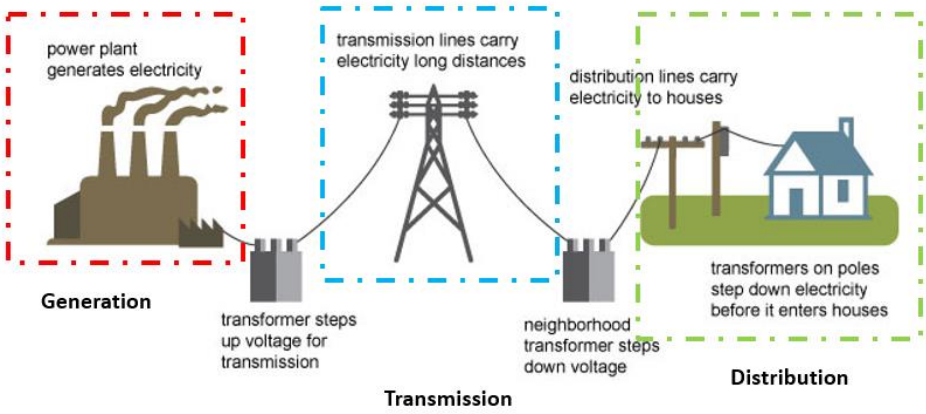

Figure 2.1. Power system composition

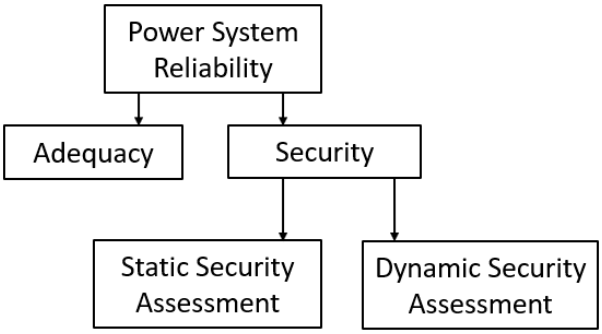

Figure 2.2. Power system reliability classification

The terms reliability, security, and stability are often confused and used interchangeably in research literature. It is important to distinguish between these terms and one must be careful as to when to use these terms. The major differences between these three significant aspects of power system are outlined below [7]:

1. "Reliability is the overall objective in power system design and operation. To be reliable, the power system must be secure most of the time. To be secure, the system must be stable, and must also be secure against other contingencies that would not be classified as stability problems e.g., damage to equipment such as an explosive failure of a cable, fall of transmission towers, due to ice loading or any other weather event. Moreover, a system may be stable following a contingency, but insecure because of post-fault system conditions, thereby causing line overloads or voltage violations.

2. System security may be further differentiated from stability in terms of the resulting consequences. For example, two systems may both be stable with equal stability margins, but one may be comparatively more secure because the consequences of instability have less severity.

3. Security and stability are time-varying attributes. This implies that they can be estimated by studying the performance of the power system, under a specified set of conditions. On the other hand, reliability is a function of the time-average performance of the power system, i.e., it can only be estimated by considering of the behavior of the system, over an appreciable period."

Power systems are generally operated to deliver continuous power supply that maintains the stability. However, due to unwanted events, such as lightning, weather or any other unpredictable events, short circuits between the phase wires of the transmission lines or between a phase wire and the ground may occur. This is commonly known as a fault. Due to a fault, one or more generators may be severely disturbed, resulting in an imbalance between generation and load. If the fault perseveres and is not cleared in a pre-specified time, it may result in severe damages to the equipment, which, in turn, may cause a power loss and ultimately, a power outage. Consequently, protective equipment is installed to sense faults and isolate faulted parts of the power system, as soon as possible, before the fault energy is proliferated to the rest of the network [1].

Power system stability is very imperative for maintaining continuous supply of power [9]. Instability of a power system can occur in many diverse scenarios, majorly contingent on the system configuration and operating mode. One of the key stability problems is to maintain synchronism, especially, when the power system relies on synchronous machines. The dynamics of generator rotor angles and power-angle relationships influence this aspect. Another instability problem that may frequently occur is voltage collapse, which is mostly associated with load behavior [1]. Voltage collapse is generally described as the

process by which the sequence of events accompanying voltage instability causes a blackout or abnormally low voltages in a large portion of the power system [7].

Since the early 20th century, power system stability has been documented as a noteworthy issue in secure power system planning and operation [10-11]. Most of the blackouts caused by power system instability have demonstrated the significance of this phenomenon [12-13]. Historically, transient stability has been the leading stability issue in most power networks. However, with the introduction of innovative technologies and increasing load demands, voltage and frequency stability have also gained importance. This has necessitated an understanding of the basics of power system stability.

2.2 Classification of Power System Stability

[7] has broadly classified power system stability into three major kinds: frequency, voltage, and rotor angle. This is pictorially shown in Figure 2.3. A brief description of this classification follows.

Frequency Stability

Frequency stability is the ability of a power system to maintain steady frequency after a severe system stress causes a substantial disparity between generation and load. It relies on the ability to maintain equilibrium between system generation and demand, with minimum inadvertent loss of load. Instability can manifest itself in the shape of sustained frequency swings, which subsequently cause tripping of generating units and loads. Frequency stability is determined by the overall response of the system (or each island if the system splits into islands). The governor response, also known as primary frequency control, ranges from few seconds to tens of minutes. Frequency stability analysis during this

duration is termed as short-term analysis. In the long-term analysis, central automatic generation control (AGC) directs specific generators to adjust their outputs to satisfy frequency constraints. It is important to note that frequency stability cannot be classified as small signal (small disturbance) or large signal (large disturbance) because it can occur due to any disturbance.

Voltage Stability

Voltage stability is the ability of a power system to maintain steady voltages at all buses in the system, after being subjected to a disturbance from a given initial operating condition. It depends on the ability to maintain equilibrium between system demand and system generation. Instability may manifest itself in the shape of a continuing decrease or increase of voltages at some or all buses. The culprit for voltage instability is typically the loads. Large disturbance voltage stability is the ability of a power system to maintain steady voltages after the occurrence of large disturbances, such as three-phase faults. The inherent features of system and load are major determinants of this ability. The period of interest of short-term, large disturbance, voltage stability typically ranges from a few seconds to tens of minutes. Small disturbance voltage stability refers to the ability of the system to maintain steady voltages when subjected to small perturbations, such as small changes in system load. Long term voltage stability constitutes of slower acting equipment, such as tap-changing transformers, thermostatically controlled loads, and generator current limiters. The period of interest may extend to several or many minutes, and long-term simulations are required to assess the dynamic performance of the system.

Rotor Angle Stability

Rotor angle stability is the ability of synchronous machines in a power system to maintain synchronism when a disturbance is applied. It depends on the ability to maintain/restore equilibrium between electromagnetic torque and mechanical torque of each synchronous machine in the system. The change in electromagnetic torque of a synchronous machine following a disturbance can be resolved into two components: (1) synchronizing torque component, in phase with rotor angle deviation, and (2) damping torque component, in phase with the speed deviation. Instability can result when the angular swing of generators causes a loss in synchronism. Small disturbance rotor angle stability is concerned with the ability of the power system to maintain synchronism under small disturbances, such as minor load variations. The disturbances are sufficiently small that linearization of system equations is permissible for analysis. On the other hand, large-disturbance rotor angle stability, or commonly called transient stability, focuses on the ability of the power system to maintain synchronism when a severe disturbance, such as a three-phase short circuit on a transmission line, is applied. The resulting system response consists of large excursions of generator rotor angles and is governed by the nonlinear power-angle relation. Transient stability relies, for the most part, on the initial operating state of the system and the severity of the disturbance. Instability is normally manifested in the form of aperiodic angular separations which are due to inadequate synchronizing torque. The time range for transient stability studies is about 3 to 5 seconds after the disturbance has occurred. It may range up to 10-20 seconds, for much larger systems with dominant inter-area swings. In addition to fault type, fault location, and system load, fault clearing time (FCT) and critical clearing time (CCT) are significant parameters in assessing transient stability [14]. FCT is the time

at which fault is cleared after fault occurrence, whereas, CCT is the maximum FCT after which the system becomes transiently unstable [14].

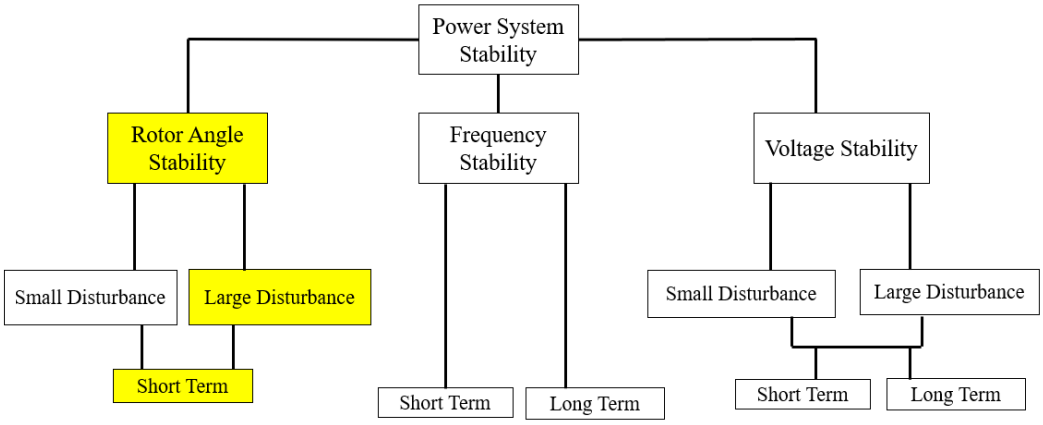

Figure 2.3. Power system stability classification (highlighted ones are considered in this work)

2.3 Historical Review of Power System Stability Problems

This section provides a brief historical review of problems related to power system stability [1]. "The methods of power system stability analysis were largely influenced by the development of computational tools, stability theories, and power system control technologies. Thus, it is indispensable to present a review of the history of the subject, to better comprehend the methods used in industries, regarding system stability and how these developments relate to the proposed approach in this research.

Power system stability is a multifaceted problem that has challenged power system engineers for many years. It was first recognized as a significant problem in 1920s [15]. The first field tests on the stability on a practical power system were conducted in 1925 [16-17]. The early stability problems were related to remote power plants, feeding load centers over long transmission lines. With slow exciters and non-continuously acting

voltage regulators, power transfer capability was frequently impeded by steady-state and transient instability, due to inadequate synchronizing torque [18].

In the early years, graphical methods such as Equal Area Criterion (EAC) and the power circle diagrams were established. These methods were successfully applied to early systems that could be represented as two-machine systems. As the systems became larger and more interconnected, the intricacy of the systems grew and thus, the stability problems became more complex. This annulled the treatment of the systems as two-machine systems. A noteworthy step towards the enhancement of stability computations was the development of the network analyzer in 1930, which was capable of power flow analysis of multi-machine power systems [9]. A network analyzer, fundamentally, is a scaled model of an AC power system, with adjustable resistors, inductors, and capacitors, to represent the transmission network and loads, voltage sources, whose magnitudes and angles are adjustable, and meters to measure voltages, currents, and power anywhere in the network. However, system dynamics still had to be solved by hand, by solving the swing equations, using step-by-step numerical integration. During this period, classical models were used for the swing equations. These models represented the generators as fixed transient reactances, with a fixed power supply behind these reactances.

In the early 1950s, electronic analog computers were used for analysis of special problems, which necessitated detailed modeling of the synchronous machine, excitation system, and speed governor. In addition, during that period, expansion of digital computers was seen, and specifically, around 1956, the first digital program, for power system stability analysis was developed. In the 1960s, most of the power systems in the USA and Canada were combined as part of one of two large interconnected systems, one in the East and the other

in the West. In 1967, low capacity high-voltage direct current (HVDC) ties were also established between the east and west systems. Today, the power systems in the United States and Canada form virtually one large system. This interconnection between the two systems results in operating economy and increased reliability, though, it increased the complication of stability problems and augmented the impacts of instability [9].

Until recently, most industry effort has been focused on transient stability [9]. Powerful transient stability simulation programs have been developed that are capable of modeling large multifaceted systems, using detailed models. In the early 1990s, the emphasis was on small signal stability, which led to the growth of special study techniques, such as modal analysis, using eigenvalue techniques. In the 1970s and 1980s, frequency stability problems were experienced following chief system upsets. This led to an examination of the primary causes of such problems, and to the development of long-term dynamic simulation programs, to support in their analysis. In 1983, guidelines were established, for improving power plant response, during major frequency disturbance events.

Recently, power systems are being operated under increasingly stressed conditions, due to the dominant trend to make the most of existing facilities. Increased competition, open transmission access, and construction and environmental constraints are determining the operation of electric power systems, which present greater challenges for secure system operation. This is evident from the increasing number of major power-grid blackouts that have been experienced in recent years, such as Northeast USA-Canada blackout of August 14, 2003. Planning and operation of modern power systems require a cautious deliberation of all kinds of system instability. Momentous advances have been made in recent years, in providing better tools and techniques, to examine instability in power systems. Figure 2.4

presents a timeline view for historical appearances of different power system stability problems [19].”

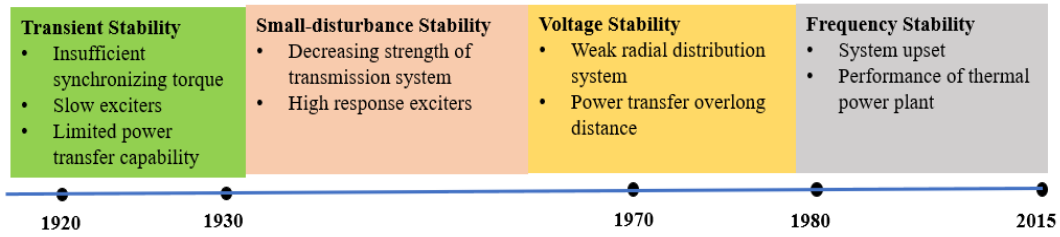

Figure 2.4. Historical appearances of power system stability problems

2.4 Applications of Power System Stability Studies

A power system must maintain its integrity after disturbances and must have the ability to withstand a wide variety of faults, for a reliable service. However, limited by economic and technical limitations, power systems, in practice, can only be designed to be stable, for selected disturbances, based on their probability of occurrence and severity. Power system stability related studies can ensure secure operation of transmission networks by:

1. Ensuring proper selection and deployment of protective and emergency control facilities.
2. Obtaining power system stability limits and ensuring operation stays within these limits.

Power system stability studies provide good references for the system operators, when monitoring system components from potential instability hazards. One of the tasks for a power system operator is to make sure that the system plants are operating under acceptable conditions and output reliable electric power, after being subjected to credible events under heavily stressed scenarios. There are several roles of system operators according to [20]. For example, the ISO (Independent System Operator) is responsible for planning and operation of the network, and TNSO and DNO (Transmission Network System Operator,

Distribution Network Operator) should be responsible for their own portion of the network [19]. The stable and economical operation of a power system within security limit is of great interest to these operators. Two types of studies, operational studies which focus on short-term secure and reliable operation of the network, and planning studies, which deals with long-term market-profiting operation of the network, are considered by operators and planners. Table 2.1 presents the applications of stability studies in different areas and time frames of system analysis [19]. Power system transient, small signal and voltage stability problems are equally important for all areas of the network and are of interest of both short-term operational studies and long-term planning studies. The frequency stability is usually of concern for operators, responsible for the whole system, and when the operational studies are considered.

Table 2.1. Power system stability aspects and their implications in system-wide regime

Power System Phenomenon	**Interest of ISO (whole network)**	**Interest of TNSO, DNO (part of the network)**	**Operational studies (short term)**	**Planning studies (long term)**
Transient stability	✓	✓	✓	✓
Small signal stability	✓	✓	✓	✓
Voltage stability	✓	✓	✓	✓
Frequency stability	✓	x	✓	x

2.5 Significance of Stability for Power System Security

Evaluation of the operating state of a power system is mandatory for the system operator in taking suitable decisions, and thereby, keeping the electrical properties within satisfactory limits. Also, analysis of the system performance under various conditions and phenomena helps designing the control systems. The power system can be seen as a "black box", with a well-defined function, to which consumers are connected. It should ensure the continuity of supply with electrical energy to the consumers irrespective of its operating conditions. In order to perform this function, the power system must be designed so that to survive any disturbance, e.g., short-circuits followed by a line tripping, tripping of any element without a fault, etc. [21].

The number of possible events that theoretically may occur is infinite and therefore, designing the power system to withstand all possible events or combination of events is literally impossible. The practice is to evaluate the power system state using one or more indexes for all possible events, and consequently, rank the events in descending order of the combined index values. The indices penalize violation of transmission capacity limit, voltage limits, stability limits, etc. Countermeasures are designed for those disturbances that the power system cannot survive. These measures are designed to counteract all critical disturbances starting from the top of the list [21].

As mentioned before, security is used to indicate if the power system is able to withstand disturbances. The power system security may be defined as the ability to withstand any kind of disturbance without interruption of the power supply service. It can be implied that the power system is fully reliable if it is secure at all times. Security may also be associated

to the term robustness, which reflects the way in which the power system can withstand disturbances. As mentioned before, the security of a power system is defined as its ability to overcome contingencies, such a severe fault, and deliver the energy to customers in abnormal conditions. A secure system is also a stable one—both during and after a contingency. Therefore, stability assessment of the system is essential to ensure its security. In other words, a secure system is the one that is stable for all credible contingencies. However, for the stability of the system, only the current system condition needs to be checked (without any contingency assessment) [22].

2.5.1 Power System Security States

"The national grid codes stipulate that the state quantities must remain within acceptable ranges for any disconnection of an element (line, transformer, generator, etc.). This is commonly known as the (*N*-1) security criteria. The operating conditions vary continuously, and the power system moves from one state to another, as indicated by S_i in Figure 2.5. Transition to one state or another depends on the random events that may occur or on the decision taken by the system operator. In the 'NORMAL' state, all parameters are within acceptable ranges and the power system is stable and secure. Furthermore, from this state, disconnection of any element can bring no harm to the power system. However, significant changes, for instance, sudden load increase or extreme weather conditions make the system vulnerable to disconnection on an element. and the power system may enter in the 'ALERT' state. Figure 2.5 illustrates a classification of possible states of the power system, depending on the event that may occur [21].

When the power system enters in the 'ALERT' state, immediate corrective actions, such as tap changing, active and reactive power control of generators, etc., must be taken in

order to restore the normal operation. The restoration process may take shorter or longer time depending on the dynamics of the corrective actions. If, during this transition, a contingency occurs, the system can enter in an 'EMERGENCY' state, in which there are a large number of bus voltage limits violations or exceeding of branch ampacity. In this state, extreme remedial actions, such as load shedding, tripping of transmission lines, disconnections of generators, etc., can still be taken and system restoring to a normal operation is possible. If the contingency is too severe, the power system may become instable and eventually enters the 'COLLAPSE' state [21].

It must be noted that violation of an operating constraint may not necessarily mean that the power system becomes instable. Due to high loading, the voltage in the system nodes can become too low and the loads demand more current which, in time, may overload the transmission lines. This will finally jeopardize the electrical network integrity, limits the transmission capacity and eventually, leads to instability conditions [21]."

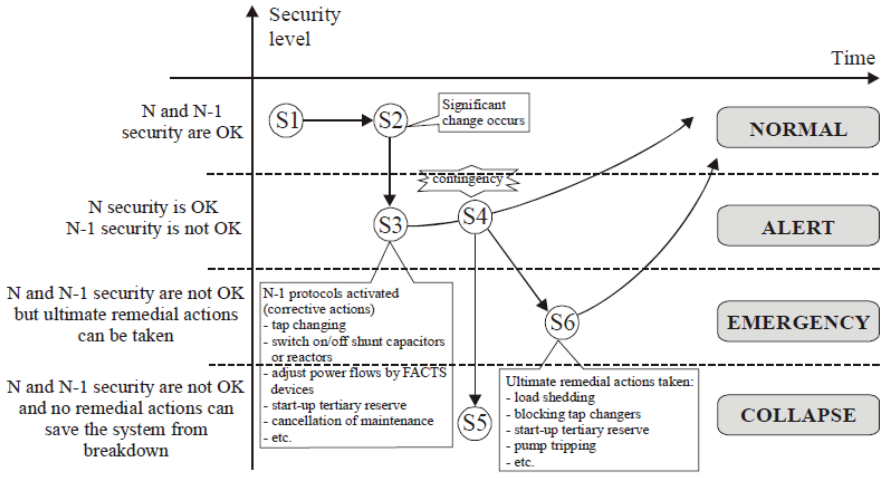

Figure 2.5. Power system security states

2.6 Factors affecting Transient Stability

There are various factors which affect transient stability. They are outlined below [23-24].

1. Pre-and-post-disturbance system state, such as the generators loading, before the fault and the generator outputs, during the fault, system load, and network topology. The higher the loading before the fault, the more likely is the system to be less stable during faults.
2. The duration, location and type of the fault determine the amount of kinetic energy gained by synchronous generator (SG). Longer fault duration allows generator rotors to gain more kinetic energy during the fault. At a specific limit, the gained energy may not be dissipated after the fault clearance. This gained energy may cause instability. Moreover, in general, three phase faults near generator terminals are more likely to cause instability, as compared to other kinds of faults.
3. System relaying, and protection have remarkable significance in system stability. The kinetic energy picked by the SG is directly proportional to fault duration. Therefore, faster fault clearing time results in a greater stability margin.
4. System reactance: the power transfer capacity is inversely proportional with the transmission reactance. The transfer capability in pre-fault conditions can be increased by decreasing the series reactances of the system. A decrease in the system reactances can be attained by the addition of parallel transmission lines and/or use of transformers with low leakage reactance voltage [25].
5. System voltage: The transfer capacity increases proportionally to the square of the system voltage. An increase in system voltage increases the difference between the initial rotor angle and critical clearing angle, thereby, allowing the generator rotation through a large angle deviation, before reaching the critical clearing angle.

It is understandable that increasing the system voltage is not applicable to a large existing power system, as the components of the power system are designed for a certain voltage level and must be replaced, before increasing the voltage level.

6. System inertia: During transient events, inertia provided by SGs counteracts the changes in frequency, and therefore helps with maintaining transient stability. Thus, a reduction in inertia increases the chance of a system to become unstable [26].

2.7 Deterministic and Probabilistic Transient Stability

Traditionally, deterministic criterion has been used for transient stability evaluation for power system planning and operation [27-28]. This method is generally considered for a single operating condition, commonly known as the worst-case scenario. In most cases, the (*N*-1) contingency principle is used. This means that individual system components are removed one by one for the analysis. The worst-case scenario then gets transformed to numerous extreme operating conditions, together with several most critical contingencies, for which the system should be designed to withstand. While this worst-case approach has served the industry well; however, in a competitive environment, the utilities will need to know the level of risk, associated with their observed criteria, so that they adjust their service quality based on the expectation of consumer i.e., the acceptable level of risk and corresponding price [29]. The conventional transient stability assessment follows a step-by-step procedure in which the factors such as the load, fault types, fault locations, etc., are selected in advance, usually in accordance with the worst-case philosophy [29].

Furthermore, to guarantee that the most severe disturbance is selected, the contingency types and locations are normally provided in advance.

In the deterministic approach, the contingencies are selected based on the probability. Thus, to some extent the probability of events is considered. Though, once the contingencies are selected, they are treated as equally probable and operation limits are introduced based on the influence of contingencies. In the probabilistic approach, these operating limits are computed by considering both consequence (severity) and probability [30]. The deterministic approach has at least the following three drawbacks [31].

1. "Only consequences of contingencies are evaluated, but probabilities of occurrence of contingencies are ignored. Even if the consequence of a selected contingency is not very severe, system risk could still be high, if its probability is relatively large. Conversely, if the probability of an outage event is extremely small, the contingency analysis of such an event may result in an uneconomic operational decision.
2. All uncertain factors that exist in real life (such as uncertainty of load variations, variability of renewable generation, random failures of system components, fuzzy factors in parameters or input data, errors in real-time information, volatility of power demand on the market, etc.) are ignored in the deterministic analysis. This can lead to results, biased from the reality.
3. The deterministic approach is based on pre-selected worst cases. In implementation, however, the actual worst case may be missed [31]."

Moreover, as the result of deterministic stability analysis is binary (stable or unstable), therefore, the transient instability risk cannot be quantified. As a matter of fact, the electric

power sectors need to know the risk level to take actions to upsurge the system security. Therefore, examining the system transient stability by applying risk assessment has become a critical research technique [32]. A pictorial representation of a typical framework for deterministic transient stability study is shown in Figure 2.6.

With the current drift towards competitive and deregulated electricity market environment, the power utilities are required to guarantee, besides a safe reliability level, an economical operational efficiency. They are forced to maximize the utilization of their existing facilities. Therefore, some operators consider exploring the operating areas, beyond the traditional operating limits, where the system is vulnerable to costly outages, in order to see whether the incurred risk weighs against the potential economic benefits of violating the limits. In these situations, a probabilistic assessment approach becomes tremendously beneficial [33]. The probabilistic studies consider the stochastic and probabilistic nature of the real power system. It considers the probability distribution of one or more uncertain parameters, and hence, reflects the actual system in a better manner. Although, it has been long established that deterministic studies may not sufficiently characterize the full extent of system dynamic behavior, the probabilistic approach has not been extensively used in the past in power system studies, mainly due to lack of data, limitation of computational resources, and mixed response from power utilities and planners [28, 30-31]. Probabilistic approaches are mainly appropriate, for the examination of a system, with randomness and uncertainty, which are obviously the main features of future power networks.

In the past several years, there has been a considerable increase in connections of intermittent and stochastic, power electronics interfaced renewable energy generation sources. These uncertainties, coupled with load uncertainties, are becoming one of the

crucial characteristics of modern power systems. The transient stability assessment of such systems using traditional deterministic methodology is swiftly becoming inappropriate and thus, unique probabilistic assessment methods are desirable, and are being established [34]. Although, the deterministic approaches result in highly secured power systems, but they do not consider the probability of operating conditions. Consequently, apart from the high cost due to conservative designs, the chief disadvantage with the deterministic assessment techniques is that they treat all security problems to have equal risk [35]. The rising power system uncertainties has motivated the application of probabilistic methodologies, for transient stability assessment. It is, thus, of great significance to propose a risk-based approach, for overcoming the shortcomings of the deterministic approach. The probabilistic analysis can provide a more inclusive, coherent, and realistic measure of the system stability level; consequently, this type of assessment can provide a profound understanding of system stability problems, as compared to the deterministic stability assessment. A pictorial representation of a typical framework for probabilistic transient stability (PTS) study is shown in Figure 2.7.

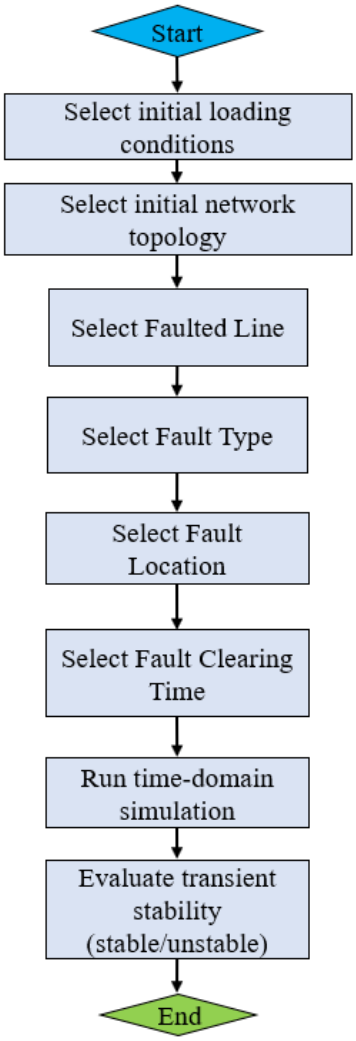

Figure 2.6. Framework for deterministic transient stability assessment

2.8 Transient Stability Assessment Methods

There are three main methods for assessing transient stability [33, 36-37], as elaborated in Figure 2.8. A brief overview of these methods is discussed below.

Time-domain Simulation Method

While formulating the mathematical model of a power system with its components, the most traditional way to observe its dynamic behavior, after a disturbance, is by numerically

integrating a set of differential equations, over a certain time. This method is also known as assessment by time domain simulation (TDS). Since almost any component can be encompassed in this mathematical model, there are no modeling boundaries.

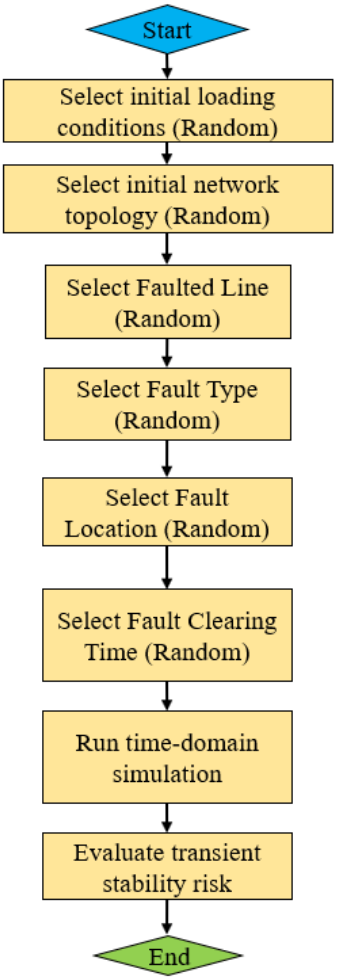

Figure 2.7. Framework for probabilistic transient stability assessment

The TDS approach is one of the traditional approaches used for transient stability assessment (TSA) [38]. It is the most accurate, and time-consuming category. This method requires the whole system detailed model and accurate information about disturbances, to solve the nonlinear differential–algebraic equations, and is based on numerical integration [39-41]. It is typically used to assess transient stability status, and to provide detailed

operational information of the faulted system. Although, this approach is a popular method due to its high accuracy, but it is inappropriate for online TSA, chiefly due to its extensive computation effort. Moreover, this method does not have any criterion to indicate the system stability before the fault clearance [42-43]. One of the primary works in this area [44], by determining the coherent group of the generators, relies on developing a low order equivalent model of the power system, to forecast stability status. To further simplify, the piecewise constant current load approximation technique was used. [44] used the swing equation, in a simplified two machine system, to envisage the rotor angle, about 0.5 second later. [45] proposed an implicitly decoupled active-reactive power (PQ) integration approach, to decrease the computation load.

Direct Methods

Direct methods are also conventional TSA approaches, which substitute the numerical integration of the post-fault system equations, by a stability standard [46]. Direct methods of TSA include equal area criterion (EAC) approach, extended equal area criterion (EEAC) technique, and energy function technique, based on the transient energy function (TEF) technique. A direct method, based on TEF, is conducted by constructing Lyapunov functions such that transient energy and system stability can be determined, while the EEAC technique (an extension of EAC approach) is the graphical solution, used to determine the stability of single machine infinite bus (SMIB), as elucidated in [47]. Since the EEAC technique considers SMIB, therefore, the system needs to be oversimplified, therefore, it cannot be applied to complex interconnected power systems. Similarly, the TEF technique presented a restriction of only capable of forecasting the stability, based on the first swing and if the second swing presents itself and go unstable, then the stability

examination will be inaccurate [48]. The TEF method is based on Lyapunov's second theory. The main advantage of Lyapunov's second theory is that the Lyapunov function is invariant from systems nonlinear equations [49]. In the literature, several functions were proposed as Lyapunov functions [50]. Determining an appropriate Lyapunov function for a bulk power system, effort in computing the level of the kinetic and potential energies, and its inefficiency are some of the disadvantages of the TEF method [51-52]. Another downside of this class is the evaluation of critical energy [53]. Although, these methods have a low computational burden, they give conservative results, since it is problematic to determine the exact energy function, and it is complicated to precisely define Lyapunov functions [36]. The accuracy and reliability of the TEF modeling is another challenge [52]. The TEF-based methods are difficult to implement, especially due to many potential function terms of the TEF of the system. Also, these approaches require postfault data for TSA, and hence, they are not suitable for TSA for online applications [43].

The first work, in the class of direct approaches, was published in 1989, which used a quasi-unstable equilibrium point (QUEP) approach [54]. This approach, by using generating unit clustering procedure, decreased the complexity of power system. It also computed stability sensitivity coefficients, for use in constrained economic dispatch. In [50], after fault occurrence, the total energy of the power system was observed incessantly. Also, the total energy function, based on Lyapunov's direct method, was computed. If this energy was greater than the stability limit, the out-of-step condition was forecasted. This approach, by using the center of oscillations, streamlines the power system into a two-machine system. In [55], the stability margin was assessed using the TEF technique. In the off-line stage, various fault conditions and post-fault topologies were inspected, and the corresponding

modes were stored in a look-up table. In the online step, after fault clearance, the normalized kinetic energy was compared with the look-up tables, to determine the most probable ones. Eventually, the unstable equilibrium point (UEP) and the associated control action were assessed. This technique required an observable system, with adequate number of phasor measurement units (PMUs).

Hybrid Method

The hybrid method integrates the TDS and the TEF methods of investigating transient stability, in which the TDS computes the real system trajectory, while the TEF constructs a stability index of the dynamic security assessment (DSA), as stated in [56]. Hybrid methods are also inadequate for online DSA, as prediction is dependent on the UEP, due to the TEF idea used, as explained in [56]. Another drawback is that the system representation becomes complex, as they use TEF concept. Research work associated with hybrid methods can be found in [41, 57-60].

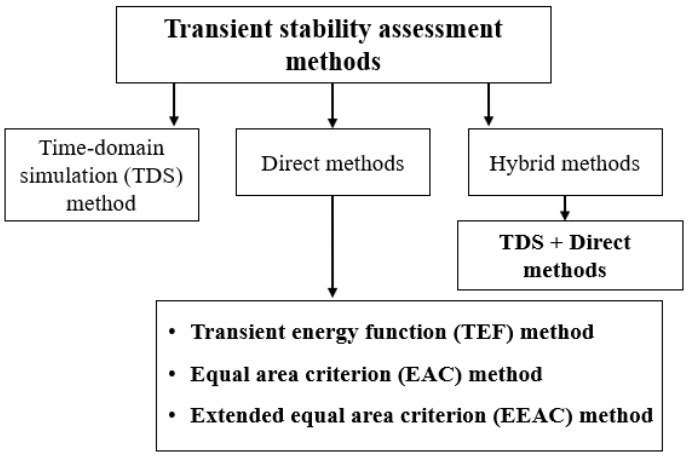

Figure 2.8. Methods for transient stability assessment of a power system

2.9 Swing Equation and Transient Stability

A power system, typically, consists of various synchronous machines, operating synchronously, under all operating conditions. The swing equation describes the rotational dynamics of a synchronous machine and is used in transient stability analysis to characterize that dynamic. “Under normal operating conditions, the relative position of the rotor axis and the resultant magnetic field axis is fixed. The angle between the two is commonly known as the power angle, load angle or torque angle. During any disturbance, the rotor decelerates or accelerates, with respect to the synchronously rotating air gap magnetomotive force, creating relative motion. The equation describing the relative motion is known as the swing equation, which is a non-linear second order differential equation that describes the swing of the rotor of synchronous machine [61-62].”

“The goal of transient stability study is to determine if the generator rotor can return to constant speed, after a disturbance. Using the simple equivalent model of SG, the equation representing the SG rotor motion is given as

$$J\frac{d^2\theta}{dt^2} = T_m - T_e \tag{2.1}$$

where J is generator’s moment of inertia (kgm^2), θ is the angular displacement of the rotor with respect to a stationary axis on generator stator (in radians), t is the time (s), T_m is the input mechanical torque (Nm), and T_e is the output electrical torque (Nm).

If the generator’s internal friction losses and the heating losses are neglected, to maintain synchronous speed under ideal operation situation, the input mechanical torque T_m and the output electrical torque T_e must be equal. When the input mechanical torque is greater than the output electrical torque, the generator rotor will accelerate and vice versa. Figure 2.9 is a simplified diagram of a SG. It illustrates the stator, rotor, input mechanical torque and

output electrical torque. ω_m denotes the synchronous speed of generator. In a 60 Hz power system, it equals to 120pi rad/s, i.e., 3600 RPM (revolutions per minute).

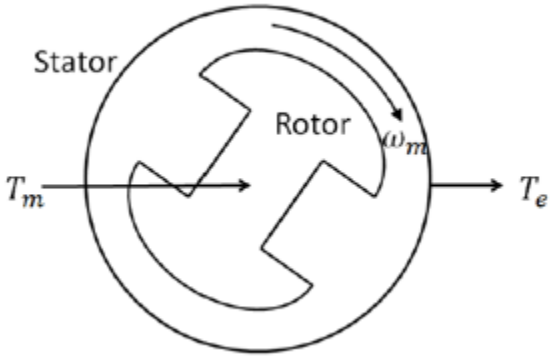

Figure 2.9. Simplified diagram of the SG

In power systems, most of the SGs are driven by the steam turbine. The input torque T_m of this kind of generator is controlled by the turbine governor. The governor adjusts the amount of steam entering the steam turbine, according to the generator output power. The output torque T_e is the equivalent torque which relates to the power fed into the power system. It reflects the instantaneous power system operation status. Due to the physical nature of the steam turbine, the generator input torque cannot be adjusted immediately. After the disturbance, because of the slow response speed of the input torque, when the output torque is less than the input torque, it is possible that the generator will gain enough energy to keep its rotor accelerating forever. This is the inherent nature of power system transient stability – the balance of generator input and output torque (power). The generator output torque cannot be directly obtained because only the generator electric power output can be measured. The electric power equals the torque multiplies the angular velocity. When the generator is synchronous with the power grid, the angular velocity is called

synchronous speed, ω_m. The relation between generator electric power, P_G, and the output torque, T_G, is illustrated using (2.2) [61]."

$$P_G = T_G \times \omega_m \tag{2.2}$$

Substituting (2.2) in (2.1),

$$J\omega_m \frac{d^2\theta}{dt^2} = P_m - P_e \tag{2.3}$$

where P_m and P_e denote input mechanical power and output electrical power, respectively.

The angle θ in (2.3) is measured with respect to a stationary reference axis on the stator. This means its value is increasing continuously with time. The most common way of describing the change of generator rotor angle is to use the synchronous speed as the reference. Therefore, (2.4) defines the generator rotor angle displacement, with respect to the synchronous speed.

$$\delta_m = \theta - \omega_m t \tag{2.4}$$

where δ_m denotes the electrical angle for distinguishing θ and δ_m.

The second order derivative of (2.4) with respect to time is

$$\frac{d^2\delta_m}{dt^2} = \frac{d^2\theta}{dt^2} \tag{2.5}$$

Substituting (2.5) in (2.3) produces the generator swing function used for transient stability studies, i.e.,

$$M\frac{d^2\delta_m}{dt^2} = P_m - P_e \tag{2.6}$$

Where M is known as the moment of inertia of the rotor mass.

In (2.6), P_m can be measured at the prime mover of the generator, and P_e is the electrical power output which is computed by the power flow equation. Equation (2.6) is known as the swing equation which describes the rotational dynamics of a synchronous machine for transient stability studies.

Figure 2.10 illustrates the behavior of a synchronous machine for stable and unstable situations [25, 63]. "In Case 1, the rotor angle increases to a maximum, then decreases and oscillates with decreasing amplitude until it reaches a steady state. This case is considered transient stable. In Case 2, the rotor angle continues to increase steadily (due to insufficient synchronizing torque) until synchronism is lost. This type of transient instability is referred to as first-swing instability. In Case 3, the system is stable in the first swing, but becomes unstable because of growing oscillations, as the end state is approached. This form of instability occurs, when the post-fault steady-state condition is itself is small-signal unstable, i.e., it occurs due to lack of sufficient damping torque in the post fault system condition."

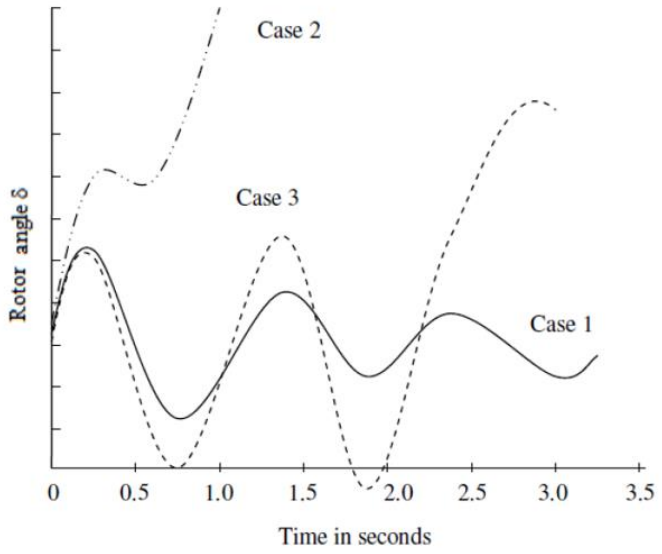

Figure 2.10. Rotor angle response to a transient disturbance

2.10 Computational Approaches for Probabilistic Transient Stability

This section briefly describes different computational approaches [64-66], which are generally applied for PTS analysis. These methods include Monte Carlo (MC), Quasi-Monte Carlo [67], Markov chain Monte Carlo [68], point estimate (PE) method [69], cumulant-based method [70], probabilistic collocation approach [71], convolution method [72], first-order-second-moment method [73], and unscented transformation [74]. This research uses MC simulation and therefore, it is discussed firstly in detail. However, a brief description of other methods is also provided.

Monte Carlo Method

The MC approach is usually considered as the standard approach for probabilistic simulation. In this method, there is a large number of random sampling of system uncertainties to obtain a large data set (i.e., a numerical solution), for determining the distribution of an unknown probabilistic entity, i.e., output probability density functions (PDFs). The accuracy of the output increases with the number of simulations. The MC simulation involves these key steps: outlining a domain of possible inputs, generating random input samples from the input probability distribution over the domain, executing deterministic simulations and analysis for each sample, and eventually, empirically analyzing the results. The MC method involves the repeated sampling of system uncertainties [66]. A large data set can be retrieved from these samplings, and the distribution of an unknown probabilistic entity can be determined. There are two essential theorems of statistics behind the application of MC method in uncertainty-related analysis: the Weak Law of Large Numbers [75], and the Central Limit Theorem [76]. The Weak

Law of Large Numbers states that the sample average converges in probability towards the expected value, and the Central Limit Theorem states that the properly normalized sum of independent random variables (RVs) tends toward a normal distribution. The simulation procedure generally involves inputs domain definition, probability distribution generation, performing deterministic computations on the inputs and aggregation of the results. The advantage of MC simulation is that the method is very flexible and virtually limitless for analysis. It is easily expandable and deployable. However, the simulation time can be quite long with large sampling database, as random samples of size N are generated following the PDFs of the input variables, and the accuracy of the samplings is highly related to the MC run-time. The MC stopping rule is needed to ensure that sufficient number of simulations are run to ensure required accuracy of the results. The outputs of this approach are estimate values rather than the exact values and the stopping rule can help to determine the minimum number of simulations required to achieve a specified confidence level. The numerical steps for obtaining the minimum number of simulations are given by (2.7) to (2.12) [76].

1. Take N samples and record $X_1,\ldots, X_N$.

2. Continue sampling until $\frac{T_{N+1}}{N} \le \frac{\varepsilon^2{}_N}{\beta^2{}_N}$, with $N+1 \to N$, and record X_N for N^{th} sample.

Calculate:

$$\overline{X_N} = \frac{1}{N}\sum_{i=1}^{N} X_i \tag{2.7}$$

$$T_N = \sum_{i=1}^{N} (X_i - \overline{X_N})^2 \tag{2.8}$$

3. Stop sampling, then:

$$\lim_{\varepsilon_r \to 0} \frac{N}{n_0(|\mu|\varepsilon_r, \delta)} = 1 \tag{2.9}$$

$$\lim_{\varepsilon_r \to 0} pr(\left|\overline{X_N} - \mu\right| \le |\mu|\varepsilon_r) = 1 - \delta \tag{2.10}$$

$$\lim_{\varepsilon_r \to 0} \frac{E(N)}{n_0(|\mu|\varepsilon_r, \delta)} = 1 \tag{2.11}$$

$$\beta = \phi^{-1}(1 - \frac{\delta}{2}) \tag{2.12}$$

In the above equations, n_0 is the initial iteration value, ε_r is the relative error for the samples, β is the maximum uncertainty level that should be achieved, *N* is the actual sample size needed for the required accuracy ε_r with the coverage probability $1-\delta$, as ε_r tends towards 0, $\overline{X}$ is the mean value of the obtained result, μ is the real mean value of the studied variables, ϕ^{-1} is the inverse function of the normal distribution with δ =1 and $\mu = 0$. The expression for the relative error is given by

$$\varepsilon_r = \frac{\phi^{-1}(1 - \frac{\delta}{2}) \times \sqrt{\frac{\sigma_N}{N}}}{E(N)} \tag{2.13}$$

where σ_N denotes the variance of the obtained result, and *E(N)* is the mean value of the samples. The relative error is computed in each simulation and is compared with a target relative error in MC simulation. Therefore, the simulation can be terminated when required confidence level is achieved. The MC method is very flexible and virtually limitless for analysis. However, for a larger test system, with a huge number of uncertain variables, it can consume a lot of computational resources.

Quasi Monte Carlo (QMC) Method

The principle of Quasi-Monte Carlo (QMC) approach is similar to the standard MC method; however, a distinct approach is used to generate the sample sets. While the MC method produces samples using a pseudorandom sequence to accurately sample from the input distributions, QMC modifies this sampling, deliberately, to more efficiently cover the desired portion of the input domain. This is usually attained using quasi-random (low-discrepancy) sequences, such as Halton or Sobol sequences, to generate samples that are equidistant in the input domain rather than equiprobable [77].

Markov Chain Monte Carlo (MCMC)

Markov Chain Monte Carlo (MCMC) approach produces samples from a probability distribution, by constructing a Markov Chain with a target distribution. Consequently, the Markov Chain produced can be used as a sample of the desired distribution. The convergence of the entire process progresses with the sample size. The MCMC is frequently used when it is problematic to directly sample the target distribution. The Markov Chain production replaces the sample generation portion of a standard MC approach, after which the individual deterministic simulations are executed, and results collected and analyzed. The key steps of an MCMC simulation are initializing an arbitrary starting sample, producing candidate points, repeating the sample generation until the convergence criteria are fulfilled, and obtaining targeted variable set, along with the desired distribution.

Point Estimate Method

The PE methods use a small number of specified point values to represent the distributions of network uncertainties. Consequently, the point values for different uncertainties are

combined in different permutations to form concentrations. The system model is then evaluated for these different concentrations, and the output values are combined, with concentration weightings, to approximate the moments of the output distribution. Different PE methods use different numbers of points, such as $2n$, $2n+1$, and $4n+1$ (where n is the number of uncertain parameters).

Cumulant-based Method

A cumulant is a statistical measure of a distribution, i.e., an alternative to the moment of the distribution. The input cumulants are mapped to the output cumulants, through sensitivity functions describing the input–output behavior. The cumulants of the system output can be described simply by a sum of the cumulants of the independent input uncertainties. The cumulant is simple to calculate, and the output cumulants can then be used to determine the output moments. This delivers an analytical solution for obtaining output variation based on the input uncertainty. These methods are tremendously dependent on the accuracy of this relationship and are frequently only valid when the uncertainty is minor and the input–output relationship is roughly linear.

Probabilistic Collocation Approach

The probabilistic collocation method (PCM) expresses the system model output as a polynomial function of the uncertain parameter set. The fundamental concept is to use a small number of sample points, to generate a computationally cheap function, that can be used to replace the intensive computation burden, of the full power system study, during further repeated sampling. Polynomial functions of increasing order and complexity can be used to capture high-order interactions, though at a larger computational cost. Uncertainties, generally, must be pre-ranked to classify critical parameters, as the number

of samples required to generate the function grows exponentially, with the number of considered uncertain variables.

Convolution Method

In this method, linearization methods are applied to represent line load flows as a linear combination of input variables. Assuming the independence of these variables, the PDFs of probabilistic input variables can be obtained. The main issue of this method is to compute the equivalent discrete function. When this function, characterized by "*u*" impulses, is convolved with another, having "*v*" impulses, the resulting function will have "*u* times *v*" impulses. The key issue related to this approach is that it demands a large amount of storage and computation time, especially for large systems [78-79].

First-Order-Second-Moment Method

The first-order second-moment (FOSM) method is based on a first order Taylor series estimation of a linearized function. This approach requires only expected values and covariances of RVs. One chief benefit of this approach is that it allows the evaluation of uncertainty in the output parameters, without previous knowledge of the PDFs of input RVs. Though, the approach is not suitable when the expected value of the input RV is in close vicinity to the global extremum of the function. In this scenario, the computed uncertainty of the RV may vary extensively from the actual one [78-79].

Unscented Transformation

The method is based on computing the statistics of a RV, which goes through a set of nonlinear transformations. This approach is time-efficient and can easily be applied to problems constituted of correlation, among multiple uncertain input parameters. However, the run time of the approach is directly linked to the number of uncertain input RVs [78-

79].

2.11 Probabilistic Factors in Transient Stability

There are various factors which are involved in PTS assessment of power systems, such as fault type, fault location, load, and FCT. Suitable PDFs are used to model these factors. The modeling approaches are described below [80].

Normally, shunt faults, such as three-phase (LLL), double-line-to-ground (LLG), line-to-line (LL) and single-line-to-ground (LG) short circuits, are considered for evaluating PTS. A probability mass function (PMF) is normally used to model the fault type. Based on past system statistics, a usual practice is to select the probability of LLL, LL, LLG, and LG short circuits, as 0.05, 0.1, 0.15 and 0.7, respectively [81].

The probability distribution of fault location on a transmission line is usually assumed to be uniform. This means that the fault can occur with equal probability at any line of the test system and at any point along the line [34]. Generally, very limited historical information is available on the locations of the faults on the lines; thus, it is reasonable to assume that nearly every location on a line has the same probability to be struck by a fault. The procedure of fault clearing consists of three stages: fault detection, relay operation and breaker operation. If the primary protection and breakers are 100% reliable, the clearing time is the only uncertain factor. A normal (Gaussian) PDF is generally used to model this time [34]. The PDFs, or PMFs, in case of discrete RVs, used for faulted line, fault type, fault location (on the line), and FCT are shown in Figures 2.11-2.14, respectively. To incorporate the uncertainty of loads, a normal PDF was used. Let $f(X_i)$ denote the PDF for load at i^{th} bus, i.e.,

$$f(X_i) = \frac{1}{\sqrt{2\pi\sigma_i^2}} e^{\frac{-(X_i-\mu_i)^2}{2\sigma_i^2}} \tag{2.14}$$

where μ_i and σ_i denotes the mean and standard deviation of the forecasted peak load for i^{th} bus, respectively.

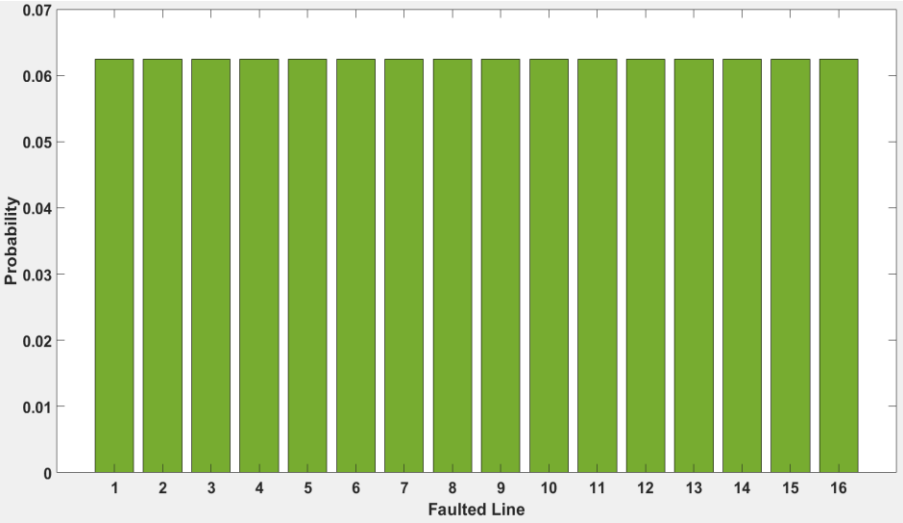

Figure 2.11. PMF for faulted line

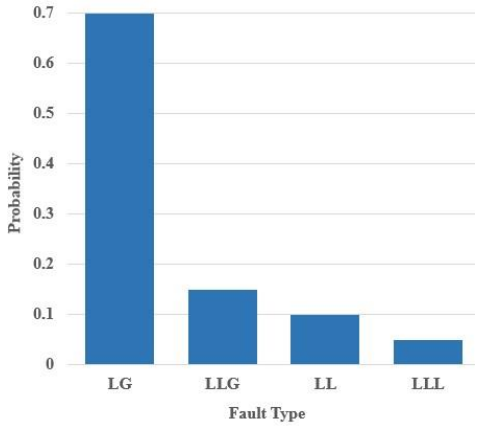

Figure 2.12. PMF for fault type

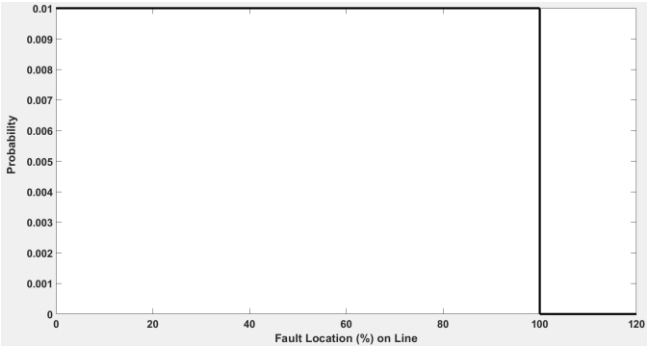

Figure 2.13. Uniform PDF for fault location (on the line)

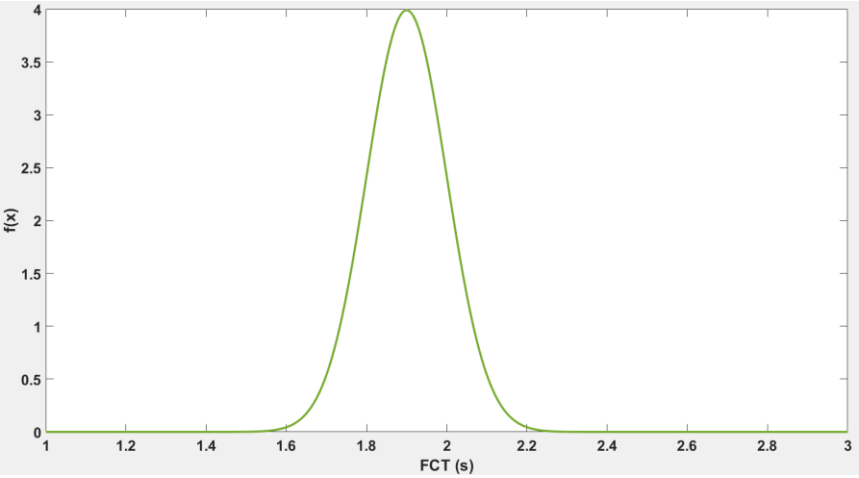

Figure 2.14. Normal (Gaussian) PDF for FCT

2.12 Quantification Index for Probabilistic Transient Stability

To quantify the output for PTS approach, different researchers have proposed various indices. These indices include *probability of instability of different lines* [82], *transfer limit calculation* [29], *generation rejection requirement* [29], *probability of system instability* [83], *maximum rotor angle deviation* [84], and *expected frequency of transient instability* [85]. In recent years, Transient Stability Index (*TSI*) has been used to quantify the transient stability of a system consisting of synchronous generators [86-87]. It is the most commonly used index for describing the transient stability of a power system with synchronous

generators. This index is based on the maximum rotor angle separation between any two synchronous generators, after the fault has occurred. Mathematically, it is given by

$$TSI_i = \frac{360 - \delta_{\max i}}{360 + \delta_{\max i}} \tag{2.15}$$

where $\delta_{\max i}$ is the post-fault maximum rotor angle separation (in degrees) between any two synchronous generators in the system at the same time (for a fault on i^{th} line). A negative *TSI* value specifies that the power system is unstable. For instance, Figure 2.15 illustrates the variation of maximum rotor angle difference, δ_{max}, for typical stable and unstable situations in a power system.

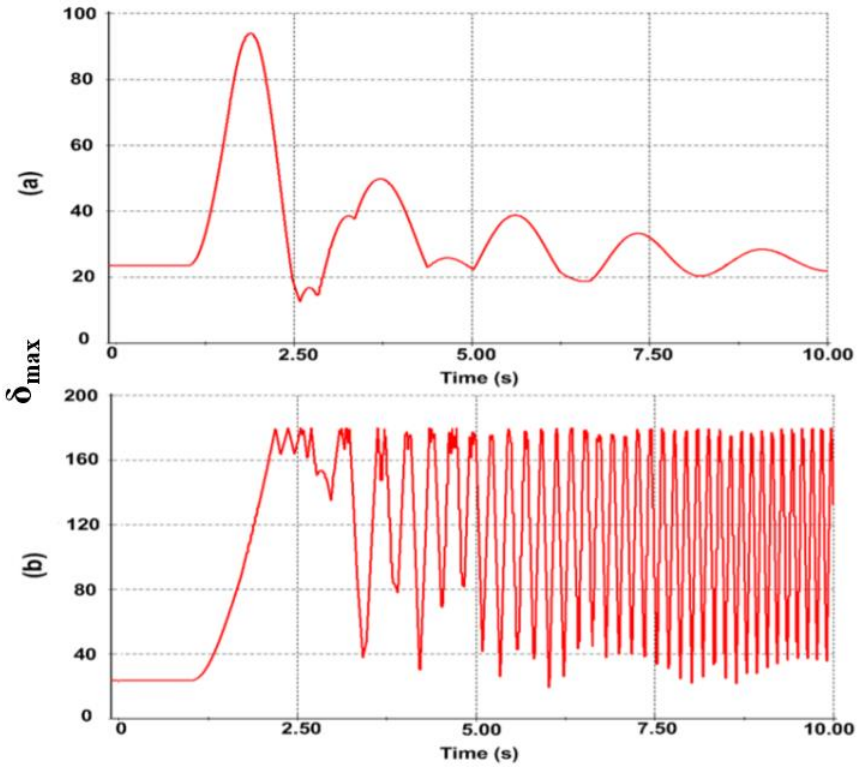

Figure 2.15. Maximum rotor angle difference: (a) stable case (b) unstable case

The *TSI* provides a swift indication of the transient stability status of the system (for a fault on any line, at any point, for any FCT and for any load). Therefore, this index is used in this paper to quantify the PTS status. Let S_i represent the PTS status indicator for the i^{th} iteration of MC simulation. Mathematically,

$$S_i = \begin{cases} 1\ , \text{ if } TSI_i < 0 \text{ (unstable)} \\ 0, \text{ if } TSI_i > 0 \quad \text{(stable)} \end{cases} \tag{2.16}$$

Therefore, if the system is transiently stable, for i^{th} MC sample, value of S_i will be 0; otherwise, it will be 1. This information will be used for training the machine learning (ML) algorithms.

2.13 Modeling of Generation Sources

This section briefly describes the modeling of generation sources including, SG and wind generator, specifically doubly fed induction generator (DFIG).

2.13.1 Synchronous Generator

Standard 6th order model is used for modeling all SGs. In this model, four windings are considered, two on the *q*-axis and two on the *d*-axis. However, the network and stator transients are neglected. All synchronous machines are equipped with (*TGOV1*) turbine governor, (*IEEEX1*) exciter and (*STAB1*) power system stabilizer. In this model of SG, the field coil on the direct axis (*d*-axis) and damper coil on the quadrature axis (*q*-axis) are considered. The associated mathematical model, for a standard 6th order SG, is given by the following equations [88].

$$T_J \frac{d\omega}{dt} = M_m - M_c - D(\omega - \omega_o) \tag{2.17}$$

$$\frac{d\delta}{dt} = \omega - \omega_o \tag{2.18}$$

$$T'_{do}\frac{dE'_q}{dt} = E_{fd} - (x_d - x'_d)I_d - E'_q \tag{2.19}$$

$$T''_{do}\frac{dE''_q}{dt} = -E''_q - (x'_d - x''_d)I_d + E'_q + T''_{do}\frac{dE'_q}{dt} \tag{2.20}$$

$$T'_{qo}\frac{dE'_d}{dt} = -E'_d + (x_q - x'_q)I_q \tag{2.21}$$

$$T''_{qo}\frac{dE''_d}{dt} = -E''_d - (x'_q - x''_q)I_q + E'_d + T''_{qo}\frac{dE'_d}{dt} \tag{2.22}$$

$$E'_q + x'_d I_d - R_a I_q = v_q \tag{2.23}$$

$$E'_d - x'_q I_q - R_a I_d = v_d \tag{2.24}$$

where T_J is generator inertia constant; M_m is mechanical torque; M_c is electromagnetic torque; D is damping coefficient; δ is rotor angle; ω is rotor speed; ω_o is synchronous speed; E_{fd} is field voltage; E'_d and E'_q are d and q axis components of transient electric potential; E''_d and E''_q are d and q axis components of subtransient electric potential; x_d , x'_d, x''_d, x_q, x'_q, x''_q are d and q axis synchronous reactance, transient reactance and subtransient reactance, respectively; T'_{do} , T'_{qo} are d and q axis transient time constants; R_a is armature resistance, v_d and v_q are d and q axis components of stator terminal voltage, and T''_{do}, T''_{qo} are d and q axis subtransient time constants.

2.13.2 Wind Generation

The dynamic behavior of wind generation is significantly different from the conventional SGs. There are four main types of wind generators. Although, DFIG wind generator is used in this research, a brief overview of other kinds is also presented [89].

Fixed Speed Induction Generator (Type 1)

In the early stages of wind power development, most wind farms were equipped with fixed speed wind turbines and induction generators. A fixed speed wind generator is typically equipped with a squirrel cage induction generator whose speed variations are limited, as shown in Figure 2.16. Power can only be controlled using pitch angle variations. Since induction machines have no reactive power control capabilities, fixed or variable power factor correction systems are typically required for compensating the generator reactive power demand [90]. The turbine speed is fixed (or nearly fixed) to the electrical grid's frequency and produces real power when the turbine shaft rotates faster than the electrical grid frequency creating a negative slip (positive slip and power is motoring convention) [89]. This kind of wind generator is usually referred to as a constant/fixed speed wind generator, as its rotor speed often fluctuates within a very small range, such as 1% to 2% of the rated speed [89].

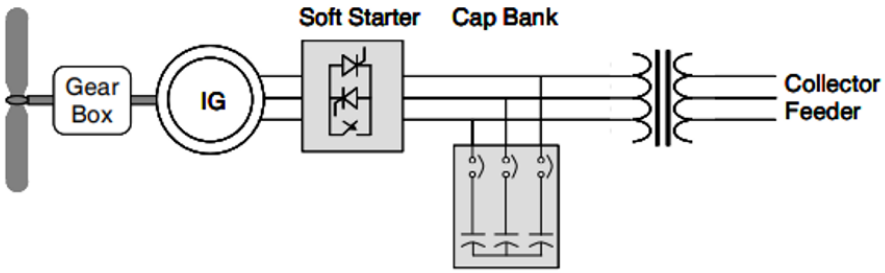

Figure 2.16. Typical configuration of a Type 1 wind generator

Induction Generator with Variable Rotor Resistance (Type 2)

This kind of wind generator is also known as wound rotor induction generator. In such generators, wound rotor induction generators are connected directly to the step-up transformer in a fashion like type 1, with regards to the machine's stator circuit, but also include a variable resistor in the rotor circuit, as illustrated in Figure 2.17. This can be achieved with a set of resistors and power electronics, external to the rotor, with currents flowing between the resistors and rotor via slip rings. On the other hand, the resistors and electronics can be mounted on the rotor, eliminating the slip rings—commonly known as the Weier design. The variable resistors are connected into the rotor circuit softly and can be used to control the rotor currents fairly swiftly to keep constant power, even during gusting conditions, and can affect the machine's dynamic response during disturbances [89].

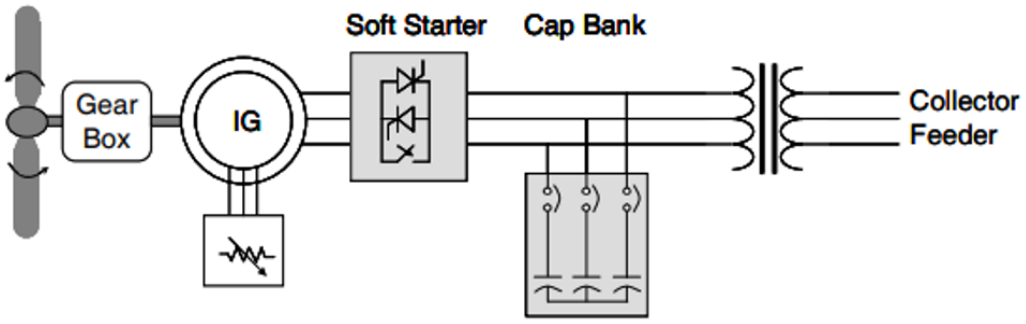

Figure 2.17. Typical configuration of a Type 2 wind generator

Generally, type 1, and type 2 generators are simple and economical, and require low maintenance, but they suffer from various drawbacks such as poor voltage control ability, large starting inrush, absence of speed control (very limited in type 2), high mechanical stress, and poor zero-voltage ride through capability [91].

Doubly Fed Induction Generator (Type 3)

The DFIG is currently the most commonly installed wind turbine in power systems [92]. They are gaining approval and acceptance these days for several reasons. The main reason being their ability to fluctuate their operating speed, typically +/- 30%, around the synchronous speed. It enhances the type 2 design to the next level, by adding variable frequency AC excitation (instead of resistance) to the rotor circuit. The additional rotor excitation is supplied via slip rings by a current regulated, voltage-source converter, which can adjust the rotor currents' magnitude and phase nearly instantaneously. The stator is directly connected to the grid and the rotor is fed from a back-to-back AC/DC/AC converter set, as shown in Figure 2.18. The rotor side converter (RSC) controls the wind turbine output power and the voltage measured at the grid side. The grid side converter (GSC) regulates the DC bus voltage and interchange reactive power with the grid, allowing the production or consumption of reactive power [92]. "A small amount power injected into the rotor circuit can affect a large control of power in the stator circuit. This is a major advantage of the DFIG — a great deal of control of the output is available with the presence of a set of converters that typically are only 30% of the rating of the machine. In addition to the real power that is delivered to the grid from the generator's stator circuit, power is delivered to the grid through the grid-connected inverter, when the generator is moving faster than synchronous speed. When the generator is moving slower than synchronous speed, real power flows from the grid, through both converters, and from rotor to stator. These two modes, made possible by the four-quadrant nature of the two converters, allows a much wider speed range, both above and below synchronous speed by up to 50%, although narrower ranges are more common. The greatest advantage of the DFIG, is that it offers the benefits of separate real and reactive power control, much like a traditional

SG, while being able to run asynchronously. The field of industrial drives has produced and matured the concepts of vector or field-oriented control of induction machines. Using these control schemes, the torque producing components of the rotor flux can be made to respond fast enough that the machine remains under relative control, even during significant grid disturbances [89].”

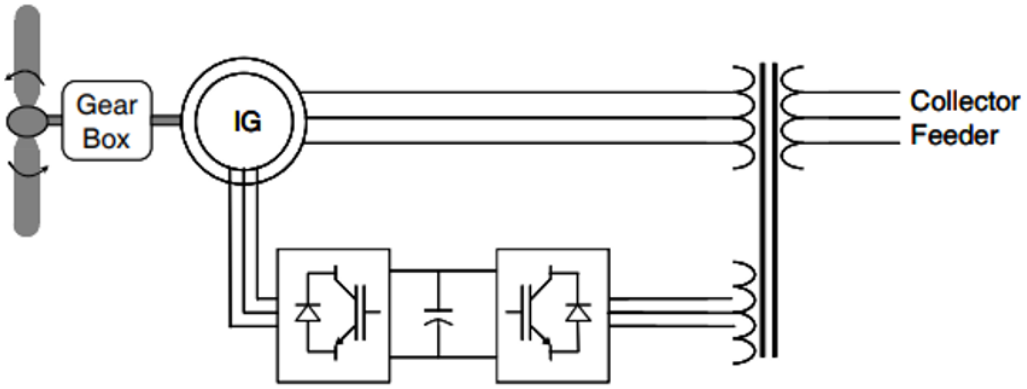

Figure 2.18. Typical configuration of a Type 3 wind generator

These wind generators have a good conversion efficiency, decoupled control of active and reactive power, and can provide voltage/frequency support. However, they require regular maintenance of slip rings and brush assembly, and there may be severe stresses on rotor and gearbox, especially during unbalanced faults [91].

A reduced 3rd order model, which neglects the stator transients, is used to represent DFIGs, in this research. The model has a structure like that proposed by WECC [93] and IEC [94]. Ignoring the stator current dynamics, the mathematical equations governing the DFIG model are as follows [95].

$$\frac{1}{\omega_b}\frac{de_d}{dt} = \frac{-1}{T_o}[e_d - (X - X')i_{qs}] + s\omega_s e_q - \omega_s \frac{L_m}{L_m + L_r} v_{qr} \tag{2.25}$$

$$\frac{1}{\omega_b}\frac{de_q}{dt} = \frac{-1}{T_o}[e_q - (X - X')i_{ds}] - s\omega_s e_d + \omega_s \frac{L_m}{L_m + L_r} v_{dr} \tag{2.26}$$

$$\frac{d\omega_r}{dt} = \frac{1}{2H_g}[K_{tw}\theta_{tw} + D_{tw}(\omega_t - \omega_r) - (e_d i_{ds} + e_q i_{qs})] \tag{2.27}$$

$$v_{ds} = -r_s i_{ds} + X' i_{qs} + e_d \tag{2.28}$$

$$v_{qs} = -r_s i_{qs} - X' i_{ds} + e_q \tag{2.29}$$

$$P_w = v_{ds} i_{ds} + v_{qs} i_{qs} - v_{dr} i_{dr} - v_{qr} i_{qr} \tag{2.30}$$

$$Q_w = v_{qs} i_{ds} - v_{ds} i_{qs} \tag{2.31}$$

where e_d, e_q are d and q components of internal voltage; P_w, Q_w are active and reactive power of DFIG absorbed by the network; X, X' are open-circuit and short-circuit reactance; T_o is the transient open-circuit time constant; H_g is generator inertia constant; ω_b, ω_s, ω_r are system base speed, synchronous speed and rotor speed, respectively; s is generator slip; θ_{tw} is the shaft twist angle (radians); K_{tw}, D_{tw} are the shaft stiffness and mechanical damping coefficients; v_{ds}, v_{qs} are stator voltages; v_{dr}, v_{qr} are rotor voltages; i_{ds}, i_{qs} are stator currents; i_{dr}, i_{qr} are rotor currents; r_s is stator resistance; L_m, L_r are magnetizing and rotor inductances.

Asynchronous or SG with Full Converter Interface (Type 4)

The type 4 wind generator (Figure 2.19) offers a great deal of flexibility in design and operation, as the output of the rotating machine is sent to the grid, through a full-scale back to-back frequency converter. The turbine can rotate at its optimal aerodynamic speed. In addition, the gearbox may be removed, such that the machine spins at the slow turbine speed and generates an electrical frequency, well below that of the grid. This is no issue for a type 4 turbine, as the inverters convert the power, and offer the possibility of reactive

power supply to the grid. The rotating machines of this type have been constructed as wound rotor synchronous machines, similar to conventional generators, found in hydroelectric plants, with control of the field current and high pole numbers, as permanent magnet synchronous machines, or as squirrel cage induction machines. However, based upon the ability of the machine side inverter to control real and reactive power flow, any type of machine could be used. Developments in power electronic devices and controls in the last decade have made the converters receptive and efficient. However, the power electronic converters must be sized to pass the full rating of the rotating machine, plus any capacity to be used for reactive compensation [89].

The type 4 wind generators generally offer maximum flexibility, due to the fully controllable converter interface. They offer decoupled control of active and reactive power and due to absence of slip rings, there is very limited maintenance required. The major drawback, however, is their high cost, as they require the full rated power converters for grid connection [91].

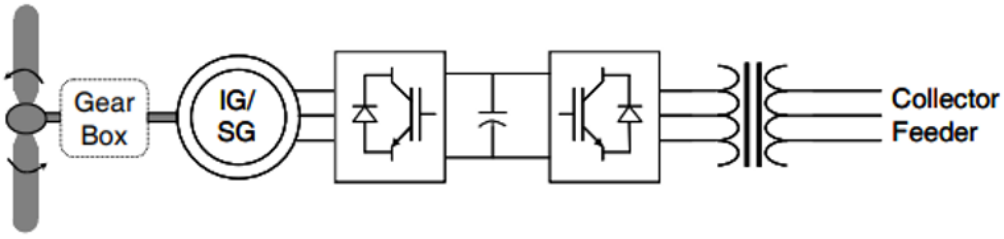

Figure 2.19. Typical configuration of a Type 4 wind generator

2.14 DFIG Control

DFIG can be operated in two different reactive power control modes: constant power factor control (PQ mode) and voltage control (PV mode). In the power factor control mode, the

reactive power from the turbine is controlled to match the active power production at a fixed ratio. When terminal voltage control is activated, the reactive power production is controlled to attain a target voltage at a specified bus [92]. In the constant power factor operation mode, the reactive power is not exchanged between the wind farm and the system. In the constant voltage operation mode, since the power factor of the wind turbine generator is adjustable, the voltage of the generator terminal can be quickly recovered through the regulation capacity of reactive power. Therefore, the PV mode produces a better transient stability. Various research [92, 96-100] have established the superiority of PV mode over PQ mode, for transient stability assessment. This is because, in voltage control mode, wind generators regulate the voltage level at their point of common coupling (PCC) to a set point. When there is a fault in the power system, the rotor of SGs accelerates. Once the fault is cleared, wind generators improve power system voltages, by injecting reactive power into the power system, to regulate the voltage level, at their PCC, to the set point. The increase in power system voltages enhances power system transfers, therefore allowing SGs to inject larger amounts of the kinetic energy stored in their rotor, resulting in the rotors of generators to decelerate quicker, which enhances transient stability [96, 101]. Thus, in this work, the DFIGs are equipped with terminal voltage control, i.e., the DFIG can exchange reactive power with the grid, to achieve a target voltage at the bus, at which DFIG is connected. This is illustrated in Figure 2.20. The point of common coupling is taken as the reference point. The voltage at the point of common coupling, denoted by V_{PCC}, is detected, and is compared with the reference voltage, V_{REF}, to compute the reactive power demanded by grid, Q_{out}.

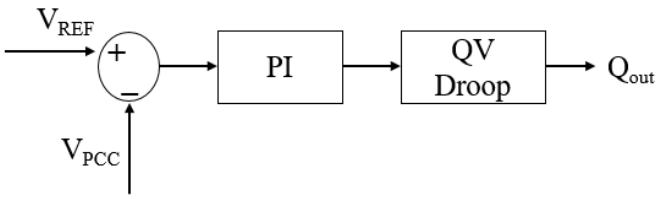

Figure 2.20. Terminal voltage control of DFIG

2.15 Literature Review: Probabilistic Transient Stability

This section will review some literature, pertinent to probabilistic transient stability (PTS). A significant amount of literature is available in the domain of PTS. [102] proposed a probabilistic framework, for power system transient stability assessment, with high renewable generation penetration. The presented framework enables comprehensive calculation of transient stability of power systems, with abridged inertia. A major drawback of the work is considering only three-phase line faults. [103] presented a study of the effects of some important power system parameters on transient stability. The parameters considered for this assessment include fault location, load increment, machine damping factor, fault clearing time (FCT) and generator synchronous speed. The work considers only three-phase line faults. In [104], the abridged version of an altered single machine infinite bus (SMIB) system, with a doubly fed induction generator (DFIG)-based wind farm integration, is analyzed, considering the transient features of the DFIG-based wind farm, in the diverse periods of a fault. The assessment specifies that the performance of synchronous generator (SG) can either be enhanced or depreciated with DFIG integration. Only a three-phase fault is considered on a specific line, which clears after a pre-selected time. [105] presented the results of a PTS assessment, conducted on the large-scale system of B.C. Hydro, including a generation rejection study, on the Peace system and a transfer

limit study on the Columbia system. In this paper, B.C. Hydro's historical statistics on the probabilistic states of load level factor, fault type, fault location, fault clearing, and automatic reclosing were used in a Monte Carlo (MC) formulation to produce sample states for the case studies.

[106] illustrated the incorporation of probabilistic analysis in transient stability of a practical power system, by applying it to a multimachine configuration. [107] provided an analytical algorithm, based on transient energy margin, for the online PTS assessment of existing or forecasted operating conditions. [108] presented two approaches, for computing the probability of transient instability. These methods are based on Bayesian theory and Cartesian products. [109] presented a stochastic-based method, to assess the PTS index of the power system, including the wind farm and the Superconducting Magnetic Energy Storage (SMES). Uncertain factors include both sequence of disturbance in power grid and stochastic generation of the wind farm. [110] proposed a stochastic-based approach, to determine the PTS indices of a power system, incorporating wind farms. In this scenario, researches were conducted on a hypothetical test system, considering the uncertainties of the factors, associated with power system operation, namely fault type, fault location, fault impedance, fault clearing process, system parameters, operating conditions, and high-speed reclosing. [111] illustrated the method of bisection to analytically evaluate the PTS indices. [112] provided a probabilistic approach to assess the transient stability of a wind farm. It extended and illustrated a rudimentary procedure, for calculating the probability of transient stability, for each transmission line, and for the overall system.

A PTS assessment method, based on MC simulation, was proposed in [113]. Two instability probability indices were also defined as indicators of the overall system stability

and the severity of individual component fault. In [114], a probabilistic assessment technique, based on quasi-MC method, to analyze the transient stability of power system, incorporating wind power, was presented. Two indices were introduced, to assess the transient stability of power system. [115] proposed an analytical assessment method, for transient stability assessment of multimachine power systems, under stochastic continuous disturbances. In the suggested technique, a probability measure of transient stability was presented and analytically solved by stochastic averaging. [116] proposed a method, for obtaining PTS assessment, by using distribution functions, based on location, fault type, and sequence. [117] proposed the use of MC simulations, in the computation of probabilistic measures, for the transient stability problem. [118] suggested approximate methods, for evaluating probabilistic transient instability, and classifying critical stability areas, for system planning. In [119-120], an approach was developed, for obtaining a stability index, for individual lines, and for the overall system, for numerous fault types. The impact of clearing times and reclosing times was also investigated for critical lines. In [121], an approach was presented to assess the distribution of the probability of instability. In [122], conditional probabilities were used in the evaluation of probabilistic transient instability. [123] proposed a real-time approach for computing probabilistic critical clearing time (CCT) which is applicable to PTS assessment. The goal of the proposed approach was to offer a low computational burden and high accuracy to calculate the probabilistic density function (PDF) of CCT in two stages.

[124] proposed a method for power system PTS assessment, considering the wind farm uncertainties and correlations. Specifically, the inverse Nataf transformation based three-point estimation method and the Cornish-Fisher expansion were combined, to deal with

the uncertainties, and the correlations amongst numerous wind farms. In [125], two-point estimate method was used, to determine the maximum relative rotor angles' probability distribution functions, for a given fault, with uncertain load demands and clearing time. A probabilistic approach, to assess the transient stability of power systems, with increased penetration of wind and photovoltaic (PV) generation, was presented in [126]. The impact on transient stability, due to the intermittent behavior of Distributed Energy Resources (DERs), and their dynamic response, when a disturbance happens, was examined. An analytical approach, for probabilistic dynamic security assessment (DSA) of power systems, incorporating wind farms, was presented in [127]. The probability of transient stability, given a specific fault and uncertainties of output power of wind farm and load was analytically computed.

2.16 Literature Review: Risk-based Transient Stability

The product of probability of an unforeseen event and its impact is commonly known as risk, which is generally mathematically defined as (2.32) [128-131].

$$Risk = \sum_{i} \mathrm{P_r}(E_i) \times Sev(E_i) \tag{2.32}$$

where E_i is the i^{th} event (contingency) and $\mathrm{P_r}(E_i)$ is its probability. $Sev(E_i)$ quantifies the impact of E_i.

The risk is the system's exposure to failure and is generally determined by considering both the probability of occurrence of an event and the impact of the event. The deterministic stability assessment introduces operating limits, based on the impact of contingencies. However, based on the risk-based approach, these operating limits are calculated by using the weighted sum of risk components of all the contingencies, considering both the

probability and the impact [129]. A simple example can be used to outline the significance of using risk in power systems. Consider two contingencies (*C1* and *C2*), along with their probability of occurrence and the corresponding severity (impact), as outlined in Table 2.2. If decision-making is assumed to be based on deterministic criteria, *C2* is found to be more severe as its impact is greater than *C1*; however, if risk-based (considers both probability and impact) decision-making is used, the converse is true.

Table 2.2. Risk values for two different contingencies

Contingency	**Probability**	**Impact**	**Risk**
C1	0.05	20	1
C2	0.02	30	0.6

Power system risk evaluation generally consists of the following four tasks [132].

1. Determining component outage models.
2. Selecting system states and calculating their probabilities.
3. Evaluating the consequences of selected system states.
4. Calculating the risk indices.

A power system consists of various components, including generators, transmission lines, transformers, circuit breakers, switches, etc. Generally, component outages are the main cause of a system failure state. The first task in system risk evaluation is to determine component outage models. Component failures are divided into two groups: independent and dependent outages. Each group can be further classified according to the outage modes. In most cases, only repairable forced outages are considered, whereas in some cases,

planned outages are also modeled. Aging failures are generally not included in the traditional risk evaluation.

The second task is to select system failure states and calculate their probabilities. There are two rudimentary methods for selecting a system state: state enumeration and MC simulation. In general, if complex operating conditions are not considered and/or the failure probabilities of components are quite small, the state enumeration techniques are more efficient. When complex operating conditions are involved, and/or the number of severe events is relatively large, MC methods are often preferable.

The third task is to perform the analysis for system failure states and evaluate their consequences. Depending on the network under study, the assessment could be related to simple power balance, optimal power flow (OPF), or even transient stability evaluation. As discussed before, risk is a combination of probability and consequence. With the information obtained in the second and third tasks, an index denoting system risk can be established. There are several conceivable risk indices for various purposes. Most of them are essentially the mean (average) value of a random variable. The expected indices serve as the risk indicators that reflect various factors, including component capacities and outages, load profiles and forecast uncertainties, and system topologies and operational conditions [132].

Risk-based approach describes possibility of contingency by probability, and the corresponding impact (or consequence) by severity function. The product of this probability and associated severity is termed as risk. In risk-based stability assessment, the risk index consists of each possible contingency occurrence probability and the associated impact [133]. The first attempt toward risk-based transient stability was proposed in [134]

and [135], where the notion of limiting operating point functions was used. These functions return the limiting generation level for any fault type and fault location. [136] used risk-based approach, to analyze the transient stability of power networks, incorporating wind farms. The proposed methodology of transient instability risk assessment is based on the MC method and eventually, an inclusive risk indicator, based on angle and voltage stability, is devised. The work considered only three phase line faults.

[137] presented a distributed computing approach for transient stability analysis, in terms of measuring critical clearing time and the overall risk index, for various uncertainties. The work considered only a three phase to ground fault. [138] presented a method to determine the risk of transient stability. It described the application of rotor trajectory index (RTI), to assess the severity of power systems, when it was subjected to a three-phase fault. The RTI was suggested as an index used to represent severity of transient instability. [139] focused on risk of transient instability. A procedure was suggested to evaluate the potential loss of synchronism of a generator, in terms of probability and consequences. A transient risk assessment method, based on trajectory sensitivity, was presented in [140]. The contingency cost was considered as minimum control cost to move the system from instability to stability, and the sensitivities of relative rotor angle, with respect to output of generators and dispatchable load, were utilized to consider the transient stability constraints in the OPF. Some other research work associated with risk-based transient stability can be found in [141-143]. **The major shortcoming in these works include not considering all the fault events (faulted line, fault type, fault location, fault clearing time) randomly, i.e., only some of the events are considered random variables, while others are considered as deterministic.**

In recent years, there has been an increased deployment of renewable energy in the U.S. electricity grid. In 2020, around 834 billion kW-hours of electricity (about 21% of total electricity production) produced in the United States were from renewable energy sources [144]. The chief motivating factors for the increased renewable generation are the reduced cost of electricity production and the state-level renewable energy portfolio standards. Most of the renewable energy penetration is in the form of type 3 wind turbine generators. The increased penetration of converter-based generation can have a substantial consequence on the transient stability of a power system. This kind of generation is integrated to the grid with no rotating mass and inertia, as they are interfaced using power converters. Power system inertia is the term given to the store of kinetic energy, found in by the rotating mass of traditional steam and diesel generators, which turn in synchronicity with each other, and are coupled to the power system, delivering a steady system frequency [145]. This reduced inertia, due to wind generation integration, implies that it is more difficult to hamper a rate of change in system variables, such as system frequency, bus voltage and generator rotor angle. Most of the research efforts [146-150] in this area are centered on transient stability analysis. These studies show that increased renewable penetration can have both valuable and harmful effects on system stability. Due to the altered dynamics of the system due to increased converter-interfaced generation, it is indispensable that the power system reliability standards be revisited [151]. Some significant work associated with PTS, incorporating wind generation, can be found in [110, 112, 114, 124, 126, 127, 136, 152-157].

From the literature review, it is established that most works on PTS assessment assume three-phase faults. Although, the assumption may be suitable for a

deterministic analysis since it is the most severe fault; however, other faults cannot be ignored, as their probabilities are higher, and must be included in probabilistic stability assessment. [158-165] indicated that risk-based instability approach is an open area of research and requires further work. Moreover, the impact of reduced inertia systems (i.e., higher wind penetration), on power system transient stability, is of great implication [166-168].

2.17 Literature Review: Transient Stability Improvement Methods

Several variables, such as fault type, fault location, FCT, system impedance, system inertia, system loading, network topology, and system voltage, can impact the transient stability [169]. Thus, an extensive range of techniques for improvement of transient stability can be found in the literature. According to [169], the techniques of transient stability improvement are aimed to realize one or more of the following effects:

(a) Reducing the impact of disturbance or severity of the fault.

(b) Increase of the synchronization forces to support the restoration of steady-state operation after a disturbance.

(c) Reduction of the acceleration or deceleration power through control of the prime mover to meet the equilibrium of mechanical and electrical power.

(d) Applying artificial load to synchronous generation to reduce accelerating power by increasing electrical power.

Effects related to (a) can be attained by faster fault clearing times, through high-speed breakers, thereby, reducing the fault severity, by reducing the clearing angle. An increase

of the synchronization forces mentioned in (b) can be achieved using Flexible Alternating Current Transmission Systems (FACTS), e.g., voltage support at a long transmission line; techniques related to (c) and (d) deals with the re-establishment of the equilibrium, between the mechanical and electrical power, thus, reducing the acceleration/deceleration power on the shaft of the synchronous machine.

An extensive assortment of methods for enhancement of transient stability can be found in the literature [169-170]. Among these methods, the braking resistor [171], (FACTS) devices [172], Superconducting Fault Current Limiter (SFCL) [173], Static VAR Compensator (SVC) [174], SMES [175], and high-speed circuit breakers (CBs) [175] are quite common. Braking resistor consists of a dummy load connected in parallel with the SGs. An artificial electrical load is applied during transient disturbance to increase the electrical power and to re-establish the equilibrium [169]. However, for application of braking resistor, a step-down transformer must be connected to the generator terminal, which adds cost to the system. Although, FACTS devices can regulate both active and reactive powers, their application will incur more cost in the power system because of their complex structure and control system. The SFCL requires cooling system to maintain its superconductivity. Therefore, its application is also expensive. Likewise, the SMES is an active and reactive power controlling device, but it is expensive. Some other methods for transient stability improvement include SG redispatch [177], load shedding [178], fast excitation system [179], fast valving [180], and virtual inertia [181]. A comparative analysis of these methods is beyond the scope of this research; however, a brief description for each one is provided [169].

Synchronous Generator (SG) Redispatch

One of the most effective actions, to upsurge the transient stability margin, is to redispatch generators, to reduce their active power set point, which implies generators are operated further away from the stability limit. Due to cost efficiency, the dispatch of generators is usually determined using OPF calculations, and transient stability constraints are derived using a time-domain sensitivity analysis [177, 182-183]. In [177], a fourth-order Taylor expansion is used to speed up the solutions of OPF computations, including transient stability constraints. [182] proposed to derive linearized transient stability constraints, outside the OPF calculations, to reduce the complexity of the OPF formulation. In [183], time-domain simulations were combined with pre-assessment contingency filtering and a fast re-dispatch estimation, to decrease the computational burden of the stability assessment.

Load Shedding

Commonly, load shedding is associated with frequency regulation, such as under frequency load shedding, to prevent a power system from collapse, due to generation shortfall. However, load shedding can also be used to enhance transient stability of power systems, e.g., decrease the loading of generators, by reducing the load. [178] proposed a coordination of generation rescheduling and load shedding to enhance transient stability. If the generation rescheduling is unable to rectify the issue, load shedding is done to ensure that the system is within the defined security boundaries. In [184], the power system stability problem was dealt, by reducing transients, using a load shedding scheme. A relay coordination scheme was suggested, based on under-voltage/overcurrent, to shed the less priority load, in an interconnected transmission network.

Fast Excitation System

A noteworthy improvement in transient stability can be attained with fast responding excitation systems of SGs, by increasing the field current to increase the internal machine voltage. This increases the electrical power, during the fault, which decreases the acceleration area and hence, leads to an increased (CCT). High-speed excitation systems are not very effective for bolted three-phase faults at the generator terminals since the voltage drops to zero. However, they are very effective for faults occurring further away from the generator, where the voltage at the generator bus is greater than zero [185].

Another method, in this regard, is a discontinuous excitation control, which is referred to as transient stability excitation control, where the terminal voltage is kept near the maximum permissible value, over the entire positive swing of the rotor angle and returns to normal operation mode, after the first swing. The use of fast excitation systems may compromise the damping of local plant oscillations, which requires the excitation system to be supplemented with a power system stabilizer [186]. [187-188] used nonlinear excitation controllers. The controllers are designed robustly and can operate over a wide range of operating conditions. The simulation results showed boosted performance, under transient conditions, through increased damping of oscillations, after the disturbance. [189] presented the design of the voltage source converter (VSC) excitation system and an integrated prototype using digital signal processor (DSP). It was demonstrated that the VSC excitation system drastically improved the transient stability limit.

Fast Valving

Fast valving of the turbine is an effective practice, to improve transient stability, by rapidly plummeting the mechanical power during the fault [190]. According to [191], the notion of fast valving was first introduced in 1925. During faults, the electrical power drops to a

reduced value and creates a difference between mechanical and electrical power, causing an acceleration of the machine. To counter that imbalance, generators with fast valving capability, rapidly lower the mechanical power, which is applied to the machine, to decrease the acceleration power to a minimum. One of the restrictions of fast valving is that it can only be applied to thermal generating units [192]. A fast-valving scheme of steam turbines, using two parallel valves, for enhancing the transient stability, was presented in [193].

Virtual Inertia

The provision of virtual inertia has recently gained more significance, as the penetration of converter-based renewable energy systems is increasing. Virtual inertia is typically referred to frequency stability, but it has also a considerable impact on transient stability. Although, the installation of virtual inertia devices does not affect the inertia of SGs, it may re-route some active power flows in the network, thereby, changing the CCT in certain situations. At times of high renewable generation penetration, the available rotational inertia is reduced because conventional synchronous machines are shut down. Virtual inertia can be provided by units which have stored additional energy, either in rotational or chemical form. Wind turbines of type 3 and 4, and battery storages equipped with an inertia control algorithm, are appropriate for virtual inertia provision. [194-196] proposed an application of virtual inertia control of DFIG wind turbines, using a derivative controller, which uses the frequency as an input signal, to adjust the active power set point of the machine, based on the rate of change of frequency (ROCOF).

Braking Resistor

The idea to use a braking resistor for transient stability improvement can be seen from a similar viewpoint as fast valving, with the difference that it acts on the electrical power, instead of the mechanical power. An artificial electrical load is applied during transient disturbance to upsurge the electrical power and to re-establish the equilibrium, or at least minimize the difference, between mechanical and electrical power. In [197-198], the authors proposed two variants of braking resistor: one with a thyristor rectifier and one with a diode rectifier and chopper. A fuzzy logic-controlled braking resistor was presented in [199]. However, for the practical application of braking resistors, cautious deliberations must be made regarding installation costs, torsional stress on the shaft and other additional adverse impacts which may occur.

Variable Series Compensation

Thyristor-controlled series capacitor (TCSC) and static synchronous series compensator (SSSC) are capable to act on the power system in a serial manner, contrary to shunt devices. Variable series devices can be used in preventive, as well as, in emergency control. Instead of line reinforcement or installation of additional lines, TCSCs offer a robust substitute to improve transient stability, by optimizing the transmission impedance [200]. The TCSC decreases the effective series reactance and, thus, decreases the angular separation with the power transfer being constant. In [201], a fuzzy logic based TCSC damping controller was used, in a SMIB system, to augment the power system transient stability. [202] presented an approach for enhancing the transient stability using SSSC.

Variable Shunt Compensation

FACTS, with reactive power capability for voltage control, at selected points of the power system can contribute to enhance the transient stability, by increasing the synchronization

power flow among the generators [190]. SVC and static synchronous compensator (STATCOM) are capable to control the voltage/reactive power at their connection point. SVC comprises a mechanical reactor and capacitor, a thyristor-controlled reactor and capacitor, and a harmonic filter. A standard proportional integral (PI) controller for SVCs, which used the voltage measurement as input to control the reactive power, was introduced in [203]. [204] presented the application of a fuzzy logic controlled Static Compensator STATCOM to enhance the transient stability of a power system.

Fault Current Limiter (FCL)

A fault current limiter (FCL), also known as fault current controller (FCC), is a device which restricts the forthcoming fault current, when a fault occurs, without complete disconnection. Resistive, inductive, or combined designs of FCLs are generally used to improve the transient stability during faults. The resistive type is effective in consuming the acceleration energy of generators during faults, whereas, the inductive type suppresses the voltage drop. The FCLs used in power systems can be grouped into two main groups, namely superconducting FCLs [205-207] with highly nonlinear response to temperature, current and magnetic field variations, and bridge-type FCLs, based on solid-state devices [208-210], which are either insulated-gate bipolar transistor (IGBT)- or thyristor-controlled.

Magnetic Energy Storage

A SMES system is a very efficient storage device capable of storing large amounts of energy. SMES is a device, consisting of a superconducting coil, in which AC power is converted into DC by an AC-DC converter and stored in superconducting coil, in the form of magnetic energy. Its storage efficiency is nearly 90%. As SMES can be charged or

discharged very quickly, through the semiconductor AC-DC converter system, it can be applied for both active and reactive power compensation, for power system stabilization [211-212]. The chief disadvantage of the SMES technology is the need of a large amount of power, to keep the coil at low temperatures, combined with the high overall cost, for the employment of such a unit. Moreover, above a specific field strength, known as the critical field, the superconducting state is destroyed [213-214]. [215] discussed artificial neural network (ANN) controlled SMES unit, for enhancement of transient stability of a power system, under various system operating conditions and different fault conditions.

High Speed Circuit Breakers

Synchronous generators accelerate and pick up kinetic energy during severe faults. The kinetic energy, which is picked up by the generator, is directly proportional to the fault duration. Therefore, it is desired to clear faults as fast as possible, i.e., the shorter the fault duration, the smaller the severity of the disturbance [190]. The tripping times of modern high-speed CBs are around two to three cycles for high-voltage and one cycle for low- and medium-voltage CBs [216-217]. [176] proposed the coordinated operation of optimal reclosing of CBs and SVC, for enhancing the transient stability, of a multi-machine power system. [170] proposed the coordinated operation of the bridge type fault current limiter (BFCL) and optimal reclosing of CBs to improve the transient stability of a multi-machine power system. [218] investigated the enhancement of transient stability of a two-area system, using solid state CB, which is capable of fast switching, in case of faults and thus, controlling the real and reactive power flows, in a faulted transmission line. [219] presented the optimal reclosing of CBs, for the distributed generation (DG) connected IEEE 9-bus system. The optimal reclosing technique is derived using the total load angles of the DGs.

Both transient and permanent faults, at different points in the power system model, were considered.

Although, [170, 218-219] dealt with transient stability enhancement using CBs, but these works consider the deterministic approach, implying the transient stability analysis is performed deterministically, i.e., for specific fault types, fault locations, and FCTs. Although, the concept of risk was applied for transient stability improvement, using generation redispatch/load shedding and TCSC, in [178] and [220], respectively; it has not been attempted for a decision-making framework involving CBs (according to the best of author's knowledge).

As mentioned before, conventional approaches for transient stability assessment are inadequate for transient stability assessment, specifically online, due to various drawbacks discussed, such as huge computational burden and modeling limitations. Moreover, the evaluation of transient instability risk is computationally rigorous due to repetitive time domain simulations. ML approaches, with their useful features of pattern recognition, learning abilities, and rapid prediction of system security states, provide a good substitute. Thus, the ML algorithms are applied, in this research, to tackle this issue, as these algorithms have widely been suggested for reducing the computation efforts [221-227]. Moreover, ML algorithms have a major advantage of incorporating significant inputs only, whereas time-domain method generally requires complete description of the system model [221-222].

2.18 Replacement of Circuit Breakers

Utilities generally use deterministic approaches for CB replacement as they are using worst-case conditions (peak load, three phase faults, etc.) to comply with NERC TPL standards. In other words, utilities use deterministic approaches because they must follow Standards (ANSI, IEEE, etc.) in which probabilistic approaches are unacceptable. Probabilistic approaches require information about the probability distribution of the input variables, and thereby, they violate the industry standards (NERC TPL-001-4) of using specific deterministic values/parameters. Majority of utilities replace CBs based on aging, repair, tests, maintenance, manufacturer no longer supporting CB model, and upgradation (replacing oil filled breakers with SF6 breakers). The interrupting capability for symmetrical current rated CBs is generally assessed using the latest version of IEEE/ANSI standard C37.010. Existing literature uses asset management strategies (mainly based on maintenance and aging) to replace CBs. A brief overview follows.

In [228], optimum maintenance strategies of CBs were predicted depending on their ages. It presented an optimization approach of maintenance strategy for CB in a transmission and distribution system. In this approach, a present and future CB performance was estimated by using monitoring data, and impacts of performance in each CB on the system were evaluated. By minimizing the impacts, suitable maintenance procedures and timing were derived. In order to maintain availability of power system, [229] established a refreshment regulation of high voltage CBs. The refreshment criteria were based on technical, age, and risk consideration. It discussed parameters used for technical consideration, as well as failure occurrence and criticality of substation towards the system for risk consideration. Consequently, the parameters of technical and risk consideration

were applied to obtain risk criteria which were used as priority scale for refreshment program.

[230] dealt with the model of ageing behavior of different CB types. Furthermore, the model simplified the condition control of pieces of equipment, and provided the asset manager with information about the yearly capital and operational expenditures. Capital expenditures included new installation costs in consequence of replacement and grid enlargement, and operational expenditures comprised of maintenance costs, which can be subdivided in inspection and overhaul costs as well as in minor and major failure costs. [231] proposed risk assessment for CB utilization by using failure modes, effects, and criticality analysis (FMECA). The CB components were divided into three main parts such as live part, operating mechanism and insulation, and control device. FMECA was categorized into four groups such as finance, safety, environment and efficiency. Finally, criticality was categorized into 4 levels such as low level, medium level, high level and highest risk level. Data collection and analysis were performed. The results of the criticality analysis can assist in making maintenance resource management decisions in order to avoid damage to the system with efficient usage of available or limited resources.

[232] suggested a reliable life calculation method for SF6 CBs. The reliability evaluation model was established based on Fault Tree Analysis (FTA) method. From FTA, all the bottom events which cause top event (failure of the SF6 CB) were determined. To estimate the reliability by inspection results and test results of the CBs, Health Index (HI) was presented to compute the failure probability of each bottom event. Consequently, the reliability of the top event was obtained. Collecting the failure probability and the age data of different SF6 CBs with the same type in power system, the data was applied to estimate

the reliability function of two parameters Weibull distribution by the least square method. Finally, the reliable life of the SF6 CB was estimated after inversion calculation by reliability function.

[233] presented methods for evaluation on grease degradation of aged generator CBs and for establishing equipment replacement criteria with asset management applied to power transmission equipment. Replacement criteria were organized with asset management based on risk evaluation. [234] proposed a system developed at American Electric Power (AEP) for monitoring CBs. It presented an overview of the results of field trials that showed the maintenance reductions gained by predicting required maintenance instead of scheduling it. The system discussed provided the ability to monitor real-time trip and close coil assembly performance by recording mechanical and electrical characteristics including trip coil current and operate time during CB operations. Recorded trip and close coil characteristics were used to diagnose armature misalignments, lubrication problems, and interwinding short circuits to assist with maintenance and ensure future operations.

In [235], an algorithm was proposed using branch and bound method to search an optimal replacement scheduling of obsolete equipment in aged primary substations. The developed tool can efficiently find an optimal solution from huge combinations of replacement schedules. The objective function to be minimized was the net present value of the sum of operation and maintenance cost, replacement cost, and the reliability cost in case of CB failure for each substation during the specified time frame. To solve this problem under the constraints such as annual budget ceiling, number of replacement targets per annum, and replacement time frame for each substation, the branch and bound method was applied. A probabilistic maintenance model was implemented in [236]. The model parameters used

were mean time in each stage, inspection rate of each stage, and probabilities of transition from one stage to others. Sensitivity analysis of model parameters was conducted to establish cost-effective maintenance process. The analysis covered mean time to first failure, probability of failure, maintenance cost, inspection cost, and failure cost.

[237] proposed the study of the variable failure rate of high-voltage CB by using the Weibull distribution method to set time-based maintenance schedule properly. Power CB was installed in the substation in varied configurations. The case study conducted was compared with, main and transfer bus arrangement, and breaker-and-a-half bus arrangement. After getting mean time between failure, the parts can be categorized, and consequently, the time-based maintenance schedule can be set. The results from the table of maintenance schedule by categorizing each part of power CB with the new replacement subcomponent group showed that the failure rate of power CB can be diminished and it helped to plan maintenance schedule. A probabilistic maintenance model for CBs was suggested in [238]. Information collected during inspection tests was analyzed and the condition of the breaker was defined. Maintenance action was taken according to the condition of the breaker. Monte Carlo simulation was used to implement the model. Maintenance cost and time to failure of each transformer and CB was also incorporated in the analysis. Some other work dealing with replacement of CB based on aging/maintenance can be found in [239-244].

From the literature review, it is deduced that existing practices of replacement of CBs are mainly based on asset management and short circuit studies (to determine the interrupting capacity of the CB). The maintenance is considered if the point of concern is aging of the breaker and the decision must be made between replacing the breaker based on

maintenance. However, the decision-making in this proposed research is based solely on FCTs. It implicitly assumes the same "maintenance status" for all CBs. Moreover, aging/maintenance issue in CB is a "static" problem (as opposed to the "transient" problem dealt in this research). As mentioned in Chapter 1, there are sufficient evidences to demonstrate that risk-based transient stability approaches are a potential future work and hence, this research applies it to a specific problem (CB replacement), which has not been attempted before.

2.19 Marginal Transient Stability

[245-249] uses the method of transient energy function (TEF) to evaluate the marginal transient stability. In [245], an on-line dynamic contingency screening and ranking approach integrated with fast potential energy boundary surface (PEBS) method and corrected hybrid method to enhance efficiency of system transient stability assessment for a large complex power system was presented. The PEBS approach re-ranked the must-run contingencies according to their critical clearing time evaluated at this level. Ignoring highly stable contingencies, the marginal stable and unstable contingencies were further examined in the corrected hybrid method to detect their stability behavior and operation limits.

[246] proposed an iterative algorithm for determining parameter values that resulted in marginal stability of a system. The algorithm was based on Gauss-Newton solution of a nonlinear least-squares problem. Gradient information was provided by trajectory sensitivities. In [247], a new approach called marginally unstable injection (MUI) for developing a more accurate transient stability index was proposed using the concept of

TEF. An energy margin for transient stability assessment based on a reduced energy function was also formulated.

In [248], stability limits in terms of plant generation limits, load changes, or network configuration changes were derived using analytical sensitivity approach of the energy margin. A detailed development of the analytical sensitivity procedure was also suggested. [249] presented the development and evaluation of an analytical method for the direct determination of transient stability. The method developed was based on the analysis of transient energy and considered the nature of the system disturbance and the impact of transfer conductances on the system behavior. The approach also predicted critical clearing times for first swing transient stability.

The above-mentioned papers use the concept of unstable equilibrium points (UEPs) located on the boundary of potential energy surface. The system energy at this boundary is known as the critical energy. The TEF method is used to control these UEPs to acquire the required stability level. If the system energy exceeds the critical energy, system goes into unstable region. In other words, the energy required by the post-fault unstable system to reach the first instance of stability is known as the marginal stability. The theoretical background of TEFs can be found in [250-251]. Although, the TEF methods are not computationally extensive, but they do not yield accurate results when compared to time domain approach. Also, it is complex to integrate renewable generation with their associated controllers [252-253]. Their main drawback is their high intricacy in the following situations: (1) considering differential-algebraic equation models of power systems, (2) dealing with the detailed models of the system's components, (3) when a large number of system's parameters must be considered for the sensitivity analysis [252-253]. The time-domain

approach yields accurate results and is a widely recognized approach to describe power system transient behavior [254]. Thus, the concept of marginal transient stability using time domain approach has been explored in this research.

2.20 Machine Learning: Overview and Background

ML is widely regarded as the subset of artificial intelligence (simulation of human intelligence in machines, which are programmed to think like humans and mimic their actions), as outlined by Figure 2.21. The term ML was invented, in 1959, by Arthur Samuel, an American pioneer, in the domain of artificial intelligence [255]. ML basically is an application of artificial intelligence that provides systems the ability to automatically learn and enhance from experience without being explicitly programmed [226, 255-256]. In fact, the ML performs data analysis, using a set of instructions, through a variety of algorithms, for decision making and/or predictions [257]. Laborious designing and programming of algorithms are essential to be conducted, for ML, to implement diverse functionalities, such as, classification, clustering, and regression. Deep learning (DL) is a class of ML algorithms that uses multiple layers to progressively extract higher-level features from the raw input. For example, in image processing, lower layers may identify edges, whereas higher layers may identify the concepts relevant to a human being, such as digits, letters or faces [258]. It is majorly used for speech recognition, computer vision (high-level understanding from digital images or videos), medical image analysis, and natural language processing. There are several architectures used in DL such as deep neural networks, deep belief networks, recurrent neural networks, long short-term memory, and

convolutional neural networks. The DL generally requires huge processing power and massive data [258]. The focus of this work is, however, on ML.

ML differs from traditional programming, in a very distinct manner. In traditional programming, the input data and a well written and tested program is fed into a machine to produce output. When it comes to ML, input data along with the output is fed into the machine during the learning phase, and it works out a program for itself. This is illustrated in Figure 2.22 [259].

During the last decade, ML, and DL has demonstrated promising contributions to many research and engineering areas, such as data mining [260], medical imaging [261], communication [262], multimedia [263], geoscience [264], remote sensing classification [265], real-time object tracking [266], computer vision-based fault detection [267], and so forth. The integration of advanced information and communication technologies, specifically Internet of Things (IoT), in the power grid infrastructures, is one of the main steps towards the smart grid. Since the vital capability of IoT devices is their capability to communicate data to other devices in a more pervasive fashion, and hence a massive amount of data is made available at the control centers. Such meaningfully enhanced system condition awareness and data availability demands for ML-based solutions and tools to conduct efficient data processing and analysis, to encourage the system operational management and decision-making [225]. Therefore, ML has been applied in various fields of power system, such as load forecasting [268], fault diagnosis [269], substation monitoring [270], reactive power control [271], unit commitment [272], maintenance scheduling [273], wind power prediction [274], energy management [275], load restoration

[276], solar power prediction [277], state estimation [278], transient stability assessment [279], economic dispatch [280], and electricity price forecasting [281].

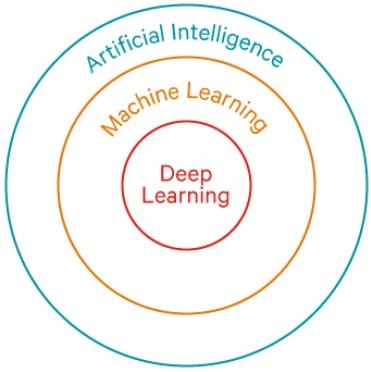

Figure 2.21. ML as a subfield of artificial intelligence

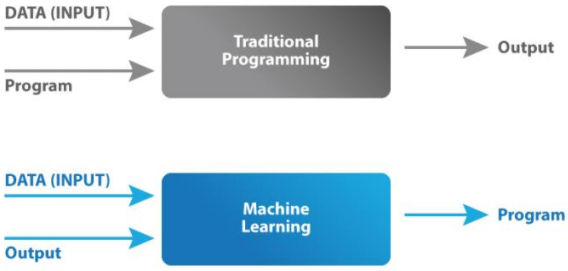

Figure 2.22. Traditional programming vs. ML

2.20.1 Steps of Machine Learning

There are seven main steps of successfully implementing ML. They are outlined below and illustrated in Figure 2.23 [282].

Gathering Data

The first and the most significant step of ML is gathering data. This step is very critical, as the quality and quantity of data gathered will directly determine how good the predictive

model will turn out to be. The data collected is then tabulated, and is commonly called as the training or learning data.

Data Preparation

After the training data is gathered, the next step of ML is data preparation, where the data is loaded into a suitable place and then, prepared for use in ML training. Here, the data is first put all together and consequently, the order is randomized as the order of data should not affect what is learned. This is also a good chance to do any visualizations of the data, as this will help see if there are any pertinent relationships between the different variables, and presence of any data imbalances or anomalies. Also, at this stage, the data must be divided into two parts. The first part, that is used in training the model, will be most of the dataset and the second will be used for the evaluation (validation and testing) of the performance of the trained model.

Model Selection

The subsequent step that follows in the workflow is choosing a model among the many that researchers and data scientists have created over the years. There are different algorithms for different tasks. Some are appropriate for image data, others for sequences (such as text, or music), some for numerical data, others for text-based data. A selection should be made based on the task required.

Training

After the above-mentioned steps are completed, the next step involves training, where the data is used to incrementally improve the ability of the model to predict. The training process involves initializing some random values for the model, predicting the output with those values, then comparing it with the model's prediction and eventually, adjusting the

values such that they match the predictions that were made formerly. This process then repeats, and each cycle of updating is called one training step.

Evaluation

Once training is complete, evaluation is performed. This is where the testing dataset comes into play. Evaluation allows the testing of the model against data that has never been seen and used for training and is meant to be representative of how the model might perform in the real world.

Hyperparameter Tuning

Once the evaluation is over, any further improvement in the training process is possible by tuning the parameters. There were a few parameters that were implicitly assumed when the training was done. Another parameter included is the learning rate that defines how far the line is shifted during each step, based on the information from the previous training step. These values are significant in the accuracy of the training model, and how long the training will take. For complicated models, initial conditions play a significant role in the determination of the outcome of training. Differences can be seen depending on whether a model starts off training with values initialized to zeroes versus some distribution of values. These parameters are commonly referred to as hyperparameters. The tuning of these parameters depends on the dataset, model, and the training process.

Prediction

ML is fundamentally using data to answer questions. Prediction is the final step where you get to answer few questions. This is the point where the value of ML is realized. The model gains independence from human interference and thus, draws its own conclusion, based on

its data sets and training process. Here, eventually, the trained model can be used to predict the outcome for any desired inputs.

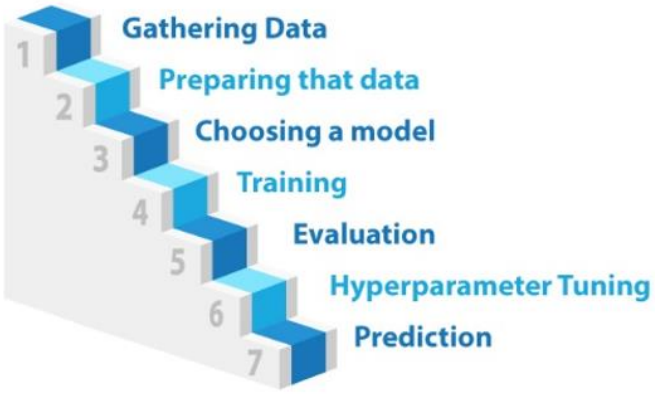

Figure 2.23. Seven steps of ML

2.21 Classification of Machine Learning

ML is generally classified into three broad types [225], as shown in Figure 2.24. A brief description of each type is given below.

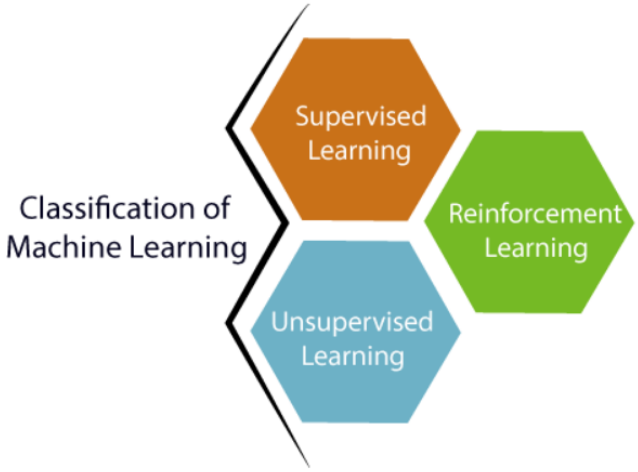

Figure 2.24. Types of ML

Supervised Learning (SL)

In supervised machine learning (SML), the aim is to learn a mapping between the inputs to outputs based on a given labeled set of input/output pairs in the training set. In this kind of learning, each example is a pair consisting of an input object (typically a vector) and a desired output value. A supervised learning (SL) algorithm examines the training data and generates an inferred function, which can be used for mapping new examples. Some common SL algorithms include artificial neural network (ANN), support vector machine (SVM), decision trees, Naïve Bayes, and k-nearest neighbor (kNN) [283]. The generic framework for SL is illustrated in Figure 2.25.

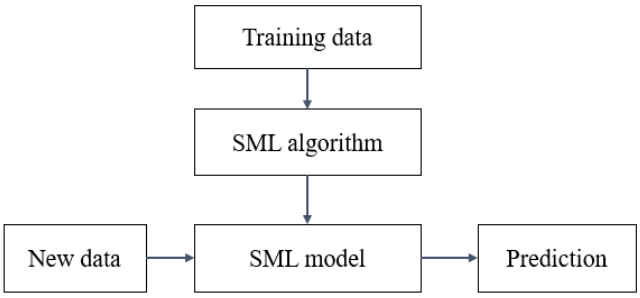

Figure 2.25. SL generic framework

Unsupervised Learning (UL)

In unsupervised learning (UL), the training of an algorithm is conducted, using information that is neither labeled nor classified, such that the algorithm may cluster the information based on similarity or difference. In contrast to the SL that makes use of labeled data, UL, also known as self-organization, allows for modeling of probability densities over inputs.

The goal of UL is to discover hidden patterns in unlabeled data. Some of the most common algorithms used in UL include clustering and anomaly detection [283]. The generic framework for UL is illustrated in Figure 2.26.

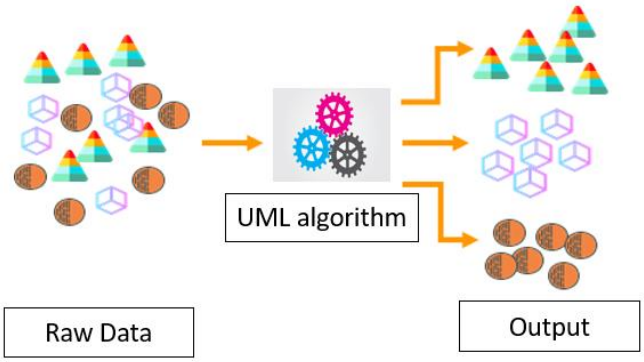

Figure 2.26. UL generic framework

Reinforcement Learning (RL)

Reinforcement learning (RL) is an iterative process to predict the next optimal step to perform a task to get a final reward. In each stage, the deep learning agent receives an award when it moves in the direction of the goal. RL is suitable for training a computer to drive a vehicle or playing a game against an opponent [224]. Basically, in RL, an agent interacts with its environment and adapts its actions, based on the reward received in response to its actions [293]. RL differs from SL in the sense that it does not need labelled input/output pairs be presented, and does not need sub-optimal actions to be explicitly corrected. Instead, the emphasis is on finding a balance between exploration (of uncharted territory) and exploitation (of current knowledge) [294]. A few significant terms associated with RL are defined below [293].

a) Agent: the program, trained with the goal of doing the job specified.

b) Environment: the world, real or virtual, in which the agent performs actions.

c) State: the observation, the agent does on the environment, after performing an action.

d) Action: A move, that the agent performs on the environment, based on its observation.

e) Reward: The feedback the agent receives, based on the action it performed. If the feedback is positive, it receives a reward and if the feedback is negative, it receives a penalty.

In RL process, the environment gives the agent a state. The agent chooses an action and receives a reward from the environment along with the new state. This learning process continues, until the goal is achieved. The generic framework for RL is illustrated in Figure 2.27.

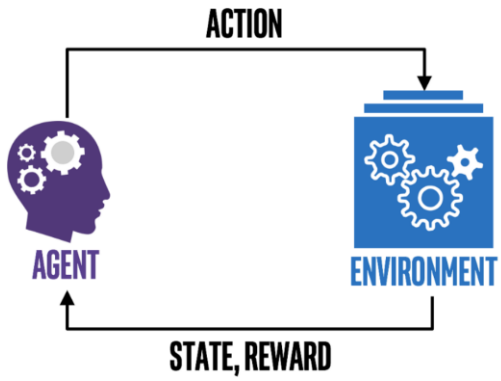

Figure 2.27. RL generic framework

The focus of the present work is on SML algorithms. As previously mentioned, commonly used SMLs include ANN, SVM, decision trees, random forest, and Naïve Bayes. However, amongst these algorithms, the present work will focus on ANN and SVM.

2.22 Machine Learning Regression and Classification

The two most common applications of ML include regression and classification [285]. These two are often confused and used interchangeably, however, a distinction must be made between the two. Both applications deal with predicting a quantity. Regression is the task of predicting a continuous quantity. The core goal of regression problems is to estimate a mapping function, based on the input and output variables. Some common types of regression are linear, logistic, polynomial. Linear regression establishes a relationship between dependent variable and one or more independent variables, using a best fit straight line (also known as regression line). Logistic regression is used when the dependent variable is dichotomous (binary). Logistic regression estimates the parameters of a logistic model and is a form of binomial regression. Logistic regression focuses on the data that has two possible criterions and the relationship between the criterions and the predictors. A regression equation is a polynomial regression equation if the power of independent variable is more than one. In this regression technique, the best fit line is not a straight line. It is rather a curve that fits into the data points.

Classification is the task of predicting discrete class labels. There are two kinds of classification: binary and multi. Binary classification refers to those classification tasks that have two class labels, and multi classification refers to the tasks with more than two class labels. Some algorithms can be used for both classification and regression, with small modifications, such as decision trees and ANNs. Some algorithms cannot, or cannot easily be used for both problem types, such as linear regression is used only for regression predictive modeling, and logistic regression for classification predictive modeling. Most prominently, the manner classification and regression predictions are evaluated differs significantly and does not overlap; for example, classification predictions can be evaluated

using accuracy metric, whereas regression predictions cannot. In a similar fashion, regression predictions can be evaluated using root mean squared error, whereas classification predictions cannot. The difference between classification and regression is graphically illustrated, using Figures 2.28 and 2.29.

Generally, for ML classification and regression tasks, the entire data set is divided into three parts: training, validation, and testing. A training dataset is used during the learning process and is used to fit the parameters (e.g., weights, biases) of the model. A validation dataset is used to tune the hyperparameters (i.e., the architecture) of a classifier. A test dataset is independent of the training dataset, but that follows the same probability distribution as the training dataset. Thus, a test set is used only to assess the performance (i.e., generalization) of a fully specified model [286].

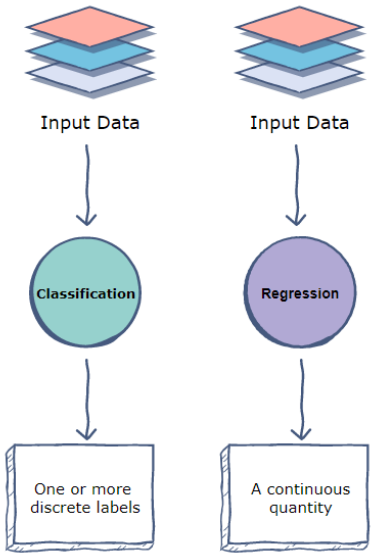

Figure 2.28. Difference between regression and classification

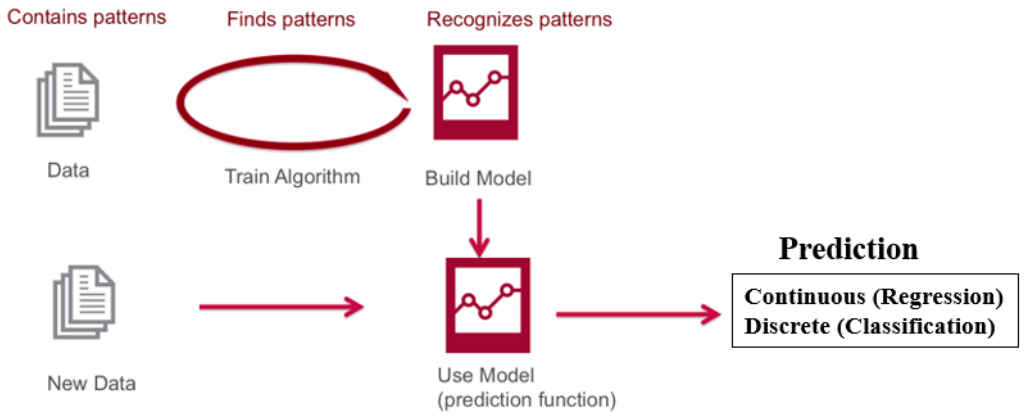

Figure 2.29. Application of ML in prediction (classification and regression)

2.23 Bias and Variance

A constant challenge in ML is differentiating between underfitting and overfitting. These terms decide how closely your model follows the actual patterns of the dataset. To comprehend underfitting and overfitting, bias and variance must be elaborated [287].

Bias, essentially, refers to the gap between the predicted value and the actual value. In the case of high bias, the predictions are likely to be skewed in a certain direction away from the actual values. Variance describes how dispersed your predicted values are. Bias and variance can be greatly understood by examining the visual representation of shooting targets, as shown in Figure 2.30. Shooting targets are not a visual chart used in ML, but it greatly aids to clarify bias and variance. Imagine that the center of the target (the bull's eye), perfectly predicts the correct value of the model. The dots marked on the target then represent an individual realization of the model, based on the training data. In certain cases, the dots will be densely positioned close to the bull's eye, ensuring that predictions made by the model are close to the actual data. In other cases, the training data will be scattered

across the target. The more the dots deviate from the bull's eye, the higher the bias and the less accurate the model will be in its overall predictive ability. In the first target, we can see an example of low bias and low variance. Bias is low because the hits are closely aligned to the center and there is low variance because the hits are densely positioned in one location. The second target (located on the right of the first row) shows a case of low bias and high variance. Although, the hits are not as close to the bull's eye as the previous example, they are still near to the center and bias is therefore relatively low. However, there is high variance because the hits are spread out from each other [287].

The third target (located on the left of the second row) represents high bias and low variance and the fourth target (located on the right of the second row) shows high bias and high variance. Ideally, we want a situation where there is low variance and low bias. In reality, though, there is more often a trade-off between optimal bias and variance. Bias and variance both contribute to error, but it is the prediction error that we want to minimize, not bias or variance, specifically [287].

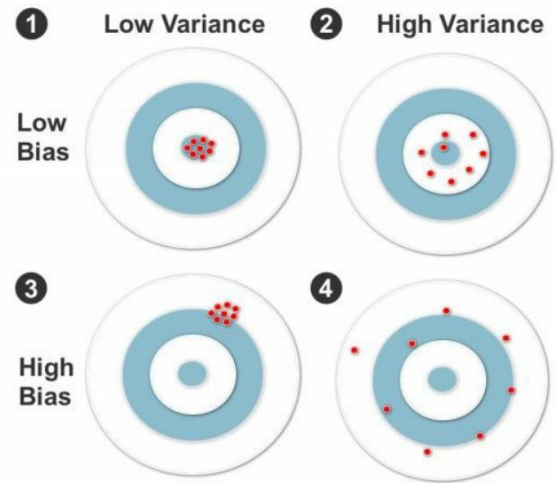

Figure 2.30. Shooting targets used to represent bias and variance

As evident in Figure 2.31, we can see two lines moving from left to right. The purple line represents the test data, and the other line below represents the training data. From the left, both lines begin at a point of high prediction error, due to low variance and high bias. As they move from left to right, the converse occurs: high variance and low bias. This leads to low prediction error in the case of the training data and high prediction error for the test data. In the center of the chart, there is an optimal balance of prediction error between the training and test data. This is commonly known as bias-variance trade-off [287].

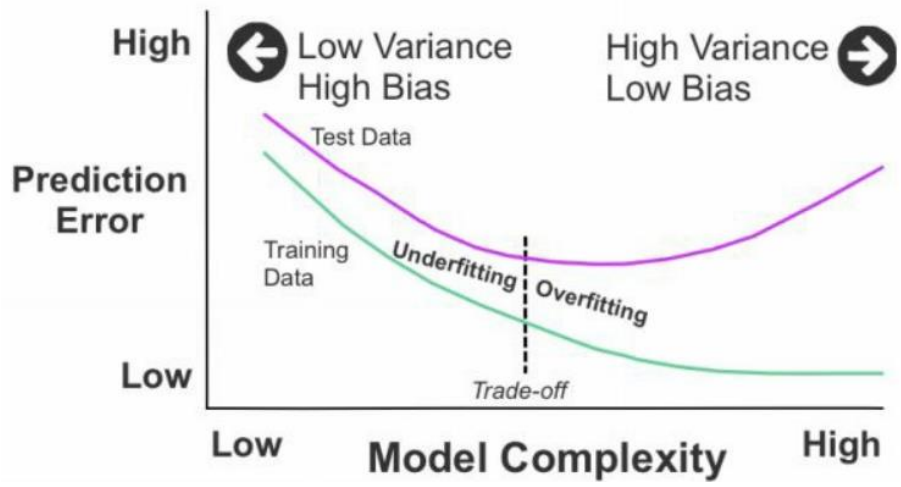

Figure 2.31. Model complexity based on prediction error

2.24 Overfitting and Underfitting

Two of the most well-known issues in ML are overfitting and underfitting [288-289]. These terms refer to the insufficiencies that the model's performance might suffer from. This means that knowing "how off" the model's predictions is a matter of knowing how close it is to overfitting or underfitting. A model that generalizes well is a model that is neither

underfit nor overfit. Generalization is the ability of the model to give sensible outputs to sets of input that it has never seen before. Let us assume, we have the dataset (shown in Figure 2.32), for which the ML model is required.

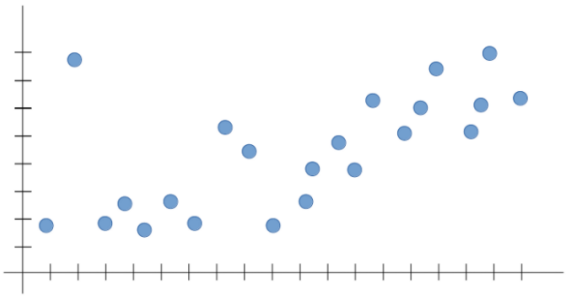

Figure 2.32. Example dataset

Training the linear regression model in this example is all about minimizing the total distance (i.e., cost) between the line we are trying to fit and the actual data points. This goes through multiple iterations, until we find the relatively "optimal" configuration of our line within the data set. This is precisely where overfitting and underfitting occur. In linear regression, we would like our model to follow a line, as shown in Figure 2.33.

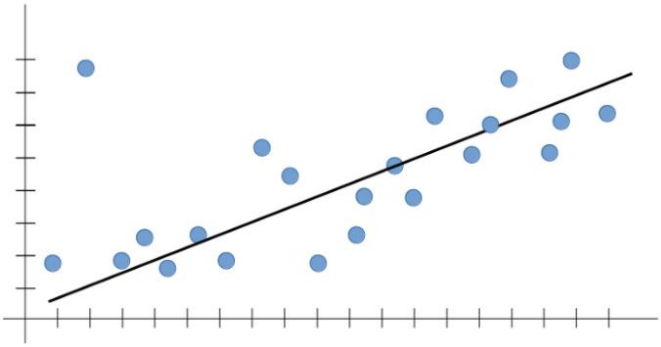

Figure 2.33. Desired ML model

Even though the overall cost is not minimal (i.e., there is a better configuration, in which the line could yield a smaller distance to the data points), the line above fits within the trend very well, making the model reliable. Let us say, we want to infer an output for an input

value that is not currently resident in the data set (i.e., generalize). The line above could give a very likely prediction for the new input, as, in terms of ML, the outputs are expected to follow the trend, observed in the training set.

When the training algorithm is run on the example data set, the overall cost (i.e., distance from each point to the line) can become smaller with more iterations. Leaving this training algorithm run for a long duration, results in minimal overall cost. However, this means that the line will be fit into all the points (including noise), capturing secondary patterns that may not be required for the generalization of the model. Referring to our example, if we leave the learning algorithm running for long, it could end up fitting the line, as illustrated in Figure 2.34.

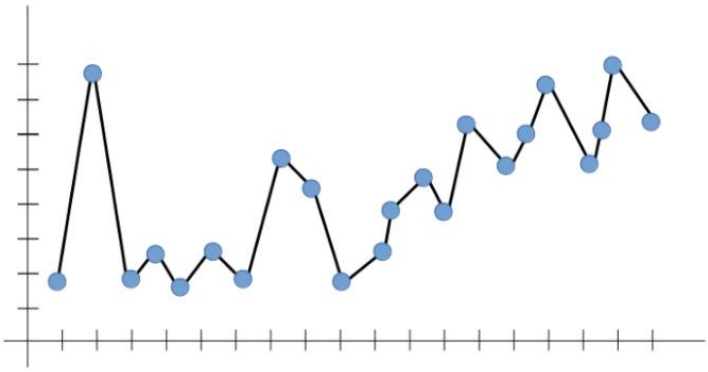

Figure 2.34. Overfitting in ML model

The core of an algorithm like linear regression is to capture the dominant trend and fit our line within that trend. In Figure 2.34, the algorithm captured all trends — but not the dominant one. If we want to test the model on inputs that are beyond the line limits, we have (i.e., generalize), what would that line look like? There is really no way to tell. Therefore, the outputs are not reliable. If the model does not capture the dominant trend

that we can see (positively increasing, in our case), it cannot predict a likely output, for an input that it has never seen before, thereby, confronting the purpose of ML.

Overfitting is the case where the overall cost is small, but the generalization of the model is unreliable. This is due to the model learning “too much” from the training data set. We always want to find the trend, not fit the line to all the data points. Overfitting (or high variance) leads to more bad than good. There is no use of such a model that has learned very well from the training data but still is unable to make reliable predictions for new inputs.

We want the model to learn from the training data, but we do not want it to learn too much (i.e., too many patterns). One solution could be to stop the training earlier. However, this could lead the model to not learn enough patterns from the training data, and possibly not even capture the dominant trend. This case is called underfitting, as illustrated by Figure 2.35. Underfitting is the case where the model has “not learned enough” from the training data, resulting in low generalization and unreliable predictions.

It must be noted that underfitting (or high bias) is just as bad for generalization of the model as overfitting. In high bias, the model might not have enough flexibility in terms of line fitting, resulting in a simplistic line that does not generalize well.

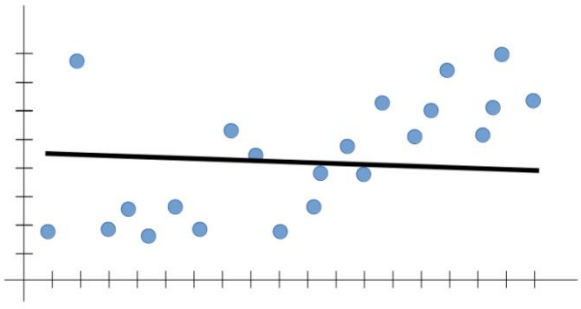

Figure 2.35. Underfitting in ML model

As the goal of SL is to approximate an unknown function by using a dataset of samples, it is a common problem that the model either adapts too well to the input data or is unable to approximate the unknown function because of lack in model capacity. This is undesirable, as we want the model to learn the general patterns found in the input space, and not adapt too much to the noise in the data samples. If the model is unable to approximate the function due to lack of model capacity, it is termed as underfitting. If the model adapts too well to the training dataset, and ends up memorizing the data samples, it is termed as overfitting [288]. There are multiple ways to deal with overfitting and underfitting of models, which can roughly be divided into two categories: data augmentation and model tuning. The process of making the model more robust is called generalization, and methods from both categories are often used to reduce the risk of overfitting or underfitting. The goal of generalization is to reduce the estimated generalization error, the model loss, when presented with new, unseen samples.

The method of early stopping [290] is often used to counter overfitting. When a learning algorithm is trained iteratively, the performance of each iteration of the model can be estimated. Up until a certain number of iterations, new iterations improve the model. After that point, however, the model's ability to generalize can deteriorate, as it begins to overfit the training data. Early stopping means the stopping of the training process before the learner passes that point. In other words, early stopping rules provide guidance as to how many iterations can be run before the learner begins to overfit. This is illustrated in Figure 2.36. On the other hand, underfitting can be prevented by increasing the training time of the model, and by increasing the number of input features.

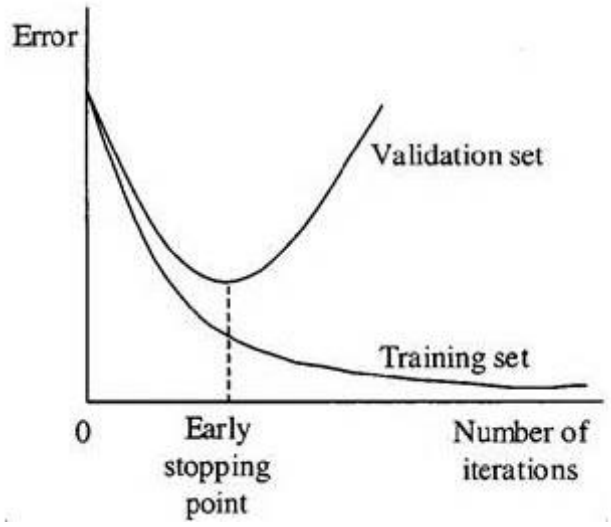

Figure 2.36. The method of early stopping

2.25 Cross Validation: Overview and Classification

This section presents a brief overview and different kinds of cross validation, which is commonly used in ML algorithms [291]. Cross validation is a method for validating the model efficiency by training it on the subset of input data and testing it on unseen subset of the input data, i.e., test data. It is a procedure used to verify how a statistical model generalizes to an independent dataset. Usually, an error estimation for the model is made after training, better known as evaluation of residuals. In this process, a numerical estimate of the difference in predicted and original responses is conducted, also called the training error. However, this only gives an idea about how well the model does on data used to train. The problem with residual evaluations is that they do not give an indication of how well the learner will do when it is asked to make new predictions for data it has not already seen. One way to overcome this problem is to not use the entire data set when training a

learner. Some of the data is removed before training begins. Then, when training is done, the data that was removed, can be used to test the performance of the learned model on new data. This is the basic idea for a whole class of model evaluation methods, commonly called cross validation. Cross validation is a model evaluation method that is generally better than evaluation of residuals. Some commonly used types of cross validation methods are described below.

Holdout Cross Validation

The holdout method is the simplest kind of cross validation. In this approach, the data set is split into two sets, called the training set and the testing set. The function approximator fits a function using the training set only. Afterwards, the function approximator is asked to predict the output values for the data in the testing set (it has never seen these output values before). The errors it makes, are accrued, to determine the mean absolute test set error, which is used to evaluate the model. The benefit of holdout method is that it is typically superior to the residual method and takes no longer to compute. It is because it only needs to be run once. However, its evaluation can have a high variance. The evaluation may depend profoundly on which data points end up in the training set and which end up in the test set, and thus, the evaluation may be significantly different, depending on how the division is made. This is illustrated in Figure 2.37.

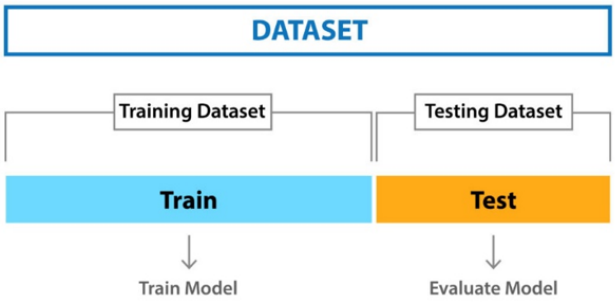

Figure 2.37. Holdout cross validation

Monte Carlo Cross Validation

Monte Carlo validation divides the data randomly into train and test set, and this process is repeated multiple times. The results are averaged over all splits. The drawback of this method is that some observations may never be chosen, whereas some might be selected multiple times. This is illustrated in Figure 2.38.

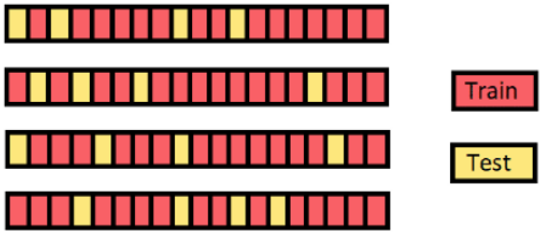

Figure 2.38. Monte Carlo cross validation

K-Fold Cross Validation

K-fold cross validation is the most commonly used cross validation approach and provides one way to improve the holdout method. The data set is divided into *k* subsets, and the holdout method is repeated *k* times. Each time, one of the *k* subsets is used as the test set and the other *k-1* subsets are put together to form a training set. Then, the average error

across all *k* trials is computed. The benefit of this method is that it matters less how the data gets divided. Every data point gets to be in a test set exactly once and gets to be in a training set (*k*-1) times. The variance of the resulting estimate is reduced, as *k* is increased. The disadvantage of this method is that the training algorithm must be rerun from scratch, *k* times, which means it takes *k* times as much computation to make an evaluation. In other words, this approach attempts to maximize the use of the available data for training and consequently, testing a model. It can also prevent over fitting, while training the data [292]. This is illustrated in Figure 2.39.

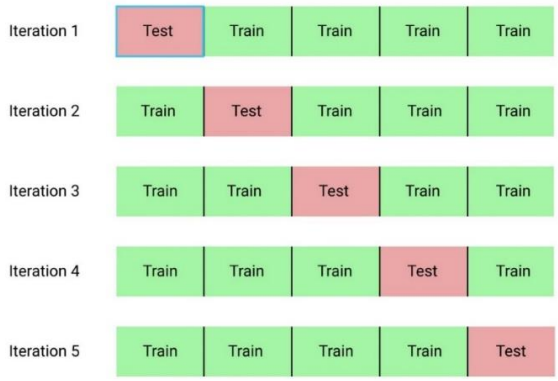

Figure 2.39. Graphical representation of *K*-fold cross-validation (for *K*=5)

2.26 Performance Evaluation Metrics

This section will discuss the commonly used performance evaluation metrics for ML regression and classification tasks.

2.26.1 Regression Metrics

There are various metrics used to evaluate the results of the regression. A brief description of these metrics is presented below [293]. In these metrics, *N* denotes total data points, y_i

denotes actual (target) value, $\hat{y}$ denotes the predicted value, and $\bar{y}$ denotes the mean value of *y*.

Coefficient of Regression, R

It is also known as the Pearson correlation coefficient or the bivariate correlation. This metric measures linear correlation between two variables. It has a value between +1 and −1. A value of +1 is total positive linear correlation, 0 is no linear correlation, and −1 is total negative linear correlation. Mathematically, it is given by,

$$R(y_i, \hat{y}) = \frac{\text{cov}(y_i, \hat{y})}{\sigma_{yi}\sigma_{\hat{y}}} \tag{2.33}$$

where $\text{cov}(y_i, \hat{y})$ is the covariance between y_i and $\hat{y}$; σ_{yi} and $\sigma_{\hat{y}}$ denotes the standard deviation of actual and predicted values, respectively.

Coefficient of Determination, R-squared

This metric measures how much of variability in dependent variable can be explained by the model. It is the square of *R* and hence, is called *R*-squared. It estimates the proportion of the variance for a dependent variable that is explained by an independent variable or variables in the regression model. Its value is between 0 to 1 and a larger value indicates a better fit between predicted and actual value. Mathematically,

$$R^2 = 1 - \frac{SS_{res}}{SS_{tot}} = 1 - \frac{\sum_{i=1}^{N}(y_i - \hat{y})^2}{\sum_{i=1}^{N}(y_i - \bar{y})^2} \tag{2.34}$$

where SS_{res} and SS_{tot} denotes the residual sum of squared errors and total sum of squared errors of the regression model, respectively.

MSE

The mean squared error (MSE) is calculated by the sum of square of prediction error, i.e., the actual output minus predicted output, divided by the total number of data points. It gives an absolute number on how much the predicted values deviate from the actual values. Mathematically,

$$MSE = \frac{1}{N}\sum_{i=1}^{N}(y_i - \hat{y})^2 \tag{2.35}$$

RMSE

Root mean square error (RMSE) is the square root of MSE. It is used more commonly than MSE because firstly, sometimes MSE value can be too big to compare easily. Secondly, MSE is calculated by the square of error, and thus, square root brings it back to the same level of prediction error and makes it easier for interpretation purposes. Mathematically,

$$RMSE = \sqrt{\frac{1}{N}\sum_{i=1}^{N}(y_i - \hat{y})^2} \tag{2.36}$$

MAE

Mean absolute error (MAE) is similar to MSE. However, instead of the sum of square of errors in MSE, MAE is taking the sum of absolute value of errors. Mathematically,

$$MAE = \frac{1}{N}\sum_{i=1}^{N}|\, y_i - \hat{y}\,| \tag{2.37}$$

Compared to MSE or RMSE, MAE is a more direct representation of sum of error terms. MSE gives larger penalization to a larger prediction error by squaring it, while MAE treats all errors the same.

2.26.2 Classification Metrics

There are various metrics used to evaluate the results of the regression. A brief description of these metrics is presented below.

Confusion Matrix

The confusion matrix [294] is one of the most intuitive and simplest approach, used for determining the correctness and accuracy of the model. It is used for a classification problem, where the output can be of two or more types of classes. An example will better help in illustrating the concept. The first step is to assume a label to the target variable, i.e., say, 1, for a person having cancer, and 0, for a person not having cancer. Now, the problem is identified, the confusion matrix, is a table with two dimensions ("actual" and "predicted") and sets of "classes" in both dimensions. In the presented example, the actual classifications are columns and predicted ones are rows, as illustrated in Figure 2.40.

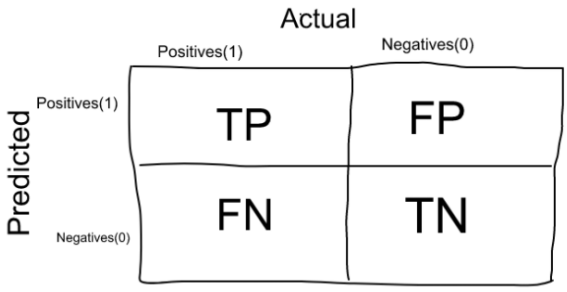

Figure 2.40. Confusion matrix illustration

The confusion matrix is not a performance measure as such, but almost all the performance metrics are based on this matrix and the numbers inside it. Some significant terms associated with confusion matrix are defined below.

True Positives (TP): True positives are the cases when the actual class of the data point was 1 (True) and the predicted is also 1 (True), for instance, the case where a person is

actually having cancer (1) and the model classifying his case as cancer (1) comes under true positive.

True Negatives (TN): True negatives are the cases when the actual class of the data point was 0 (False) and the predicted is also 0 (False), for instance, the case where a person not having cancer and the model classifying his case as not cancer comes under true negatives.

False Positives (FP)*:* False positives are the cases when the actual class of the data point was 0 (False) and the predicted is 1 (True). False is because the model has predicted incorrectly and positive because the class predicted was a positive one (1), for instance, a person not having cancer and the model classifying his case as cancer comes under false positive. They are also known as type I errors.

False Negatives (FN)*:* False negatives are the cases when the actual class of the data point was 1 (True) and the predicted is 0 (False). False is because the model has predicted incorrectly and negative because the class predicted was a negative one (0), for example, a person having cancer and the model classifying his case as no cancer, comes under false negative. They are also known as type II errors.

Classification Accuracy (CA)

Classification Accuracy (CA) is a commonly used classification performance metric [295]. It is calculated as the number of all correct predictions divided by the total number of the data points. The ideal value of *CA* is 1, whereas the worst is 0. It is mathematically defined as

$$CA = \frac{TP + TN}{TP + TN + FP + FN} = \frac{TP + TN}{N} \tag{2.38}$$

where *TP*, *TN* denote the correctly predicted data, and *FP*, *FN* denote incorrectly predicted data, respectively. *N* denotes total data points.

Classification Error (CE)

Classification error (*CE*) represents the number of incorrect predictions from the total number of the data points. The closer it is to zero, the better. Mathematically,

$$CE = \frac{FP + FN}{TP + TN + FP + FN} = \frac{FP + FN}{N} = 1 - CA \tag{2.39}$$

Sensitivity

Sensitivity (or recall) is a measure of actual positives which are correctly identified [295]. Mathematically,

$$S1 = \frac{TP}{TP + FN} \tag{2.40}$$

where $S1$ denotes sensitivity.

Specificity

Specificity is the proportion of truly negative cases that were classified as negative [295]. Mathematically,

$$S2 = \frac{TN}{TN + FP} \tag{2.41}$$

where $S2$ denotes specificity.

Precision

In the simplest terms, precision is the ratio between the true positives and all the positives [296]. Mathematically,

$$P = \frac{TP}{TP + FP} \tag{2.42}$$

where P denotes precision.

F1-score

As evident from (2.40) and (2.42), recall $S1$ and precision P cannot be simultaneously improved; increasing one leads to the decrease of the other one, and vice versa. To counter this, the metric F_1-score is normally used. The F_1-score is a single metric which relates both P and $S1$ through their harmonic mean. This score lies between 0 and 1; with 1 being ideal and 0 being the worst. In simple words, the F_1-score tries to find the balance between precision and recall. Mathematically, it is given by,

$$F_1 = \frac{2 \times P \times S_1}{P + S_1} \tag{2.43}$$

Receiver Operating Characteristic (ROC) curve

A receiver operating characteristic (*ROC*) curve is a graphical plot that establishes the diagnostic ability of a binary ML classifier [297]. In this plot, the true positive rate ($S1$) is plotted against the false positive rate (1-$S2$), as exemplified by Figure 2.41. A classification ML model with perfect discrimination has a *ROC* plot that passes through the upper left corner (100% sensitivity, 100% specificity), i.e., its area under curve (*AUC*) is equal to 1. *AUC* is one of the most widely used metrics for classification. *AUC* of a classifier is equal to the probability that the classifier will rank a randomly chosen positive example higher than a randomly chosen negative example. The closer the *AUC* is to 1, the greater the classification accuracy, i.e., the blue curve in Figure 2.41 should be as close as possible to the top left corner. The closer the curve comes to the 45-degree diagonal of the ROC space, the less accurate is the classifier.

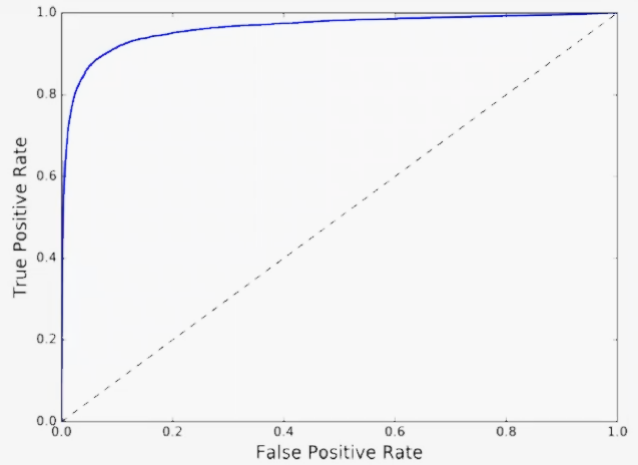

Figure 2.41. Illustration of the *ROC* curve

2.27 Artificial Neural Networks: Background and Overview

The development of ANNs was inspired by the studies of the central nervous system of the human, where the nodes and the interaction within themselves, are to mimic the brains neurons and their synaptic connections. By introducing a training data set to the network, the synaptic weights are iteratively strengthened, until the response of the network follows the output data, like the learning process in the biological brain [298]. ANNs are powerful processing tools, enfolding the ability of learning from experience. From a general viewpoint, ANNs are a data-driven, black box technique, aiming at learning and modeling the input-output relationship, of a given process, from the knowledge of a set of input-output measurements only. Neural networks are nature-inspired techniques. Specifically, ANNs aim at mirroring the functionality and the learning capabilities of the human brain. Thus, they are organized as a network of atomic computational units called artificial neurons, each of them performing a simple and rudimentary processing of its inputs and consequently, propagating the resulting output to the other neurons [299]. Thus, ANNs

have been applied, with promising performances, in various black box modeling tasks, involving classification [300-302] and function approximation [303-305]. Generally, ANNs have three layers: input, hidden, and output. The input layer contains the initial data which is fed into the neural network; the output layer produces the results for the given inputs, and the hidden layer is an intermediate layer between input and output layer, where all the required computation is done, i.e., the hidden layer performs nonlinear transformations of the inputs entering the network [303-304].

2.28 Brief History of ANNs

This section will briefly overview the historical timeline of ANNs, as illustrated in Figure 2.42 [306-308]. Although, the study of the human brain is thousands of years old, the first step towards neural networks took place in 1943, when Warren McCulloch, a neurophysiologist, and Walter Pitts, a young mathematician, wrote a paper on the working principle of neurons. Strengthening the concept of neurons and their working principle was formulated into a book, *The Organization of Behavior*, written by Donald Hebb, in 1949. It concluded that neural pathways are reinforced each time that they are used.

"In the 1950s, N. Rochester, from the International Business Machines (IBM) research laboratories, led the first effort to simulate a neural network. In 1956, the Dartmouth Summer Research Project on artificial intelligence provided a boost to both artificial intelligence and neural networks. This stimulated research in artificial intelligence and in the much lower-level neural processing part of the brain. In 1957, J. Neumann recommended emulating simple neuron functions, by using telegraph relays or vacuum tubes.

In 1958, F. Rosenblatt began working on the perceptron. He was absorbed with the operation of the eye of a fly. Much of the processing which tells a fly to flee is done in its eye. The perceptron, which resulted from this research, was built in hardware and is the oldest neural network still in use today. A single-layer perceptron was found to be valuable in classifying a continuous-valued set of inputs into one of two classes. The perceptron computes a weighted sum of the inputs, subtracts a threshold, and passes one of two possible values out as the result.

In 1959, B. Widrow and M. Hoff developed models they called ADALINE and MADALINE. These models were named for their use of Multiple ADAptive LINear Elements. MADALINE was the first neural network to be applied to a real-world problem. It is an adaptive filter which eliminates echoes on phone lines. Growth on neural network research ceased due fear, unfulfilled claims, etc. until 1981. This caused respected voices to evaluate the neural network research. The result was to pause much of the funding. This period of underdeveloped growth lasted through 1981.

In 1982, several events caused a renewed interest. J. Hopfield presented a paper to the national Academy of Sciences. His approach was not to simply model brains, but to create useful devices. With lucidity and mathematical investigation, he showed how such networks could work and what they could do. By 1985, the American Institute of Physics began what has become a yearly meeting - Neural Networks for Computing. By 1987, the IEEE first International Conference on Neural Networks drew more than 1,800 attendees.

In 1997, A recurrent neural network framework, Long Short-Term Memory (LSTM) was proposed by Schmidhuber and Hochreiter.

In 1998, Y. LeCun published *Gradient-Based Learning Applied to Document Recognition.*

Several other steps have been taken since then to get us to where we are now; today, discussions regarding ANNs are predominant. Presently, most neural network development is simply proving that the principal works."

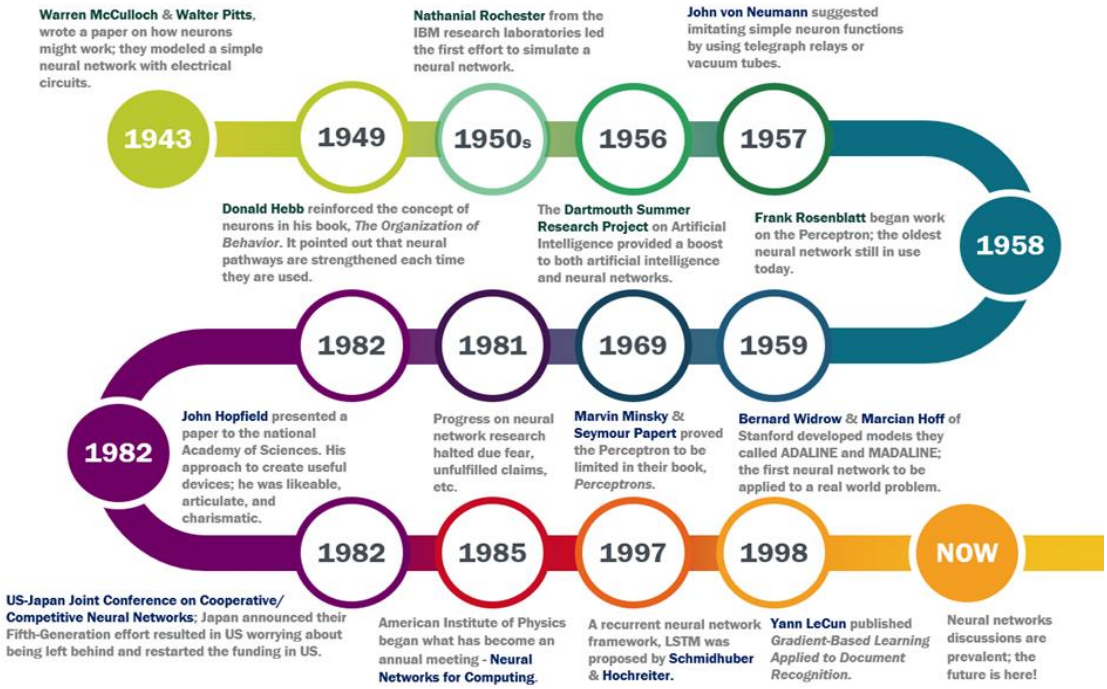

Figure 2.42. Pictorial view of history of neural networks

2.29 Components of ANN

This section will briefly describe various components of a typical ANN [309-310].

Neurons

ANNs are comprised of artificial neurons which are theoretically derived from biological neurons. Each artificial neuron has inputs and produces a single output, which can be directed to numerous other neurons. The inputs can be the feature values of a sample of external data, such as images or documents, or they can be the outputs of other neurons. The outputs of the final output neurons of the neural net achieve the task, such as image recognition. To determine the output of the neuron, the weighted sum of all the inputs is

computed, weighted by the weights of the connections from the inputs to the neuron. Then, a bias term is added to this sum. This weighted sum is occasionally called the activation. This weighted sum is then passed through a (usually nonlinear) activation function for output generation.

Connections and Weights

"The network consists of connections, each connection providing the output of one neuron as an input to another neuron. Each connection is assigned a weight that represents its relative importance. A given neuron can have multiple input and output connections."

Activation Function

Activation functions are functions used in neural networks to compute the weighted sum of input and biases, which is used to decide whether a neuron can be fired or not. Activation function can be either linear or non-linear depending on the function it represents.

Layers

"The neurons are typically organized into multiple layers, especially in deep learning. Neurons of one layer connect only to neurons of the immediately preceding and immediately following layers. The layer that receives external data is the input layer. The layer that produces the ultimate result is the output layer. In between them are zero or more hidden layers. Single layer and unlayered networks are also used. Between two layers, multiple connection patterns are possible. They can be fully connected, with every neuron in one layer connecting to every neuron in the next layer. They can be pooling, where a group of neurons in one layer connect to a single neuron in the next layer, thereby reducing the number of neurons in that layer. Neurons with only such connections form a directed acyclic graph and are known as feedforward networks. Alternatively, networks

that allow connections between neurons in the same or previous layers are known as recurrent networks."

Hyperparameter

A hyperparameter is a constant parameter whose value is set before the learning process begins. The values of parameters are derived via learning. Examples of hyperparameters include number of neurons, the number of hidden layers and batch size.

Loss Function

The loss function (or a cost function) is one of the most significant component of the ANN. It essentially represents the prediction error of neural network, and the method to calculate the loss is called loss function. The loss function simply computes the absolute difference between the predicted and the actual value.

2.30 Types of ANN

This section briefly describes various kinds of ANN. There are two broad types of ANN: static and dynamic [311]. The static ANNs, such as multilayer perceptron neural network (MLPNN), are characterized by memoryless node equations; on the contrary, dynamic ANNs, such as recurrent neural network (RNN), are described by differential equations. ANN has various kinds as depicted by Figure 2.43 [312]. A brief description follows.

Feedforward Neural Network (FNN)

It is the simplest form of neural networks, where input data travels in one direction only, passing through artificial neural nodes, and exiting through output nodes. In feedforward neural network (FNN), hidden layers may or may not be present, input and output layers are always present. They can be further categorized as a single-layered (no hidden layer)

or multilayered (at least one hidden layer) FNN. This kind of ANNs have some advantages, for instance, they are less intricate, easy to design, and maintain. Moreover, they are fast and speedy (one-way propagation), and are highly responsive to noisy data; however, they cannot be applied for deep learning applications, due to absence of dense layers.

Convolutional Neural Network (CNN)

Convolutional neural network (CNN) contains a three-dimensional arrangement of neurons, instead of the standard two-dimensional array. The first layer is called a convolutional layer. Each neuron in the convolutional layer only processes the information from a small part of the visual field. Input features are taken in batch-wise, like a filter. The network understands the images in parts, and can compute these operations multiple times, to complete the full image processing. Processing involves conversion of the image from *RGB* (red, green, blue) or *HIS* (hue, saturation, intensity) scale to grey-scale. Advancing the changes in the pixel value aids in detecting the edges, and therefore, images can be classified into different categories. They are commonly used for deep learning with few parameters and therefore, they require few parameters to learn as compared to fully connected layer. Their main advantage is that they automatically detect the significant features, without any human supervision. For instance, given many pictures of pigs and horses, it can learn the key features for each class by itself. However, they are comparatively intricate to design and maintain, and are comparatively slow (depending on the number of hidden layers). Moreover, CNN requires a huge dataset to process and train the network.

Radial Basis Function Neural Network (RBFNN)

A radial basis function (RBF) neural network is a network which uses radial basis functions as activation functions (these functions are used to determine the output of a neuron in an ANN). The output of the network is a linear combination of radial basis functions of the inputs and neuron parameters. Radial basis function neural network (RBFNN) consists of an input vector, followed by a layer of RBF neurons and an output layer, with one node per category. Classification using RBFNN is generally performed by measuring the input's similarity to data points from the training set, where each neuron stores a prototype. Compared to MLPNN, the training phase is faster, due to absence of backpropagation (BP) learning. Moreover, RBFNN have advantages of easy design, good generalization, and strong tolerance to input noise. However, the classification is slow, in comparison to MLPNN, as every node in the hidden layer must compute the RBF function, for the input sample vector.

Recurrent Neural Network (RNN)

A recurrent neural network (RNN) is a class of ANNs, where connections between nodes form a directed graph, along a temporal sequence. This allows it to exhibit temporal dynamic behavior. Primarily derived from FNNs, RNNs can use their internal state (memory) to process variable length sequences of inputs. They have a crucial advantage to process inputs of any length. Moreover, they can use their internal memory for processing the arbitrary series of inputs, which is not the case with FNNs. Moreover, model size does not increase with input size in the case of RNNs. However, due to its recurrent nature, the computation process of RNN is slow, it cannot consider any future input for the current state, and the training procedure of RNN models can be complicated. Moreover, it cannot process very long sequences if using tanh as an activation function.

Modular Neural Network (MNN)

"A modular neural network (MNN) is an ANN, characterized by a series of independent neural networks, moderated by some intermediary. Each independent neural network serves as a module and operates on separate inputs, to achieve some subtask of the task, the network aims to perform. [313] The intermediary takes the outputs of each module and processes them to produce the net output of the network. The intermediary only accepts the modules' outputs—it does not respond to, nor otherwise signal, the modules. Moreover, the modules do not interact with each other. The possible neuron (node) connections increase quadratically, as nodes are added to a network. Computation time depends on the number of nodes and their connections; any increase has drastic consequences for processing time. Assigning specific subtasks to individual modules reduce the number of necessary connections. However, each module can be trained independently and thereby, can precisely accomplish its simpler task. This means the training algorithm and the training data can be implemented more quickly."

The nonexistence of wide research into learning and formation techniques for neural modularity makes it hard for practitioners to proficiently deploy the technique. Also, there is still a substantial gap regarding stimulation of problem decomposition in modular networks, so that their topological modularity may also become a full functional modularity [314].

Multilayer Perceptron Neural Network (MLPNN)

A multilayer perceptron (MLP) is a class of FNN. This work focuses on MLPNN as it is the simplest and most commonly used ANN [315]. The MLPNN was first developed in early 1970s [316]. This kind of neural network constitutes an input layer (comprising of

input neurons), an output layer (comprising of output neurons), and one or more hidden layers (comprising of hidden neurons). The numbers of neurons in both input and output layers depend on the kind of problem, whereas the numbers of neurons in the hidden layers are arbitrary and are generally selected by trial-and-error approach. Figure 2.44 demonstrates a generic MLPNN with a single hidden layer. The layers in the MLPNN are interconnected by links, which are related with weights that command the impact on the information passing through them [317]. In this network, the flow of information is unidirectional (from input to output through the hidden layer). A learning algorithm determines the weights.

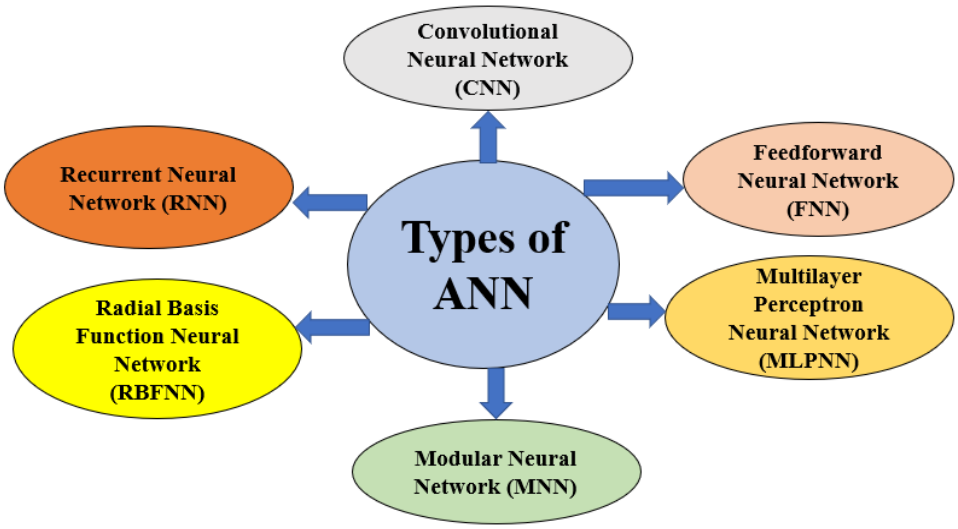

Figure 2.43. Types of ANN

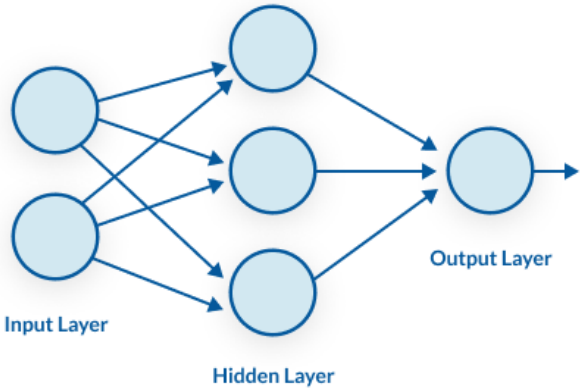

Figure 2.44. A typical MLPNN with one hidden layer

2.31 Network Training: Feedforward

Consider a basic neural network structure, as shown in Figure 2.45 [317]. The input layer consists of x_0 to x_D input nodes. The input parameters are x_1 to x_D and the input bias is x_0. When moving from the input layer to the hidden layer, the raw data, x, is linearly combined using a pre-activation equation, a_j, given by

$$a_j = \sum_{i=1}^{D} w_{ji}^{(1)} x_i + w_{j0}^{(1)} \tag{2.44}$$

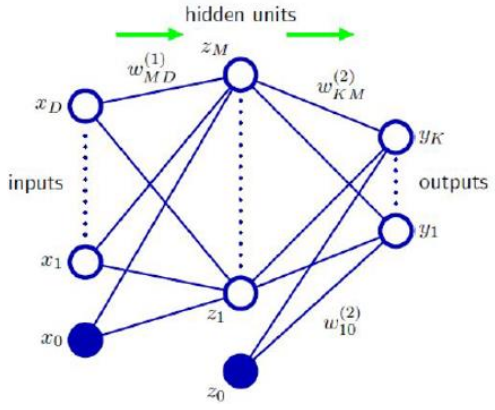

Figure 2.45. Basic structure of a neural network

After each nodal input, a_j, is calculated, the activation is calculated, for each node, in the hidden layer z_1 to z_M, using a chosen activation function. For this example, both the hidden layers and output layer activation functions are the same. The variables in the pre-activation equation are as follows:

1. x_i: Input matrix (includes the bias term x_0).

2. $w_{ji}^{(1)}$: Weights are defined for each connection between an input node and the successive layer's nodes.

3. $w_{j0}^{(1)}$: Weights defined exclusively for each connection between the bias (x_0) and hidden layer nodes.

Now that the input data from the input layer has been transformed using (2.44), a_j becomes the input, going into each node, within the hidden layer. The hidden layer activation function is the next mathematical transformation and will occur at each node, within the hidden layer. For this example, the hidden layer activation function is arbitrarily defined as the sigmoid function, given by,

$$z_j = \sigma(a_j) = \frac{1}{1+e^{-a_j}} \tag{2.45}$$

Once the hidden layer activation function output z_j is computed, the process can be considered as repeating. Now, the output of the first hidden layer is the input to the next layer, which, in this example, is the output layer.

$$a_k = \sum_{j=1}^{D} w_{kj}^{(2)} z_j + w_{k0}^{(2)} \tag{2.46}$$

The activation function can be different for each layer, but for this example the activation function will be the same for both the hidden layer and the output layer. Because this is the

output layer, the output of the output layer's activation function represents the predicted targets, given by,

$$y_k = \sigma(a_k) = \frac{1}{1+e^{-a_k}} \tag{2.47}$$

To finish describing the feed forward mathematical process, for the entire neural network, shown in Figure 2.45, the overall output of the FNN example can be described by

$$y_k(x, w) = \sigma\left(\sum_{j=1}^{D} w_{kj}^{(2)} \sigma\left(\sum_{i=1}^{D} w_{ji}^{(1)} x_i + w_{j0}^{(1)}\right) + w_{k0}^{(2)}\right) \tag{2.48}$$

The above explanation is a basic example of how a FNN can be built and used.

2.32 Network Training: Backpropagation Algorithm

BP algorithm is extensively used as a learning algorithm for MLPNN. It is based on a gradient descent technique [317]. It comprises of the re-iteration of two main phases, known as the forward and the backward ones. During the forward step, the input samples belonging to the training set are fed to the network, and the related outputs are assessed. During the backward phase, the predicted outputs are compared with the expected ones and the resulting errors are fed back to the network to update all the weights. The latter step is performed by minimizing an appropriate error function by means of the gradient descent algorithm. Usually, a differentiable function must be considered, and common examples include mean squared error (MSE), or root mean squared error (RMSE), which are used in function approximation problems [299]. Now that one pass through a FNN has been explained (Section 2.31), the next step is where the ML occurs. The overall output, which was calculated, y_k, is compared to the training outputs, which correspond to the

training inputs, x_i. Depending on how the FNNs predicted targets, $y(x_n, w)$, compared with the actual training data targets, t_n, the weights and biases will be adjusted. This is done by using an error function: as with the activation function, many different error functions can be used depending on the application, but for this example a simple sum-of-squares error function is used, given by

$$E(w) = \frac{1}{2}\sum_{n=1}^{N}\{y(x_n, w) - t_n\}^2 \tag{2.49}$$

where N indicates the total number of data observations.

The error is first calculated at the output layer and propagates backwards, through the FNN all the way back to the input layer. By taking the derivative of the error function with respect to the weights, the error function provides information on how to adjust the weights and biases from their initial values to a value, which will improve the output of the FNN, y_k, to be closer to the actual target values. One parameter which influences the quickness of error minimization is the learning rate hyperparameter. The learning rate describes the magnitude of the adjustment step size for improved weight and biases. A simple flowchart [318] elaborating the basics steps of BP algorithm is shown in Figure 2.46. The detailed mathematics of this algorithm is beyond the scope of this research; however, an eager reader may refer to [319] for associated mathematical functions involved.

2.33 Output of a Single Neuron

The artificial neuron was firstly theorized by McCulloch and Pitts in [320], and a first real implementation of an ANN was proposed by Rosenblatt, in [321], with the perceptron. The

procedure to determine the output of any single neuron in the MLPNN is described as follows [322]. Let x_1, x_2, x_3, ... x_n, and w_{i1}, w_{i2},.....w_{in} be the inputs and corresponding weights, associated with the neuron, respectively. Let b_k be the bias (constant). Let y be the output of the neuron. This is illustrated in Figure 2.47. Mathematically,

$$y = \Phi[\sum_{j=1}^{n}(x_j w_{ij}) + b_k] \tag{2.50}$$

where Φ denotes the activation function. It must be noted that bias is a significant parameter in the MLPNN which is used to offset the output, along with the weighted sum of the inputs to the neuron. Moreover, it allows to shift the activation function to either right or left [322].

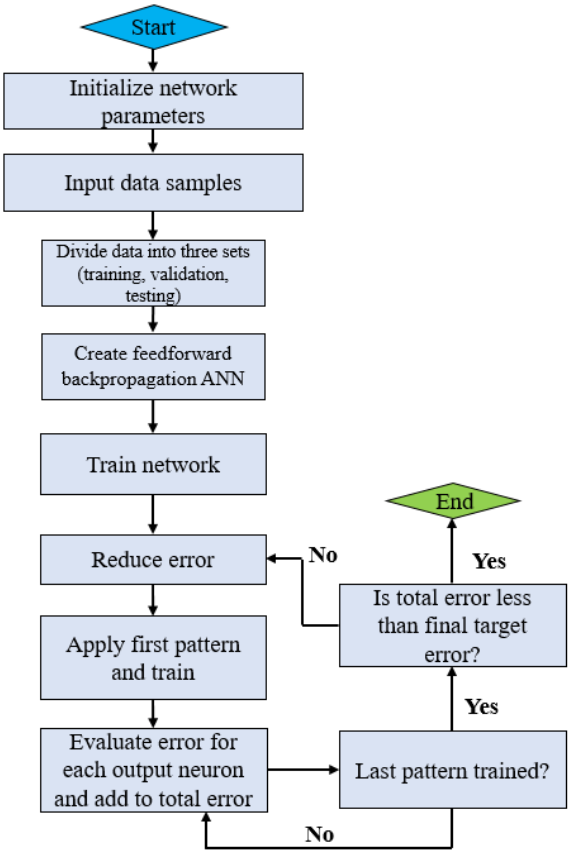

Figure 2.46. Backpropagation (BP) algorithm

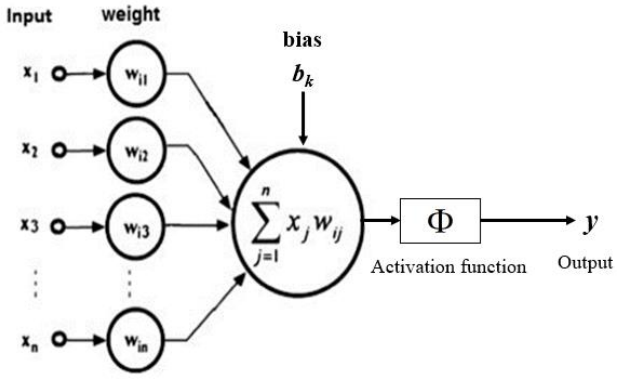

Figure 2.47. Output evaluation process of a single neuron

2.34 ANN Model Construction

There is no recognized technique regarding how to determine the best number of neurons to use in an ANN. In fact, the number of hidden neurons and layers depends on the number of inputs, outputs, sample points, complexity of data [323]. The most common way to decide the configuration is by running a series of tests repeatedly, where the number of hidden neurons is modified until the best configuration is determined.

At the beginning of each training, the synaptic weights are assigned a randomly set starting value, which means that unless the starting values are saved, the chance that the exact network is repeated twice is tremendously small. There is also a risk that the starting values are far from the minima and the learning algorithm gets stuck in a local minimum, the result being a network with poor performance. It is, thus, important to train the network iteratively, with the established configuration, to ensure that the network is not the product of a poor learning cycle. The general framework to train the network is illustrated in Figure 2.48, where training data is fed into the model who gives its response and compares it to the actual value supplied by the training data.

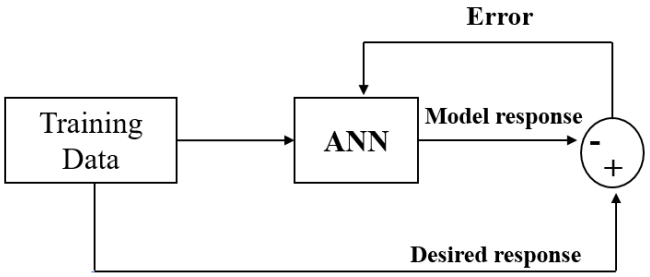

Figure 2.48. ANN training framework

2.35 Activation Functions in ANNs

This section briefly describes common activation (or transfer) functions, used in ANN training [324-325]. An activation function is used to determine the output of neural network, like, say, yes or no. It maps the resulting values in between 0 to 1 or -1 to 1, etc. (depending upon the function). Activation functions also have a key impact on the neural network's ability to converge and the convergence speed, or in some cases, activation functions might prevent neural networks from converging in the first place. The activation functions can be basically divided into two major kinds, linear and nonlinear, as described below.

2.35.1 Linear or Identity Activation Function

The simplest activation function is known as the linear activation, where no transform is applied, as illustrated in Figure 2.49. A network comprised of only linear activation functions is quite easy to train but cannot learn complex mapping functions. Linear activation functions are still used in the output layer for networks that predict a continuous quantity (e.g., regression problems). As evident, the output of such functions will not be confined between any range. Therefore, it does not help with the complexity or various

parameters of usual data, that is fed to the neural networks. Moreover, it is not possible to use BP (gradient descent) to train the model—the derivative of the function is a constant and has no relation to the input. Thus, it is not possible to go back and understand which weights in the input neurons can provide a better prediction.

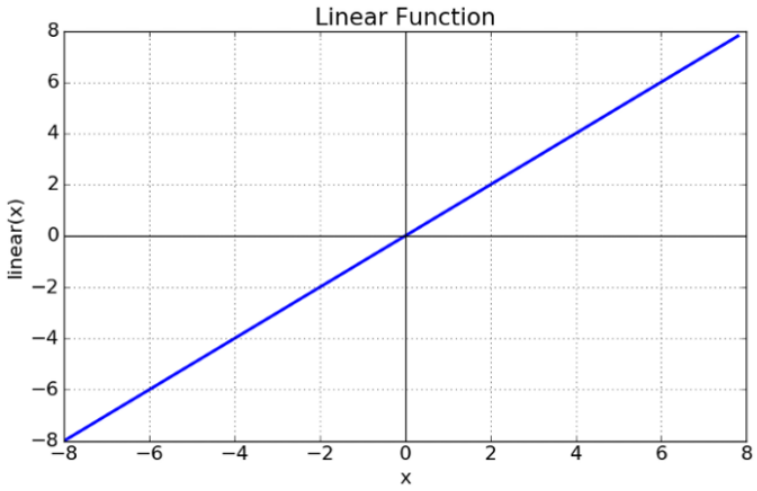

Figure 2.49. Linear activation function

2.35.2 Nonlinear Activation Function

The Nonlinear activation functions are the most commonly used activation functions in ANNs. All modern neural network models use this type of activation function. These functions allow the model to create complex mappings between the network's inputs and outputs, which are essential for learning and modeling complex data, such as images, video, audio, and data sets, which are non-linear or have high dimensionality. Nonlinearity helps to makes the graph, something similar to Figure 2.50. Thus, it makes it easy for the model to generalize or adapt, with variety of data and to differentiate between the output. The main benefit of using nonlinear functions is that they allow BP because they have a derivative function, which is related to the inputs. Moreover, they allow "stacking" of

multiple layers of neurons to form a deep neural network. Multiple hidden layers of neurons are required to learn complex data sets with high accuracy levels.

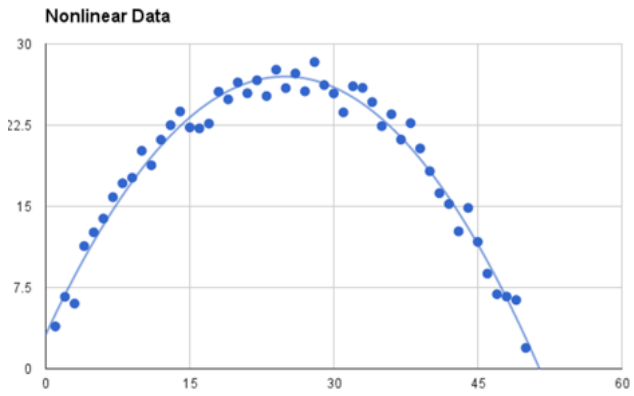

Figure 2.50. Nonlinear activation function

Nonlinear activation functions are further divided into various kinds described below.

Sigmoid or Logistic Activation Function

The sigmoid (or logistic/Fermi) function curve resembles a *S*-shape, as shown in Figure 2.51. The main reason to use sigmoid function is because it lies between 0 to 1. Therefore, it is specifically used for models, where it is required to predict the probability as an output. Since probability exists only between the range of 0 and 1, sigmoid is the right choice. Another benefit of using this activation function is that it is differentiable, i.e., the slope of the sigmoid curve at any two points can be determined. The function is monotonic (a function which is either entirely non-increasing or non-decreasing). The monotonicity criterion helps the neural network to converge easier into a more accurate classifier. Although, the major advantage of using this activation function is that it is simple to apply for classification; however, it has a disadvantage that it gives rise to the problem of "vanishing gradient" because its output is not zero centered. Therefore, a large change in

the input of the sigmoid function will cause a small change in the output. Hence, the derivative (gradient) becomes small. In other words, when inputs become very small or very large, the sigmoid function saturates at 0 and 1. In this case, its derivative is very close to zero. Thus, in this case, it has almost no gradient to propagate back through the network. A small gradient implies that the weights and biases of the initial layers will not be updated effectively in each epoch (iteration through the process of providing the network with an input and updating the network's weights). Since these initial layers are quite crucial in identifying the essential elements of the input data, it can cause an overall inaccuracy in the entire network.

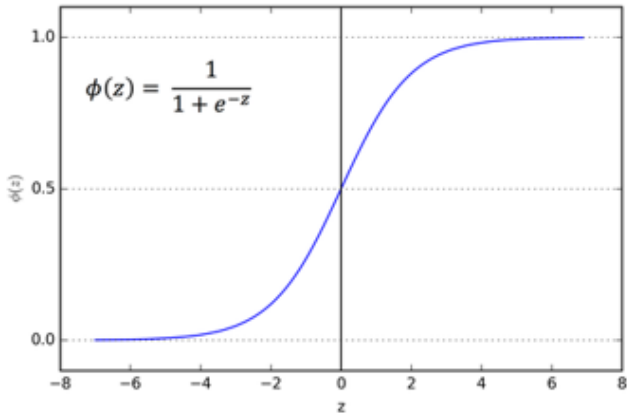

Figure 2.51. Sigmoid activation function

Hyperbolic tangent (Tanh) Activation Function

Tanh is also like logistic sigmoid but better. The range of the tanh function is from -1 to 1. Tanh is also sigmoidal (*S*-shaped), as illustrated by Figure 2.52.

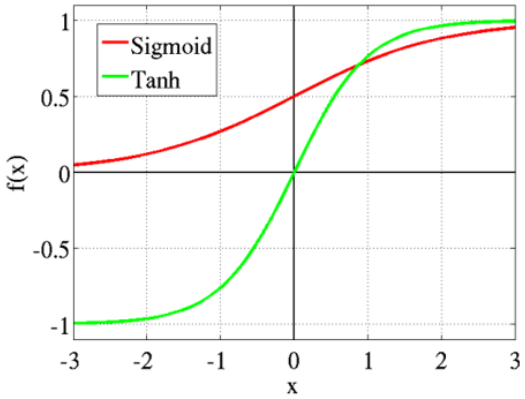

Figure 2.52. Tanh vs. sigmoid activation functions

The tanh function is differentiable and monotonic and is mainly used classification between two classes. Both tanh and sigmoid activation functions are commonly used in FNNs. However, similar to sigmoid activation function, when input becomes large (negative or positive) the tanh function saturates at -1 or +1, with the derivative extremely close to zero. Therefore, like sigmoid function, tanh activation function suffers from vanishing gradients problem, and comes under the class of saturating activation functions.

ReLU (Rectified Linear Unit) Activation Function

The rectified linear unit (ReLU) activation function is a piecewise linear function that will output the input directly if it is positive, otherwise, it will output zero, as illustrated by Figure 2.53. It is used in almost all the CNNs and deep learning algorithms. As evident, the ReLU is half rectified (from bottom). *R*(*z*) is zero when *z* is less than zero and *R(z)* is equal to *z* when *z* is above or equal to zero. Although, the function and its derivative both are monotonic**,** but the issue is that all the negative values become zero immediately, which reduces the ability of the model to fit or train from the data correctly. This implies that any negative input given to the ReLU activation function turns the value into zero, immediately

in the graph, which in turns affects the resulting graph by not mapping the negative values correctly.

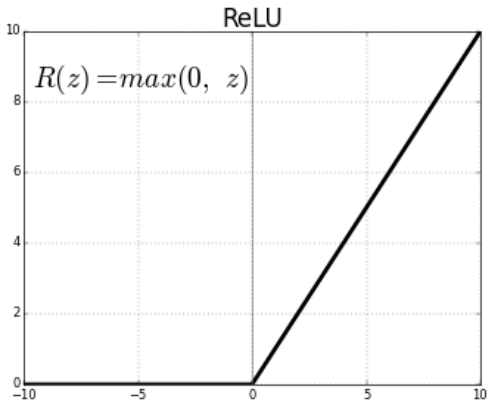

Figure 2.53. ReLU activation function

Softmax Activation Function

The softmax function is a more generalized logistic activation function, which is widely used for multiclass classification. The softmax function, also known as softargmax or normalized exponential function, is a generalization of the logistic function to multiple dimensions. It is used in multinomial logistic regression and is often used as the last activation function of a neural network to normalize the output of a network to a probability distribution, over predicted output classes, based on Luce's choice axiom. The softmax regression is a form of logistic regression that normalizes an input value into a vector of values that follows a probability distribution whose total sums up to 1. The output values are between the range [0,1], which is beneficial because it allows to avoid binary classification and accommodate as many classes or dimensions as possible in the neural network model. This is why softmax is sometimes referred to as a multinomial logistic regression. It is mathematically given by

$$\sigma(\vec{z})_i = \frac{e^{z_i}}{\sum_{j=1}^{K} e^{z_j}} \tag{2.51}$$

where

σ : softmax

$\vec{z}$: input vector

e^{z_i} : standard exponential function for input vector

K : number of classes in the multi-classifier

e^{z_j} : standard exponential function for output vector

2.36 Review of Related Work: ANN Application to Transient Stability

Application of ANN to power system is an area of rising interest; the chief reason being the ability of ANN to process and learn intricate nonlinear relations [326]. Moreover, they possess the ability of parallel processing of data. Recently, transient stability assessment (TSA) problem has been approached using pattern recognition techniques, with some promising results. In these techniques, a relation mapping is established between the input features and the output results of a stability assessment, based on many offline simulations. ANNs have been widely applied to create this relation mapping by numerous research [319, 327-329].

[319] used ANNs to predict critical clearing time (CCT) for a small test power system. [327] used an individual transient energy function (TEF) approach to predict energy margin and stability. [328] devised an integrated approach of unsupervised and supervised learning for TSA. [330] proposed a fast pattern recognition and classification method for states of dynamic security. In [331], ANNs were used to predict stability of a system consisting of

227 buses and 53 generators. [332] applied the recurrent RBF and the MLPNN for predicting rotor angles and angular velocities of synchronous machines. [333] used ANN to classify system stability status for various contingencies. In [334], the nonlinear mapping relation between the transient energy margin and the generator power, at different fault clearing time (FCT), was established by using the multilayer FNN. Lyapunov's direct method, based on the system dynamic equivalents, was used as a fast method to obtain the training set for the ANN. [335] presented a novel ANN-based global online fault detection, pattern classification, and relaying detection scheme, for synchronous generators (SGs) in interconnected electric utility networks. The online ANN based relaying scheme classified fault existence and fault type as either transient stability or loss of excitation, and the allowable CCT, and loss of excitation type as either open circuit or short circuit condition. An innovative two-layer, fuzzy hyperrectangular composite neural network was presented, in [336], to provide real-time transient stability prediction, for high-speed control in power systems. In [337], investigation was carried out for the improvement of power system stability, by utilizing auxiliary controls for controlling high voltage direct current (HVDC) power flow. The current controller model and the line dynamics were considered in the stability analysis. Transient stability analysis was done on a multi-machine system, where a neural network controller was developed to enhance the stability of the power system.

[338] discussed the issue of ANN input dimension reduction. Two different methods, for TSA application, were discussed and compared for efficiency and accuracy. [339] described a neural network-based, adaptive pattern recognition approach, for estimation of the CCT. [340] proposed an application of ANN, for contingency screening and ranking of a power system, with respect to transient stability. [341] suggested a method of TSA, by

adaptive pattern recognition which makes use of an ANN. [342] aimed to examine the use of ANNs, in the analysis of the transient stability of a power system (determination of CCT for short-circuit faults type, with transmission line outage), using a supervised FNN.

In [343], a multilayer feedforward ANN is employed for the online TSA of a power system. [344] used RBFNN as a control scheme, for the unified power flow controller (UPFC), to improve the transient stability performance of a multimachine power system. [345] focused on validating the accuracy of ANN for assessing the transient stability of a single machine infinite bus system. The fault CCT, obtained through ANN, was compared with the results, obtained through the traditional equal area criterion (EAC) method. The multilayer FNN concept was applied to the test system. Some other significant work, associated with ANN applications to transient stability, can be found in [346-353].

Based on the detailed literature review and to the best of author's knowledge, there exists no research works on probabilistic transient stability (PTS), which uses ANN-based ML approach, considering the uncertainties of load, faulted line, fault type, fault location (on the line), and FCT. Although, [354] used ANN for probabilistic dynamic security assessment (DSA), but the approach only considers the uncertainty of load (ignoring other uncertainties of fault type, fault location, and FCT). Moreover, [347] specifically mentions the potential of ANN for online DSA. In addition, [355-361] strongly indicate that ML is a promising and upcoming approach for online DSA. Thus, one of the main contributions of this research is to predict (classify) PTS status using an ANN-based ML approach.

2.37 Support Vector Machine: Background and Overview

A SVM is a supervised learning algorithm that can use given data to solve certain problems by attempting to convert them into linearly separable problems [360]. SVM, which is also known as maximum margin classifier, can be used for both classification and regression problems. It was first introduced by Vapnik [360-361] and was elaborated by Schölkopf et al. [362]. Although, ANN is the most commonly used ML method for transient stability classification, it generally requires an extensive training process and an intricate design procedure. Moreover, ANN usually performs well for interpolation but not so well for extrapolation, which reduces its generalization ability. They are more susceptible to becoming trapped in a local minimum. Although, majority of ML algorithms can overfit, if there is a dearth of training samples, but ANNs can also overfit if training goes on for a very long duration [361]. On the other hand, in the recent years, SVM classifiers have received a huge attention from power systems researchers because of producing single, optimum and automatic sparse solution by simultaneously minimizing both generalization and training error and separating data by the large margin at high dimensional space [364-365]. Due to some of these downsides of ANN, it becomes essential to develop a more efficient classifier for transient stability status prediction. SVM does not suffer from these drawbacks and has the following advantages over ANN [366]: (1) less number of tuning parameters, (2) less susceptibility to overfitting, and (3) the complexity is dependent on number of support vectors (SVs) rather than dimensionality of transformed input space.

SVM classifiers depend on training points, which lie on the boundary of separation between different classes, where the evaluation of transient stability is important. A decent theoretical progress of the SVM, due to its basics built on the Statistical Learning Theory (SLT) [360], made it possible to develop fast training methods, even with large training

sets and high input dimensions [367-369]. This useful characteristic can be applied to tackle the issue of high input dimension and large training datasets in the PTS problem. The basic implementation of an SVM, commonly known as a hard margin SVM, requires the binary classification problem to be linearly separable. This is frequently not the case in practical problems, and therefore, SVM provides a kernel trick to resolve this issue. The strength of the SVM algorithm is based on the use of this kernel trick to transform the input space into a higher dimensional feature space. This allows for defining a decision boundary that linearly separates the classes. The SVM algorithm attempts to find that decision boundary or hyperplane with the highest distance from each class [366]. The hyperplane can be mathematically defined as follows [370],

$$(w^T x) + b = 0 \tag{2.52}$$

where w is the weight vector, x is the sample feature vector and b is the bias value. The samples that assist the algorithm to define the optimal hyperplane are those that lie closest to it, and they are known as SVs. The kernel function plays a significant role in SVM classification [371]. The kernel function is applied on each data instance to map the original non-linear data points into a higher-dimensional space in which they become linearly separable. An SVM classifier minimizes the generalization error by optimizing the relation between the number of training errors and the so-called Vapnik-Chervonenkis (VC) dimension (this dimension is a measure of the capacity of a set of functions that can be learned by a statistical binary classification algorithm. It is defined as the cardinality of the largest set of points that the classification algorithm can shatter). This is achieved by following the method of structural risk minimization (SRM) which states that the

classification error expectation of unseen data is bounded by the sum of a training error rate and a term that depends on the VC dimension [370]. Compared to empirical risk minimization (ERM)-based formulation (which is used by most ML algorithms, including ANN), the SRM-based formulation allows the SVM to prevent overfitting problems, by defining an upper bound, on the expected risk. A formal theoretical bound exists for the generalization ability of an SVM, which depends on the number of training errors (*t*), the size of the training set (*N*), the VC dimension associated to the resulting classifier (*h*), and a chosen confidence measure for the bound itself (η) [370, 372-373]:

$$R < \frac{t}{N} + \sqrt{\frac{h(\ln(\frac{2N}{h}) + 1) - \ln(\frac{\eta}{4})}{N}} \tag{2.53}$$

The risk (or classification error expectation) *R* represents the classification error expectation over all the population of input/output pairs, even though the population is only partially known. This risk is a measure of the actual generalization error and does not require prior knowledge of the probability distribution of the data. SLT derives inequality (2.53) to mean that the generalization ability of an SVM is measured by an upper limit of the actual error given by the right-hand side of (2.53), and this upper limit is valid with a probability of 1- η ($0 < \eta < 1$). As *h* increases, the first summand of the upper bound (2.53) decreases and the second summand increases, such that there is a balanced compromise between the two terms (complexity and training error), respectively [370]. The SVMs used for binary classification problems are based on linear hyperplanes to separate the data, as shown in Figure 2.54. The hyperplane (represented by dotted line in Figure 2.54) is determined by an orthogonal vector *w* and a bias *b*, which identify the points that satisfy

$(w^T x) + b = 0$. By determining a hyperplane which maximizes the margin of separation, denoted by ρ, it is instinctively anticipated that the classifier will have an improved generalization ability. The hyperplane having the largest margin on the training set can be completely determined by the points that lie closest to the hyperplane. Two such points are x_1 and x_2 as shown in in Figure 2.54 (b), and they are known as SVs because the hyperplane (i.e., the classifier) is completely dependent on these vectors. Consequently, in their simplest form, SVMs learn linear decision rules as

$$f(x) = sign(w^T x + b) \tag{2.54}$$

so that (w, b) are determined as to correctly classify the training examples and to maximize ρ. For linearly separable data, as shown in Figure 2.54, a linear classifier can be found such that the first summand of bound (2.53) is zero.

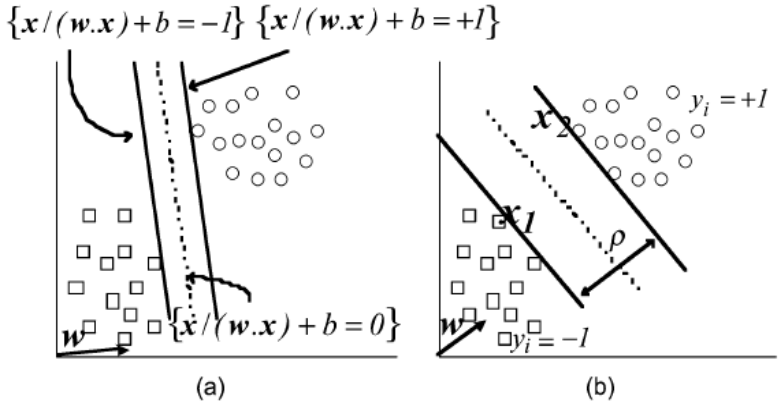

Figure 2.54. SVM (maximum margin) classifier

It is always possible to scale w and b such that

$$w^T x + b = \pm 1 \tag{2.55}$$

for the SVs, with

$$w^T x + b > +1 \text{and} \quad w^T x + b < -1 \tag{2.56}$$

for non-SVs.

Using the SVs x_1 and x_2 of Figure 2.54, and (2.55), the margin ρ can be calculated as

$$\rho = \frac{w^T}{||w||}(x_2 - x_1) = \frac{2}{||w||} \tag{2.57}$$

where $||w||$ is the Euclidean Norm of w. For linearly separable data, the VC dimension of SVM classifiers can be evaluated as

$$h < \min\left\{n, \frac{4D^2}{\rho^2}\right\} + 1 = \min\left\{n, D^2 \,||\, w^2 \,||\right\} + 1 \tag{2.58}$$

where n is the dimension of the training vectors and D is the minimum radius of a ball which contains the training points. Thus, the risk (2.53) can be reduced by decreasing the complexity of the SVM, that is, by increasing the margin of separation ρ, which is equivalent to reducing $||w||$. In practice, as the problems are not probable to be detachable by a linear classifier, thus, the linear SVM can be extended to a nonlinear version by mapping the training data to an expanded feature space using a nonlinear transformation:

$$\Phi(x) = (\phi_1(x), ..., \phi_m(x)) \in R^m \tag{2.59}$$

where $m > n$. Then, the maximum margin classifier of the data for the new space can be determined. With this method, the data points which are non-separable in the original space may become separable in the expanded feature space. The next step is to estimate the SVM by minimizing $|w||$ (i.e., maximizing ρ)

$$V(w) = \frac{1}{2} w^T . w \tag{2.60}$$

subject to the constraint that all training patterns are correctly classified, i.e.,

$$y_i \cdot \{w^T \cdot \Phi(x_i) + b\} \geq 1, \quad i = 1, ..., N \tag{2.61}$$

Though, depending on the kind of nonlinear mapping (2.59), the samples of training data may not be linearly separable. In this case, it is not possible to find a linear classifier that satisfies all the conditions given by (2.60). Thus, instead of (2.60), a new cost function is optimized, i.e.,

$$\begin{aligned} &\min V(w, \varepsilon) = \frac{1}{2} w^T . w + C \sum_{i=1}^{N} \varepsilon_i \\ &s.t. \;\; y_i \cdot \{w^T \cdot \Phi(x_i) + b\} \geq 1 - \varepsilon_i \;\; \text{for } i = 1, ..., N \\ &\qquad\qquad \varepsilon_i \geq 0 \qquad\qquad \text{for } i = 1, ..., N \end{aligned} \tag{2.62}$$

where N non-negative slack variables ε_i are introduced to allow training errors (i.e., training patterns for which $y_i \cdot \{w^T \cdot \Phi(x_i) + b\} \geq 1 - \varepsilon_i$ and $\varepsilon_i > 1$) and allow for some misclassification. By minimizing the first summand of (2.62), the complexity of the SVM is reduced, and by minimizing the second summand of (2.62), the number of training errors is decreased. C is a positive penalty factor (also known as regularization factor or soft margin parameter) which decides the tradeoff between the two terms. In case it is small, the separating hyperplane is more focused on maximizing the margin (at the expense of larger classification mistakes), as shown in Figure 2.55 (A), while the number of misclassified points is minimized for larger C values (at the expense of keeping the margin small and tendency to overfit the data), as shown in Figure 2.55 (B), The minimization of the cost function (2.62) leads to a quadratic optimization problem with a unique solution.

The nonlinear mapping (2.59) is indirectly obtained by the kernel functions, which correspond to inner products of data vectors in the expanded feature space $K(a,b) = \Phi(a)^T \cdot \Phi(b),\ a,b \in R^n$ [370, 373].

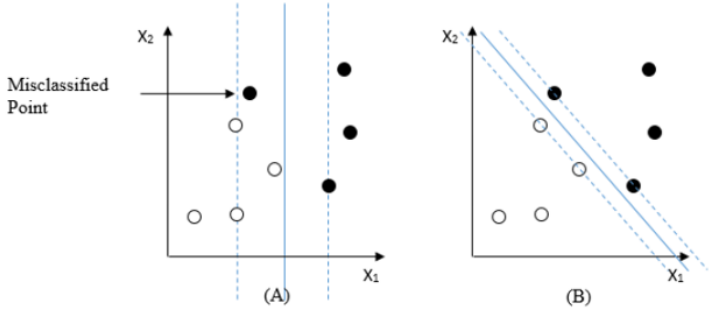

Figure 2.55. Trade-off between maximum margin and minimum training error

2.38 Types of SVM

There are two main types of SVM [374-375]. They are described below.

Linear SVM

Linear SVM is used for data that are linearly separable i.e., for a dataset that can be categorized into two categories by utilizing a single straight line, as shown in Figure 2.56. Such data points are termed as linearly separable data, and the classifier used is described as a linear SVM classifier.

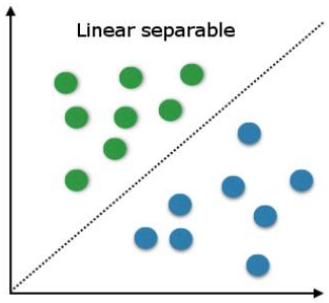

Figure 2.56. Linear SVM

Nonlinear SVM

Nonlinear SVM is used for data that are non-linearly separable data i.e., a straight line cannot be used to classify the dataset, as illustrated by Figure 2.57. For this, something known as a kernel trick is used that sets data points in a higher dimension, where they can be separated using planes or other mathematical functions. Such data points are termed as non-linear data, and the classifier used is termed as a nonlinear SVM classifier.

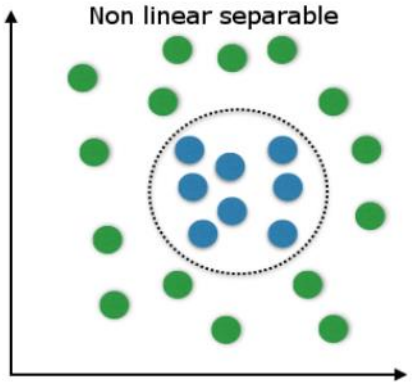

Figure 2.57. Nonlinear SVM

2.39 Kernel Functions in SVM

In certain applications, the data set classes can be deeply overlapping, which makes it impossible to perform a linear classification in the feature space, even by introducing slack variables. The solution for these applications can be obtained by applying Cover's theorem [376]. It stipulates that it is highly probable to solve a nonlinear classification problem, using linear classifiers, by projecting the input set into a higher dimensional space using a nonlinear transformation function [376]. It is evident that the equation of the optimal hyperplane and the decision rule are function of the inner product of the SVs and the new input vector. By mapping the input set into a higher dimensional space, it is required to

compute the high dimensional inner product of their transformation, which requires a good knowledge of the mapping function. According to the Hilbert-Schmidt theory for inner products in high dimensional spaces, computing $\langle \phi(x_i), \phi(x_j) \rangle$ is equivalent to computing a symmetric function $K(x_i, x_j)$ satisfying Mercer's theorem [377]. Here, *K* is called the kernel function. Its main advantage is that it does not require any knowledge on the mapping function. Therefore, the use of the function *K* is commonly referred to as the kernel trick.

The kernel function is what is applied on each data instance to map the original non-linear observations into a higher-dimensional space in which they become separable. The kernel functions return the inner product between two points in a suitable feature space. The function of kernel is to take data as input and transform it into the required form using the transformation ϕ (as illustrated by Figure 2.58). Commonly used kernel functions include the linear, polynomial, sigmoid, Gaussian RBF, and Laplace RBF, as shown in Table 2.3. The choice of the kernel function depends essentially on the data set and in certain cases several trials must be performed before choosing the appropriate one. The Gaussian kernel generally is preferred over others because it has the ability of mapping samples nonlinearly into a higher dimensional space, and therefore, unlike linear kernel, it can tackle the scenario when the relationship between class labels and attributes is nonlinear. Although, sigmoid kernel performs like a Gaussian kernel for certain parameters, but there are some parameters for which the sigmoid kernel is not the dot product of two vectors, thus, it is invalid. Moreover, as compared to polynomial kernel, it has few hyperparameters

(parameters whose values are used to control the learning process) [371]. Thus, this research work uses a Gaussian RBF kernel, which is mathematically given by,

$$K(a,b) = e^{-\gamma\|a-b\|^2},\ \gamma > 0,\ \gamma = \frac{1}{2\sigma^2} \tag{2.63}$$

where γ denotes the kernel parameter of the SVM classifier and σ is the width of the Gaussian function.

The hyperparameters *C* and γ impact how sparse and easily separable the training data are in the expanded feature space. Consequently, these parameters decide the complexity and training error rate of the resulting SVM classifier. These parameters must be optimized for achieving the best performance for the SVM classifier.

Table 2.3. Common SVM kernels

Kernel	**Equation**
Polynomial	$K(x_i, x_j) = (x_i \cdot x_j)^{\text{degree}}$
Gaussian RBF	$K(x_i, x_j) = e^{-\gamma\|x_i - x_j\|^2}$
Linear	$K(x_i, x_j) = x_i \cdot x_j + \text{constant}$
Laplace RBF	$K(x_i, x_j) = e^{-\frac{\|x_i - x_j\|}{\sigma}}$

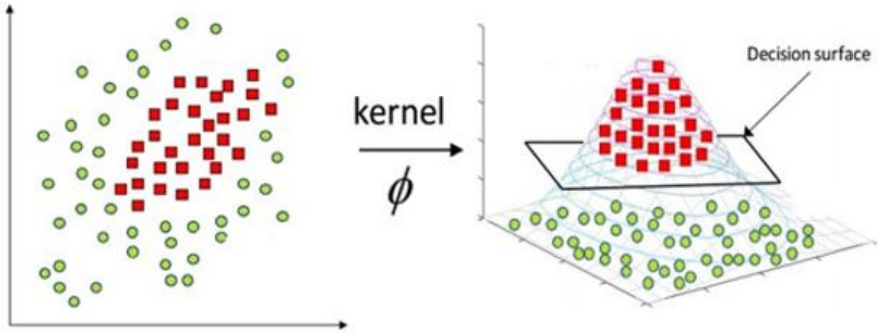

Figure 2.58. Illustration of kernel function

2.40 Methods Used in SVM Optimization

There are various methods which are used to optimize the hyperparameters of SVM: *C* and γ. This section briefly reviews these methods.

Grid Search

The conventional way of performing hyperparameter optimization has been grid search, or a parameter sweep, which is merely an exhaustive searching through a manually specified subset of the hyperparameter space of a learning algorithm. A grid search algorithm must be guided by some performance metric, typically measured by cross validation on the training set or evaluation on a hold-out validation set. Since the parameter space of the learner may include real-valued or unbounded value spaces for certain parameters, manually set bounds and discretization may be essential before applying grid search. A major drawback is that the grid search suffers from the curse of dimensionality [378].

Random Search

Random search replaces the exhaustive enumeration of all combinations by selecting them randomly. This can be simply applied to the discrete spaces and can also be generalized to continuous and mixed spaces. It can outclass grid search, especially when only a small

number of hyperparameters affects the final performance of the ML algorithm. In this case, the optimization problem is said to have a low intrinsic dimensionality. The chance of finding the optimal parameter is relatively higher in random search because of the random search pattern where the model might end up being trained on the optimized parameters, without any aliasing. Random search works best for lower dimensional data, since the time taken to determine the right set is less with less number of iterations. However, the drawback of random search is that it yields high variance during computing. This is because the selection of parameters is completely random, and no intelligence is used to sample these combinations [379].

Bayesian Optimization

Bayesian optimization is a global optimization method for noisy black-box functions. It is a sequential search framework that includes both exploration and exploitation and can be considerably more efficient than either grid search or random search. Applied to hyperparameter optimization, Bayesian optimization builds a probabilistic model of the function mapping, from hyperparameter values to the objective evaluated on a validation set. By iteratively evaluating a promising hyperparameter configuration, based on the current model, and then updating it, Bayesian optimization aims to gather observations revealing as much information as possible about this function, and the location of the optimum. It attempts to balance exploration (hyperparameters for which the outcome is most uncertain) and exploitation (hyperparameters expected close to the optimum). In practice, Bayesian optimization has been proven to obtain better results in fewer evaluations compared to grid search and random search, due to the capability to reason

about the quality of experiments before they are run [380]. In short, Bayesian optimization jointly tunes more parameters with fewer experiments and find better values [380].

2.41 Review of Related Work: SVM Application to Transient Stability

This section will review the application of SVM in transient stability prediction, specifically classification. Recently, SVM has been applied to power system transient stability classification problem. A SVM-based transient stability classifier was trained in [370] and its performance was compared with a MLP classifier. [363] devised a multiclass SVM classifier for TSA classification. [381] suggested a SVM classifier to predict the transient stability status, using voltage variation trajectory templates. [366] trained a binary SVM classifier, with combinatorial trajectories inputs, to predict the transient stability status. [382] employed the SVM to rank the SGs, based on transient stability severity, and consequently, classified them into vulnerable and nonvulnerable machines. [383] proposed two SVMs, using Gaussian kernels, for classifying the post-fault transient stability status of the system. [384] presented a SVM-based approach, for transient stability detection, using post-disturbance signals, from the optimally located distributed generations. [385] proposed a multi-SVM power system TSA method, based on relief algorithm. Firstly, the proposed method selected numerous feature subsets, with different size based on relief algorithm; then, used these selected feature subsets for SVM training, and eventually, these trained SVMs were integrated to evaluate the transient stability of power system.

[386] focused on the prediction of the transient stability of power systems, using only pre-fault and fault duration data, measured by wide area measurement system (WAMS). In the

suggested approach, the time-synchronized values of voltage and current, generated by SGs, were measured using phasor measurement units (PMUs), installed at generator buses, and given as input to the proposed algorithm, to extract a proper feature set. Then, the proposed feature set was applied to (SVM) classifier, to predict the transient stability status. In [387], a different time series forecasting algorithm, using SVM, was proposed, which utilized synchronized phasor data, to provide fast transient stability swings prediction, for the use of emergency control. In [388], a conservative prediction model, for power system transient stability, was suggested, targeting at enhancing accuracy, for predicting the unstable cases. The model was recognized as an ensemble learning model, using multiple SVMs as sub-learning machines. Some other relevant work dealing with SVM-based transient stability prediction can be found in [372, 389-395].

Based on the detailed literature review and to the best of author's knowledge, there exists no research work on PTS which uses SVM-based ML approach, considering the uncertainties of load, faulted line, fault type, fault location (on the line), and FCT. Moreover, [396] specifically mentions the potential of SVM for online transient stability assessment. In addition, [355-359] strongly indicate that ML is a promising and upcoming approach for online DSA. Thus, one of the main contributions of this research is to predict (classify) PTS status using an SVM-based ML approach.

2.42 References

[1] H. H. A. Marhoon, "Adaptive online transient stability assessment of power systems for operational purposes," *Ph.D. Dissertation,* University of New Orleans, USA, 2015.

[2] M. P. Bhavaraju, R. Billinton, R. E. Brown, J. Endrenyi, W. Li, A. P. Meliopoulos, and C. Singh, "IEEE tutorial on electric delivery system reliability evaluation," *IEEE Power Engineering Society General Meeting*, 2005.

[3] R. Billinton and E. Khan, "A security based approach to composite power system evaluation," *IEEE Transactions on Power Systems*, vol. 7, no. 1, pp. 65-72, Feb. 1992.

[4] R. Billinton and S. Aboreshaid, "Security evaluation of composite power systems", *IET Proceedings- Generation, Transmisson, and Distribution*, vol. 142, no. 5, pp. 511-516, Sep. 1995.

[5] The North American Reliability Council, "NERC Planning Standards," approved by NERC Board of Trustees, Sep. 1997.

[6] K. Hua, A. Vahidnia, Y. Mishra, and G. Ledwich, "Efficient probabilistic contingency analysis through a stability measure considering wind perturbation," *IET Generation, Transmission & Distribution*, vol. 10, no. 4, pp. 897–905, 2016.

[7] P. Kundur et al., "Definition and classification of power system stability," *IEEE Transactions on Power Systems*, vol. 19, no. 3, pp. 1387-1401, May 2004.

[8] P. Zhang, S. T. Lee, and D. Sobajic, "Moving toward probabilistic reliability assessment methods," *International Conference on Probabilistic Methods Applied to Power Systems*, 2004, pp. 906-913.

[9] P. Kundur, *Power System Stability and Control*, McGraw-Hill, Inc., New York, 1994.

[10] C. P. Steinmetz, "Power control and stability of electric generating stations," *Transactions of the American Institute of Electrical Engineers*, vol. XXXIX, pp. 1215–1287, Jul. 1920.

[11] AIEE Subcommittee on interconnections and stability factors, “First report of power system stability,” *Transactions of the American Institute of Electrical Engineers*, vol. 56, no. 2, pp. 261-282, Feb. 1937.

[12] G. S. Vassell, “Northeast blackout of 1965,” *IEEE Power Engineering Review*, vol. 11, no. 1, pp. 4–8, Jan. 1991.

[13] S. F. B. Shakil, N. Husain, M. D. Wasim, and S. Junaid, “Improving the voltage stability and performance of power networks using power electronics based FACTS controllers,” *International Conference on Energy Systems and Policies*, 2014, pp. 1-6.

[14] A. M. A. Haidar, M. W. Mustafa, F. A. F. Ibrahim, and I. A. Ahmed, “Transient stability evaluation of electrical power system using generalized regression neural networks,” *Applied Soft Computing*, vol. 1, no. 4, pp. 3558-3570, Jun. 2011.

[15] P. W. Sauer and M. A. Pai, *Power System Dynamic and Stability*, Prentice-Hall, Inc. New Jersey, 1998.

[16] R. Wilkins, “Practical aspects of system stability,” *Journal of the A.I.E.E.*, vol. 45, no. 2, pp. 142-151, Feb. 1926.

[17] R. D. Evans and C. F. Wagner, “Further studies of transmission system stability,” *AIEE Transactions*, pp. 51-80, 1926.

[18] E. Csanyi, Historical Review of Power System Stability Problems, Dec. 2010. [Online]. Available: http://electrical-engineering-portal.com/historical-review-of-power-system-stability-problems

[19] K. N. Hasan, “Application of probabilistic methods in power systems stability assessment-a review identifying future research needs,” Faculty of Engineering and Physical Sciences, The University of Manchester, Apr. 2015.

[20] CIGRE Working Group, “Review of the current status of tools and techniques for risk-based and probabilistic planning in power systems,” C4.601 on Power System Security Assessment, Technical report, Sep. 2010. [Online]. Available: https://www.researchgate.net/publication/272482844_Review_of_the_Current_Status_of_Tools_and_Techniques_for_Risk-Based_and_Probabilistic_Planning_in_Power_Systems

[21] S. S. Venkata, M. Eremia, and L. Toma, Background of power system stability, in *Handbook of Electrical Power System Dynamics: Modeling, Stability, and Control*, NJ, USA: John Wiley & Sons, Inc, 2013.

[22] B. Shakerighadi et al., “A new guideline for security assessment of power systems with a high penetration of wind turbines,” *Applied Sciences*, vol. 10, no. 9, pp. 1-16, May 2020.

[23] M. Pertl, T. Weckesser, M. Rezkalla, and M. Marinelli, “Transient stability improvement: a review and comparison of conventional and renewable-based techniques for preventive and emergency control,” *Electrical Engineering*, vol. 100, pp. 1701–1718, Oct. 2017.

[24] A. Hoballah, “Power system stability assessment and enhancement using computational intelligence,” *Ph.D. Dissertation*, University of Duisburg-Essen, Germany, 2011.

[25] P. Kundur, *Power System Stability and Control*, McGraw-Hill, Inc., New York, 1994.

[26] U. Agrawal, J. O. Brien, and A. Somani, “A study of the impact of reduced inertia in power systems,” *Proceedings of the 53rd Hawaii International Conference on System Sciences*, 2020, pp. 3010-3017.

[27] P. Zhang, “Probabilistic vs deterministic power system stability and reliability assessment,” in *Emerging Techniques in Power System Analysis*. Springer, 2010.

[28] W. Li, *Risk assessment of power systems: models, methods, and applications*, New York: John Wiley & Sons, 2005.

[29] E. Vaahedi, W. Li, T. Chia, and 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.

[30] A. Dissanayaka, U. D. Annakage, B. Jayasekara, and B. Bagen, “Risk-based dynamic security assessment,” *IEEE Transactions on Power Systems*, vol. 26, no. 3, pp. 1302-1308, Aug. 2011.

[31] W. Li and J. Zhou, “Probabilistic reliability assessment of power system operations,” *Electric Power Components and Systems*, vol. 36, no. 10, pp. 1102-1114, Sep. 2008.

[32] L. Miao, J. Fang, and J. Wen, “Transient stability risk assessment of power systems incorporating wind farms,” *Journal of Modern Power Systems and Clean Energy*, vol. 1, no. 2, pp. 134-141, Sep. 2013.

[33] V. E. V. Acker, “Transient stability assessment and decision-making using risk,” *Ph.D. Dissertation*, Iowa State University, 2000.

[34] P. N. Papadopolous, “Probabilistic framework for transient stability assessment of power systems with high penetration of renewable generation,” *IEEE Transactions on Power Systems*, vol. 32, no. 4, pp. 3078-3088, Jul. 2017.

[35] W. Li and P. Choudhury, “Probabilistic planning of transmission systems: Why, how and an actual example,” *IEEE Power and Energy Society General Meeting - Conversion and Delivery of Electrical Energy in the 21st Century*, 2008, pp. 1-8.

[36] A. Sabo and N. I. A. Wahab, “Rotor angle transient stability methodologies of power systems: a comparison,” *IEEE Student Conference on Research and Development*, 2019, pp. 1-6.

[37] A. R. Sobbouhi and A. Vahedi, “Transient stability prediction of power system; a review on methods, classification and considerations,” *Electric Power Systems Research*, vol. 190, pp. 1-16, Jan. 2021.

[38] K. W. Chan, C. H. Cheung, and H. T. Su, “Time domain simulation based transient stability assessment and control,” *International Conference on Power System Technology*, 2002, pp. 1578-1582.

[39] Y. Li and Z. Yang, “Application of EOS-ELM with binary jaya-based feature selection to real-time transient stability assessment using PMU data,” *IEEE Access*, vol. 5 , pp. 23092–23101, 2017.

[40] M. Li, A. Pal, A.G. Phadke, and J. S. Thorp, “Transient stability prediction based on apparent impedance trajectory recorded by PMUs,” *International Journal of Electrical Power & Energy Systems*, vol. 54, pp. 498–504, 2014.

[41] Y. Tang, F. Li, Q. Wang, and Y. Xu, “Hybrid method for power system transient stability prediction based on two-stage computing resources,” *IET Generation, Transmission & Distribution*, vol. 12, no. 8, pp.1697–1703, 2018.

[42] D. You, K. Wang, L. Ye, J. Wu, and R. Huang, “Transient stability assessment of power system based on support vector machine,” *International Journal of Electrical Power and Energy Systems*, vol. 44, no. 1, pp. 318-325, Jan. 2013.

[43] M. Tajdinian, A. R. Seifi, and M. Allahbakhshi, “Transient stability of power grids comprising wind turbines: new formulation, implementation, and application in real-time assessment,” *IEEE Systems Journal*, vol. 13, no. 1, pp. 894-905, Mar. 2019.

[44] S. Rovnyak, C.W. Liu, J. Lu, W. Ma, and J. Thorp, “Predicting future behavior of transient events rapidly enough to evaluate remedial control options in real-time,” *IEEE Transactions on Power Systems*, vol. 10, no. 3, pp. 1195–1203. Aug. 1995.

[45] C. W. Liu and J. S. Thorp, “New methods for computing power system dynamic response for real-time transient stability prediction,” *IEEE Transactions on Circuits and Systems I: Fundamental Theory and Appications.*, vol. 47, no. 3, pp. 324–337, Mar. 2000.

[46] M. A. Pai, *Energy function analysis for power system stability*. Kluwer; 1989.

[47] E. Abreut, B. Wang, and K. Sun, “Semi-analytical fault-on trajectory simulation and its application in direct methods,” *IEEE Power and Energy Society General Meeting*, 2017, pp. 1–5.

[48] T. Liu, Y. Liu, L. Xu, J. Liu, J. Mitra, and Y. Tian, “Non-parametric statistics-based predictor enabling online transient stability assessment,” *IET Generation, Transmission & Distribution*, vol. 12, no. 21, pp. 5761–5769, 2018.

[49] D. P. Wadduwage, C. Q. Wu, and U. D. Annakkage, “Power system transient stability analysis via the concept of Lyapunov Exponents,” *Electric Power Systems Research*, vol. 104, pp. 183–192, 2013.

[50] E. Farantatos, R. Huang, G. J. Cokkinides, and A. P. Meliopoulos, "A predictive generator out-of-step protection and transient stability monitoring scheme enabled by a distributed dynamic state estimator," *IEEE Transactions on Power Delivery*, vol. 31, no. 4, pp. 1826–1835, 2016.

[51] A. Sharifian and S. Sharifian, "A new power system transient stability assessment method based on type-2 fuzzy neural network estimation," *International Journal of Electrical Power and Energy Systems,* vol. 64, pp. 71–87, 2015.

[52] H. D. Chiang, "Direct methods for stability analysis of electric power systems: theoretical foundation, BCU methodologies, and applications," *John Wiley and Sons*, 2010.

[53] Protection system response to power swings, NERC System Protection and Control Subcommittee, Aug. 2013.

[54] W. W. Lemmon, K.R.C. Mamandur, and W.R. Barcelo, "Transient stability prediction and control in real-time by QUEP," *IEEE Transactions on Power Systems*, vol. 4, no.2, pp. 627–642, 1989.

[55] P. Bhui ad N. Senroy, "Real-time prediction and control of transient stability using transient energy function," *IEEE Transactions on Power Systems,* vol. 32, no. 2, pp. 923–934, Mar. 2017.

[56] E. A. Frimpong, P. Y. Okyere, and J. Asumadu, "On-line determination of transient stability status using MLPNN," *IEEE PES PowerAfrica*, 2017, pp. 23-27.

[57] G. A. Maria, C. Tang, and J. Kim, "Hybrid transient stability analysis (power systems)," *IEEE Transactions on Power Systems,* vol. 5, no. 2, pp. 384-393, May 1990.

[58] L. Le-Thanh, T. Tran, and O. Devaux, "Hybrid methods for transient stability assessment and preventive control for distributed generators," *IEEE Power and Energy*

Society General Meeting - Conversion and Delivery of Electrical Energy in the 21st Century, 2008, pp. 1-6.

[59] S. Ma, "A measurement-simulation hybrid method for transient stability assessment and control based on the deviation energy," *International Journal of Electrical Power & Energy Systems*, vol. 115, pp. 1-9, Feb. 2020.

[60] M. H. Haque, "Hybrid method of determining the transient stability margin of a power system," *IEE Proceedings - Generation, Transmission and Distribution*, vol. 143, no. 1, pp. 27-32, Jan. 1996.

[61] Z. Wang, "On-line transient stability studies incorporating wind power," *Ph.D. Dissertation*, Clemson University, USA, 2012.

[62] A. R. Bergen, and V. Vittal, *Power Systems Analysis*, Prentice-Hall, Inc, New Jersey, 2nd edition, 2000.

[63] H. H. A. Marhoon, "Adaptive online transient stability assessment of power systems for operational purposes," *Ph.D. Dissertation*, University of New Orleans, USA, 2015.

[64] K. N. Hasan, R. Preece, and J. V. Milanovic, "Existing approaches and trends in uncertainty modelling and probabilistic stability analysis of power systems with renewable generation," *Renewable and Sustainable Energy Reviews*, vol. 101, pp. 168-180, Mar. 2019.

[65] J. V. Milanovic,"Probabilistic stability analysis: the way forward for stability analysis of sustainable power systems," *Philosophical Transactions*, vol. 375, no. 2100, pp. 1-22, Jul. 2017.

[66] B. Qi, "Identification of influential parameters affecting power system voltage and angular stability analysis," *Ph.D. Dissertation*, The University of Manchester, UK, 2019.

[67] H. Huazhang, C. Y. Chung, and K. W. Chan, "Quasi-Monte Carlo based probabilistic small signal stability analysis for power systems with plug-in electric vehicle and wind power integration," *IEEE Transactions on Power Systems*, vol. 28, no.3, pp. 3335–3343, 2013.

[68] Y. Fan, X. Zai, H. Qian, X. Yang, L. Liu, and Y. Zhu, "Transient stability analysis of power system based on bayesian networks and main electrical wiring," *Asia-Pacific Power and Energy Engineering Conference*, 2009, pp. 1-4.

[69] R. Preece, K. Huang, and J. V. Milanović, "Probabilistic small-disturbance stability assessment of uncertain power systems using efficient estimation methods," *IEEE Transactions on Power Systems*, vol. 29, no. 5, pp. 2509-2517, Sep. 2014.

[70] S. Q. Bu, W. Du, and H. F. Wang, "Investigation on probabilistic small-signal stability of power systems as affected by offshore wind generation," *IEEE Transactions on Power Systems*, vol. 30, no. 5, pp. 2479-2486, Sep. 2015.

[71] D. Han, J. Ma, A. Xue, T. Lin, and G. Zhang, "The uncertainty and its influence of wind generated power on power system transient stability under different penetration," *International Conference on Power System Technology*, 2014, pp. 675-680.

[72] R. N. Allan, A. M. L. Da Silva, and R. C. Burchett, "Evaluation methods and accuracy in probabilistic load flow solutions," *IEEE Transactions on Power Apparatus and Systems*, vol. PAS-100, no. 5, pp. 2539-2546, May 1981.

[73] C. Wan, Z. Xu, Z. Y. Dong, and K. P. Wong, "Probabilistic load flow computation using first-order second-moment method," *IEEE Power and Energy Society General Meeting*, 2012, pp. 1-6.

[74] M. Aien, M. Fotuhi-Firuzabad, and F. Aminifar, "Probabilistic load flow in correlated uncertain environment using unscented transformation," *IEEE Transactions on Power Systems*, vol. 27, no. 4, pp. 2233-2241, Nov. 2012.

[75] J. C. Helton and F. J. Davis, "Latin hypercube sampling and the propagation of uncertainty in analyses of complex systems," *Reliability Engineering & System Safety*, vol. 81, pp. 23-69, 2003.

[76] G. Fishman, *Monte Carlo: concepts, algorithms, and applications*. Springer Science & Business Media, 2013.

[77] R. Preece and J. V. Milanović, "The probabilistic collocation method for dealing with uncertainties in power system small disturbance studies," *IEEE Power and Energy Society General Meeting*, 2012, pp. 1-7.

[78] U. Shahzad and S. Asgarpoor, "Probabilistic evaluation of line loading and line active power in an active distribution network using numerical and analytical approaches," *North American Power Symposium (NAPS)*, 2018, pp. 1-6.

[79] M. Aien, "A comprehensive review on uncertainty modeling techniques in power system studies," *Renewable and Sustainable Energy Reviews*, vol. 57, pp. 1077-1089, May 2016.

[80] K. Jayashree and K. S. Swarup, "A distributed computing environment for probabilistic transient stability analysis," *16th National Power Systems Conference*, 2010, pp. 329-334.

[81] M. Abapour, "Probabilistic transient stability assessment for on-line applications," *International Journal of Electrical Power & Energy Systems*, vol. 42, no. 1, pp. 627-634, Nov. 2012.

[82] S. Aboreshaid, R. Billinton, and M. Fotuhi-Firuzabad, "Probabilistic transient stability studies using the method of bisection," *IEEE Transactions on Power Systems*, vol. 11, no. 4, pp. 1990-1995, Nov. 1996.

[83] R. Billinton and 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.

[84] H. Ahmadi and H. Ghasemi, "Maximum penetration level of wind generation considering power system security limits," *IET Generation, Transmission & Distribution*, vol. 6, no. 11, pp. 1164–1170, Oct. 2012.

[85] S. O. Faried, R. Billinton, and S. Aboreshaid, "Probabilistic evaluation of transient stability of a wind farm" *IEEE Transactions On Energy Conversion*, vol. 24, no. 3, pp. 733-739, Sep. 2009.

[86] P. N. Papadopoulos and J. V. Milanović, "Impact of penetration of non-synchronous generators on power system dynamics," *IEEE Eindhoven PowerTech*, 2015, pp. 1-6.

[87] B. Qi, K. N. Hasan, and J. V. Milanović, "Identification of critical parameters affecting voltage and angular stability considering load-renewable generation correlations," *IEEE Transactions on Power Systems*, vol. 34, no. 4, pp. 2859-2869, Jul. 2019.

[88] Y. N. Zhu and J. X. Jin, "Simulation analysis of a synchronous generator system under fault conditions," *IEEE International Conference on Applied Superconductivity and Electromagnetic Devices*, 2013, pp. 180-184.

[89] E. H. Camm, M. R. Behnke, and O. Boldado, "Characteristics of wind turbine generators for wind power plants," *IEEE Power & Energy Society General Meeting*, 2009, pp. 1-5.

[90] A. G. Abo-Khalil, “Impacts of wind farms on power system stability, modeling and control aspects of wind power systems,” IntechOpen, Mar. 2013. [Online]. Available: https://www.intechopen.com/books/modeling-and-control-aspects-of-wind-power-systems/impacts-of-wind-farms-on-power-system-stability

[91] R. Moxley, Wind energy and protection. [Online]. Available: https://na.eventscloud.com/file_uploads/685732b97917a6e6b078629077fcc88e_WindEnergyr12019.pdf

[92] R. R. Londero, J. P. A Vieira, and C.M. Affonso, “Comparative analysis of DFIG based wind farms control mode on long-term voltage stability,” Advances in Wind Power, IntechOpen, Nov. 2012. [Online]. Available: https://www.intechopen.com/books/advances-in-wind-power/comparative-analysis-of-dfig-based-wind-farms-control-mode-on-long-term-voltage-stability

[93] WECC wind power plant dynamic modeling guide, WECC Renewable Energy Modeling Task Force, Apr. 2014. [Online].
Available: https://www.wecc.org/Reliability/WECC%20Wind%20Plant%20Dynamic%20Modeling%20Guidelines.pdf

[94] Wind Turbines - Part 27-1: Electrical Simulation Models - Wind Turbines, IEC 61400-27-1, Oct. 2015 [Online]. Available: https://infostore.saiglobal.com/preview/is/en/2015/i.s.en61400-27-1-2015.pdf?sku=1841002

[95] S. Xia, Q. Zhang, S. T. Hussain, B. Hong, and W. Zou, “Impacts of integration of wind farms on power system transient stability,” *Applied Sciences*, vol. 8, no. 8, pp. 1-16, Aug. 2018.

[96] M. Edrah, K. L. Lo, and O. Anaya-Lara, “Impacts of high penetration of DFIGwind turbines on rotor angle stability of power systems,” *IEEE Transactions on Sustainable Energy*, vol. 6, no. 3, pp. 759-766, Jul. 2015.

[97] M. Edrah, “Impact of DFIG based wind farms on transient stability of power systems,” *3rd International Conference on Automation, Control, Engineering and Computer Science*, 2016, pp. 1-6.

[98] E. Vittal, M. O’Malley, and A. Keane, “Rotor angle stability with high penetrations of wind generation,” *IEEE Transactions on Power Systems*, vol. 27, no. 1, pp. 353-362, Feb. 2012.

[99] J. Xiang, D. Chen, and M. Yang, “The impact of wind power in wind-thermal hybrid centralized transmission on transient angle stability of power system,” *IEEE 3rd International Electrical and Energy Conference*, 2019, pp. 562-567.

[100] E. Vittal, P. Cuffe, and A. Keane, “Transient stability impacts from distribution connected wind farms,” *IEEE Power and Energy Society General Meeting*, 2012, pp. 1-5.

[101] S. Ncwane and K. A. Folly, “A review of the impact of integrating wind generation on transient stability,” *International SAUPEC/RobMech/PRASA Conference*, 2020, pp. 1-6.

[102] P. N. Papadopolous, “Probabilistic framework for transient stability assessment of power systems with high penetration of renewable generation,” *IEEE Transactions on Power Systems*, vol. 32, no. 4, pp. 3078-3088, Jul. 2017.

[103] J. U. Agber, P. E. Odaba, and C. O. Onah, “Effect of power system parameters on transient stability studies,” *American Journal of Engineering Research,* vol. 4, no. 2, pp. 87-94, 2015.

[104] Q. Wang, A. Xue, T. Bi, and Y. Zheng, “Impact of DFIG-based wind farm on transient stability of single machine infinite bus system,” *2013 IEEE PES Asia-Pacific Power and Energy Engineering Conference (APPEEC),* pp. 1-5.

[105] E. Vaahedi, W. Li, T. Chia, and 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.

[106] R. Billinton and 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.

[107] M. Abapour and M. Haghifam, “Probabilistic transient stability assessment for on-line applications,” *International Journal of Electrical Power & Energy Systems,* vol. 42, no. 1, , pp. 627-634, Nov. 2012.

[108] A. A. Irizarry-Rivera, J. D. McCalley, and V. Vittal, “Computing probability of instability for stability constrained electric power systems,” *Electric Power Systems Research,* vol. 42, no. 2, pp. 135-143, Aug. 1997.

[109] J. Fang, “Probabilistic assessment of power system transient stability incorporating SMES,” *Physica C: Superconductivity,* vol. 484, pp. 276-281, Jan. 2013.

[110] S. O. Faried, R. Billinton, and S. Aboreshaid, “Probabilistic evaluation of transient stability of a power system incorporating wind farms,” *IET Renewable Power Generation,* vol. 4, no. 4, pp. 299–307, Jul. 2010.

[111] S. Aboreshaid, R. Billinton, and M. Fotuhi-Firuzabad, "Probabilistic transient stability studies using the method of bisection," *IEEE Transactions on Power Systems*, vol. 11, no. 4, pp. 1990-1995, Nov. 1996.

[112] S. O. Faried, R. Billinton, and S. Aboreshaid, "Probabilistic evaluation of transient stability of a wind farm" *IEEE Transactions On Energy Conversion*, vol. 24, no. 3, pp. 733-739, Sep. 2009.

[113] X. Zhao and J. Zhou, "Probabilistic transient stability assessment based on distributed DSA computation tool," *IEEE 11th International Conference on Probabilistic Methods Applied to Power Systems*, 2010, pp. 685-690.

[114] S. Chen, L. Shi, L. Yao, Y. Li, and B. Yang, "Transient stability assessment of power system incorporating wind power using quasi-monte carlo method," *TENCON IEEE Region 10 Conference*, 2015, pp. 1-4.

[115] P. Ju, H. Li, C. Gan, Y. Liu, Y. Yu, and Y. Liu, "Analytical assessment for transient stability under stochastic continuous disturbances," *IEEE Transactions on Power Systems*, vol. 33, no. 2, pp. 2004-2014, Mar. 2018.

[116] P. M. Anderson and 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.

[117] K. J. Timko, A. Bose, and 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

[118] R. Billinton, P. R. S. Kuruganty, and 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.

[119] R. Billinton and 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.

[120] R. Billinton and 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.

[121] F. F. Wu, Y. 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.

[122] Y. Hsu and C. Chang, “Probabilistic transient stability studies using the conditional probability approach,” *IEEE Transactions on Power Systems*, vol. 3, no. 4, pp. 1565-1572, Nov. 1988.

[123] M. Tajdinian, “Calculating probability density function of critical clearing time: novel Formulation, implementation and application in probabilistic transient stability assessment,” *International Journal of Electrical Power & Energy Systems,* vol. 103, pp. 622-633, Dec. 2018.

[124] Z. Yue, “Probabilistic transient stability assessment of power system considering wind power uncertainties and correlations,” *International Journal of Electrical Power & Energy Systems,* vol. 117, pp. 1-11, May 2020.

[125] A. Karimishad and T. T. Nguyen, “Probabilistic transient stability assessment using two-point estimate method,” *8th International Conference on Advances in Power System Control, Operation and Management*, 2009, pp. 1-6.

[126] P. N. Papadopoulos, A. Adrees, and J. V. Milanovicć, “Probabilistic assessment of transient stability in reduced inertia systems,” *IEEE Power and Energy Society General Meeting (PESGM)*, 2016, pp. 1-5

[127] Y. Liu, “An analytical approach to probabilistic dynamic security assessment of power systems incorporating wind farms,” *Energy Procedia*, vol. 142, pp. 224-229, Dec. 2017.

[128] J. McCalley, S. Asgarpoor, L. Bertling, R. Billinton, H. Chao, J. Chen, J. Endrenyi, R. Fletcher, A. Ford, C. Grigg, G. Hamoud, D. Logan, A.P. Meliopoulos, M. Ni, N. Rau, L. Salvaderi, M. Schilling, Y. Schlumberger, A. Schneider, and C. Singh, “Probabilistic security assessment for power system operations,” *Proc. IEEE Power Energy Soc. Gen. Meet.*, 2004, pp. 1-9.

[129] A. Dissanayaka, U. D. Annakage, B. Jayasekara and B. Bagen, “Risk-based dynamic security assessment,” *IEEE Transactions on Power Systems*, vol. 26, no. 3, pp. 1302-1308, Aug. 2011.

[130] D. H. Nguyen and M. Negnevitsky, “A probabilistic approach for power system security assessment,” *Proceedings of the 22nd AUPEC*, 2012, pp. 1-6.

[131] S. Datta and 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.

[132] W. Li, Risk assessment of power systems: models, methods, and applications, New York: John Wiley & Sons, 2005.

[133] H. Wan, J. D. McCalley, and V. Vittal, "Risk based voltage security assessment," *IEEE Transactions on Power System,* vol. 15, no. 4, pp. 1247-1254, Nov. 2000.

[134] A. A. Irizarry-Rivera, "Risk-based operating limits for dynamic security constrained electric power systems," *Ph.D. Dissertation*, Iowa State University, USA, 1996.

[135] J. D. McCalley, A. A. Fouad, V. Vittal, A. A. Irizarry-Rivera, B. L. Agrawal, and 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.

[136] L. Miao, J. Fang, and J. Wen, "Transient stability risk assessment of power systems incorporating wind farms," *Journal of Modern Power Systems and Clean Energy*, vol. 1, no. 2, pp. 134-141, Sep. 2013.

[137] K. Jayashree and K. S. Swarup, "A distributed computing environment for probabilistic transient stability analysis," *16th National Power Systems Conference*, 2010, pp. 329-334.

[138] E. B. Elghali, "Risk of transient stability using rotor trajectory index as severity function," *Indonesian Journal of Electrical Engineering and Computer Science*, vol. 6, no. 3, pp. 591-601, Jun. 2017.

[139] V. V. Acker, J. D. McCalley, V. Vittal, and J. A. Pecas Lopes, "Risk-based transient stability assessment," *PowerTech Budapest 99*. Abstract Records. (Cat. No.99EX376), 1999, pp. 235.

[140] W. Bingdong, L. Zhe, and J. Hongjie, “Transient instability risk assessment based on trajectory sensitivity,” *IEEE PES Innovative Smart Grid Technologies*, 2012, pp. 1-6.

[141] S. X. Wang, B. M. Zhang, and Q. Guo, “Transient security risk assessment of global power system based on time margin,” *Proceedings of the Chinese Society of Electrical and Electronics Engineers*, vol. 25, no. 15, pp. 51–55, 2005 (in Chinese).

[142] W. Wang, J. A. Mao, and L. Z. Zhang, “Risk assessment of power system transient security under market condition,” *Proceedings of the Chinese Society of Electrical and Electronics Engineers*, vol. 29, no. 1, pp. 68–73, 2009 (in Chinese).

[143] X. D. Liu , Q. Y. Jiang, and Y. J. Cao, “Transient security risk assessment of power system based on risk theory and fuzzy reasoning,” *Electric Power Automation Equipment*, vol. 29, no. 2, pp. 15–20, 2009 (in Chinese).

[144] M. Francis, Renewables became the second-most prevalent U.S. electricity source in 2020, U. S. Energy Information Administration, Jul. 2021. [Online]. Available: https://www.eia.gov/todayinenergy/detail.php?id=48896

[145] M. Rezkalla, M. Pertl, and M. Marinelli, “Electric power system inertia: requirements, challenges and solutions,” *Electr. Eng*, vol. 100, pp. 2677–2693, 2018.

[146] S. Eftekharnejad, V. Vittal, G. T. Heydt, B. Keel, and J. Loehr, “Impact of increased penetration of photovoltaic generation on power systems,” *IEEE Transactions on Power Systems*, vol. 28, no. 2, pp. 893-901, May 2013.

[147] D. Gautam, V. Vittal, and T. Harbour, “Impact of increased penetration of DFIG-based wind turbine generators on transient and small signal stability of power systems,” *IEEE Transactions on Power Systems*, vol. 24, no. 3, pp. 1426-1434, Aug. 2009.

[148] D. Gautam, L. Goel, R. Ayyanar, V. Vittal, and T. Harbour, "Control strategy to mitigate the impact of reduced inertia due to doubly fed induction generators on large power systems," *IEEE Transactions on Power Systems,* vol. 26, no. 1, pp. 214-224, Feb. 2011.

[149] M. A. Chowdhury, S. Weixiang, N. Hosseinzadeh, and H. R. Pota, "Transient stability of power system integrated with doubly fed induction generator wind farms," *IET Renewable Power Generation,* vol. 9, no. 2, pp. 184-194, 2015.

[150] Y. Zhang, A. A. Chowdhury, and D. O. Koval, "Probabilistic wind energy mod-eling in electric generation system reliability assessment," *IEEE Transactions on Industry Applications,* vol. 47, no. 3, pp. 1507-1514, May-June 2011.

[151] S. Datta, "Risk-based dynamic security assessment of the electricity grid with high penetration of renewable generation," *Ph.D. Dissertation,* Arizona State University, USA, 2017.

[152] K. Hua, A. Vahidnia, Y. Mishra, and G. Ledwich, "Efficient probabilistic contingency analysis through a stability measure considering wind perturbation," *IET Generation, Transmission & Distribution,* vol. 10, no. 4, pp. 897–905, 2016.

[153] H. Dong, M. Jin, X. Ancheng, and L. Tao, "The uncertainty and its influence of wind generated power on power system transient stability under different penetration," *POWERCON,* 2014, pp. 675–680.

[154] S. Xia, X. Luo, K. W. Chan, M. Zhou, and G. Li, "Probabilistic transient stability constrained optimal power flow for power systems with multiple correlated uncertain wind generations," *IEEE Transactions on Sustainable Energy,* vol. 7, no. 3, pp. 1133-1144, Jul. 2016.

[155] P. Wang, Z. Zhang, Q. Huang, and W. Lee, "Wind farm dynamic equivalent modeling method for power system probabilistic stability assessment," *IEEE Industry Applications Society Annual Meeting*, 2019, pp. 1-7.

[156] C. Chen, W. Du, S. Q. Bu, and H. F. Wang, "Probabilistic rotor angular stability considering renewable power generations — A survey," *International Conference on Renewable Power Generation (RPG 2015)*, 2015, pp. 1-6.

[157] T. R. Ayodele, A. A. Jimoh, J. L. Munda. and J. T. Agee, "The impact of wind power on the transient stability of a power system using probabilistic approach," *IEEE Africon*, 2011, pp. 1-6.

[158] W. Li and J. Zhou, "Probabilistic reliability assessment of power system operations," *Electric Power Components and Systems*, vol. 36, no. 10, pp. 1102-1114, Sep. 2008.

[159] K. N. Hasan, R. Preece, and J. V. Milanovic, "Existing approaches and trends in uncertainty modelling and probabilistic stability analysis of power systems with renewable generation," *Renewable and Sustainable Energy Reviews*, vol. 101, pp. 168-180, Mar. 2019.

[160] J. V. Milanovic, "Probabilistic stability analysis: the way forward for stability analysis of sustainable power systems," *Philosophical Transactions*, vol. 375, no. 2100, pp. 1-22, Jul. 2017.

[161] A white paper on the incorporation of risk analysis into planning processes, EPRI Report, 2015. [Online]. Available: https://pubs.naruc.org/pub.cfm?id=536DCF19-2354-D714-5117-47F9BA06F062

[162] V. E. Van, "Transient stability assessment and decision-making using risk," *Ph.D. Dissertation*, Iowa State University, USA, 2000.

[163] W. Li and P. Choudhury, "Probabilistic planning of transmission systems: Why, how and an actual example," *IEEE Power and Energy Society General Meeting - Conversion and Delivery of Electrical Energy in the 21st Century*, 2008, pp. 1-8.

[164] W. Li, "Framework of probabilistic power system planning," *CSEE Journal of Power and Energy Systems*, vol. 1, no. 1, pp. 1-8, Mar. 2015.

[165] J. Choi, T. Tran, A. A. El-Keib, R. Thomas, H. Oh, and R. Billinton, "A method for transmission system expansion planning considering probabilistic reliability criteria," *IEEE Transactions on Power Systems*, vol. 20, no. 3, pp. 1606-1615, Aug. 2005.

[166] E. Rakhshani, D. Gusain, and V. Sewdien, "A key performance indicator to assess the frequency stability of wind generation dominated power system," *IEEE Access*, vol. 7, pp. 130957-130969, Sep. 2019.

[167] J. Fang, H. Li, Y. Tang, and F. Blaabjerg, "On the inertia of future more-electronics power systems," *IEEE Journal of Emerging and Selected Topics in Power Electronics*, vol. 7, no. 4, pp. 2130-2146, Dec. 2019.

[168] G. Papaefthymiou and K. Dragoon, "Towards 100%renewable energy systems: Uncapping power system flexibility," *Energy Policy*, vol. 92, pp. 69-82, 2016.

[169] M. Pertl, T. Weckesser, M. Rezkalla, and M. Marinelli, "Transient stability improvement: a review and comparison of conventional and renewable-based techniques for preventive and emergency control," *Electrical Engineering*, vol. 100, pp. 1701–1718, Oct. 2017.

[170] M. A. H. Sadi and M. H. Ali, "Transient stability enhancement by bridge type fault current limiter considering coordination with optimal reclosing of circuit breakers," *Electric Power Sysems Research*, vol. 124, pp. 160-172, Jul. 2015.

[171] M. H. Ali, T. Murata, and J. Tamura, "The effect of temperature rise of the fuzzy logic controlled braking resistors on transient stability," *IEEE Transactions on Power Systems*, vol. 19, no. 2, pp. 1085–1095, May 2004.

[172] M. H. Haque, "Evaluation of first swing stability of a large power system with various FACTS devices," *IEEE Transactions on Power Systems*, vol. 23, no. 3, pp.1144–1151, Aug. 2008.

[173] M. A. H. Sadi and M. H. Ali, "Combined operation of SFCL and optimal reclosing of circuit breakers for power system transient stability enhancement," *Proceedings of IEEE Southeastcon*, 2013, pp. 1-6.

[174] N.G. Hingorani, L. Gyugyi, and M. El-Hawary, "Understanding FACTS: concepts and technology of flexible AC transmission systems," New York, NY, USA: IEEE Press, 2000.

[175] M. H. Ali, T. Murata, and J. Tamura, "A fuzzy logic-controlled superconducting magnetic energy storage (SMES) for transient stability augmentation," *IEEE Transactions on Control Sysems Technology*, vol. 15, no. 1, pp. 144–150, Jan. 2007.

[176] M. A. H. Sadi and M. H. Ali, "Combined operation of SVC and optimal reclosing of circuit breakers for power system transient stability enhancement," *Electric Power Systems Research*, vol. 106, pp. 241-248, Jan. 2014.

[177] A. Alam and E. B. Makram, "Transient stability constrained optimal power flow," *IEEE Power Engineering Society General Meeting*, 2006, pp. 1-6.

[178] Z. Wang, X. Song, H. Xin, D. Gan, and K. P. Wong, "Risk-based coordination of generation rescheduling and load shedding for transient stability enhancement," *IEEE Transactions on Power Systems*, vol. 28, no. 4, pp. 4674-4682, Nov. 2013.

[179] J. Hossain, A. Mahmud, N. K. Roy, and H. R. Pota, "Enhancement of transient stability limit and voltage regulation with dynamic loads using robust excitation control," *International Journal of Emerging Electric Power Systems*, vol. 14, no. 6, pp. 561-570, Oct. 2013.

[180] G. G. Karady and M. A. Mohamed, "Improving transient stability using fast valving based on tracking rotor-angle and active power," *IEEE Power Engineering Society Summer Meeting*, 2002, pp. 1576-1581.

[181] T. Santhoshkumar and V. Senthilkumar, "Transient and small signal stability improvement in microgrid using AWOALO with virtual synchronous generator control scheme," *ISA Transactions*, vol. 104, pp. 233-244, Sep. 2020.

[182] T.T. Nguyen, V. L. Nguyen, and A. Karimishad, "Transient stability-constrained optimal power flow for online dispatch and nodal price evaluation in power systems with flexible AC transmission system devices," *IET Generation, Transmission & Distribution*, vol. 5, no. 3, pp. 332-346, Mar. 2011.

[183] M. Pertl, "A decision support tool for transient stability preventive control," *Electric Power Systems Research*, vol. 147, pp. 88-96, 2017.

[184] M. F. N. Khan, M. O. Iqbal, and R. Hanif, "Optimal load shedding management for ensuring power system stability," *IEEE International Conference on Power and Renewable Energy (ICPRE)*, 2016, pp. 460-464.

[185] C. J. Mozina, "Power system instability — What relay engineers need to know," *64th Annual Conference for Protective Relay Engineers*, 2011, pp. 103-112.

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

[187] J. M. Ramirez, F. V. Arroyave, and R. E. Correa Gutierrez, "Transient stability improvement by nonlinear controllers based on tracking," *International Journal of Electrical Power and Energy Systems*, vol. 33, no. 2, pp. 315-321, 2011.

[188] M. A. Mahmud, H. R. Pota, M. Aldeen, and M. J. Hossain, "Partial feedback linearizing excitation controller for multimachine power systems to improve transient stability," *IEEE Transactions on Power Systems*, vol. 29, no. 2, pp. 561-571, 2014.

[189] Z. Chen, C. Mao, D. Wang, J. Lu, and Y. Zhou, "Design and implementation of voltage source converter excitation system to improve power system stability," *IEEE Transactions on Industry Applications*, vol. 52, no. 4, pp. 2778-2788, July-Aug. 2016.

[190] P. Kundur et al., "Definition and classification of power system stability," *IEEE Transactions on Power Systems*, vol. 19, no. 3, pp. 1387-1401, May 2004.

[191] R. H. Park, "Fast turbine valving," *IEEE Transactions on Power Apparatus and Systems*, vol. 92, no. 3, pp. 1065-1073, May 1973.

[192] R. Patel, T. S. Bhatti, and D. P. Kothari, "Improvement of power system transient stability using fast valving: a review," *Electric Power Components and Systems*, vol. 29, no. 10, pp. 927-938, 2001.

[193] F. F. Hassan, R. Balasubramanian, and T. S. Bhatti, "Fast valving scheme using parallel valves for transient stability improvement," *IEE Proceedings - Generation, Transmission and Distribution*, vol. 146, no. 3, pp. 330-336, May 1999.

[194] M. Marinelli and S. Massucco, "Analysis of inertial response and primary power-frequency control provision by Doubly Fed Induction Generator wind turbines in a small power system," *17th Power Systems Computation Conference*, 2011, pp. 1-7.

[195] M. F. M. Arani and E. F. El-Saadany, “Implementing virtual inertia in DFIG-based wind power generation,” *IEEE Transactions on Power Systems,* vol. 28, no. 2, pp. 1373-1384, May 2013.

[196] E. Rakhshani and P. Rodriguez, “Analysis of derivative control based virtual inertia in multi-area high-voltage direct current interconnected power systems,” *IET Generation, Transmission & Distribution,* vol. 10, no. 6, pp. 1458-1469, Apr. 2016.

[197] R. Saluja and M.H. Ali, “Novel braking resistor models for transient stability enhancement in power grid system,” *IEEE PES Innovative Smart Grid Technologies Conference (ISGT),* 2013, pp. 1-6.

[198] R. Saluja, S. Ghosh, and M. H. Ali, “Transient stability enhancement of multi-machine power system by novel braking resistor models,” *IEEE Southeastcon*, 2013, pp. 1-6.

[199] M. H. Ali, T. Murata, and J. Tamura, “Augmentation of transient stability by fuzzy-logic controlled braking resistor in multi-machine power system,” *IEEE Russia Power Tech*, 2005, pp. 1-7.

[200] R. Grunbaum and J. Pernot, “Thyristor-controlled series compensation: a state of the art approach for optimization of transmission over power links,” *Technical report,* ABB, 2001.

[201] S. Khan and S. Bhowmick, “A fuzzy TCSC controller for transient stability improvement,” *Annual IEEE India Conference (INDICON),* 2015, pp. 1-5.

[202] A. K. Jadhav and V. A. Kulkarni, “Improvement in power system stability using SSSC based damping controller,” *International Conference on Smart Electric Drives and Power System (ICSEDPS),* 2018, pp. 162-166.

[203] M. O. Brien and G. Ledwich, “Static reactive-power compensator controls for improved system stability,” *IEE Proceedings C Generation, Transmission and Distribution*, vol. 134, no. 1, pp. 38-42, 1987.

[204] A. Ghafouri, M. R. Zolghadri, M. Ehsan, O. Elmatboly, and A. Homaifar, “Fuzzy controlled STATCOM for improving the power system transient stability,” *North American Power Symposium*, 2007, pp. 212-216.

[205] Y. Goto, K. Yukita, H. Yamada, K. Ichiyanagi, and T. Matsumura, “A study on power system transient stability due to introduction of superconducting fault current limiters,” *International Conference on Power System Technology*, 2000, pp. 275-280.

[206] G. Y. Yokomizu, K. Yukita, K. Mizuno, K. Ichiyanagi, Y. Yokomizu, and T. Matsumura, “Experimental studies on power system transient stability due to introduction of superconducting fault current limiters,” *IEEE Power Engineering Society Winter Meeting*, 2000, pp. 1129-1134.

[207] M. Tsuda, Y. Mitani, K. Tsuji, and K. Kakihana, “Application of resistor based superconducting fault current limiter to enhancement of power system transient stability,” *IEEE Transactions on Applied Super-conductivity*, vol. 11, no. 1, pp. 2122-2125, 2001.

[208] M. T. Hagh, S. B. Naderi, and M. Jafari, “Application of non-superconducting fault current limiter to improve transient stability,” *IEEE International Conference on Power and Energy*, 2010, pp. 646-650.

[209] M. Tarafdar Hagh, M. Jafari, and S. B. Naderi, “Transient stability improvement using non-superconducting fault current limiter,” *Power Electronic & Drive Systems & Technologies Conference (PEDSTC)*, 2010, pp. 367-370.

[210] M. Ashraf, H.Sadi, and M. H. Ali, "Transient stability enhancement of multi-machine power system by parallel resonance type fault current limiter," *North American Power Symposium (NAPS)*, 2015, pp.1-6.

[211] R. Khanna, G. Singh, and T. K. Nagsarkar, "Power system stability enhancement with SMES," *International Conference on Power, Signals, Controls and Computation*, 2012, pp. 1-6.

[212] G. Rabbani, J. B. X. Devotta, and S. Elangovan, "Application of simultaneous active and reactive power modulation of SMES unit under unequal mode for power system stabilization," *IEEE Transaction on Power Systems*, vol. 14, no. 2, pp. 547-552, May 1999.

[213] P. Mukherjee and V. V. Rao, "Superconducting magnetic energy storage for stabilizing grid integrated with wind power generation systems," *Journal of Modern Power Systems and Clean Energy*, vol. 7, pp. 400–411, 2019.

[214] P. Tixador, "Superconducting magnetic energy storage: status and perspective," *IEEE/CSC & ESAS European Superconductivity News Forum*, no. 3, January 2008. [Online]. Available:
https://snf.ieeecsc.org/sites/ieeecsc.org/files/CR5_Final3_012008.pdf

[215] R. Khanna, G. Singh, and T. K. Nagsarkar, "Artificial neural network based SMES unit for transient stability improvement," *Joint International Conference on Power Electronics, Drives and Energy Systems*, 2010, pp. 1-7.

[216] Schneider Electric, "Circuit breaker characteristic trip curves and coordination," *Technical report*, Schneider Electric, Cedar Rapids, 2001.

[217] Siemens, "High-Voltage Circuit Breakers," *Technical report*, Siemens, 2012.

[218] K. Murugan, “Improvement of transient stability of power system using solid state circuit breaker,” *American Journal of Applied Sciences*, vol. 10, no. 6, pp. 563-569, Jun. 2013.

[219] A. R. Adly, R. A. El Sehiemy, and A. Y. Abdelaziz, “Optimal reclosing time to improve transient stability in distribution system,” *CIRED - Open Access Proceedings Journal*, vol. 2017, no. 1, pp. 1359-1362, Oct. 2017.

[220] M. Abapour, P. Aliasghari, and M. Haghifam, “Risk-based placement of TCSC for transient stability enhancement,” *IET Generation, Transmission & Distribution*, vol. 10, no. 13, pp. 3296-3303, Oct. 2016.

[221] P. V. Gomes and J. T. Saraiva, “A two-stage strategy for security-constrained AC dynamic transmission expansion planning,” *Electric Power Systems Research*, vol. 180, pp. 1-10, Mar. 2020.

[222] A. Karami, “Transient stability assessment of power systems described with detailed models using neural networks,” *International Journal of Electrical Power & Energy* Systems, vol. 45, no. 1, pp. 279-292, Feb. 2013.

[223] L. Hassan, “Current state of neural networks applications in power system monitoring and control,” *International Journal of Electrical Power & Energy Systems*, vol. 51, pp. 134-144, Oct. 2013.

[224] S. M. Miraftabzadeh, F. Foiadelli, M. Longo, and M. Pasetti, “A survey of machine learning applications for power system analytics,” *IEEE International Conference on Environment and Electrical Engineering*, 2019, pp. 1-5.

[225] M. S. Ibrahim, “Machine learning driven smart electric power systems: current trends and new perspectives,” *Applied Energy*, vol. 272, pp. 1-19, Aug. 2020.

[226] O. A. Alimi, K. Ouahada, and A. M. Abu-Mahfouz, “A review of machine learning approaches to power system security and stability,” *IEEE Access*, vol. 8, pp. 113512-113531, Jun. 2020.

[227] W. Li and J. Zhou, “Probabilistic reliability assessment of power system operations,” *Electric Power Components and Systems*, vol. 36, no. 10, pp. 1102-1114, 2008.

[228] Y. Suzuki, H. Kojima, N. Hayakawa, F. Endo, and H. Okubo, “Optimization of asset management in high voltage substation based on equipment monitoring and power system operation,” *IEEE International Symposium on Electrical Insulation*, 2010, pp. 1-5.

[229] M. R. Pahlevi, W. F. Praditama and B. L. Daniel, “Justification for circuit breaker refreshment program in PLN trans-JBTB based on technical condition and impact criteria,” *International Conference on High Voltage Engineering and Power Systems (ICHVEPS)*, 2017, pp. 472-475.

[230] L. Asgarieh, G. Balzer, and A. J. Gaul, “Circuit-breaker estimation with the aid of ageing models,” *Proceedings of the 10th International Conference on Probabilistic Methods Applied to Power Systems*, 2008, pp. 1-6.

[231] L. Promseela, C. Suwanasri, S. Saribut, T. Suwanasri, and R. Phadungthin, “Risk assessment for power circuit breaker by using failure modes, effects and criticality analysis*,” International Conference on Power, Energy and Innovations (ICPEI)*, 2020, pp. 149-152.

[232] Z. Guo, Y. Yang, H. Liu, Y. Wang, and L. Zhou, “Reliable life calculation of circuit breaker based on FTA and reliability function,” *IEEE International Conference on High Voltage Engineering and Application (ICHVE)*, 2016, pp. 1-4.

[233] M. Urayama et al., “Evaluation of deterioration in aged GCBs and study on equipment replacement criteria,” *Transmission & Distribution Conference & Exposition: Asia and Pacific*, 2009, pp. 1-4.

[234] J. M. Byerly, C. Schneider, R. Schloss, and I. West, “Real-time circuit breaker health diagnostics,” *70th Annual Conference for Protective Relay Engineers (CPRE)*, 2017, pp. 1-6.

[235] H. Tanaka, H. Magori, T. Niimura, and R. Yokoyama, “Optimal replacement scheduling of obsolete substation equipment by branch & bound method,” *IEEE PES General Meeting*, 2010, pp. 1-6.

[236] S. Natti, M. Kezunovic, and C. Singh, “Sensitivity analysis on the probabilistic maintenance model of circuit breaker,” *International Conference on Probabilistic Methods Applied to Power Systems*, 2006, pp. 1-7.

[237] S. Lipirodjanapong, C. Suwanasri, and T. Suwanasri, “The reliability evaluation of configuration of bus arrangement in substations by the variable failure rate of power circuit breaker under time-based maintenance,” *International Conference on Condition Monitoring and Diagnosis (CMD)*, 2016, pp. 705-708.

[238] S. Natti, P. Jirutitijaroen, M. Kezunovic, and C. Singh, “Circuit breaker and transformer inspection and maintenance: probabilistic models,” *International Conference on Probabilistic Methods Applied to Power Systems*, 2004, pp. 1003-1008.

[239] N. U. A. Wardani and S. Naswil, “Transmission asset lifecycle management in PLN TJBB,” *International Conference on High Voltage Engineering and Power Systems (ICHVEPS)*, 2017, pp. 95-98.

[240] K. K. Leung, “Replacement of oil-insulated circuit-breakers/switches by vacuum counterparts*,” International Conference on Advances in Power System Control, Operation and Management*, 1991, pp. 207-210.

[241] R. M. A Velasquez, J. V. M. Lara, and A. Melgar, “Reliability model for switchgear failure analysis applied to ageing,” *Engineering Failure Analysis*, vol. 101, pp. 36-60, Jul. 2019.

[242] J. Bowen and T. Burse, “Medium voltage replacement breaker projects,” Record of Conference Papers. Industry Applications Society Forty-Seventh Annual Conference, Petroleum and Chemical Industry Technical Conference (Cat. No.00CH37112), 2000, pp. 233-244.

[243] Maintenance of Power Circuit Breakers. [Onine]. Available: https://www.usbr.gov/power/data/fist/fist3_16/FIST_3-16_(06-2020).pdf

[244] Asset Management Plan. [Online]. Available: https://www.aer.gov.au/system/files/PWC%20-%2014.7%20AMP%20High%20Voltage%20Circuit%20Breakers%20-%2028%20February%202018.pdf

[245] T. S. Chung and F. Zhong, “On-line dynamic security assessment in energy management system,” *International Conference on Power System Technology Proceedings* (Cat. No.98EX151), 1998, pp. 1398-1401.

[246] I. A. Hiskens “Iterative computation of marginally stable trajectories” *International Journal of Nonlinear Robust Control*, vol. 14, pp. 911-924, 2004.

[247] K. W. Cheung, “An improved transient stability index for dynamic security assessment using a marginally unstable injection (MUI) approach,” *IEEE Power*

Engineering Society Winter Meeting. Conference Proceedings (Cat. No.00CH37077), 2000, pp. 931-936.

[248] V. Vittal, E-Z. Zhou, C. Hwang, and A.A. Fouad, "Derivation of stability limits using analytical sensitivity of the transient energy margin," *IEEE Transactions on Power Systems*, vol. 4, no. 4, pp. 1363–1372, Nov. 1989.

[249] T. Athay, R. Podmore, and S. Virmani, "A practical method for direct analysis of transient stability," *IEEE Transactions on Power Apparatus and Systems*, vol. PAS-98, pp. 573–584, 1979.

[250] D. Fang, T. S. Chung, Y. Zhang and W. Song, "Transient stability limit conditions analysis using a corrected transient energy function approach," *IEEE Transactions on Power Systems*, vol. 15, no. 2, pp. 804-810, May 2000.

[251] M. A. Chowdhury, N. Hosseinzadeh, W. X. Shen, and H. R. Pota, "Comparative study on fault responses of synchronous generators and wind turbine generators using transient stability index based on transient energy function," *International Journal of Electrical Power and Energy Systems*, vol. 51, pp. 145-152, Oct. 2013.

[252] K. Hua, A. Vahidnia, Y. Mishra, and G. Ledwich, "Efficient probabilistic contingency analysis through a stability measure considering wind perturbation," *IET Generation, Transmission, & Distribution*, vol. 10, no. 4, pp. 897–905, 2016.

[253] J. Geeganage, U. D. Annakkage, T. Weekes, and B. A. Archer, "Application of energy-based power system features for dynamic security assessment," *IEEE Transactions on Power Systems*, vol. 30, no. 4, pp. 1957–1965, Jul. 2015.

[254] Y. Wang, V. Vittal, M. Abdi-Khorsand, and C. Singh, "Probabilistic reliability evaluation including adequacy and dynamic security assessment," *IEEE Transactions on Power Systems*, vol. 35, no. 1, pp. 551-559, Jan. 2020.

[255] A. Samuel, "Some studies in machine learning using the game of checkers," *IBM Journal of Research and Development*, vol. 3, no. 3, pp. 210-229, 1959.

[256] E. Alpaydin, *Introduction to Machine Learning*, 4th ed., MIT Press, 2020.

[257] E. Hossain, I. Khan, F. Un-Noor, S. S. Sikander, and M. S. H. Sunny, "Application of big data and machine learning in smart grid, and associated security concerns: a review," *IEEE Access*, vol. 7, pp. 13960-13988, 2019.

[258] L. Deng and D. Yu, "Deep learning: methods and applications," *Foundations and Trends in Signal Processing*, vol. 7, no. 3, pp. 197–387, 2013.

[259] V. Advani, What is Machine Learning? How Machine Learning Works and future of it? [Online]. Available: https://www.mygreatlearning.com/blog/what-is-machine-learning/

[260] S. A. Kiranmai and A. J. Laxmi, "Data mining for classification of power quality problems using WEKA and the effect of attributes on classification accuracy," *Protection and Control of Modern Power Systems*, vol. 3, no. 29, pp. 1-12, 2018.

[261] S. Khan, J. Huh, and J. C. Ye, "Deep learning-based universal beamformer for ultrasound imaging," arXiv preprint, arXiv:1904.02843, 2019.

[262] Z. M. Fadlullah and F. Tang, "State-of-the-art deep learning: evolving machine intelligence toward tomorrow's intelligent network traffic control systems," *IEEE Communications Surveys & Tutorials*, vol. 19, no. 4, pp. 2432-2455, 2017.

[263] N. Zhao, H. Zhang, R. Hong, M. Wang, and T. Chua, "Videowhisper: toward discriminative unsupervised video feature learning with attention-based recurrent neural networks," *IEEE Transactions on Multimedia*, vol. 19, no. 9, pp. 2080-2092, Sep. 2017.

[264] I. Giannakis, A. Giannopoulos, and C. Warren, "A machine learning-based fast-forward solver for ground penetrating radar with application to full-waveform inversion," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 57, no. 7, pp. 4417-4426, Jul. 2019.

[265] P. Liu, H. Zhang, and K. B. Eom, "Active deep learning for classification of hyperspectral images," *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, vol. 10, no. 2, pp. 712-724, Feb. 2017.

[266] D. Han, J. Lee, J. Lee. and H. Yoo, "A low-power deep neural network online learning processor for real-time object tracking application," *IEEE Transactions on Circuits and Systems I: Regular Papers*, vol. 66, no. 5, pp. 1794-1804, May 2019.

[267] X. Li, Q. Yang, Z. Lou, and W. Yan, "Deep learning-based module defect analysis for large-scale photovoltaic farms," *IEEE Transactions on Energy Conversion*, vol. 34, no. 1, pp. 520-529, Mar. 2019.

[268] H. S. Hippert, C. E. Pedreira, and R. C. Souza, "Neural networks for short-term load forecasting: a review and evaluation," *IEEE Transactions on Power Systems*, vol. 16, no. 1, pp. 44-55, Feb. 2001.

[269] G. Rolim, J. G. Zurn, "Interpretation of remote backup protection for fault section estimation by a fuzzy expert system," *IEEE PowerTech Conference*, 2003, pp. 312-315.

[270] B. D. Russell and K. Watson, “Power substation automation using a knowledge-based system - justification and preliminary field experiments,” *IEEE Transactions on Power Delivery*, vol. 2, pp. 1090-1097, Oct. 1987.

[271] C. Liu and K. Tomsovic, “An expert system assisting decision-making of reactive power/voltage control,” *IEEE Transactions on Power Systems*, vol. 1, pp. 195-201, Aug. 1986.

[272] K. P. Wong and H. N. Cheung, “Thermal generator scheduling algorithm based on heuristic guided depth-first search,” *IEE Proceedings C - Generation, Transmission and Distribution*, vol. 137, pp. 33-43, Jan. 1990.

[273] K. Nara, T. Satoh, and K. Maeda, “Maintenance scheduling by expert system combined with mathematical programming,” *Proceedings on Expert Systems Application to Power Systems*, 1991, pp. 385-390.

[274] M. Negnevitsky, P. Mandal, and A. K. Srivastava, “Machine learning applications for load, price and wind power prediction in power systems,” *International Conference on Intelligent System Applications to Power Systems*, 2009, pp. 1-6.

[275] L. Duchesne, E. Karangelos, and L. Wehenkel, “Recent developments in machine learning for energy systems reliability management,” *Proceedings of the IEEE*, vol. 108, no. 9, pp. 1656-1676, May 2020.

[276] I. Kalysh, M. Kenzhina, N. Kaiyrbekov, H. S. V. S. Kumar Nunna, A. Dadlani, and S. Doolla, “Machine learning-based service restoration scheme for smart distribution systems with DGs and high priority loads,” *International Conference on Smart Energy Systems and Technologies (SEST)*, 2019, pp. 1-6.

[277] W. Buwei, C. Jianfeng, W. Bo, and F. Shuanglei, “A solar power prediction using support vector machines based on multi-source data fusion,” *International Conference on Power System Technology*, 2018, pp. 4573-4577.

[278] A. Sayghe, O. M. Anubi, and C. Konstantinou, “Adversarial examples on power systems state estimation,” *IEEE Power & Energy Society Innovative Smart Grid Technologies Conference*, 2020, pp. 1-5.

[279] P. Mandal, “Application of an artificial neural network to study the transient stability of an alternator connected to infinite bus,” *1st International Conference on Non-Conventional Energy*, 2014, pp. 224-227.

[280] N. Kumarappan, M. R. Mohan, and S. Murugappan, “ANN approach applied to combined economic and emission dispatch for large-scale system,” *Proceedings of the International Joint Conference on Neural Networks*, 2002, pp. 323-327.

[281] S. Chai, Z. Xu, and Y. Jia, “Conditional density forecast of electricity price based on ensemble ELM and logistic EMOS,” *IEEE Transactions on Smart Grid*, vol. 10, no. 3, pp. 3031-3043, May 2019.

[282] K. Bakshi, The Seven Steps of Machine Learning, Oct. 2017. [Online]. Available: https://www.techleer.com/articles/379-the-seven-steps-of-machine-learning/

[283] M. I. A. Alfarra, “Improving solar power system’s efficiency using artificial neural network,” *M.Sc. Thesis*, Islamic University of Gaza, Palestine, 2018.

[284] L. P. Kaelbling and M. L. Littman, “Reinforcement learning: a survey,” *Journal of Artificial Intelligence Research*, vol. 4, pp. 237-285, May 1996.

[285] S. Angra and S. Ahuja, "Machine learning and its applications: A review," *International Conference on Big Data Analytics and Computational Intelligence (ICBDAC)*, 2017, pp. 57-60.

[286] B. D. Ripley, *Pattern Recognition and Neural Networks*. Cambridge, UK: Cambridge University Press, 1996.

[287] O. Theobald, Machine Learning for Absolute Beginners, 2nd Edition, 2017. [Online]. Available: https://bmansoori.ir/book/Machine%20Learning%20For%20Absolute%20Beginners.pdf

[288] V. M. Santi, "Predicting faults in power grids using machine learning methods," *M.Sc. Thesis*, The Norwegian University of Science and Technology, Norway, 2019.

[289] A. Masri, What Are Overfitting and Underfitting in Machine Learning? Jun. 2019. [Online]. Available: https://towardsdatascience.com/what-are-overfitting-and-underfitting-in-machine-learning-a96b30864690

[290] A. Karami and K. M Galougahi, "Improvement in power system transient stability by using STATCOM and neural networks," *Electrical Engineering*, vol. 101, pp. 19–33, 2019.

[291] J. Schneider, Cross Validation, 1997. [Online]. Available: https://www.cs.cmu.edu/~schneide/tut5/node42.html

[292] D. You, K. Wang, L. Ye, J. Wu, and R. Huang, "Transient stability assessment of power system using support vector machine with generator combinatorial trajectories inputs," *International Journal of Electrical Power and Energy Systems*, vol. 44, no. 1, pp. 318-325, Jan. 2013.

[293] S. Wu, 3 Best metrics to evaluate Regression Model? [Online]. Available: https://towardsdatascience.com/what-are-the-best-metrics-to-evaluate-your-regression-model-418ca481755b

[294] M. Sunasra, Performance metrics for classification problems in machine learning, Nov. 2017. [Online]. Available: https://medium.com/@MohammedS/performance-metrics-for-classification-problems-in-machine-learning-part-i-b085d432082b

[295] N. I. A Wahab, A. Mohamed, and A. Hussain, "Fast transient stability assessment of large power system using probabilistic neural network with feature reduction techniques," *Expert Systems with Applications*, vol. 38, no. 9, pp. 11112-11119, Sep. 2011.

[296] C. Sammut and G. I. Webb, Encyclopedia of Machine Learning. New York, NY, USA: Springer, 2011.

[297] Using the ROC curve to analyze a classification model. [Online]. Available: http://www.math.utah.edu/~gamez/files/ROC-Curves.pdf

[298] S. Haykin, *Neural Networks and Learning Machines, 3rd Ed.*, New Jersey: Pearson Education, 2009.

[299] M. Luzi, "Design and implementation of machine learning techniques for modeling and managing battery energy storage systems," *Ph.D. Dissertation*, The University of Rome, Italy, 2019.

[300] G. P. Zhang, "Neural networks for classification: A survey," *IEEE Transactions on Systems, Man, and Cybernetics, Part C (Applications and Reviews)*, vol. 30, no. 4, pp. 451–462, 2000.

[301] G. Aoki and Y. Sakakibara, “Convolutional neural networks for classification of alignments of non-coding RNA sequences,” *Bioinformatics*, vol. 34, no. 13, pp. i237–i244, 2018.

[302] A. Krizhevsky, I. Sutskever, and G. E. Hinton, “Imagenet classification with deep convolutional neural networks,” *Communications of the ACM*, vol. 60, no. 6, May 2017.

[303] S. Yang, T. Ting, K. Man, and S. U. Guan, “Investigation of neural networks for function approximation,” *Procedia Computer Science*, vol. 17, pp. 586–594, 2013.

[304] K. Hornik, M. Stinchcombe, and H. White, “Multilayer feedforward networks are universal approximators,” *Neural Networks*, vol. 2, no. 5, pp. 359–366, 1989.

[305] T. Thien and T. Cong, “Approximation of functions and their derivatives: a neural network implementation with applications,” *Applied Mathematical Modelling*, vol. 23, no. 9, pp. 687–704, 1999.

[306] K. Strachnyi, *Brief History of Neural Networks*, Jan. 2019. [Online]. Available: https://medium.com/analytics-vidhya/brief-history-of-neural-networks-44c2bf72eec

[307] Artificial Neural Networks Technology. [Online]. Available: http://www2.psych.utoronto.ca/users/reingold/courses/ai/cache/neural4.html

[308] History of artificial neural networks. [Online]. Available: https://en.wikipedia.org/wiki/History_of_artificial_neural_networks

[309] J. Thakur, Components of ANN, Apr. 2020. [Online]. Available: https://www.theikigailab.com/post/components-of-ann-jitender-thakur

[310] Artificial neural network. [Online]. Available: https://en.wikipedia.org/wiki/Artificial_neural_network

[311] R. Yousefian, and S. Kamalasadan, A review of neural network based machine learning approaches for rotor angle stability control. [Online]. Available: https://www.researchgate.net/publication/312091714_A_Review_of_Neural_Network_Based_Machine_Learning_Approaches_for_Rotor_Angle_Stability_Control

[312] Types of Neural Networks and Definition of Neural Network, Apr. 2020. [Online]. Available: https://www.mygreatlearning.com/blog/types-of-neural-networks/

[313] P. V. Gomes and J. T. Saraiva, "A two-stage strategy for security-constrained AC dynamic transmission expansion planning," *Electric Power Systems Research*, vol. 180, pp. 1-10, Mar. 2020.

[314] M. Amer and T. Maul, "A review of modularization techniques in artificial neural networks," *Artificial Intelligence Review*, vol. 52, pp. 527–561, 2019.

[315] J. Schmidhuber, "Deep learning in neural networks: An overview," *Neural Networks*, vol. 61, pp. 85-117, Jan. 2015.

[316] A. Karami, "Transient stability assessment of power systems described with detailed models using neural networks," *International Journal of Electrical Power & Energy Systems*, vol. 45, no. 1, pp. 279-292, Feb. 2013.

[317] R. Akeel, "Predicting weather-related transmission line failures using machine learning," *M.Sc. Thesis*, Reykjavík University, Iceland, 2019.

[318] F. Li and S. Gao, "Character recognition system based on back-propagation neural network," *International Conference on Machine Vision and Human-machine Interface*, 2010, pp. 393-396.

[319] D. J. Sobajic and Y. Pao, “Artificial neural-net based dynamic security assessment for electric power systems,” *IEEE Transactions on Power Systems*, vol. 4, no. 1, pp. 220-228, Feb. 1989.

[320] W. S. McCulloch and W. Pitts, “A logical calculus of the ideas immanent in nervous activity,” *The Bulletin of Mathematical Biophysics*, vol. 5, no. 4, pp. 115–133, 1943.

[321] F. Rosenblatt, “The perceptron: A probabilistic model for information storage and organization in the brain,” *Psychological Review*, vol. 65, no. 6, pp. 386, 1958.

[322] P. K. Olulope, K. A. Folly, S. Chowdhury, and S. P. Chowdhury, “Transient stability assessment using artificial neural network considering fault location,” *1st International Conference on Energy, Power, and Control*, 2010, pp. 67-72.

[323] D. Karlsson, “Wind turbine performance monitoring using artificial neural networks with a multi-dimensional data filtering approach,” *M.Sc. Thesis,* Chalmers University Of Technology, Sweden, 2015.

[324] S. Sharma, Activation Functions in Neural Networks, Sep. 2017. [Online]. Available: https://towardsdatascience.com/activation-functions-neural-networks-1cbd9f8d91d6

[325] 7 Types of Neural Network Activation Functions: How to Choose? [Online]. Available: https://missinglink.ai/guides/neural-network-concepts/7-types-neural-network-activation-functions-right/

[326] T. S. Dillon, and D. Niebur, *Neural Networks Applications in Power* Systems, Leicestershire, UK: CRL Publishing, 1996.

[327] M. Djukanovic, D. J. Sobajic, and Y. Pao, “Neural-net based unstable machine identification using individual energy functions,” *International Journal of Electrical Power & Energy Systems*, vol. 13, no. 5, pp. 255-262, Oct. 1991.

[328] Y. Pao and D. J. Sobajic, “Combined use of unsupervised and supervised learning for dynamic security assessment,” *IEEE Transactions on Power Systems*, vol. 7, no. 2, pp. 878-884, May 1992.

[329] Y. Lin, “Explaining critical clearing time with rules extracted from a multilayer perceptron artificial neural network,” *International Journal of Electrical Power & Energy Systems*, vol. 32, no. 8, pp. 873-878, Oct. 2010.

[330] Q. Zhou, J. Davidson, and A. A. Fouad, “Application of artificial neural networks in power system security and vulnerability assessment,” *IEEE Transactions on Power Systems*, vol. 9, no. 1, pp. 525-532, Feb. 1994.

[331] F. Aboytes and R. Ramirez, “Transient stability assessment in longitudinal power systems using artificial neural networks,” *IEEE Transactions on Power Systems*, vol. 11, no. 4, pp. 2003-2010, Nov. 1996.

[332] A. Bahbah and A. Girgis, “New method for generators’ angles and angular velocities prediction for transient stability assessment of multi-machine power systems using recurrent artificial neural network,” *Proceedings of IEEE Power Energy Society General Meeting*, 2004, pp. 117.

[333] S. Krishna and K. R. Padiyar, “Transient stability assessment using artificial neural networks,” *Proceedings of IEEE International Conference on Industrial Technology*, 2000, pp. 627-632.

[334] Y. Liangzhong, N. Yixin, and Z. Buoming, “Estimation of transient stability limits using artificial neural network,” *Proceedings of TENCON IEEE Region 10 International Conference on Computers, Communications and Automation*, 1993, pp. 87-90.

[335] A. M. Sharaf and T. T. Lie, “Neural network pattern classifications of transient stability and loss of excitation for synchronous generators,” *Proceedings of 1994 IEEE International Conference on Neural Networks*, 1994, pp. 2916-2921.

[336] M. Su, C. W. Liu, and S. Tsay, “Neural-network-based fuzzy model and its application to transient stability prediction in power systems,” *IEEE Transactions on Systems, Man, and Cybernetics, Part C (Applications and Reviews)*, vol. 29, no. 1, pp. 149-157, Feb. 1999.

[337] I. Kulkarni and S. Dash, “Development of neural network based HVDC controller for transient stability enhancement of AC/DC system,” *IEEE Students’ Conference on Electrical, Electronics and Computer Science*, 2012, pp. 1-4.

[338] S. Muknahallipatna and B. H. Chowdhury, “Input dimension reduction in neural network training-case study in transient stability assessment of large systems,” *Proceedings of International Conference on Intelligent System Application to Power Systems*, 1996, pp. 50-54.

[339] S. Wei, K. Nakamura, M. Sone, and H. Fujita, “Neural network based power system transient stability criterion using DSP-PC system,” *Proceedings of the Second International Forum on Applications of Neural Networks to Power Systems*, 1993, pp. 136-141.

[340] H. Sawhney and B. Jeyasurya, “On-line transient stability assessment using artificial neural network,” *Large Engineering Systems Conference on Power Engineering (IEEE Cat. No.04EX819)*, 2004, pp. 76-80.

[341] K. K. Sanyal, “Transient stability assessment using artificial neural network,” *IEEE International Conference on Electric Utility Deregulation, Restructuring and Power Technologies*, 2004, pp. 633-637.

[342] C. R. Minussi and M. G. Silveira, “Electric power systems transient stability analysis by neural networks,” *38th Midwest Symposium on Circuits and Systems*, 1995, pp. 1305-1308.

[343] L. Xianshu and J. Ma, “Using artificial neural network to assess transient stability of power systems,” *Proceedings of TENCON IEEE Region 10 International Conference on Computers, Communications and Automation*, 1993, pp. 91-94.

[344] S. Mishra, “Neural-network-based adaptive UPFC for improving transient stability performance of power system,” *IEEE Transactions on Neural Networks*, vol. 17, no. 2, pp. 461-470, Mar. 2006.

[345] D. M. Eltigani, K. Ramadan, and E. Zakaria, “Implementation of transient stability assessment using artificial neural networks,” *International Conference on Computing, Electrical and Electronic Engineering*, 2013, pp. 659-662.

[346] A. M. A. Haidar, M. W. Mustafa, F. A. F. Ibrahim, and I. A. Ahmed, “Transient stability evaluation of electrical power system using generalized regression neural networks,” *Applied Soft Computing*, vol. 1, no. 4, pp. 3558-3570, Jun. 2011.

[347] H. Sawhney and B. Jeyasurya, “A feed-forward artificial neural network with enhanced feature selection for power system transient stability assessment,” *Electric Power Systems Research*, vol. 76, no. 12, pp. 1047-1054, Aug. 2006.

[348] K. Omata and K. Tanomura, “Transient stability evaluation using an artificial neural network (power systems),” *Proceedings of the Second International Forum on Applications of Neural Networks to Power Systems*, 1993, pp. 130-135.

[349] E. M. Voumvoulakis, A. E. Gavoyiannis and N. D. Hatziargyriou, “Application of machine learning on power system dynamic security assessment,” *International Conference on Intelligent Systems Applications to Power Systems*, 2007, pp. 1-6.

[350] L. Chunyan, T. Biqiang, and C. Xiangyi, “On-line transient stability assessment using hybrid artificial neural network,” *2nd IEEE Conference on Industrial Electronics and Applications*, 2007, pp. 342-346.

[351] N. Kandil, S. Georges, and M. Saad, “Rapid power system transient stability limit search using signal energy and neural networks,” *IEEE International Symposium on Industrial Electronics*, 2006, pp. 1658-1661.

[352] M. Mahdi and V. M. I. Genc, “Artificial neural network based algorithm for early prediction of transient stability using wide area measurements,” *5th International Istanbul Smart Grid and Cities Congress and Fair*, 2017, pp. 17-21.

[353] M. Li and P. Jiang, “Artificial neural network classifier of transient stability based on time-domain simulation,” *2nd IEEE Advanced Information Management, Communicates, Electronic and Automation Control Conference (IMCEC)*, 2018, pp. 685-689.

[354] S. Jafarzadeh and V. M. I. Genc, “Probabilistic dynamic security assessment of large power systems using machine learning algorithms,” *Turkish Journal of Electrical Engineering & Computer Sciences*, vol. 26, no. 3, pp. 1479–1490, Jan. 2018.

[355] A. Sabo and N. I. A. Wahab, “Rotor angle transient stability methodologies of power systems: a comparison,” *IEEE Student Conference on Research and Development*, 2019, pp. 1-6.

[356] W. Li and J. Zhou, “Probabilistic reliability assessment of power system operations,” *Electric Power Components and Systems*, vol. 36, no. 10, pp. 1102-1114, Sep. 2008.

[357] J. L. Cremer and G. Strbac, “A machine-learning based probabilistic perspective on dynamic security assessment,” *International Journal of Electrical Power & Energy Systems*, vol. 128, pp. 1-15, Jun. 2021.

[358] H. Yuan, J. Tan, and Y. Zhang, “Machine learning-based security assessment and control for bulk electric system operation,” NREL, Feb. 2020. [Online]. Available: https://www.nrel.gov/docs/fy21osti/76089.pdf

[359] T. Zhang, M. Sun, J. L. Cremer, N. Zhang, G. Strbac, and C. Kang, “A confidence-aware machine learning framework for dynamic security assessment,” *IEEE Transactions on Power Systems*, pp. 1-14, Feb. 2021.

[360] V. Vapnik, *Statistical Learning Theory*. New York, NY, USA: John Wiley & Sons, 1998.

[361] V. Vapnik, *The Nature of Statistical Learning Theory*. New York, NY, USA: Springer-Verlag, 1995.

[362] B. Scholköpf, C. Burges, and A. Smola, *Advances in Kernel Methods—Support Vector Learning*. Cambridge, MA, USA: MIT Press, 1999.

[363] S. Kalyani and K. S. Swarup, “Classification and assessment of power system security using multiclass SVM,” *IEEE Transactions on Systems, Man, and Cybernetics, Part C (Applications and Reviews)*, vol. 41, no. 5, pp. 753-758, Sep. 2011.

[364] H. Eristi, "Wavelet-based feature extraction and selection for classification of power system disturbances using support vector machines," *Electric Power Systems Research*, vol. 80, no. 7, pp. 743-752, Jul. 2010.

[365] P. Janik and T. Lobos, "Automated classification of power-quality disturbances using SVM and RBF networks," *IEEE Transactions on Power Delivery*, vol. 21, no. 3, pp. 1663-1669, Jul. 2006.

[366] D. You, K. Wang, L. Ye, J. Wu, and R. Huang, "Transient stability assessment of power system using support vector machine with generator combinatorial trajectories inputs", *International Journal of Electrical Power & Energy Systems*, vol. 44, no. 1, pp. 318-325, Jan. 2013.

[367] E. Osuna, R. Freund, and F. Girosi, "Training support vector machines: an application to face detection," *Proceedings of IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, 1997, pp. 1-8.

[368] T. Joachims, "Making large-scale SVM learning practical," in *Advances in Kernel Methods - Support Vector Learning*. Cambridge, MA, USA: MIT Press, 1998.

[369] N. Cristianini and J. Shawe-Taylor, *An Introduction to Support Vector Machines and Other Kernel-Based Learning Methods*. Cambridge, UK: Cambridge University Press, 2000.

[370] L. S. Moulin, A. P. A. da Silva, M. A. El-Sharkawi, and R. J. Marks, "Support vector machines for transient stability analysis of large-scale power systems," *IEEE Transactions on Power Systems*, vol. 19, no. 2, pp. 818-825, May 2004.

[371] C. Savas and F. Dovis, “The impact of different kernel functions on the performance of scintillation detection based on support vector machines,” *Sensors*, vol. 19, no. 23, pp. 1-16, Nov. 2019.

[372] L. S. Moulin, A. P. A. da Silva, M. A. El-Sharkawi, and R. J. Marks, “Neural networks and support vector machines applied to power systems transient stability analysis,” *International Journal of Engineering Intelligent Systems for Electrical Engineering and Communications*, vol. 9, no. 4, pp. 205-211, Nov. 2001.

[373] L. S. Moulin, A. P. A. da Silva, M. A. El-Sharkawi, and R. J. Marks, “Support vector and multilayer perceptron neural networks applied to power systems transient stability analysis with input dimensionality reduction,” *IEEE Power Engineering Society Summer Meeting*, 2002, pp. 1308-1313.

[374] P. Vadapalli, Support Vector Machines: Types of SVM, Dec. 2020. [Online]. Available: https://www.upgrad.com/blog/support-vector-machines/

[375] Support Vector Machine Algorithm. [Online]. Available https://www.javatpoint.com/machine-learning-support-vector-machine-algorithm

[376] T. M. Cover, “Geometrical and statistical properties of systems of linear inequalities with applications in pattern recognition,” *IEEE Transactions on Electronic Computers*, vol. EC-14, no. 3, pp. 326-334, Jun. 1965.

[377] C. M. Bishop. *Pattern recognition and machine learning*. Information science and statistics. New York: Springer, 2006.

[378] D. Chicco, “Ten quick tips for machine learning in computational biology,” *BioData Mining*, vol. 10, no. 35, pp. 1-17, Dec. 2017.

[379] J. Bergstra and Y. Bengio, “Random search for hyper-parameter optimization,” *Journal of Machine Learning Research*, vol. 13, pp. 281–305, Feb 2012.

[380] J. Snoek and R. Adams, “Practical Bayesian optimization of machine learning algorithms,” *Advances in Neural Information Processing Systems*, vol. 2, pp. 2951–2959, Dec. 2012.

[381] A. D. Rajapakse, F. Gomez, K. Nanayakkara, P. A. Crossley, and V. V. Terzija, “Rotor angle instability prediction using post-disturbance voltage trajectories,” *IEEE Transactions on Power Systems*, vol. 25, no. 2, pp. 947-956, May 2010.

[382] B. P. Soni, A. Saxena, V. Gupta, and S. L. Surana, “Assessment of transient stability through coherent machine identification by using least-square support vector machine”, *Modelling and Simulation in Engineering*, vol. 2018, pp. 1-18, May 2018.

[383] F. R. Gomez, A. D. Rajapakse, U. D. Annakkage, and I. T. Fernando, “Support vector machine-based algorithm for post-fault transient stability status prediction using synchronized measurements,” *IEEE Transactions on Power Systems*, vol. 26, no. 3, pp. 1474-1483, Aug. 2011.

[384] P. Pavani and S. N. Singh, “Support vector machine based transient stability identification in distribution system with distributed generation,” *Electric Power Components and Systems*, vol. 44, no. 1, pp. 60-71, Nov. 2015.

[385] D. Yuanhang, C. Lei, Z. Weiling and M. Yong, “Multi-support vector machine power system transient stability assessment based on relief algorithm,” *IEEE PES Asia-Pacific Power and Energy Engineering Conference (APPEEC)*, 2015, pp. 1-5.

[386] M. Shahriyari, H. Khoshkhoo, A. Pouryekta, and V. K. Ramachandaramurthy, “Fast prediction of angle stability using support vector machine and fault duration data,” *IEEE*

International Conference on Automatic Control and Intelligent Systems, 2019, pp. 258-263.

[387] N. Lin and J. Ling, "Application of time series forecasting algorithm via support vector machines to power system wide-area stability prediction," *IEEE/PES Transmission & Distribution Conference & Exposition: Asia and Pacific*, 2005, pp. 1-6.

[388] J. Liu and X. Wang, "A conservative prediction model of power system transient stability," *IEEE Power & Energy Society General Meeting (PESGM)*, 2018, pp. 1-5.

[389] A. E. Gavoyiannis, D. G. Vogiatzis, D. R. Georgiadis, and N. D. Hatziargyriou, "Combined support vector classifiers using fuzzy clustering for dynamic security assessment," *Power Engineering Society Summer Meeting*, 2001, pp. 1281-1286.

[390] S. Ye, X. Li, X. Wang, and Q. Qian, "Power system transient stability assessment based on Adaboost and support vector machines," *Asia-Pacific Power and Energy Engineering Conference*, 2012, pp. 1-4.

[391] N. I. A. Wahab, A. Mohamed, and A. Hussain, "Transient stability assessment of a power system using PNN and LS-SVM methods," *Journal of Applied Sciences*, vol. 7, no. 21, pp. 3208-3216, 2007.

[392] B. D. A. Selvi and N. Kamaraj, "Investigation of power system transient stability using clustering based support vector machines and preventive control by rescheduling generators," *IET-UK International Conference on Information and Communication Technology in Electrical Sciences*, 2007, pp. 137-142.

[393] N. G. Baltas, P. Mazidi, J. Ma, F. Fernandez, and P. Rodriguez, "A comparative analysis of decision trees, support vector machines and artificial neural networks for on-

line transient stability assessment," *International Conference on Smart Energy Systems and Technologies (SEST)*, 2018, pp. 1-6.

[394] M. Arefi and B. Chowdhury, "Ensemble adaptive neuro fuzzy support vector machine for prediction of transient stability," *North American Power Symposium (NAPS)*, 2017, pp. 1-6.

[395] E. A. Frimpong, P. Y. Okyere, and J. Asumadu, "Prediction of transient stability status using Walsh-Hadamard transform and support vector machine," *IEEE PES PowerAfrica*, 2017, pp. 301-306.

[396] B. D. A. Selvi and N. Kamaraj, "Support vector regression machine with enhanced feature selection for transient stability evaluation," *Asia-Pacific Power and Energy Engineering Conference*, 2009, pp. 1-5.

CHAPTER 3

PROBLEM FORMULATION AND PROPOSED APPROACHES

This chapter describes the problem formulation and proposed approaches. Firstly, a brief review of the softwares used is presented. Moreover, various assumptions are listed. In the remaining part, mathematical formulations, associated with the proposed approaches, are elaborated.

3.1 Brief Review of Softwares Used

This section will briefly review the two softwares used in this research work: DIgSILENT PowerFactory and MATLAB. In this research, DIgSILENT PowerFactory was primarily used for probabilistic transient stability (PTS) simulations (for obtaining training data for machine learning (ML) algorithms), whereas, MATLAB was employed for applying these ML algorithms, and consequently, evaluating their performance for regression and classification tasks.

DIgSILENT PowerFactory

DIgSILENT (Digital Simulation of Electrical Networks) PowerFactory, is principally a computer-aided engineering tool, for analyzing transmission, distribution, and industrial electrical power systems. It has been designed as an advanced integrated and interactive software package, dedicated to electrical power system, and control analysis, to accomplish the chief objectives of planning and operation optimization. The software was established by qualified engineers and programmers, with many years of experience, in both electrical

power system analysis and computer programming. The accuracy and validity of results obtained with PowerFactory has been confirmed in many implementations, by organizations involved in the planning and operation of power systems, throughout the world. To address power system analysis requirements of the users, PowerFactory was designed as an integrated engineering tool to provide an inclusive suite of power system analysis functions, within a single executable program [1]. It has applications in various fields of power system, including stability analysis, load flow analysis, short circuit studies, reliability assessment, and protection analysis. It also has an embedded coding tool, commonly known as the DIgSILENT Programming Language (DPL), which serves the purpose of offering an interface, for automating tasks in the PowerFactory program. The presence of DPL adds a new dimension to the DIgSILENT PowerFactory program by allowing the creation of new calculation functions. Such user-defined calculation commands can generally be used in all significant areas of power system analysis.

MATLAB

MATLAB is a proprietary multi-paradigm programming language and numeric computing environment, developed by MathWorks. It allows matrix manipulations, plotting of functions and data, implementation of algorithms, creation of user interfaces, and interfacing with programs written in other languages. Although, MATLAB is primarily intended for numeric computing; an additional package, Simulink, adds graphical multi-domain simulation and model-based design for dynamic and embedded systems. As of 2020, MATLAB has more than 4 million users worldwide. MATLAB users come from various backgrounds of engineering, science, and economics [2].

MATLAB has applications in various significant fields, including signal processing,

wireless communications, control systems, power systems, mathematical optimization, and code generation. MATLAB supports various ML tasks, such as regression and classification, through the availability of numerous toolboxes and built-in applications. The neural network toolbox highlights the use of neural network paradigms that build up to— or are themselves used in— engineering, financial, and other practical applications [3]. The classification learner application trains models to classify data. Using this app, supervised ML can easily be explored using various classifiers. Several tasks, including data exploration, feature selection, specification of validation schemes, training of models, and result assessment, can be effortlessly performed [4].

3.2 Assumptions

This section lists the assumptions made in this research work. They are described as follows:

1. The test transmission network (IEEE 14-bus) is assumed to be rich in wind resources, wherever these sources are connected.
2. A five-year decision-making period is considered for circuit breakers (CBs). This is the economic life of CBs.
3. All lines are equally probable (uniform distribution) to faults [5-7].
4. Fault can occur at any point of the line with equal probability [5].
5. Any kind (*LG*, *LL*, *LLG*, or *LLL*) of shunt fault (short-circuit fault) fault can occur based on the discrete probability mass function (PMF) defined [8].
6. Fault type and fault location are independent, i.e., any fault can occur on any point along any line of the network [5].

7. Fault clearing time (FCT) is assumed to follow a normal (Gaussian) probability density function (PDF) [6, 9-10].
8. Every bus load is assumed to follow a normal PDF [11-12].
9. Initial operating conditions are based on load flow [6, 9].
10. Upon fault detection, a tripping signal is sent to both ends of the line such that both breakers are tripped at the same time instance [13].
11. The CBs are assumed to be 100% reliable (i.e., stuck breaker condition is not considered) [7-8].
12. Although, for the CB replacement, and consequent ML applications, the pre-fault network topology changes are not considered, however, the impact of network topology change on P_{SYS} and R_A is determined.
13. Risk-neutrality is used for decision-making, i.e., the average (or mean) behavior of the specified performance metric (which, in this research, is the benefit-cost ratio (BCR)) is a good measure for making decision [14-15].
14. Type 3 wind generators (DFIGs) were employed to study the impact of wind generation.
15. DFIGs are equipped with terminal voltage control capability [16-18].
16. In CBA formulation, generation is always available to fulfil the load demand; thus, cost of load shedding is not incorporated.
17. Levenberg-Marquardt backpropagation algorithm was used to train the artificial neural network (ANN) [19-20].
18. Holdout cross-validation was used for ANN (70%, 15%, and 15% for training, testing, and validation, respectively) [20].

19. Logsig and tansig activation functions were used for hidden and output layers of ANN, respectively [20-21].
20. For support vector machine (SVM), Gaussian radial basis function (RBF) kernel function was used [22-23].
21. For SVM hyperparameter optimization, Bayesian optimization was used [24].
22. *K*-fold cross validation was used for SVM [22].

3.3 Methodology for Proposed Approaches

This section discusses the methodology for the two key proposed approaches: (1) ANN prediction (regression) for PTS enhancement decision making using CBs, and (2) PTS classification using ANN and SVM and consequently, comparing the resulting performance metrics. A summarized flowchart of the big picture, incorporating main research objectives (O1-O9) [as outlined in Chapter 1], and proposed approaches is shown in Figure 3.1. Moreover, specifically, the generic framework for both ML approaches is shown in Figure 3.2.

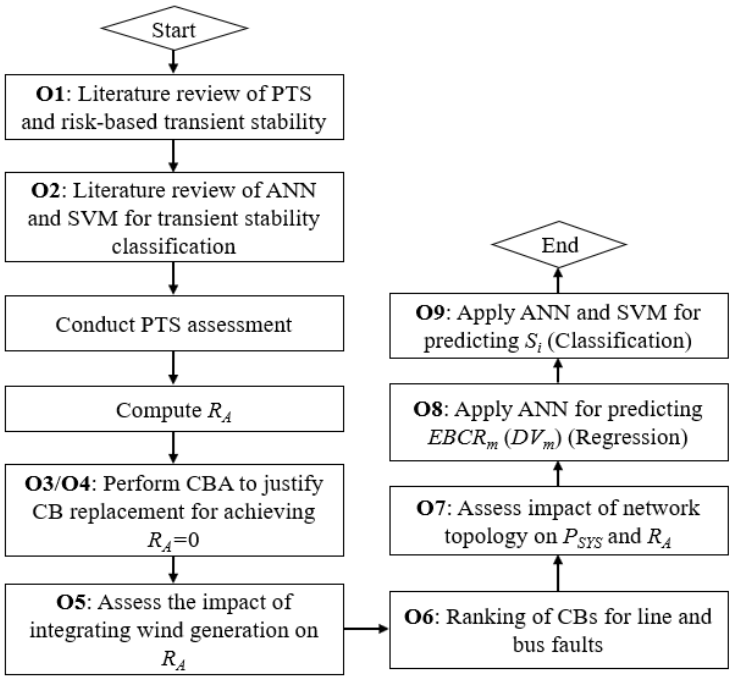

Figure 3.1. Summarized flowchart linking main objectives and proposed approaches (the big picture)

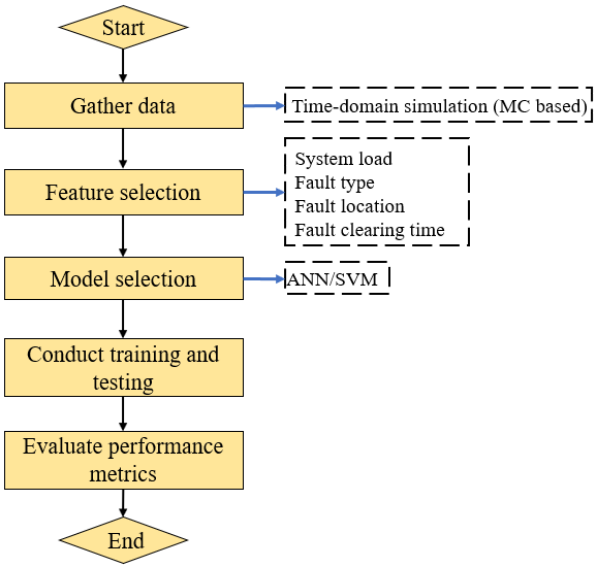

Figure 3.2. Framework for the proposed ML approaches (O8/O9)

3.3.1 ANN Prediction for PTS Enhancement Decision Making using CBs

The methodology for the first proposed approach is described in Figure 3.3. The symbol *i* indicates the sample number for the Monte Carlo (MC) simulation. A fixed system topology is used, i.e., it is assumed that before the fault, all system components (generators, transformers, etc.) are operating normally and there is no failure. In the next step, value for system load is selected. This is selected based on the normal PDF (defined for each bus load). The relevant details are discussed in Chapter 2, Section 2.11. Load flow is then run to acquire the current state of the system. After this step, faulted line, fault type, fault location, and FCT are selected, based on PDFs defined (Chapter 2, Section 2.11). The fault is created at time t=1 *s*. For each MC sample, time-domain stability simulation is run for 10 *s* to determine the outcome (transiently stable or unstable). This is determined based on maximum rotor angle difference, $\delta_{\max}$ (unstable if $\delta_{\max}$ >360). If the sample is transiently unstable, R_i (transient instability risk for i^{th} sample) is evaluated (which is consequently used for evaluation of the average risk index, R_A). The MC simulation (MCS) is stopped after N simulations (the procedure to determine N is described in Chapter 4, Section 4.2), and consequently, P_{SYS} (probability of system being transiently unstable) and, value of R_A is determined. The value of P_{SYS} is computed as the ratio of transiently unstable samples to total MC samples.

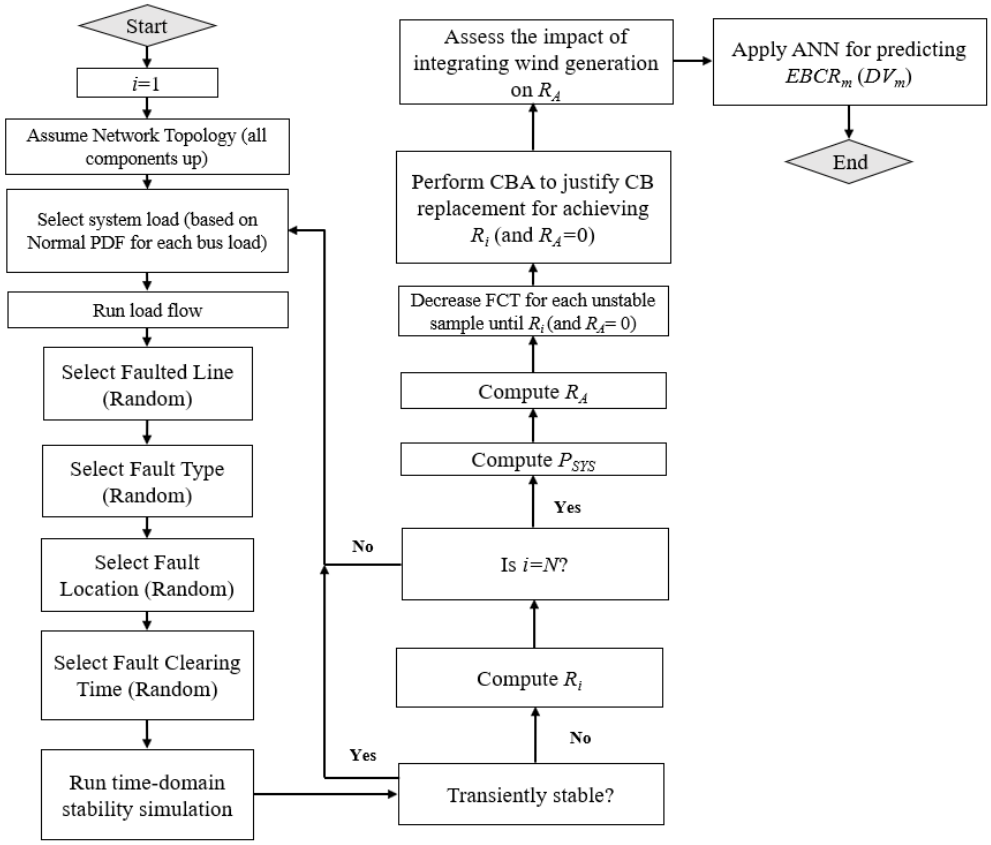

Figure 3.3. Methodology for PTS enhancement using CBs and ANN application for regression (O3, O4, O8)

In the next step, for each unstable MC sample, the FCT is reduced in steps (of 0.01 *s*) to ensure the value of R_i (and R_A) is zero. The value of R_A =0 is termed as system marginal transient stable risk. In short, the risk for the fault which is cleared at CCT is termed as marginal transient stable risk. If the system is transiently unstable for a FCT; then, the risk evaluated at the first instance of transient stability (determined by decreasing that FCT in steps) is termed as marginal transient stable risk. This is illustrated using Figure 3.4. In the next step, to achieve this risk reduction and justifying the replacement of CBs, cost benefit analysis (CBA) is conducted. In the next step, wind generation is integrated and its impact

on the average risk index, R_A, is assessed. In the final step, ANN-based ML algorithm is used to predict the expected value of BCR of m^{th} line (denoted by $EBCR_m$).

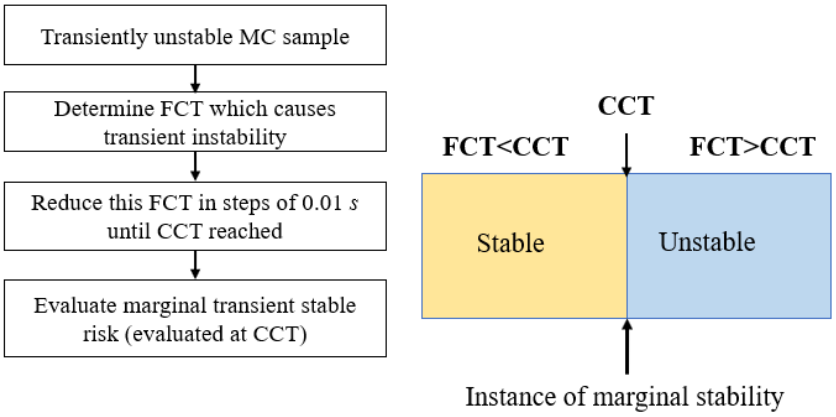

Figure 3.4. Concept of marginal stability (O3)

3.3.2 Ranking of Circuit Breakers

This subsection elaborates the approach to rank individual lines' CBs based on the value of R_A, for both line and bus faults. This will enable the planners to be extra cautious towards the critical lines/buses and corresponding CBs, for improved decision making. The approach is elaborated for line and bus faults, in Figure 3.5 and Figure 3.6, respectively. For each line/bus, 2401 MC samples (using Cochran's formula, and assuming a 95% confidence level and 2% margin of error) were used to compute the value of R_A, using time-domain simulation, considering PDFs of various uncertainties (system load, fault type, FCT, etc.). Consequently, based on the value of R_A obtained for each line/bus, the CBs were ranked. Similarly, the approach is elaborated using only three phase bus faults in Figure 3.7.

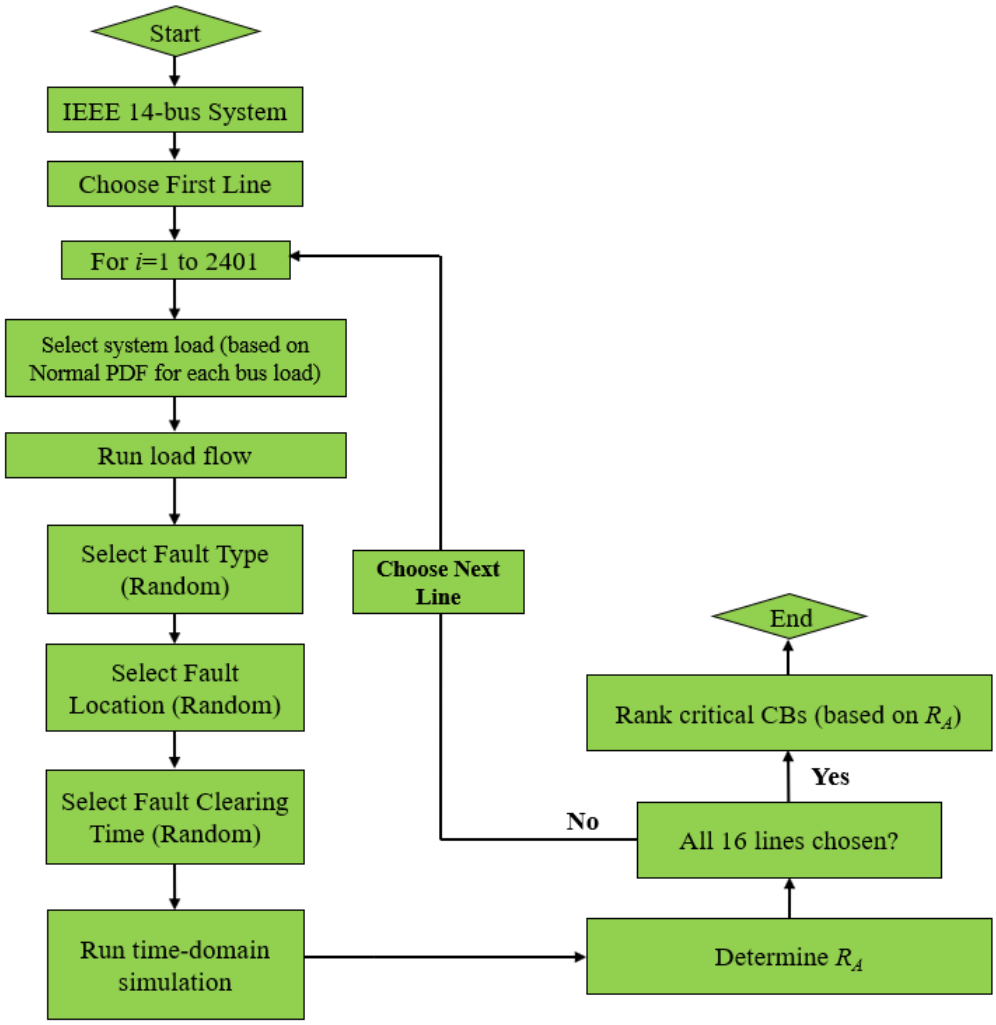

Figure 3.5. Methodology for CB ranking for line faults (O6)

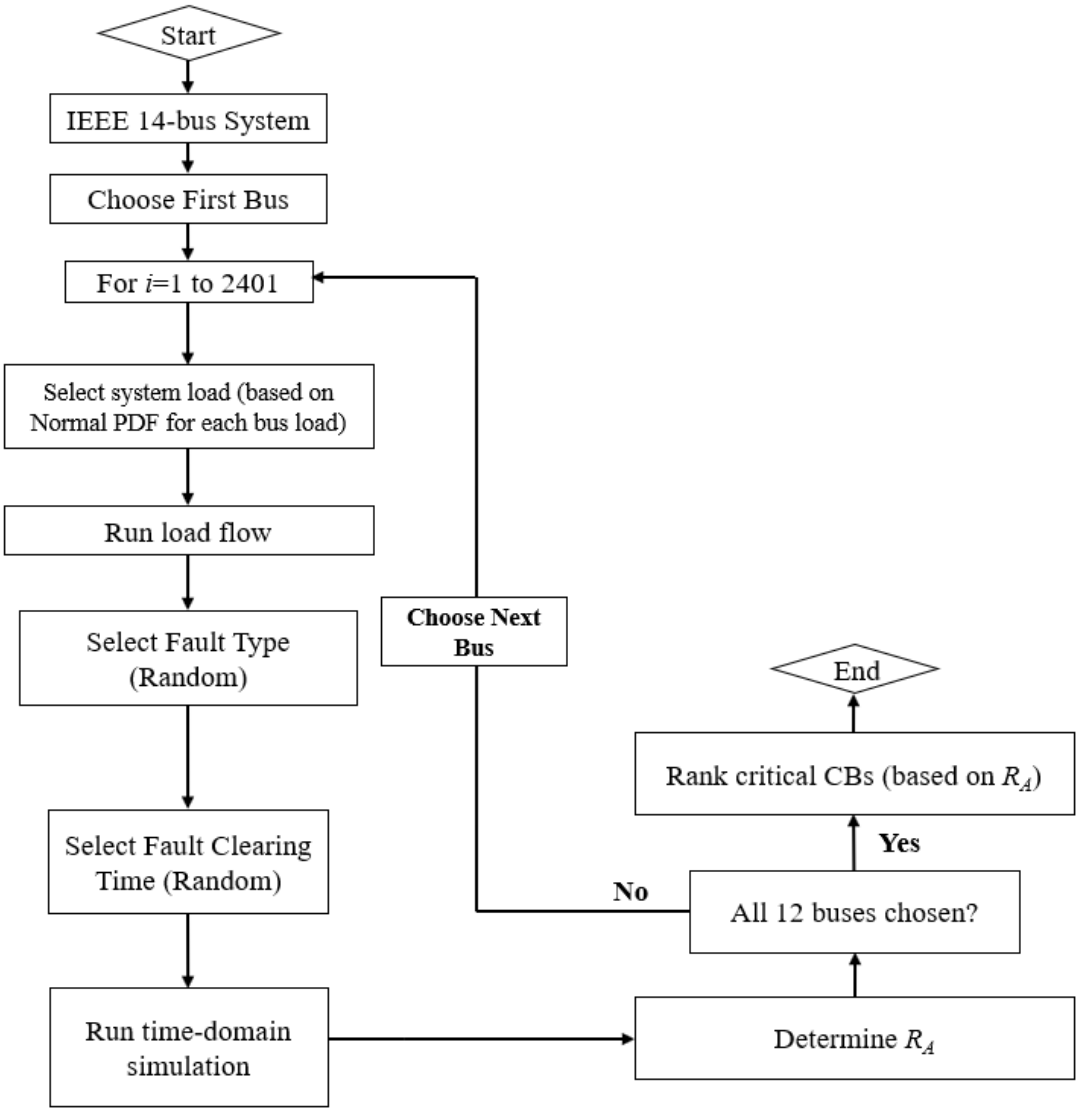

Figure 3.6. Methodology for CB ranking for bus faults (O6)

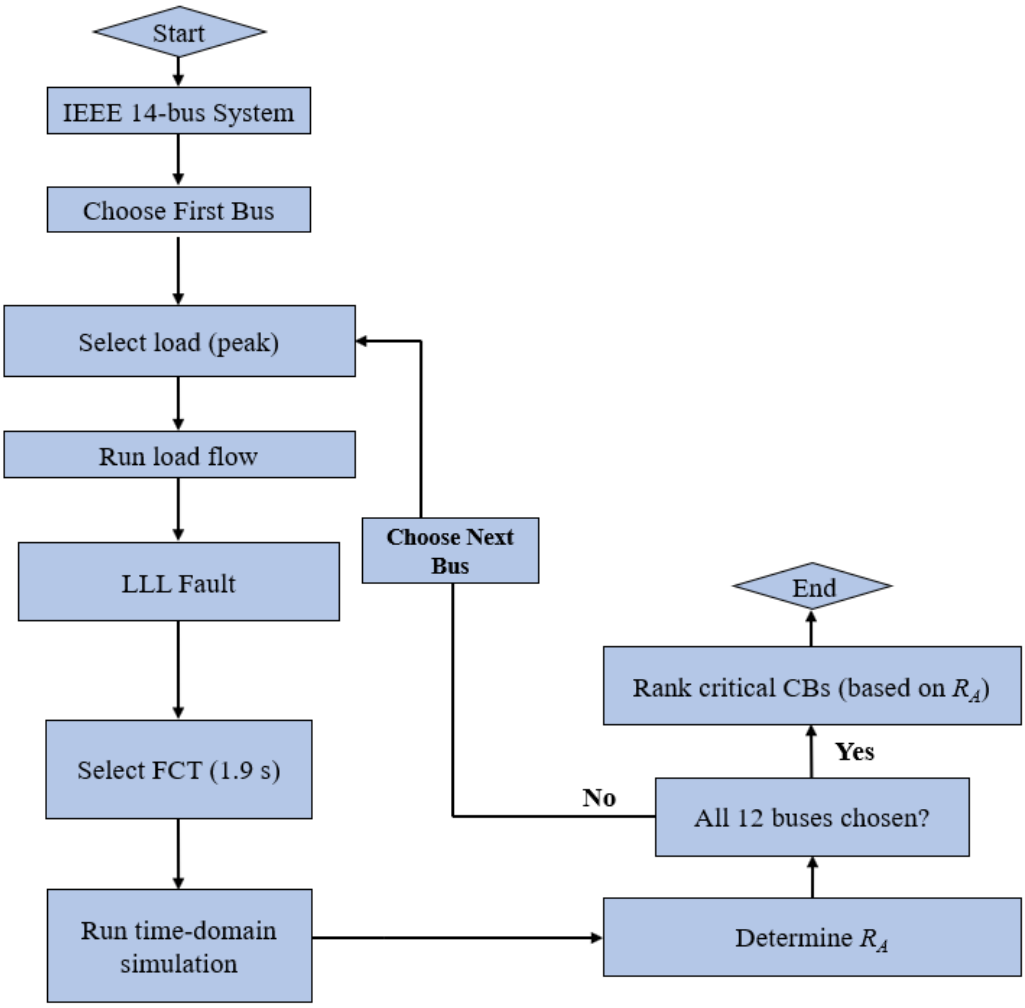

Figure 3.7. Methodology for CB ranking for LLL bus faults (O6)

3.3.3 Impact of Network Topology on P_{SYS} and R_A

Figure 3.8 elaborates the procedure to study the impact of network topology on P_{SYS} and R_A. As illustrated, in each MC sample, to simulate the variation of network topology, a single line is randomly outaged before selecting the random variables of faulted line (different from the random line already outaged), fault type, FCT, etc. The value of P_{SYS} and R_A is determined at the end of the MC simulation.

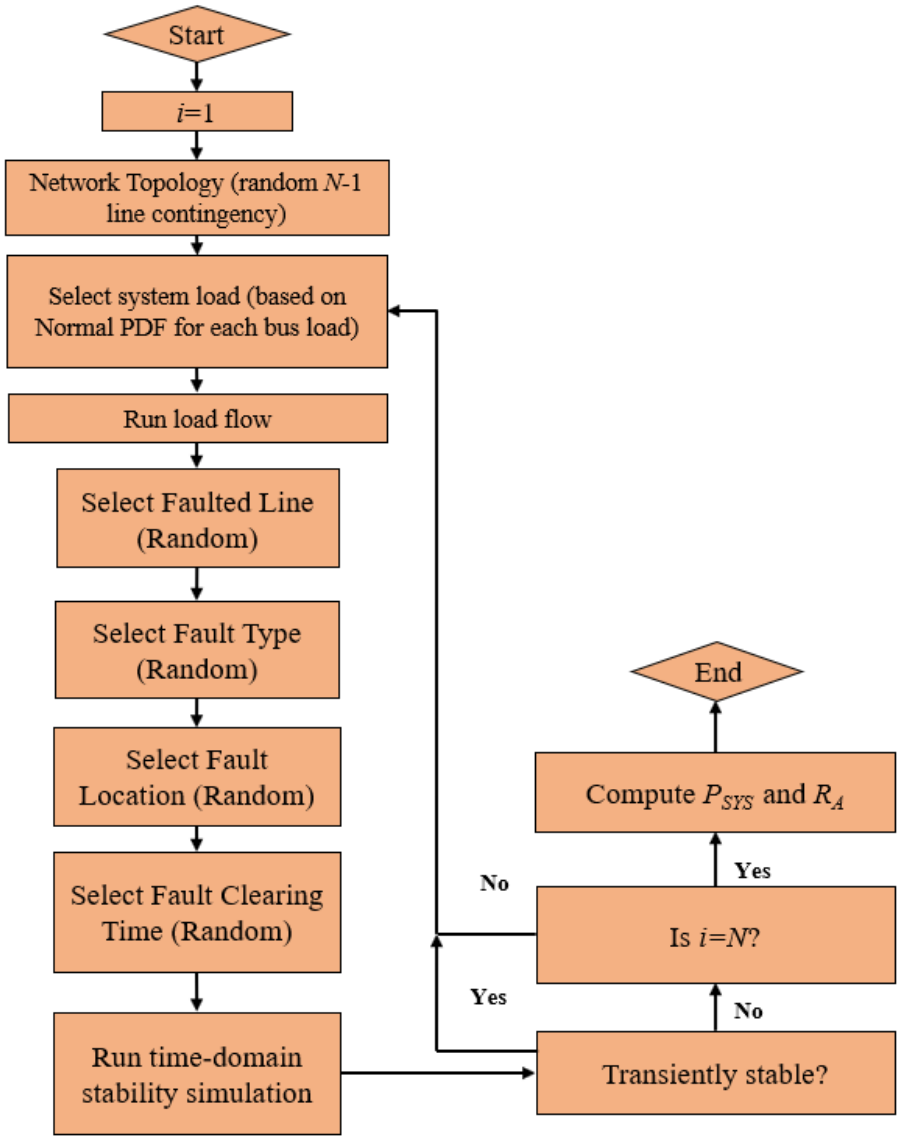

Figure 3.8. Impact of network topology on P_{SYS} and R_A (O7)

3.3.4 PTS classification using ANN and SVM

For the second proposed approach, i.e., PTS classification using ANN and SVM, the first step was feature selection, i.e., to select the input and output data, for the ANN classification model. For feature selection, time-domain simulations were conducted, based on defined input probability distributions, as elaborated by Figure 3.9. System load, fault

type, fault location, and FCT were used as inputs to the ML algorithms, and transient stability status, S_i, was selected as the output (the binary variable to be classified as transiently stable or unstable). 500 samples were used for each line to train each ML model. It must be mentioned that generally, there is no accepted rule of thumb to determine the number of samples for training the ML model; this typically depends on complexity of the problem, required performance level, and the ML algorithm used. As there are 16 lines in the system, thus, the total number of samples used for ML model were 8000 (500×16). When the MC simulation is run for all the 16 lines in the network, feature selection is performed from the resulting data obtained, to be used as training data for the ML classification model, as shown in Figure 3.10. A summarized workflow of ML application to online PTS prediction is shown in Figure 3.11. As illustrated, the first step deals with the offline mode. In this mode, time-domain simulations are conducted, considering the uncertainties of input variables in the form of PDFs (generally obtained from past historical observations). In the next step, these distributions are sampled to gather enough training data. For each sample, the PTS status is measured by a binary variable, say, x, which can take two labels (say, 1 for transiently unstable, and 0 for transiently stable). Therefore, the final training data consists of the PTS status labels and the corresponding input operating conditions. In the next step, this offline-based database is used for online PTS prediction. The ML model 'learns' the stability rules and consequently, can be used to predict the PTS status for current operating point.

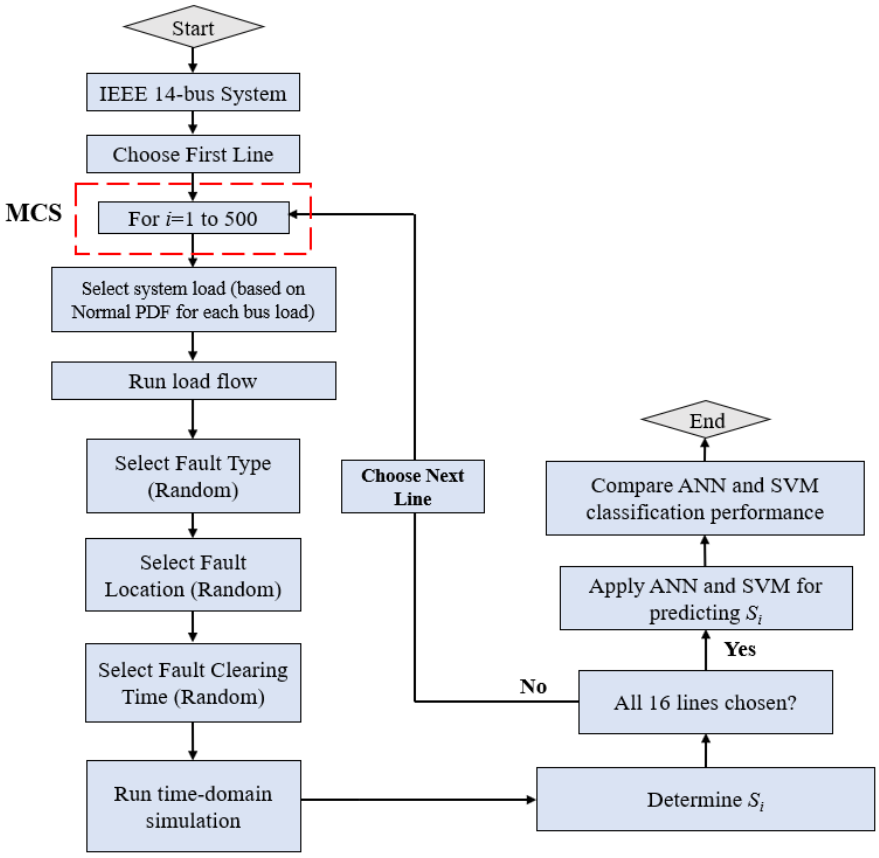

Figure 3.9. Methodology for PTS classification using ANN and SVM (O9)

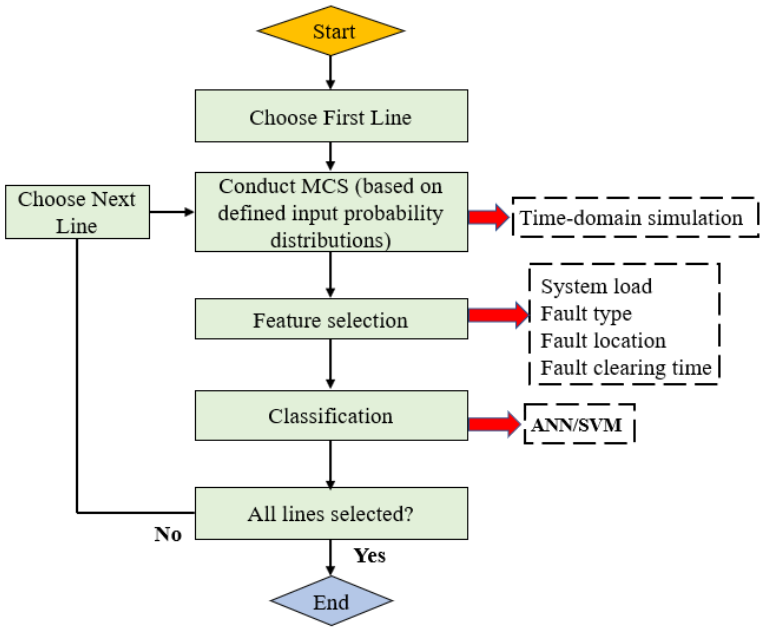

Figure 3.10. Feature selection framework for ANN and SVM (O9)

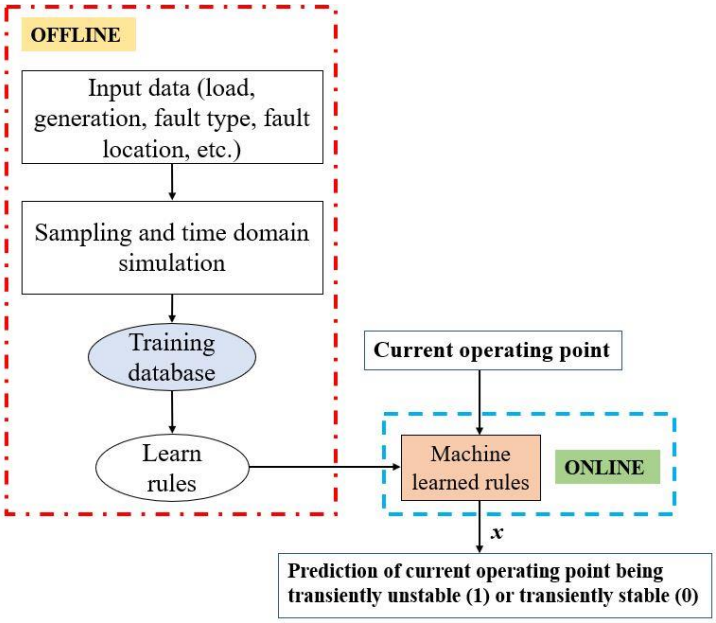

Figure 3.11. Summarized workflow of ML application for online PTS prediction (O9)

3.4 Mathematical Formulation for Transient Instability Risk

Risk-based approach describes possibility of contingency by probability, and the corresponding impact (or consequence) by severity function. The product of this probability and associated severity is termed as risk [11, 25-26]. Critical clearing time (CCT) is the maximum clearing time before which the fault must be cleared to keep the system transiently stable [27]. The risk for the fault which is cleared at CCT is termed as marginal stable risk in this research. In other words, if the system is transiently unstable for a FCT; then, the risk evaluated at the first instance of transient stability (determined by decreasing that FCT in steps) is termed as marginal transient stable risk.

Based on the conceptual framework of risk mentioned in [11, 25-26, 28-29], let R_i be the transient instability risk for i^{th} MC sample. Similarly, let R_A be the average risk index for transient instability (for the decision-making period of 5 years). Mathematically,

$$R_i = \Pr(U_i \cap F_i) \times Sev(F_i) = \Pr(F_i) \times \Pr(U_i \mid F_i) \times Sev(F_i) \tag{3.1}$$

$$R_A = \frac{\sum_{i=1}^{N} R_i}{N} = \frac{\sum_{i=1}^{N} \Pr(F_i) \times \Pr(U_i \mid F_i) \times Sev(F_i)}{N} \tag{3.2}$$

where N denotes the number of MC samples (each sample represents a faulted line).

The term $\Pr(U_i \cap F_i)$ represents the joint probability of: (i) occurrence of F_i (i^{th} fault event), and (ii) transient instability event U_i. According to conditional probability theory, this term can be written as $\Pr(U_i \cap F_i) = \Pr(U_i \mid F_i) \times \Pr(F_i)$, as reflected by (3.1).

$\Pr(U_i \mid F_i)$ is the probability of transient instability given F_i has occurred. Its value is 1 and 0 if the system is unstable and stable (for i^{th} fault event), respectively [29], i.e.,

$$\Pr(U_i \mid F_i) = \begin{cases} 1, \text{ for } \delta_{\max i} > 360 \\ 0, \text{ otherwise} \end{cases} \tag{3.3}$$

$\Pr(F_i)$ is the probability of F_i (i^{th} fault event) and can be defined mathematically as follows.

$$\Pr(F_i) = \Pr(F_{oi}) \times \Pr(F_{Li}) \times \Pr(F_{Ti}) \tag{3.4}$$

where $\Pr(F_{oi})$, $\Pr(F_{Li})$, and $\Pr(F_{Ti})$ denote the probability of fault occurrence, fault location, and fault type, respectively, for the i^{th} MC sample.

Let F_{oi} be a random variable following a uniform (PMF) [5] on the interval $\{1,2,3,\ldots..N_L\}$. Then,

$$\Pr(F_{oi}) = \begin{cases} \frac{1}{N_L}, \text{ for } 1 \le i \le N_L \\ 0, \text{ otherwise} \end{cases} \tag{3.5}$$

where N_L denotes total number of lines in the test system. Let F_{Li} be a random variable following a continuous uniform PDF [5] on the interval [0,100]. Then,

$$\Pr(F_{Li}) = \begin{cases} \dfrac{1}{N_p}, \text{ for } 0 \le i \le N_p \\ 0, \text{ otherwise} \end{cases} \tag{3.6}$$

where N_p=100. $\Pr(F_{Ti})$ is chosen based on PMF as shown in Table 3.1 [8], where, x=1, 2, 3, and 4 denote single line to ground (LG), double line to ground (LLG), line to line (LL), and three-phase (LLL) fault, respectively.

Table 3.1. Probability of fault types

x	**1**	**2**	**3**	**4**
$\Pr(F_{Ti})$	0.7	0.15	0.1	0.05

$Sev(F_i)$ quantifies the impact (severity) of F_i. Mathematically, it is given as follows

$$Sev(F_i) = \begin{cases} |TSI_i|, \text{ if } TSI_i < 0 \\ 0, \text{ if } TSI_i > 0 \end{cases}, \; 0 < |TSI_i| < 1 \tag{3.7}$$

where TSI_i denotes the transient stability index for the i^{th} MC sample, i.e.,

$$TSI_i = \frac{360 - \delta_{\max i}}{360 + \delta_{\max i}} \tag{3.8}$$

where $\delta_{\max i}$ is the post-fault maximum rotor angle difference (in degrees) between any two synchronous generators (SGs) in the system at the same time for a fault on i^{th} line [16]. A negative TSI_i indicates the system in transiently unstable for the i^{th} MC sample.

It is appropriate to model the uncertainty for each bus load forecast with a normal PDF having a mean equal to the forecasted value and an associated standard deviation [11-12]. Let $f(X_i)$ denote the PDF for load at i^{th} bus, i.e.,

$$f(X_i) = \frac{1}{\sqrt{2\pi\sigma_i^2}} e^{\frac{-(X_i - \mu_i)^2}{2\sigma_i^2}} \tag{3.9}$$

where μ_i and σ_i denotes the mean and standard deviation (10% of the mean) of the forecasted peak load for i^{th} bus, respectively. Thus, the PDF for system load, $f(X_j)$, is given by

$$f(X_j) = \frac{1}{\sqrt{2\pi\sigma_j^2}} e^{\frac{-(X_j-\mu_j)^2}{2\sigma_j^2}} \tag{3.10}$$

where μ_j and σ_j denotes the mean (259 MW) and standard deviation (11.5 MW) of the forecasted system peak load (the sum of multiple independent normally distributed random variables is normal, with its mean being the sum of the individual means, and its variance being the sum of the individual variances, i.e., the square of the standard deviation is the sum of the squares of the individual standard deviations), respectively.

The FCT is assumed to follow a normal PDF [6, 9-10], with a mean and standard deviation of 0.9 *s* and 0.1 *s*, respectively.

Let S_i represent the system transient stability status for the i^{th} MC sample. Mathematically,

$$S_i = \begin{cases} 1\,, \text{ if } TSI_i < 0 \text{ (unstable)} \\ 0, \text{ if } TSI_i > 0 \quad \text{(stable)} \end{cases} \tag{3.11}$$

Therefore, if the system is transiently stable, for i^{th} MC sample, value of S_i will be 0; otherwise, it will be 1. This information will be used for ML training for classification purpose.

Let P_{LGI}, P_{LLI}, P_{LLGI}, and P_{LLLI} denote the probability of instability for LG, LL, LLG, and LLL faults, respectively. Also, let P_{SYS} denote the probability of system instability. Mathematically,

$$P_{LGI} = \frac{N_{u1}}{N} \tag{3.12}$$

$$P_{LLI} = \frac{N_{u2}}{N} \tag{3.13}$$

$$P_{LLGI} = \frac{N_{u3}}{N} \tag{3.14}$$

$$P_{LLLI} = \frac{N_{u4}}{N} \tag{3.15}$$

$$P_{SYS} = \frac{N_u}{N} \tag{3.16}$$

where N_{u1}, N_{u2}, N_{u3}, and N_{u4} denote number of unstable samples for LG, LL, LLG and LLL faults, respectively; N_u denotes total unstable samples, irrespective of fault type.

3.5 Mathematical Formulation for Cost Benefit Assessment

The CBA is an assessment process to determine the feasibility and provide economic justification for a project investment. This decision is based on BCR (BCR>1 justifies the investment; BCR<1 does not justify the investment). In this work, CBA is applied to determine whether the existing CBs should be replaced with the faster ones, for transient stability enhancement reducing R_i (and R_A) to zero. To do this, the costs associated with R_i must be formulated. This is described below.

3.5.1 Costs of Transient Instability Risk

The cost consequence of transient instability can be evaluated by assessing the direct and indirect costs incurred due to tripping of synchronous machines. These mainly consist of two components [28, 30].

Replacement Cost

When a SG with an operation cost C_{ORIG} is tripped, a SG with a much more expensive operation cost C_{EMER} is utilized for h hours instead. The generation lost for m^{th} SG, denoted

by P_{Gm}, producing energy at an original cost of C_{ORIG} ($/MW) must be replaced for h hours, by a more expensive generation, with a cost of C_{EMER} ($/MW). Let $C_{REP}{}^{m}$ denote the replacement cost of m^{th} SG, i.e.,

$$C_{REP}{}^{m} = (C_{EMER} - C_{ORIG}) \times P_{Gm} \times h \tag{3.17}$$

It is assumed that C_{ORIG} is $20/MW and C_{EMER} is $40/MW. The downtime h is assumed to be 10 [28, 30]. Thus, the replacement cost (for i^{th} unstable MC sample), for n transiently unstable SGs, is given as

$$C_{REPi}{}^{n} = 200 \times \sum_{m=1}^{n} P_{Gm} \tag{3.18}$$

Repair and Startup Cost

When a synchronous machine is tripped due to transient instability, the tripped synchronous machines must be repaired and restarted. Let $C_{RS}{}^{i}$ denote the repair and startup cost for transient instability for i^{th} transiently unstable sample. Mathematically,

$$C_{RS}{}^{i} = 60,000 \times n_i \tag{3.19}$$

where n_i denotes the number of synchronous machines which are transiently unstable for the i^{th} MC sample [28, 30]. Thus, the cost associated with an i^{th} transiently unstable sample, C_i, is

$$C_i = C_{REPi}{}^{n} + C_{RS}{}^{i} = (200 \times \sum_{m=1}^{n} P_{Gm}) + (60,000 \times n_i) \tag{3.20}$$

In other words, the cost associated with risk R_i [$R_i = \Pr(F_i) \times Sev(F_i)$] is given by C_i. This will be the monetary benefits gained by reducing R_i to zero (marginal stability risk).

3.5.2. Costs of Circuit Breakers

Based on [31-32], CBs for high-voltage transmission systems typically have a FCT of 3 cycles (0.05 *s*). According to [33], the capital cost of such a breaker can be determined as follows: The cost consists of material cost C_M, installation cost C_I, and foundation cost C_F (elaborated in Table 3.2). Let C_{BR} be the capital cost of a single CB, with a FCT of 0.05 *s*, i.e.,

$$C_{BR} = C_M + C_I + C_F \tag{3.21}$$

A relation between FCT and C_{BR} must be formulated. It is assumed that C_M has a linear inverse relation with the FCT, i.e., $C_M \propto \frac{1}{FCT}$. Moreover, C_I and C_F are assumed to be constant. Thus, based on a FCT of 0.05 *s*, the relation between C_M and FCT can be mathematically expressed as

$$C_M = \frac{k}{FCT} \tag{3.22}$$

where k=3700 ($74,000 \times 0.05$) is the proportionality constant.

Let FCT_i denote the fault clearing time for an i^{th} unstable MC sample. Let CCT_i denote critical clearing time (the time to acquire marginal stability) for the same i^{th} unstable sample. Let C_{Bi} denote the capital cost of a single CB for reducing R_i to zero. Mathematically,

Table 3.2. Capital cost of a single CB (For a FCT of 0.05s)

Cost type	Cost value ($)
Material cost, C_M	74,000
Installation cost, C_I	14,000
Foundation cost, C_F	6,000
Capital cost, C_{BR}	94,000

$$C_{Bi} = 20,000 + \frac{k}{CCT_i} = 20,000 + \frac{3700}{CCT_i} \tag{3.23}$$

The present value of money is computed as

$$PV = \frac{FV}{(1+r)^n} \tag{3.24}$$

where *PV* denotes present value, *FV* denotes future value, *r* is interest rate (assumed as 5%), and *n*=5 years (time after which the decision-making is required). BCR is computed as

$$\text{BCR} = \frac{PV \text{ of benefits}}{PV \text{ of costs}} \tag{3.25}$$

For the i^{th} unstable MC sample, let C_i be the future monetary benefits gained by replacing two slower CBs (at both ends of a faulted line) with faster ones, and let $2C_{Bi}$ be the associated capital cost (*PV*) of the two faster CBs, which are installed in place of the slower breakers (it is assumed that upon fault detection, a tripping signal is sent to both ends of the faulted line such that both CBs are tripped at the same time instance). Thus, *BCR* for i^{th} unstable MC sample, denoted by BCR_i, is given by

$$BCR_i = \frac{C_i{}^p}{2C_{Bi}} \tag{3.26}$$

where C_i^p denotes the *PV* of C_i, and is given as

$$C_i^p = \frac{C_i}{(1+0.05)^5} = \frac{C_i}{(1.05)^5} \tag{3.27}$$

Table 3.3 displays the conceptual framework of the CBA (the values are assumed just to illustrate the concepts and associated mathematics).

Let $EBCR_m$ be the expected value of BCR for m^{th} line, i.e.,

$$EBCR_m = \frac{\sum_{m=1}^{N_m} BCR_m}{N_m} \tag{3.28}$$

where N_m denotes the number of transiently unstable MC samples for m^{th} line. For simplicity, $EBCR_m$ is denoted as DV_m (decision variable for m^{th} line), as it helps in decision-making, regarding replacement of CBs.

Table 3.3. Conceptual framework for CBA

MC sample no.	Faulted line	FCT_i *(s)*	R_i	CCT_i *(s)*	C_i ($)	C_i^p ($)	$2C_{Bi}$ ($)	$BCR_i = \frac{C_i^p}{2C_{Bi}}$
1	Line 1	0.9	10	0.7	25	19.58	30	0.652
2	Line 3	0.8	15	0.6	45	35.25	55	0.640
3	Line 5	0.9	20	0.8	34	26.63	65	0.409
4	Line 3	0.6	25	0.5	56	43.87	44	0.997
5	Line 14	0.7	30	0.6	78	61.11	33	1.851
:	:	:	:	:	:	:	:	:
:	:	:	:	:	:	:	:	:
:	:	:	:	:	:	:	:	:
2,401	Line 12	0.7	45	0.5	54	42.3	65	0.650

3.6 ANN Application for PTS Enhancement Decision Making using CBs

The generic block diagram for the proposed ANN regression approach is shown in Figure 3.12. For the regression task, the first step was to select the input and output data, for the

ANN training model. System load, fault type, fault location, and FCT were used as inputs to the ANN, and DV_m was selected as the output (the value to be predicted). This data is required for each line. To gather sufficient training data for each line, 500 MC simulations were performed for each line (based on Figure 3.3) and consequently, the number of samples which led to transient instability were used for training (for stable samples, R_A and hence, DV_m is zero). Hence, a total of 8000 (500×16) MC simulations were performed.

It must be mentioned that generally, there is no accepted rule of thumb to determine the number of samples for training the ML model; this typically depends on complexity of the problem, required performance level, and the ML algorithm used. The total unstable samples turned out to be 3,108. The random data division for training, validation, and testing was set at 2,176 (70%), 466 (15%), and 466 (15%), respectively. To quantify the performance of the regression algorithm, mean squared error (*MSE*) and root mean squared error (*RMSE*) were used.

Levenberg-Marquardt backpropagation algorithm was used to train the ANN. This algorithm gives a fast convergence and enhanced training performance. As there are 4 inputs (system load, fault type, fault location, and FCT) and one output (DV_m) for the multi-layer perceptron neural network (MLPNN), the number of neurons, in the input and output layer of MLPNN, for this problem, are 4 and 1, respectively. The number of neurons used in the hidden layer were chosen as 20 (based on trial and error approach). The tan-sigmoid activation/transfer function (this function is used to determine the output of a neuron in an ANN) was used for neurons of hidden layer and the linear activation function was used for neuron of output layer.

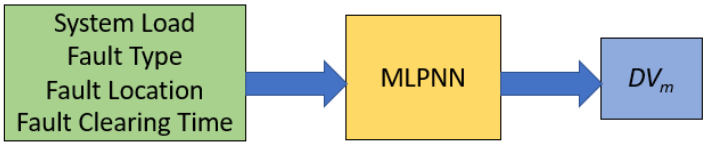

Figure 3.12. Proposed ANN approach for regression (O8)

3.7 ANN Application for Probabilistic Transient Stability Classification

The generic block diagram, for the proposed ANN classification approach, is shown in Figure 3.13. The proposed MLPNN used has four input layers (each for system load, fault type, fault location and FCT), one hidden layer, and one output layer (for S_i). Samples for training data were chosen using the MC simulation-based, time domain approach (described in Section 3.3). Batch training style was used, i.e., the weights and biases are only updated after all inputs are fed to the MLPNN. In the MATLAB environment, it is significantly faster and produces smaller errors as compared to incremental training (where the weights and biases are updated every time an input is presented to the MLPNN) [3].

For the ANN classification task, the first step was feature selection, i.e., to select the input and output data for the ANN classification model. System load, fault type, fault location, and FCT were used as inputs to the ANN, and transient stability status, S_i, was selected as the output (the binary variable to be classified as transiently stable or unstable). 500 samples were used for each line to train the ANN model. As there are 16 lines in the system, thus, the total number of samples used for ML model were 8000 (500×16). The random data division for training, validation, and testing was set at 5,600 (70%), 1,200 (15%), and 1,200 (15%), respectively. Levenberg-Marquardt backpropagation algorithm was used to train the ANN. As there are four inputs (system load, fault type, fault location, and FCT) and one output (S_i) for the MLPNN, the number of neurons, in the input and output layer

of MLPNN, for this problem, are 4 and 1, respectively. The number of neurons used in the hidden layer were chosen as 20 (based on trial and error approach). The log-sigmoid (logsig) activation/transfer function (this function is used to determine the output of a neuron in an ANN) was used for neurons of hidden layer, and the tan-sigmoid (tansig) activation function was used for neuron of output layer.

3.8 SVM Application for Probabilistic Transient Stability Classification

The block diagram for the proposed SVM framework is shown in Figure 3.14. The proposed SVM framework used has four inputs (system load, fault type, fault location and FCT), and one output (for S_i). For the SVM classification task, the first step was feature extraction, i.e., to select the most relevant input and output data for the SVM classification model. System load, fault type, fault location, and FCT were chosen as inputs, and transient stability status, S_i, was selected as the output (the binary variable to be classified as transiently stable or unstable). 500 samples were used for each line to train the SVM model. As there are 16 lines in the system, thus, the total number of samples used for SVM model were 8000 (500×16). Thus, the size of the input feature matrix was 8000×4. The Gaussian RBF kernel function was used for training the SVM, as there is ample nonlinearity amongst the data presented to the SVM classifier. The hyperparameters (C and γ) were optimized using Bayesian optimization. The data presented to SVM is randomly divided in two subsets: training subset and testing subset. The K-fold cross-validation approach [22] is used to accomplish this, as this prevents over fitting while training the data. This work used the value of K as 5, i.e., in each fold, 20% data was used for testing and 80% for training.

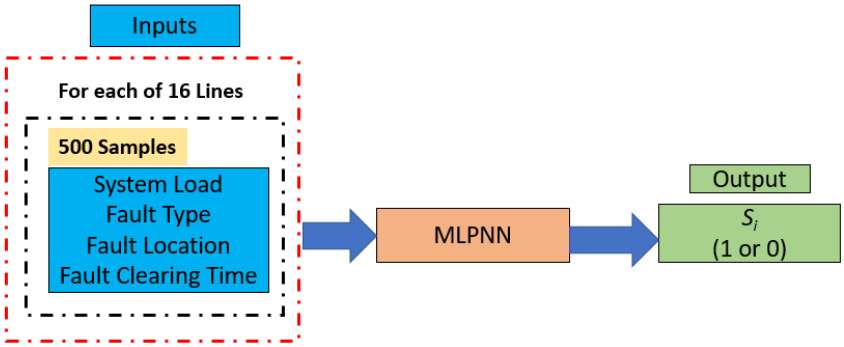

Figure 3.13. Proposed ANN approach for classification showing input features and corresponding output (O9)

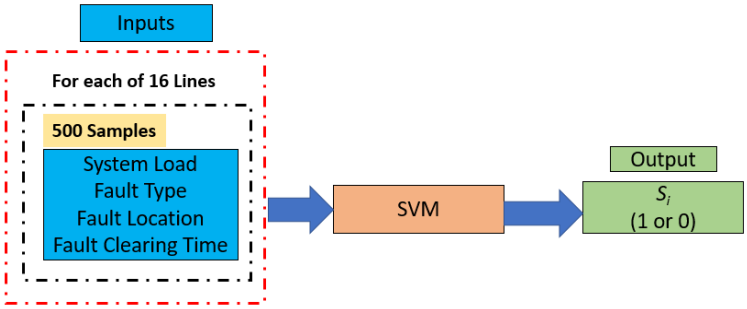

Figure 3.14. Framework for the proposed SVM classification approach showing input features and corresponding output (O9)

3.9 References

[1] DIgSILENT PowerFactory User Manual, DIgSILENT GmbH, 2018. [Online]. Available: https://www.digsilent.de/en/downloads.html

[2] The MathWorks, “Company Overview”, Feb. 2020. [Online]. Available: https://uk.mathworks.com/content/dam/mathworks/handout/2020-company-factsheet-8-5x11-8282v20.pdf

[3] M. H. Beale, M. T. Hagan, and H. B. Demuth, Neural Network Toolbox (User’s Guide). [Online]. Available: https://www2.cs.siu.edu/~rahimi/cs437/slides/nnet.pdf

[4] Classification learner. [Online].

Available: https://www.mathworks.com/help/stats/classificationlearner-app.html

[5] P. N. Papadopoulos and J. V. Milanović, "Probabilistic framework for transient stability assessment of power systems with high penetration of renewable generation," *IEEE Transactions on Power Systems*, vol. 32, no. 4, pp. 3078-3088, Jul. 2017.

[6] R. Billinton, "Probabilistic evaluation of transient stability in a multimachine power system," vol. 126, *Proceedings of the Institution of Electrical Engineers*, 1979.

[7] A. A. Irizarry-Rivera, J. D. McCalley, and V. Vittal, "Computing probability of instability for stability constrained electric power systems," *Electric Power Systems Research*, vol. 42, no. 2, pp. 135-143, Aug. 1997.

[8] M. Abapour and M. Haghifam, "Probabilistic transient stability assessment for on-line applications," *International Journal of Electrical Power & Energy Systems*, vol. 42, no. 1, pp. 627–634, Nov. 2012.

[9] L. Miao, J. Fang, J. Wen, and W. Luo, "Transient stability risk assessment of power systems incorporating wind farms," *Journal of Modern Power Systems and Clean Energy*, vol. 1, no. 2, pp. 134-141, Sep. 2013.

[10] S. O. Faried, R. Billinton, and S. Aboreshaid, "Probabilistic evaluation of transient stability of a power system incorporating wind farms," *IET Renewable Power Generation*, vol. 4, no. 4, pp. 299–307, Jul. 2010.

[11] J. McCalley et al., "Probabilistic security assessment for power system operations," *Proc. IEEE Power Energy Soc. Gen. Meet.*, 2004, pp. 1-9.

[12] R. Billinton and D. Huang, "Effects of load forecast uncertainty on bulk electric system reliability evaluation," *IEEE Transactions on Power Systems*, vol. 23, no. 2, pp. 418-425, May 2008.

[13] R. Billinton and K. Chu, "Transient stability evaluation in an undergraduate curriculum - A probabilistic approach," *IEEE Transactions on Power Systems*, vol. 1, no. 4, pp. 171-177, Nov. 1986.

[14] R. T. Clemen and T. Reilly, *Making Hard Decisions with Decision Tools*, Boston, MA, USA: Cengage Learning, 2013.

[15] J. Cervantes and F. Choobineh, "A quantile-based approach for transmission expansion planning," *IEEE Access*, vol. 8, pp. 82630-82640, 2020.

[16] M. Edrah, K. L. Lo, and O. Anaya-Lara, "Impacts of high penetration of DFIG wind turbines on rotor angle stability of power systems," *IEEE Transactions on Sustainable Energy*, vol. 6, no. 3, pp. 759-766, Jul. 2015.

[17] M. Edrah, "Impact of DFIG based wind farms on transient stability of power systems," *3rd International Conference on Automation, Control, Engineering and Computer Science*, 2016, pp. 1-6.

[18] E. Vittal, M. O'Malley, and A. Keane, "Rotor angle stability with high penetrations of wind generation," *IEEE Transactions on Power Systems*, vol. 27, no. 1, pp. 353-362, Feb. 2012.

[19] A. Karami, "Transient stability assessment of power systems described with detailed models using neural networks," *International Journal of Electrical Power & Energy* Systems, vol. 45, no. 1, pp. 279-292, Feb. 2013.

[20] A. Karami and K. M Galougahi, “Improvement in power system transient stability by using STATCOM and neural networks,” *Electrical Engineering*, vol. 101, pp. 19–33, 2019.

[21] P. K. Olulope, K. A. Folly, S. Chowdhury, and S. P. Chowdhury, “Transient stability assessment using artificial neural network considering fault location,” *1st International Conference on Energy, Power, and Control*, 2010, pp. 67-72.

[22] D. You, K. Wang, L. Ye, J. Wu, and R. Huang, “Transient stability assessment of power system using support vector machine with generator combinatorial trajectories inputs,” *International Journal of Electrical Power & Energy Systems*, vol. 44, no. 1, pp. 318-325, Jan. 2013.

[23] V. Vapnik, *The Nature of Statistical Learning Theory*. New York, NY, USA: Springer-Verlag, 1995.

[24] J. Snoek and R. Adams, “Practical Bayesian optimization of machine learning algorithms,” *Advances in Neural Information Processing Systems*, vol. 2, pp. 2951–2959, Dec. 2012.

[25] M. Negnevitsky, D. H. Nguyen, and M. Piekutowski, “Risk assessment for power system operation planning with high wind power penetration,” *IEEE Transactions on Power Systems*, vol. 30, no. 3, pp. 1359-1368, May 2015.

[26] D. D. Le et al., “A probabilistic approach to power system security assessment under uncertainty,” *IREP Symposium Bulk Power System Dynamics and Control - IX Optimization, Security and Control of the Emerging Power Grid*, 2013, pp. 1-7.

[27] P. Kundur et al., “Definition and classification of power system stability,” *IEEE Transactions on Power Systems*, vol. 19, no. 3, pp. 1387-1401, May 2004.

[28] V. E. V. Acker, “Transient stability assessment and decision-making using risk,” *Ph.D. Dissertation*, Iowa State University, 2000.

[29] S. Datta and 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.

[30] M. Abapour, P. Aliasghari, and M. Haghifam, “Risk-based placement of TCSC for transient stability enhancement,” *IET Generation, Transmission & Distribution*, vol. 10, no. 13, pp. 3296-3303, Oct. 2016.

[31] “IEEE Guide for Breaker Failure Protection of Power Circuit Breakers,” in IEEE Std C37.119-2016 (Revision of IEEE Std C37.119-2005), pp.1-73, Jul. 2016. [Online]. Available: https://ieeexplore.ieee.org/document/7509575

[32] E. Atienza and R. Moxley, “Improving breaker failure clearing times,” Schweitzer Engineering Laboratories, *Annual Protection, Automation and Control World Conference*, 2011, pp. 1-8. [Online].
Available:
https://cdn.selinc.com/assets/Literature/Publications/Technical%20Papers/6384_ImprovingBreaker_EA_20090917_Web.pdf

[33] MISO (Midcontinent Independent System Operator) Transmission Cost Estimation Guide – MTEP19, 2019. [Online]. Available:
https://cdn.misoenergy.org/20190212%20PSC%20Item%2005a%20Transmission%20Cost%20Estimation%20Guide%20for%20MTEP%202019_for%20review317692.pdf

CHAPTER 4

CASE STUDIES AND RESULTS

This chapter describes the test system used, related case studies, results, and associated discussion. A validation approach, using sensitivity analysis, is also presented.

4.1 Case Studies

The IEEE 14-bus test transmission system, as shown in Figure 4.1, was used to conduct the required simulations. Figure 4.2 shows the same system with circuit breakers (CBs) labelled for each line (which will be useful for ranking CBs in the later part of this chapter). It represents a simplified model of the transmission system in the Midwest United States. It consists of five synchronous machines, three of which are synchronous compensators used only for reactive power support. There are 11 loads in the system totaling 259 MW and 81.3 MVAr. This system has 16 transmission lines. The numerical data and parameters were taken from [1]. This system is a good choice for the present study, as it has been widely used by various researchers for studying transient stability phenomenon in power transmission systems [2-5]. As mentioned before, a normal (Gaussian) probability density function (PDF) is used to define the uncertainty in system loads. The active power of each load was assigned a mean equal to the original load active power value, as given in test system data in [1], and a standard deviation equal to 10% of the mean value. Six different cases were considered for analysis. Their description is as follows. In Case 1 (base case), the test system was used in its original format, i.e., no wind generation was present in the network and all generation consisted of conventional synchronous generators (SGs). Type

3 wind farms, i.e., doubly fed induction generators (DFIGs) are used in this research, as they are the most commonly used wind farms in power systems [6]. As mentioned before, in this work, the DFIGs are equipped with terminal voltage control, i.e., the DFIG can exchange reactive power with the grid to achieve a target voltage at the bus at which DFIG is connected. Cases 2-6 deal with the impact of wind farms. This impact is studied in two different forms: (1) replacing the existing SG with a DFIG, and (2) adding DFIGs to the existing network. Case 2 replaces the SG at bus 2 with a DFIG (with the same MW and MVA rating). Cases 3-6 integrate DFIGs at specific buses (refer to Table 4.1 for description). In Cases 3-6, to account for increased generation due to wind farms, an equal amount of load is also added at the same bus where the wind farm is connected. The active power forecast error distribution of each wind generator is represented by a normal PDF [7-10], with a mean of 40 MW and a standard deviation of 4 MW. The active power of each additional load added is a normal PDF with a mean of 40 MW and a standard deviation of 4 MW. All time-domain simulations are RMS simulations and were performed using DIgSILENT PowerFactory software [11]. For machine learning (ML) application, neural network toolbox and classification learner application of MATLAB was used [12].

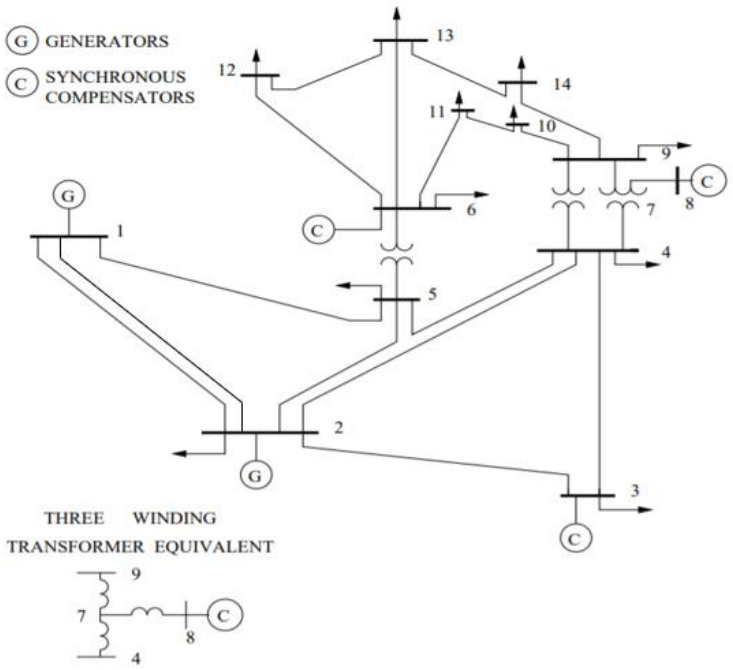

Figure 4.1. IEEE 14-bus system

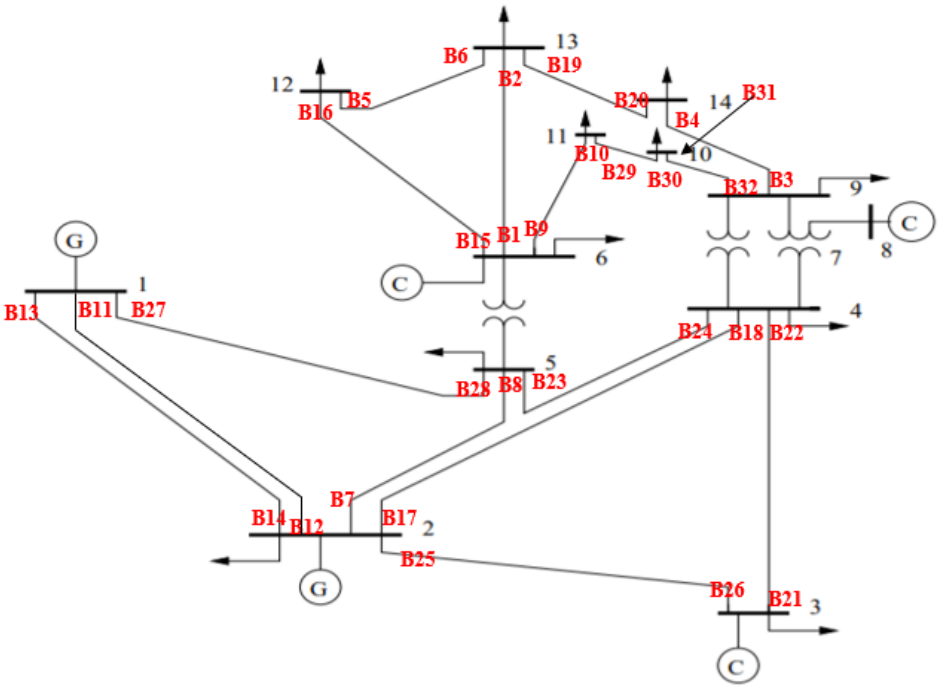

Figure 4.2. IEEE 14-bus system (with CB labels)

Table 4.1. Description of DFIG locations

Case type	Bus locations for DFIG
Case 3	2
Case 4	2, 3
Case 5	2, 3, 8
Case 6	2, 3, 8, 6

4.2 Results and Discussion

This section presents the results of the proposed research methodologies and associated discussion.

4.2.1 Risk-Based Probabilistic Transient Stability (PTS) Enhancement using Circuit Breakers and Impact of DFIG

The first step was to compute the number of samples required for performing the Monte Carlo (MC) simulation. For this purpose, Cochran's formula was used [13]. Assuming a 95% confidence level and 2% margin of error, the number of required MC samples, N, were determined to be 2,401. The MC simulation was stopped after N simulations, and consequently, P_{SYS} and value of R_A is determined. The value of P_{SYS} and R_A of comes out to be 0.37 and 0.005 %, respectively. In the next step, based on the proposed approach, described in Chapter 3, the time-domain simulations were performed for all Cases (1-6) for N MC samples. The value of R_A for each case are shown in Figure 4.3. As evident, the replacement of SG by DFIG at Bus 2 (Case 2), resulted in an increase in the value of R_A. This is because the inherent inertia of system is reduced and the DFIG does not have enough capability to provide reactive power support during fault time, as compared to SG.

Moreover, the integration of DFIG farms with increasing penetration, gradually reduced the value of R_A (Cases 3-6). This is understandable, because, in addition to the existing inherent inertia provided by SGs, DFIG provides reactive power support to the system. The results, for DV_m, are shown in Figure 4.4. The CBs on the lines whose DV_m is greater than 1 should be replaced. As evident, the replacement of SG with DFIG at Bus 2 (Case 2) causes a slight increase in DV_m values. This is due to increase in the value of R_A (as illustrated by Figure 4.3), which consequently, increases the value of C_{Ai}. Moreover, the value of DV_m gradually decreases with integration of DFIG wind farms. The reason is that with the addition of DFIG, the value of R_A decreases, which in turn, reduces the value of C_{Ai}.

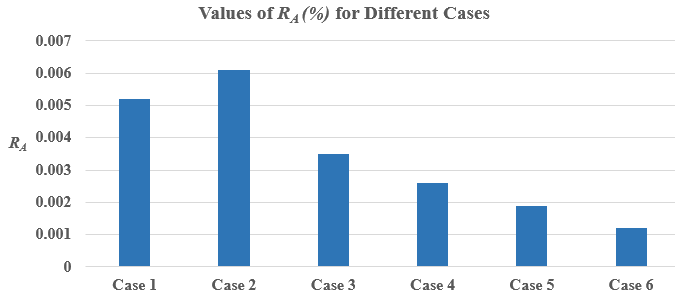

Figure 4.3. Value of R_A for different cases

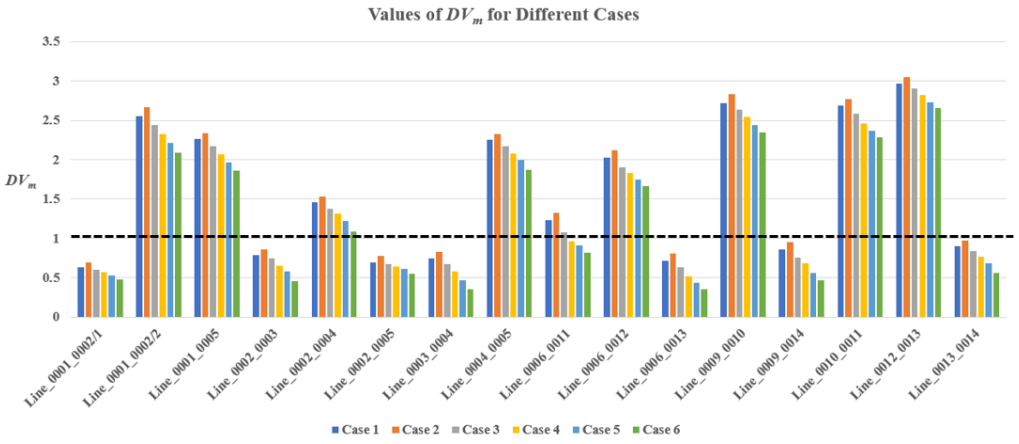

Figure 4.4. Value of DV_m for different cases

4.2.2 Comparison of Proposed Approach with Deterministic Scenarios

To compare the proposed probabilistic approach with the conventional deterministic approach, simulations were done (for Case 1 only), for each line in the network, for specific input conditions. Two different deterministic scenarios (with peak load) were considered: (1) LLL fault in the middle of line with a fault clearing time (FCT) of 0.9 *s*, and (2) a close-in LLL fault from 0.1% of the sending bus with a FCT of 0.9 *s*. The comparison, for values of DV_m, between the proposed approach and two deterministic scenarios, is graphically depicted in Figure 4.5. As evident from Figure 4.5, the scenarios for deterministic approach are conservative, and their results demonstrate that CBs on each line must be replaced, as DV_m for each of them is greater than 1; however, referring to the proposed probabilistic approach, the CBs of only 9 lines (out of 16) needs to be replaced. This implies that the proposed probabilistic approach considers the randomness of input variables for decision-making, based on risk, rather than the specific deterministic scenarios. This also verifies the fact that close-in faults (faults near sending-end bus bars) are severe than faults in the middle of line. It must be mentioned that although, this comparison is performed for base case only, it can be extended to other cases in the same manner.

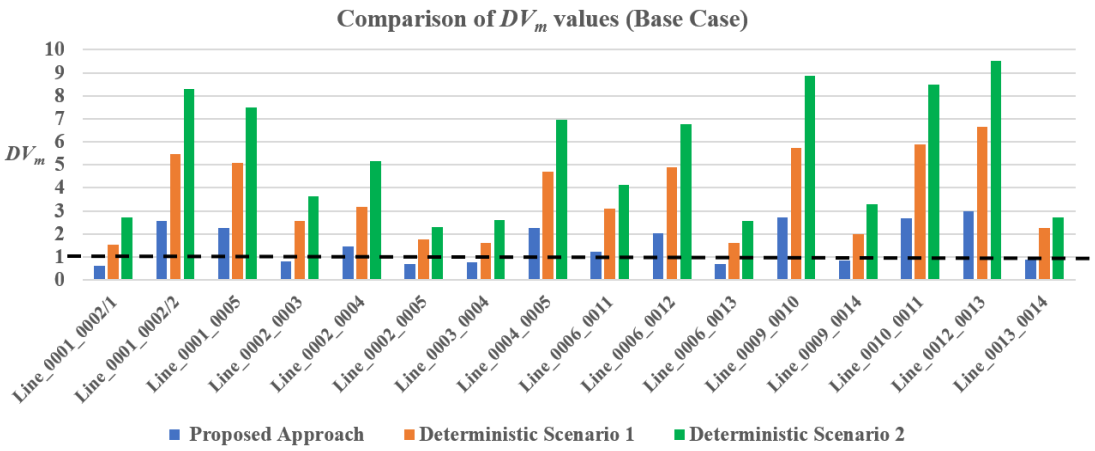

Figure 4.5. DV_m value for proposed and deterministic approaches

4.2.3 Ranking of Circuit Breakers

Based on the approach discussed in Section 3.3.2, the CBs were first ranked for line faults, as shown in Table 4.2 and Figure 4.6. As evident, CBs B1 and B2 (on line_0006_0013) are the most critical, and CBs B31 and B32 (on line_0009_0010) are the least critical. Similarly, the CBs were ranked for bus faults, as shown in Table 4.3 and Figure 4.7. It must be noted that for bus faults, the following probabilities were assumed: P_{LG}=0.007, P_{LLG}=0.0015, P_{LL}=0.001, P_{LLL}=0.0005.

As evident, CBs B1, B9, and B15 (associated with bus_0006) are the most critical, and CBs B30 and B31 (associated with bus_0010) are the least critical. This quantification is essential to power planners, as it can identify which lines/buses require more attention while planning a system. These results can aid in efficient decision making, considering the level of risk associated with CBs.

The CBs were also ranked using deterministic three phase bus faults (including other deterministic factors, such as load, FCT, etc.). The results are shown in as shown in Table

4.4 and Figure 4.8. As evident, CBs B1, B9, and B15 (associated with bus_0006) are the most critical, and CBs B30 and B31 (associated with bus_0010) are the least critical. Moreover, as expected, the value of R_A is significantly higher for the deterministic case (only three phase bus faults) as compared to the probabilistic case. This is because for the deterministic case (probability of LLL fault being 1), the impact is greater as compared to probabilistic case (which includes all faults).

Table 4.2. Ranking of CBs (based on R_A) for line faults

Priority Rank	Line	CBs	R_A (%)
1	Line_0006_0013	B1, B2	0.0058
2	Line_0009_0014	B3, B4	0.0056
3	Line_0012_0013	B5, B6	0.0056
4	Line_0002_0005	B7, B8	0.0054
5	Line_0006_0011	B9, B10	0.0053
6	Line_0001_0002/2	B11, B12	0.0052
7	Line_0001_0002/1	B13, B14	0.0051
8	Line_0006_0012	B15, B16	0.005
9	Line_0002_0004	B17, B18	0.0049
10	Line_0013_0014	B19, B20	0.0048
11	Line_0003_0004	B21, B22	0.0047
12	Line_0004_0005	B23, B24	0.0045
13	Line_0002_0003	B25, B26	0.0042
14	Line_0001_0005	B27, B28	0.0041
15	Line_0010_0011	B29, B30	0.0039
16	Line_0009_0010	B31, B32	0.0037

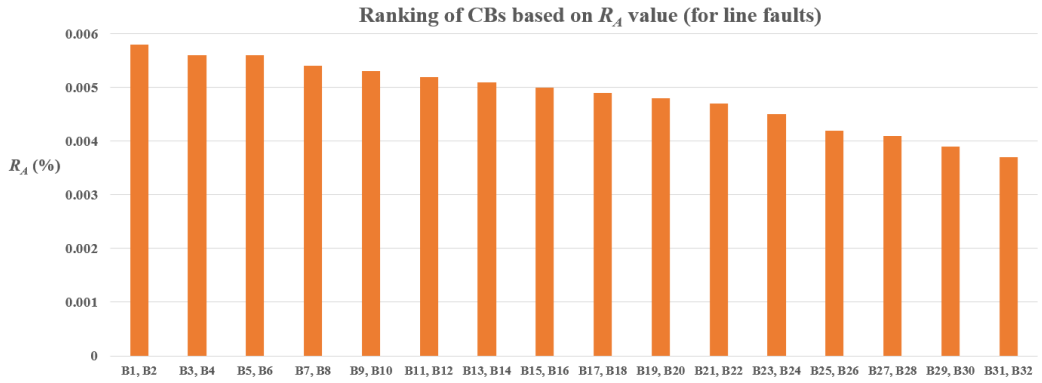

Figure 4.6. Ranking of CBs (based on R_A) for line faults

Table 4.3. Ranking of CBs (based on R_A) for bus faults

Priority Rank	Bus	CBs	R_A (%)
1	Bus_0006	B1, B9, B15	0.0061
2	Bus_0009	B3, B32	0.0059
3	Bus_00012	B5, B16	0.0058
4	Bus_0005	B8, B18, B23	0.0056
5	Bus_00011	B10, B29	0.0055
6	Bus_0001	B11, B13, B27	0.0054
7	Bus_0002	B7, B12, B14, B17	0.0054
8	Bus_0004	B18, B22, B24	0.0052
9	Bus_0013	B2, B6, B19	0.005
10	Bus_0014	B4, B20	0.0048
11	Bus_0003	B21, B26	0.0046
12	Bus_0010	B30, B31	0.0044

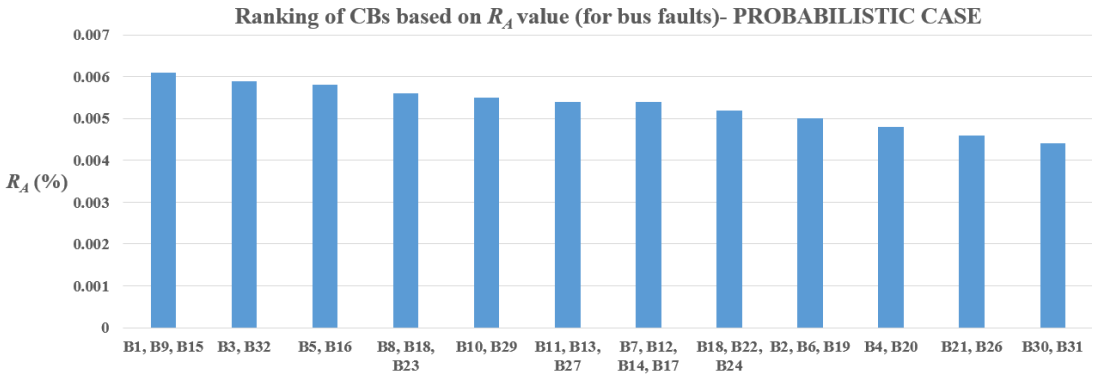

Figure 4.7. Ranking of CBs (based on R_A) for bus faults

Table 4.4. Ranking of CBs (based on R_A) for LLL bus faults

Priority Rank	Bus	CBs	R_A (%)
1	Bus_0006	B1, B9, B15	1.02
2	Bus_0009	B3, B32	0.96
3	Bus_00012	B5, B16	0.93
4	Bus_0005	B8, B18, B23	0.88
5	Bus_00011	B10, B29	0.84
6	Bus_0001	B11, B13, B27	0.81
7	Bus_0002	B7, B12, B14, B17	0.78
8	Bus_0004	B18, B22, B24	0.76
9	Bus_0013	B2, B6, B19	0.74
10	Bus_0014	B4, B20	0.71
11	Bus_0003	B21, B26	0.62
12	Bus_0010	B30, B31	0.61

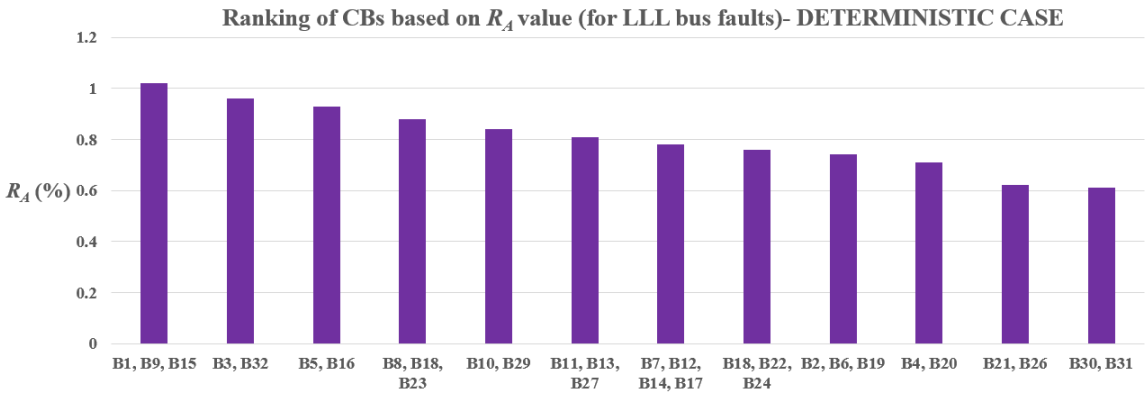

Figure 4.8. Ranking of CBs (based on R_A) for LLL bus faults

4.2.4 Computation of Instability Probability for Line and Bus Faults

The probability of instability for individual LG, LLG, LL, and LLL faults, denoted by P_{LGI}, P_{LLGI}, P_{LLI}, and P_{LLLI}, respectively, was also computed. The results are shown in Table 4.5 and Figure 4.9 for line faults. For bus faults, the results are shown in Table 4.6 and Figure 4.10. From these results, it is evident that P_{LLLI} has the maximum value, and P_{LGI} has the minimum value. This is because LLL fault is the most severe fault. Moreover, P_{LGI}, P_{LLGI}, P_{LLI}, and P_{LLLI} are approximately double for bus faults as compared to line faults. This reiterates the fact that bus faults are more severe than line faults. Also, as expected, P_{SYS} is greater for bus faults as compared to line faults.

Table 4.5. Probability of instability (for individual faults) and system instability [line faults]

P_{LGI}	P_{LLGI}	P_{LLI}	P_{LLLI}	P_{SYS}
0.34	0.36	0.38	0.39	0.37

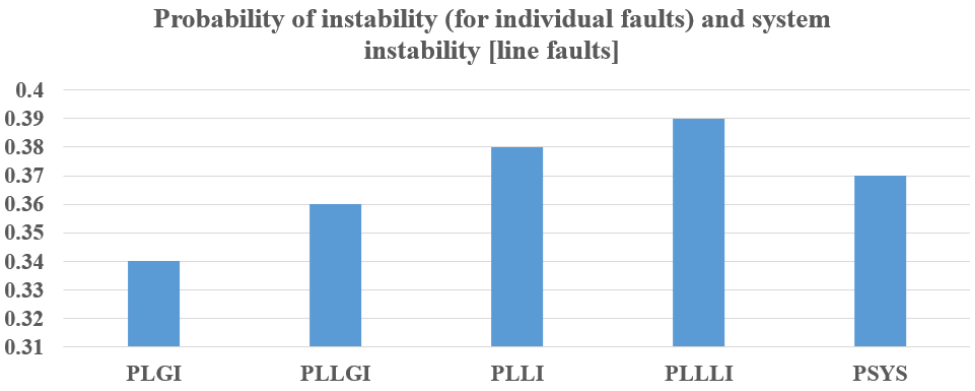

Figure 4.9. Probability of instability (for individual faults) and system instability [line faults]

Table 4.6. Probability of instability (for individual faults) and system instability [bus faults]

P_{LGI}	P_{LLGI}	P_{LLI}	P_{LLLI}	P_{SYS}
0.65	0.68	0.71	0.73	0.70

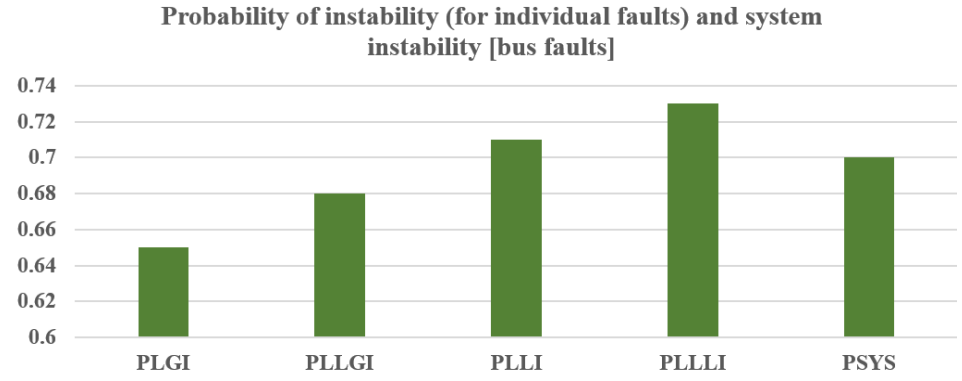

Figure 4.10. Probability of instability (for individual faults) and system instability [bus faults]

4.2.5 Impact of Network Topology on P_{SYS} and R_A

Based on the flowchart of Figure 3.7, the impact of network topology (using random *N*-1 line contingency) on P_{SYS} and R_A was studied. The results are shown in Table 4.7. As evident, the network topology considerably impacts P_{SYS} and R_A. This signifies the importance of considering network topology in probabilistic transient stability analysis.

Table 4.7. Impact of network topology change on P_{SYS} and R_A

Topology	P_{SYS}	R_A (%)
Fixed	0.37	0.005
(*N*-1) line contingency	0.52	0.0072

4.2.6 Machine Learning Approaches

In the second part, ML was applied for improving computation efficiency. This is divided into two tasks: (1) regression and (2) classification. Although, these tasks were demonstrated for the base case only, they can easily be extended to include the cases with wind integration. The results for each task are presented and discussed below.

4.3 Machine Learning Regression

The plots for coefficient of regression, are shown in Figure 4.11. For an ideal fit, the data must lie on a 45-degree line, where the predicted output values are equal to the target output values. The obtained fit is good enough for all data sets, with the values of correlation coefficient (or regression coefficient), *R*, in each case (training, validation, testing), greater than 0.97. Figure 4.12 shows the learning curve of the multilayer perceptron neural network (MLPNN). As evident, the best validation performance for the training model occurs at

epoch 161. Moreover, the value of mean squared error (*MSE)* at that point is 0.014598. It took 1 second with 167 epochs to train the ANN on an Intel Core i7 Processor with a 16 GB RAM. The training stopped after 167 epochs because the approach of early stopping (the approach to stop training at the point when performance on a validation dataset starts to worsen) was used to enhance the generalization of the trained network. The summary of various regression metrics for artificial neural network (ANN) is shown in Table 4.8.

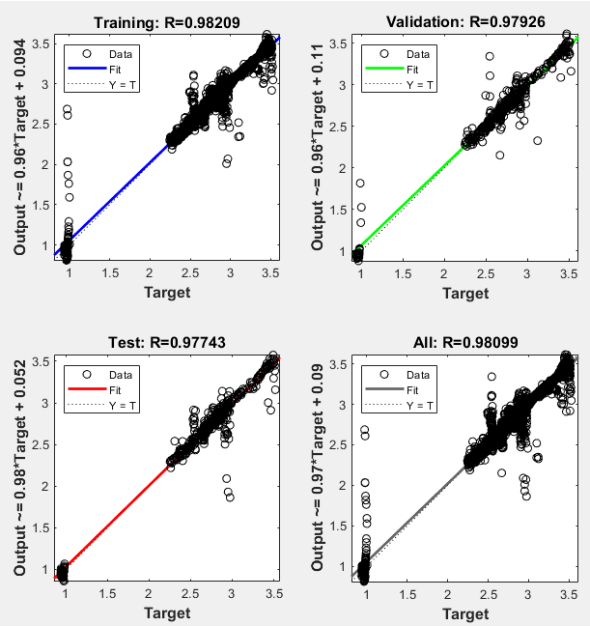

Figure 4.11. Coefficient of Regression (*R*) plot for prediction performance assessment

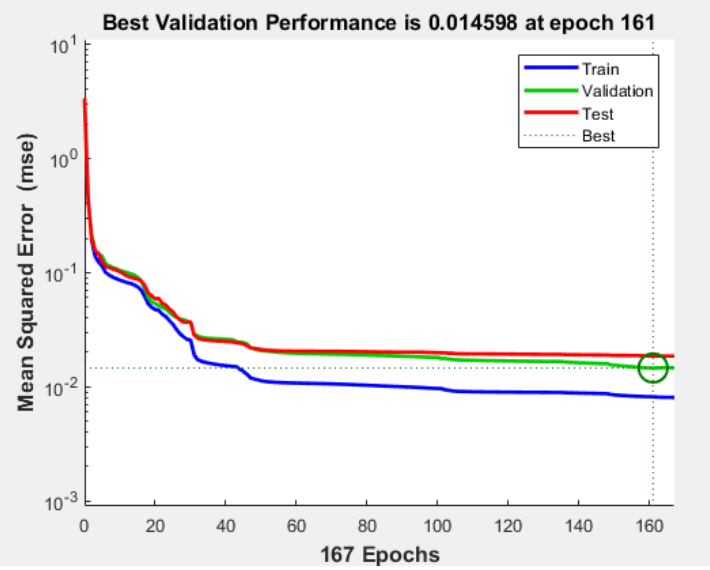

Figure 4.12. Learning curve of the MLPNN

Table 4.8. Performance metrics for ANN regression

Regression Metric	Training	Validation	Testing	All
R	0.982	0.979	0.977	0.980
R^2	0.964	0.958	0.954	0.960
MSE	0.009	0.0145	0.0185	0.042
RMSE	0.094	0.120	0.1360	0.2049

4.4 Machine Learning Classification

This section discusses the results for PTS classification using ANN and support vector machine (SVM).

4.4.1 PTS classification using ANN

The confusion matrix, for the classification of S_i, is shown in Figure 4.13. As evident, *CA* for the confusion matrix (all samples) is very high (≈98%). Moreover, as evident from

Figure 4.14, the ROC curves are very accurate (*AUC* >0.99 for all cases). Moreover, Figure 4.15 illustrates the error histogram (distribution of classification error for training, validation, and testing sets) obtained for the ANN classifier. It further verifies the excellent accuracy performance of the designed classifier. The values of various classification metrics are summarized in Table 4.9. As evident, values for metrics *CA*, *F1*, and *AUC* are in high accuracy range (≈0.97-0.99), and C_E is quite small (0.019). The training time for the ANN classifier was only 0.01 *s*. Thus, it can be inferred that the trained ML algorithm can rapidly classify the PTS status, S_i, with a high accuracy (≈98%).

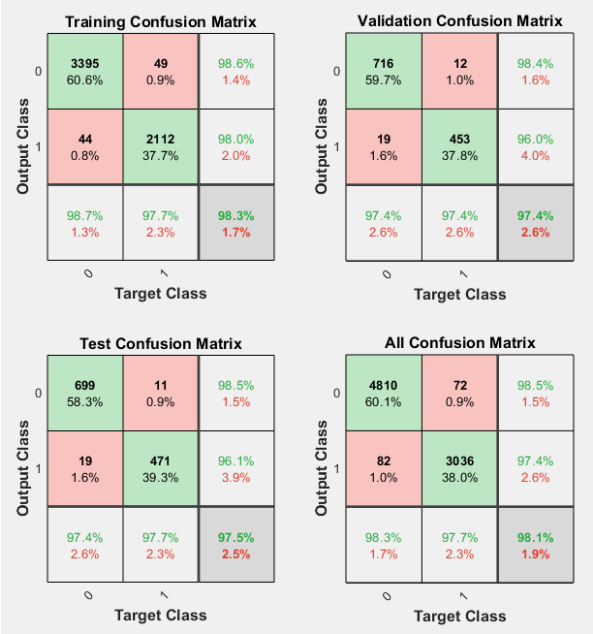

Figure 4.13. Confusion matrix for transient stability classification performance assessment

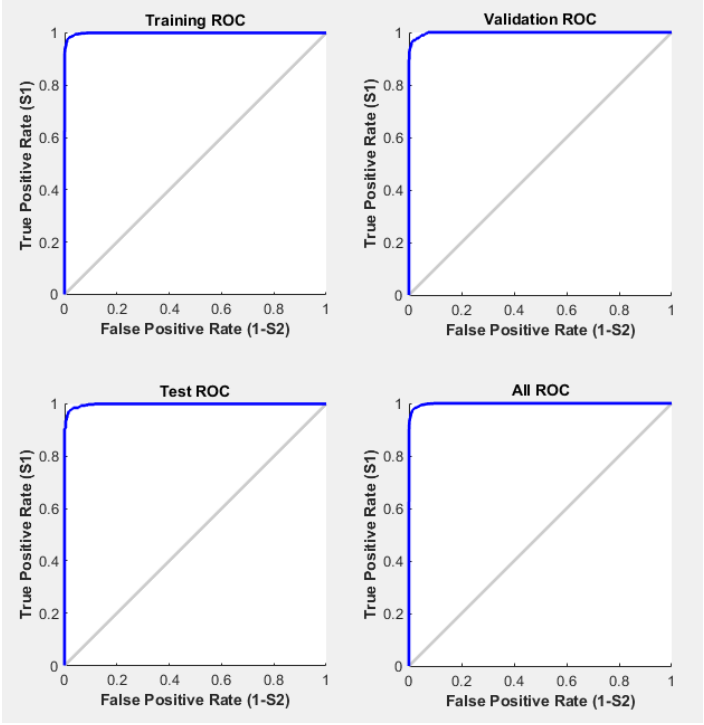

Figure 4.14. ROC curve for transient stability classification performance assessment

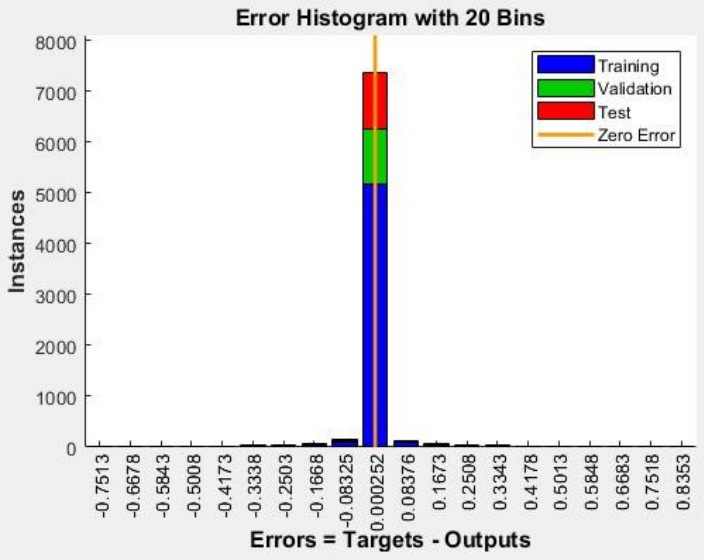

Figure 4.15. Histogram for classification error

Table 4.9. ANN classification performance assessment

Classification Metric	Training	Validation	Testing	All
CA	0.983	0.974	0.975	0.981
CE	0.017	0.026	0.025	0.019
F1	0.978	0.967	0.969	0.975
AUC	0.999	0.991	0.992	0.998

4.4.2 ANN Sensitivity Analysis

As mentioned before, logsig and tansig functions were used as activation functions, for hidden layers and output layer, respectively. The *CA* values for different activation functions (with 20 hidden neurons) are illustrated in Table 4.10. As evident, the choice of activations function, used in this work, provide the best results. Moreover, *CA* values for different data divisions (with 20 hidden neurons) are shown in Table 4.11. This verifies the fact that more training data leads to greater classification accuracy. However, one must be careful when selecting the ratios for data division. Using a small percent (≈5%) for testing and validation data may not be suitable for a small dataset, whereas, using a comparatively small percent (≈60%) of training data may generate superior results, when the dataset consider is comparatively small. Fewer training samples lead to large variance (inability to generalize to new data) in training performance, whereas, fewer testing samples lead to large variance in testing performance. Therefore, there is no formal rule for this data division, and it majorly depends on the required accuracy and the amount of data available (size of dataset).

Table 4.10. *CA* values for different activation functions

Hidden layer	Output layer	*CA* (All)
logsig	linear	0.931
tansig	linear	0.952
linear	logsig	0.598
linear	tansig	0.712
logsig	tansig	0.981
tansig	logsig	0.961

Table 4.11. *CA* values for different data divisions

Sr. No.	Training (%)	Validation (%)	Testing (%)	*CA* (All)
1	60	20	20	0.976
2	70	15	15	0.981
3	80	10	10	0.988
4	90	5	5	0.994

4.4.3 PTS Classification Using SVM

The confusion matrix and the ROC curve (for testing data), for the classification of S_i, are shown in Figure 4.16 and Figure 4.17, respectively. From Figure 4.16, it is evident that *CA* for the confusion matrix is very high, i.e., approximately 97% (59.46% + 37.23%). Moreover, as evident from Figure 4.17, the ROC curve is very accurate (*AUC* >0.99). The values of various classification metrics are summarized in Table 4.12. As evident, values for *CA* and *AUC* are in the fairly high accuracy range (>0.95), and *CE* is quite small (0.033). Once trained, the SVM classifier can be directly used to classify S_i. The training time for

the SVM classifier was only 0.03 *s*. Thus, it can be inferred that the trained SVM algorithm can rapidly classify the PTS status, S_i, with a high accuracy (≈97%).

4.4.4 SVM Sensitivity Analysis

As mentioned before, the value of *K* for *K*-fold cross-validation used was 5. To verify that it is indeed the best value, a sensitivity analysis was performed. The SVM classifier was trained for various values of *K*, and the corresponding *CA* values were determined. The results obtained are shown in Figure 4.18. As evident, increasing *K* beyond 5 does not alter the *CA*. Hence, *K*=5 is a good choice for *K*-fold cross-validation, for this work. This also validates the fact that *K*=5 and *K*=10 are generally the most commonly used values for a *K*-fold cross-validation procedure [14]. Moreover, for *K*=5, the values of *CA*, for different kernel functions, are shown in Table 4.13. As evident, Gaussian radial basis function (RBF) kernel has the highest accuracy. This also validates the reason of Gaussian RBF kernel being the most commonly used kernel function for SVM classification [15].

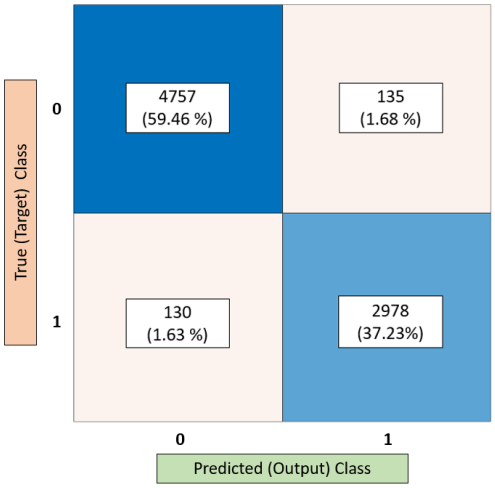

Figure 4.16. Confusion matrix for transient stability classification performance assessment

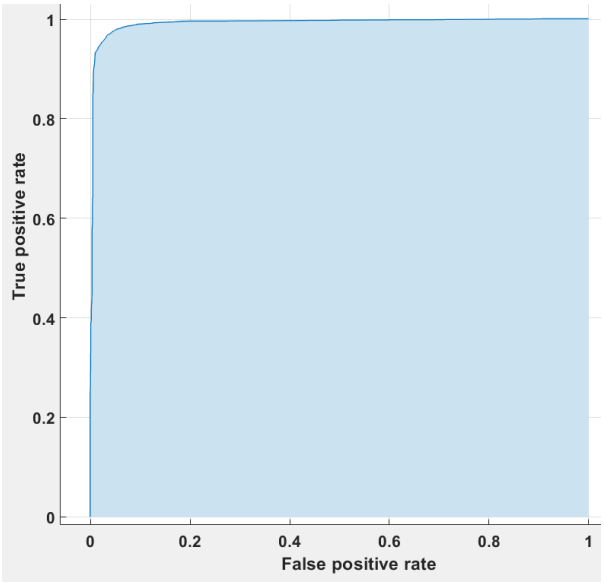

Figure 4.17. ROC curve for transient stability classification performance assessment

Table 4.12. SVM performance assessment using various classification metrics

Classification Metric	Value
CA	0.967
CE	0.033
F1	0.957
AUC	0.991

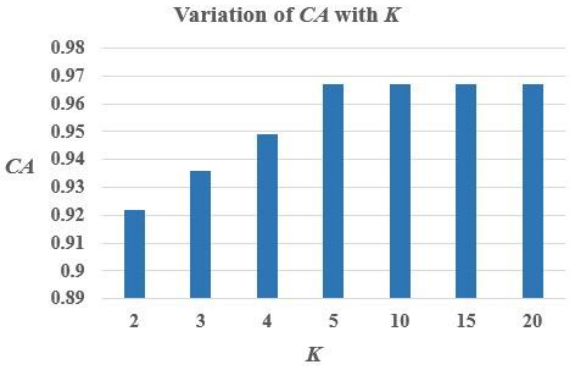

Figure 4.18. Variation of *CA* with *K*

Table 4.13. Variation of *CA* for different kernel functions

Kernel function	***CA***
Linear	0.872
Polynomial (order 2)	0.916
Polynomial (order 3)	0.829
Gaussian RBF	0.967

Thus, to sum up, the proposed ANN approach can be used to predict (regression) the decision-making variable, DV_m, incorporating various uncertain factors (load, fault type, fault location, and FCT), with a fairly high accuracy. Moreover, proposed ANN and SVM approaches can be used to predict (classify) the PTS status, incorporating various uncertain factors (system load, fault type, fault location, and FCT). These approaches have an edge over the conventional approaches, as they are computationally efficient, without sacrificing the accuracy too much. It is strongly believed that the proposed approaches can drastically contribute to advancing the existing methods of online dynamic security assessment (DSA). Comparing the performance of SVM and ANN for PTS classification, it is evident, values for all metrics for ANN surpass SVM. Therefore, assuming an enhanced and more accurate requirement for system dynamic security, ANN should be preferred over SVM, for an online transient stability prediction, incorporating various uncertainties of load, fault type, fault location, and FCT.

It must be mentioned that the demonstrated ML algorithms used in this research are system-specific and, although, they performed quite well, for the IEEE 14-bus system, it does not assure that they will still perform the same for other systems with different operating

conditions and characteristics. As the proposed approaches are scalable and, can be extended to any test system; therefore, ample testing and validation of the proposed approaches must be conducted on other standard test systems, before reaching a generic conclusion, regarding classification performance assessment of ANN and SVM.

Moreover, the proposed approaches using ML does not consider all possible risks, i.e., risks related to limited number of operating conditions (based on margin of error and confidence level), and assessment of cost impacts of transient instability. Moreover, generic restrictions exist for ML-based approaches, for instance, the training database and ML models must be reorganized when the PDFs of the input random variables, and the network topology varies over time, and consequently, the number of transient stability simulations required for training may be greater than that estimated for a fixed topology. An additional limitation regarding ANN is that there is no rule for determining the structure of ANN as it needs to be determined through repeated trial and error approach. Moreover, it requires processors with parallel processing power, according to their structure [16]. Also, the best ML approach may change depending on the application [17]. An additional limitation regarding SVM is that it is sensitive to noise (target classes overlap) and outliers (target classes deviate significantly from the rest of the classes), and consequently, does not give a good performance. Moreover, choosing the optimal kernel function is not straightforward, and may require several optimization simulations [18].

4.5 Tradeoff between *N*, *CA*, and Training Time

The trade-off between the three parameters: (1) total training samples (*N*), (2) *CA*, and (3) offline training time is a significant matter in ML modeling which needs to be discussed. The results of sensitivity analysis for this trade-off, for the proposed ANN and SVM models, are shown in Table 4.14 and Table 4.15, respectively. As evident, the classification accuracy, *CA*, and training time, increases with increasing *N*. The training time roughly doubles by doubling *N*; however, the change in *CA* is not that drastic. This type of sensitivity analysis must be performed by the power utilities, and consequently, the parameters (*CA*, *N*, training time) must be chosen carefully based on the specific requirements of utilities.

Table 4.14. Variation of *CA* values and training time with *N* (for ANN)

N	***CA* (All)**	**Training time (s)**
8,000	0.981	0.01
16,000	0.988	0.04
32,000	0.996	0.09
64,000	0.999	0.17

Table 4.15. Variation of *CA* values and training time with *N* (for SVM)

N	***CA***	**Training time (s)**
8,000	0.967	0.03
16,000	0.975	0.07
32,000	0.983	0.15
64,000	0.992	0.28

4.6 Validation Using Sensitivity Analysis

Validation means to quantify the confidence in the predictive capability of a code for a given application through comparison of calculations with a set of experimental data [19]. It is significant to validate the simulation results for a power system. Performing a comparison of simulation output data with a real power system is the best way to validate a model. A real time digital simulator also provides a good alternative because the simulator functions in real time, the power system algorithms are computed rapidly enough to unceasingly produce output conditions that realistically represent conditions in a real system. These approaches are hard to implement due to cost constraints, therefore, sensitivity analysis is often employed, as a validation technique in research literature. Sensitivity analysis characterizes how the uncertainty in the output of a mathematical model or system (numerical or otherwise) can be divided and allocated to numerous sources of uncertainty in its inputs [20]. A sensitivity analysis, essentially, regulates how different values of an independent variable affect a dependent variable. Various research [19, 21-25] have used this approach to validate different mathematical models and results. Thus, this work also used it to validate the findings. It must be noted that the sensitivity analysis is performed only on the base case, however, it can easily be extended to other cases to verify its aptness.

One of the simplest and most commonly used sensitivity analysis approaches is that of changing one-factor-at-a-time (OAT), to determine what effect this produces on the output. OAT usually involves two main steps [26]: (1) moving one input variable, keeping others at their baseline (nominal) values, then, (2) returning the variable to its nominal

value, then repeating for each of the other inputs in the same manner. In the first step, several variables were changed (one at a time) to observe the impact on the output, which in this research, is the average risk index, R_A. The PDF of other variables was kept fixed (the description of the PDFs can be found in Chapter 3), as illustrated in Table 4.16. The results (obtained for 2,401 MC samples) are shown in Tables 4.17-4.19, and, graphically, in Figures 4.19-4.21.

Table 4.16. Description of variables for performing sensitivity analysis

Case No.	Variable changed	Variables kept fixed	Description of variable (kept fixed)
1	Fault type	FCT, fault location	FCT (normal) Fault location (uniform)
2	Fault location	FCT, fault type	FCT (normal) Fault type (discrete)
3	FCT	Fault type, fault location	Fault type (discrete) Fault location (uniform)

Table 4.17. Sensitivity analysis (Case 1)

Fault Type	R_A (%)
LG	0.0064
LLG	0.0012
LL	0.00088
LLL	0.00042

Table 4.18. Sensitivity analysis (Case 2)

Fault Location (%)	R_A (%)
10	0.0056
30	0.0047
50	0.0039
70	0.0045
90	0.0054

Table 4.19. Sensitivity analysis (Case 3)

FCT	R_A (%)
0.7	0.0035
0.8	0.0042
0.9	0.0054
1.0	0.0063
1.1	0.0068
1.2	0.0075
1.3	0.0075

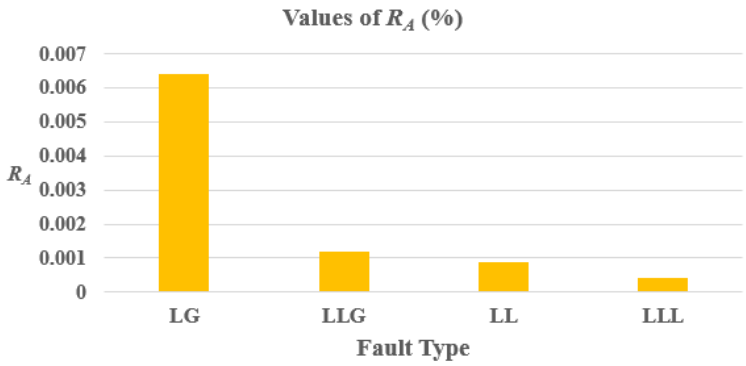

Figure 4.19. Sensitivity analysis (Case 1)

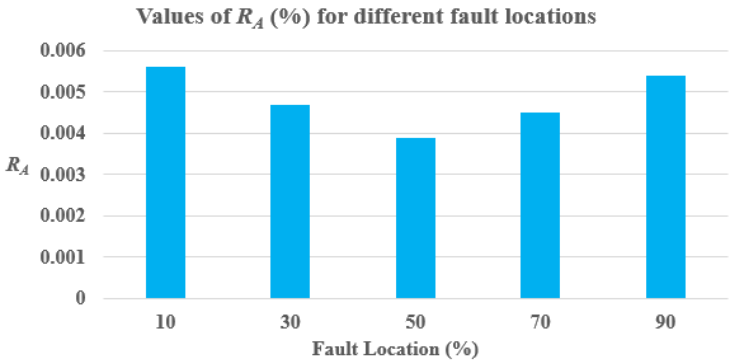

Figure 4.20. Sensitivity analysis (Case 2)

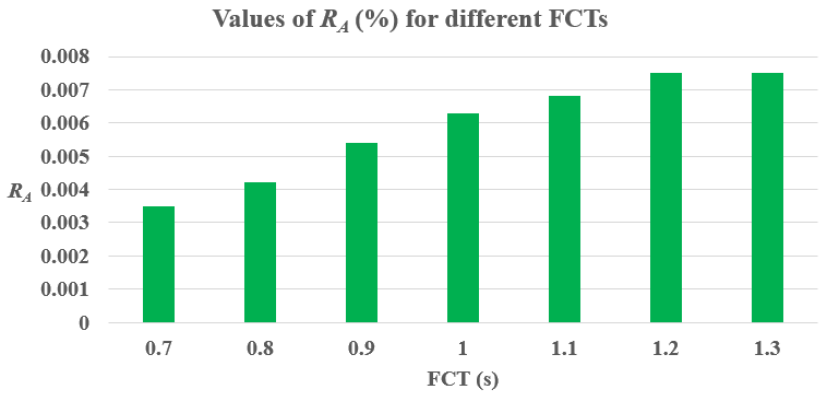

Figure 4.21. Sensitivity analysis (Case 3)

The actual results, based on MC simulation (MCS), for 6 random samples are shown in Table 4.20. Comparing these results and trends with the sensitivity analysis results, obtained above, it can easily be verified that actual results are valid. Here, R_i indicates the transient instability risk for the i^{th} *MC* sample.

Table 4.20. Actual results based on MCS

Sr. No.	Fault Type	Fault Location (%)	FCT(s)	R_i (%)
1	LG	41.5	1.08	0.00462
2	LG	4.6	0.88	0.00123
3	LL	81.9	0.95	0.00151
4	LLG	93.6	1.08	0.00867
5	LG	76.9	1.01	0.00314
6	LLL	87.2	1.04	0.00397

4.7 References

[1] Power Systems Test Case Archive. [Online].

Available: http://www.ee.washington.edu/research/pstca/pf14/pg_tca14bus.htm

[2] P. K. Iyambo and R. Tzoneva, "Transient stability analysis of the IEEE 14-bus electric power system," *AFRICON*, 2007, pp. 1-9.

[3] S. Kumar, A. Kumar, and N. K. Sharma, "Analysis of power flow, continuous power flow and transient stability of IEEE-14 bus integrated wind farm using PSAT," *International Conference on Energy Economics and Environment (ICEEE)*, 2015, pp. 1-6.

[4] N. Hashim, N. Hamzah, M. F. A. Latip, and A. A. Sallehhudin, "Transient stability analysis of the IEEE 14-bus test system using dynamic computation for power systems (DCPS)," *International Conference on Intelligent Systems Modelling and Simulation*, 2012, pp. 481-486.

[5] S. Sengupta, A. Kumar, and S. Tiwari, "Transient stability enhancement of a hybrid wind-PV farm incorporating a STATCOM," *3rd IEEE International Conference on Recent*

Trends in Electronics, Information & Communication Technology (RTEICT), 2018, pp. 1574-1580.

[6] R. R. Londero, J. P. A. Vieira, and C. M. Affonso, "Comparative analysis of DFIG based wind farms control mode on long-term voltage stability," in *Advances in Wind Power*. Rijeka, Croatia: InTech, 2012.

[7] R. Doherty and M. O. Malley, "A new approach to quantify reserve demand in systems with significant installed wind capacity," *IEEE Transactions on Power Systems*, vol. 20, no. 2, pp. 587-595, May 2005.

[8] T. K. A. Brekken, A. Yokochi, A. Jouanne, Z. Z. Yen, H. M. Hapke, and D. A. Halamay, "Optimal energy storage sizing and control for wind power applications," *IEEE Transactions on Sustainable Energy*, vol. 2, no. 1, pp. 69-77, Jan. 2011.

[9] J. Wu and B. Zhang, "Statistical distribution for wind power forecast error and its application to determine optimal size of energy storage system," *International Journal of Electrical Power & Energy Systems*, vol. 55, pp. 100-107, Feb. 2014.

[10] B. M. Hodge and M. Milligan, Wind power forecasting error distributions over multiple timescales, 2011. [Online]. Available:
https://www.nrel.gov/docs/fy11osti/50614.pdf

[11] DIgSILENT PowerFactory User Manual, DIgSILENT GmbH, 2018. [Online]. Available: https://www.digsilent.de/en/downloads.html

[12] M. H. Beale, M. T. Hagan, and H. B. Demuth, Neural Network Toolbox (User's Guide). [Online]. Available: https://www2.cs.siu.edu/~rahimi/cs437/slides/nnet.pdf

[13] Sample Size in Statistics (How to Find it): Excel, Cochran's Formula, General Tips. [Online]. Available:

https://www.statisticshowto.com/probability-and-statistics/find-sample-size/

[14] J. Brownlee, A gentle introduction to k-fold cross-validation. [Online]. Available: https://machinelearningmastery.com/k-fold-cross-validation/

[15] V. Vapnik, *The Nature of Statistical Learning Theory*. New York, NY, USA: Springer-Verlag, 1995.

[16] O. I. Abiodun, A. Jantan, A. E. Omalara, K. V. Dada, A. M. Umar, and M. U. Kiru, "Comprehensive review of artificial neural network applications to pattern recognition," *IEEE Access,* vol. 7, pp. 158820-158846, 2019.

[17] J. L. Cremer and G. Strbac, "A machine-learning based probabilistic perspective on dynamic security assessment," *International Journal of Electrical Power & Energy* Systems, vol. 128, pp. 1-15, Jun. 2021.

[18] S. Karamizadeh, S. M. Abdullah, M. Halimi, J. Shayan and M. J. Rajabi, "Advantage and drawback of support vector machine functionality," *International Conference on Computer, Communication, and Control Technology*, 2014, pp. 63-65.

[19] T. G. Trucano and L. P. Swiler "Calibration, validation, and sensitivity analysis: what's what," *Reliability Engineering & System Safety,* vol. 91, no. 10–11, pp. 1331-1357, Oct.–Nov. 2006.

[20] A. Saltelli, M. Ratto, D. Gatelli, and T. Andres, *Global Sensitivity Analysis: The Primer*. John Wiley & Sons, 2008.

[21] E. D. Smith, F. Szidarovszky, W. J. Karnavas, and A. T Bahill, "Sensitivity analysis, a powerful system validation technique," *The Open Cybernetics and Systemics Journal*, vol. 2, pp. 39-56, 2008.

[22] Power system model validation, a white paper by the NERC Model Validation Task Force of the Transmission Issues Subcommittee, Dec. 2010.

[23] J. D. Salciccioli and D. C. Marshall, *Sensitivity Analysis and Model Validation*, Springer, 2016.

[24] S. Cole and F. Promel, "Tools for validation and calibration of very large power system models," *IEEE International Energy Conference (ENERGYCON)*, 2018, pp. 1-6.

[25] R. Huang, R. Diao, and Y. Li, "Calibrating parameters of power system stability models using advanced ensemble Kalman filter," *IEEE Transactions on Power Systems*, vol. 33, no. 3, pp. 2895-2905, May 2018.

[26] J. Murphy, D. Sexton, and D. Barnett, "Quantification of modelling uncertainties in a large ensemble of climate change simulations," *Nature*, vol. 430, pp. 768–772, Aug. 2004.

CHAPTER 5

CONCLUSION AND FUTURE RESEARCH

This chapter concludes the research presented in this dissertation and provides numerous recommendations for relevant future research work. Moreover, generic recommendations for promoting risk-based decision making in power systems are outlined.

5.1 Conclusion

Power system transient stability is an integral part of power system planning and operation. Traditionally, it has been assessed using deterministic approach. Also, current North American Electric Reliability Corporation (NERC) reliability standards are deterministic and do not include any probabilistic methods. With the increasing system uncertainties, environmental pressures of incorporating green energy, and widespread electricity market liberalization (deregulation), there is a strong need to incorporate probabilistic analysis in transient stability evaluation. Moreover, conventional approaches to assess transient stability are time consuming and hence, are not suitable for online application. Thus, this dissertation presented risk-based machine learning (ML) approaches for probabilistic transient stability (PTS). The proposed approaches are based on time-domain simulation approach and ML. Time-domain simulations, using DIgSILENT PowerFactory, were used for gathering the training data, for the two ML algorithms, i.e., artificial neural network (ANN) and support vector machine (SVM). MATLAB was used to model the ML algorithms, and the IEEE 14-bus test system was used to test and validate the effective of

the proposed approaches. Numerous parameters, such as system load, fault type, fault location, and fault clearing time (FCT) were considered as random variables.

In the first part, a probabilistic cost benefit analysis was applied for decision making, regarding replacement of circuit breakers (CBs). The results obtained showed that risk-based approach has an edge over deterministic approach in the sense that it considers the uncertainty in the relevant associated variables, in the form of relevant probability density functions, as opposed to the deterministic approach (which considers the input variables as specific worst-case scenarios). The analysis also demonstrated that replacement of synchronous generators with doubly fed induction generators (DFIGs) deteriorated the transient stability. Moreover, the addition of DFIGs to an existing network improved the transient stability by decreasing the average risk of the system. The CBs were ranked for line and bus faults, based on R_A value. This ranking can be very useful for system planners for decision-making. Moreover, probability of instability (for individual faults) and system instability was evaluated was line and bus faults. The results obtained reinforces the fact that bus faults are more severe than line faults. The variation of network topology (using single line contingency) drastically impacts the value of P_{SYS} and R_A, indicating the significance of its inclusion in the PTS framework. Moreover, ANN-based regression was applied to improve computation efficiency, for predicting the value of benefit-cost ratio (BCR), for each line in the network.

In the second part, ANN and SVM were applied for online PTS prediction (classification). In addition to uncertain system load conditions, various uncertain factors such as faulted line, fault type, fault location, and FCT were considered. Time-domain simulations were used to create the data required for training the ML models. The *TSI* was used as the

indicator for the PTS status. The results obtained for the IEEE 14-bus system demonstrated that both ANN and SVM can rapidly estimate the transient stability, considering uncertainties, with a fairly high accuracy, however, ANN outperformed SVM, as its classification performance metrics were found to be superior.

The results obtained for the proposed approaches indicated a strong possibility of ANN and SVM, for online dynamic security assessment (DSA) because of their enhanced computational efficiency and high accuracy. Based on the work presented in this research and recent advancements in the domain, it is firmly believed that ANNs and SVMs have more to offer in the domain of power system security and stability. Moreover, the proposed approaches are universal in the sense that they are scalable and hence, can be extended to any system size, topology, and can incorporate various other ML approaches.

It is significant to recognize that the purpose of using the PTS assessment is not to replace the traditional deterministic criteria, which have been utilized for years in the power sector. However, the probabilistic method provides a sophisticated complement to enhance the transient stability analysis in utilities. The comprehensive acceptance of the probabilistic method is a gradual process. Jumping from an established approach directly to an evolving one would result in confusion and loss of credibility. Therefore, utilities should start to get used to the probabilistic perspective by gradually implementing probabilistic approaches into the power system decision-making processes.

5.2 Recommendations for Future Research

This section presents some recommendations for future research. They are outlined below.

Ensemble Machine Learning Approaches

Ensemble methods employ multiple learning algorithms to obtain an improved predictive performance than could be obtained from any of the constituent individual learning algorithms [1-2]. Their main advantages are performance (an ensemble allows improved predictions and can achieve improved performance than any individual contributing model) and robustness (an ensemble decreases the dispersion of the predictions and model performance). However, selecting the right ensemble approach is an exigent task, as it depends on a lot of factors, including input data characteristics, available computational resources and required prediction performance. For instance, models with high variance are likely to benefit from using bagging, whereas highly biased models, it is better to use them boosting. Moreover, it is hard to estimate the correlation between the individual algorithms employed in an ensemble. Ensemble learning approaches for online transient stability assessment are still being researched, and this is still an open area of research. Some common ensemble algorithms include bagging, boosting, Bayes optimal classifier, and Bayesian model combination.

Transient Stability of Integrated Gas and Power System

With the quick development of natural gas fired units worldwide, the interdependency of natural gas system and power system has substantially augmented. This integrated energy system, constitutes of natural gas system and power system, has benefits of low harmful emissions. Conventionally, power systems and natural gas systems have been planned discretely [3-4]. As these two energy systems are increasingly becoming interconnected, it is valuable to model their joint expansion planning, which also includes transient stability assessment. There is a need to propose a unified approach for incorporating natural gas systems in the transient stability analysis, including, but not limited to, wells, pipelines,

compressors, liquefied natural gas terminals and gas storage. In addition, economic aspects of gas, such as gas price and gas contracts, gas supply constraints, limited transmission capacity of pipeline network, etc. could also affect the natural gas supply adequacy and hence, the long-term transmission reliability and planning.

Validation Challenge

One of the real problems that the simulation modeling faces is how to validate its models [5]. A valid simulation model is a functionally accurate representation of the real system. If the real system is hard to create, the simulation model and its output data can be verified with solving the analytics model, if all conditions in the simulation model are applied to the analytical model. Performing a comparison of simulation output data with a real system is the best way to validate the proposed model. In this approach, the simulation model and its outputs are compared to the real system and its output. Although, this process of validation seems simple; however, it may present some hurdles in implementation process, as real wind and diesel generators will be required to model the power system.

Network Topology

Another key challenge is to incorporate changes of the network topology in the ML approaches. The security of the power system is highly related to the topology of the system [6], and changes in the network topology can happen frequently for various reasons, such as for maintenance purposes or unexpected component failures [7]. The impact of changing topology on transient stability rules is a substantial challenge. This is because if topological changes are not considered and transient stability rules are trained only for one specific topology, the resulting assessment of the transient stability (and security) using these rules may provide erroneous predictions, which ultimately leads to incorrect decision-making.

Therefore, a key future direction is to enhance the ML workflow, by considering changes in the system topology.

Power System Resilience: Weather and Cyber Attacks

The resilience of a power system, in general, its ability to respond to unforeseen, high impact low frequency events, and its ability to rapidly and efficiently restore to its pre-event operation state [8]. Uncertainties introduced by unforeseen, high risk events, such as natural disasters (hurricanes, earthquakes, floods, etc.), extreme weather (ice storms, heat waves, high winds, etc.), cyber-attacks, etc. can prove lethal to the power system, and hence, can have a detrimental impact on the transient stability. The central challenge in this regard is to model these phenomena, and their allied impact accurately enough, to be incorporated in a risk-based transient instability formulation.

Other Aspects of Stability

Although, classification of power system stability is an efficient way to deal with the complexities of the problem, the overall stability of the system should always be the broader aspect. Solutions to stability problems of one category should not be at the expense of another [9]. Therefore, it is indispensable to consider other aspects of the stability phenomenon, and at each aspect, from more than one viewpoint (under both small and large disturbances). As this work focused only on transient stability, the proposed risk-based approach can be extended to voltage and frequency stability (for both large and small disturbances), and consequently, suitable risk indices can be established.

Larger Test System

The study can be repeated for a larger test system. The proposed framework should be applied on large-scale power system, where the system can be divided into multiple

regions, to improve the accuracy of ANN and SVM algorithms [10], to approximate the enclosed transient stability aspects, in each region. This should be accomplished by establishing a technique to consider the impact of stability of each region on the other regions during implementation of ANN and SVM.

5.3 Recommendations for Risk-based Decision-Making

Given the unprecedented changes in the electric power industry, and the pressure to ensure system reliability at a minimum cost, transmission planning is becoming more complex than ever. Risk-based planning has a momentous potential to provide a better decision-making framework for transmission planners. However, this is still an area of active research and substantial gaps remain in terms of developing a robust probabilistic framework that planners can routinely use. Some significant recommendations in this regard are outlined below [11-12].

Closer Coordination with NERC

All the existing NERC transmission planning standards are deterministic. However, recently NERC has shown interest in considering probabilistic approaches in transmission planning and organized multiple workshops in Eastern Interconnection as well as in WECC on risk-based planning. NERC and states, in collaboration with other stakeholders can collaborate, and develop a long-term vision for establishing risk-based planning framework. Also, there are no well-defined risk criteria and indices that can be largely accepted and enforced. This is an important domain which requires noteworthy effort and coordination.

Greater Awareness about Uncertainties and Risks

Transmission planning is a demanding activity involving federal, state, local, and private entities. All the stakeholders may not be equally aware of various risks and uncertainties that are going to impact transmission planning in the future. Also, they may have different views about the future. It is also recommended that states get more actively involved in regional transmission planning processes initiated by Federal Energy Regulatory Commission (FERC) or regional planning organizations for improved alliance and wider outlook on risk-based decision-making.

Promote Research Efforts on Risk-based Planning

As mentioned before, risk-based planning requires active research and industry participation for its wider adoption. States can encourage research efforts and work closely with research organizations, universities, national labs, commercial software developers, and utility industry to ensure that research needs are addressed, and practical solutions are proposed. The research community needs to work more closely with the industry to clearly demonstrate the benefits of probabilistic methods. The industry needs to clearly communicate inadequacies in deterministic methods and areas that probabilistic methods can be most productive. Cooperative efforts of the research and industry communities are the need of the hour to face the challenges in terms of both comprehending and practical application.

5.4 References

[1] X. Chen, L. Zhang, Y. Liu, and C. Tang, “Ensemble learning methods for power system cyber-attack detection,” *IEEE 3rd International Conference on Cloud Computing and Big Data Analysis (ICCCBDA)*, 2018, pp. 613-616.

[2] R. Polikar, “Ensemble based systems in decision making,” *IEEE Circuits and Systems Magazine*, vol. 6, no. 3, pp. 21-45, 2006.

[3] C. Unsihuay, J. W. Lima, and A. C. Z. de Souza, “Integrated power generation and natural gas expansion planning,” *IEEE Lausanne Power Tech*, 2007, pp. 1404-1409.

[4] Y. Zhang, Y. Hu, J. Ma, and Z. Bie, “A mixed-integer linear programming approach to security-constrained co-optimization expansion planning of natural gas and electricity transmission systems,” *IEEE Transactions on Power Systems*, vol. 33, no. 6, pp. 6368-6378, Nov. 2018.

[5] E. Foruzan, “Control and monitoring strategies of smart grids using artificial intelligent methods,” *Ph.D. Dissertation*, The University of Nebraska–Lincoln, USA, 2019.

[6] K. A. Loparo and F. Abdel-Malek, “A probabilistic approach to dynamic power system security,” *IEEE Transactions on Circuits and Systems*, vol. 37, no. 6, pp. 787–798, 1990.

[7] B. Donnot, I. Guyon, M. Schoenauer, P. Panciatici, and A. Marot, “Introducing machine learning for power system operation support,” Sep. 2017. [Online]. Available: https://arxiv.org/pdf/1709.09527.pdf

[8] M. Mahzarnia, M. P. Moghaddam, P. T. Baboli, and P. Siano, “A review of the measures to enhance power systems resilience,” *IEEE Systems Journal*, vol. 14, no. 3, pp. 4059-4070, Sep. 2020.

[9] P. Kundur et al., “Definition and classification of power system stability,” *IEEE Transactions on Power Systems*, vol. 19, no. 3, pp. 1387-1401, Aug. 2004.

[10] A. Hoballah, “Power system stability assessment and enhancement using computational intelligence,” *Ph.D. Dissertation*, University of Duisburg-Essen, Germany, 2011.

[11] A white paper on the incorporation of risk analysis into planning processes, EPRI Report, 2015. [Online]. Available: https://pubs.naruc.org/pub.cfm?id=536DCF19-2354-D714-5117-47F9BA06F062

[12] NERC IVGTF Task 1.6: Probabilistic Methods, Technical Report, Jul. 2014. [Online]. https://www.nerc.com/comm/PC/Integration%20of%20Variable%20Generation%20Task%20Force%20I1/IVGTF%20Task%201-6_09182014.pdf

APPENDIX A

A.1 IEEE 14 bus data (input data)

A.1.1 Load demand

Load	Bus	P in MW	Q in Mvar
Load_0002	Bus_0002	21.7	12.7
Load_0003	Bus_0003	94.2	19.0
Load_0004	Bus_0004	47.8	-3.9
Load_0005	Bus_0005	7.6	1.6
Load_0006	Bus_0006	11.2	7.5
Load_0009	Bus_0009	29.5	16.6
Load_0010	Bus_0010	9.0	5.8
Load_0011	Bus_0011	3.5	1.8
Load_0012	Bus_0012	6.1	1.6
Load_0013	Bus_0013	13.5	5.8
Load_0014	Bus_0014	14.9	5.0

A.1.2 Generator dispatch

Generator	Bus	P in MW	Q in Mvar
Gen_0001	Bus_0001	N.A.	N.A.
Gen_0002	Bus_0002	40.0	N.A.
Gen_0003	Bus_0003	0.0	N.A.
Gen_0006	Bus_0006	0.0	N.A.
Gen_0008	Bus_0008	0.0	N.A.

A.1.3 Generator controller settings

Generator	Bus Type	Voltage in p.u.	Minimum capability in MVA	Maximum capability in MVA
Gen_0001	Slack	1.060	N.A.	N.A.
Gen_0002	PV	1.045	-40.0	50.0
Gen_0003	PV	1.010	0.0	40.0
Gen_0006	PV	1.070	-6.0	24.0
Gen_0008	PV	1.090	-6.0	24.0

A.1.4 Data of lines given in [1] based on 100 MVA

From Bus	To Bus	r in p.u.	x in p.u.	$q_{c/2}$ in p.u.	b in p.u.
1	2	0.01938	0.05917	0.0264	0.0528
1	5	0.05403	0.22304	0.0246	0.0492
2	3	0.04699	0.19797	0.0219	0.0438
2	4	0.05811	0.17632	0.0187	0.0374
2	5	0.05695	0.17388	0.0170	0.0340
3	4	0.06701	0.17103	0.0173	0.0346
4	5	0.01335	0.04211	0.0064	0.0128
6	11	0.09498	0.19890	0.0000	0.0000
6	12	0.12291	0.25581	0.0000	0.0000
6	13	0.06615	0.13027	0.0000	0.0000
9	10	0.03181	0.08450	0.0000	0.0000
9	14	0.12711	0.27038	0.0000	0.0000
10	11	0.08205	0.19207	0.0000	0.0000
12	13	0.22092	0.19988	0.0000	0.0000
13	14	0.17093	0.34802	0.0000	0.0000

A.1.5 Data of lines given in the PowerFactory model

Line	From Bus	To Bus	Un in kV	R in Ω	X in Ω	B in μS
Line_0001_0002/1	1	2	132.0	6.753542	20.619560	151.5152
Line_0001_0002/2	1	2	132.0	6.753542	20.619560	151.5152
Line_0001_0005	1	5	132.0	9.414187	38.862490	282.3691
Line_0002_0003	2	3	132.0	8.187537	34.494280	251.3774
Line_0002_0004	2	4	132.0	10.125090	30.722000	214.6465
Line_0002_0005	2	5	132.0	9.922968	30.296850	195.1331
Line_0003_0004	3	4	132.0	11.675820	29.800270	198.5767
Line_0004_0005	4	5	132.0	2.326104	7.337246	73.4619
Line_0006_0011	6	11	33.0	1.034332	2.166021	0.0000
Line_0006_0012	6	12	33.0	1.338490	2.785771	0.0000
Line_0006_0013	6	13	33.0	0.720374	1.418640	0.0000
Line_0009_0010	9	10	33.0	0.346411	0.920205	0.0000
Line_0009_0014	9	14	33.0	1.384228	2.944439	0.0000
Line_0010_0011	10	11	33.0	0.893524	2.091643	0.0000
Line_0012_0013	12	13	33.0	2.405819	2.176693	0.0000
Line_0013_0014	13	14	33.0	1.861428	3.789938	0.0000

A.1.6 Data of transformers given in [1] based on 100 MVA, with rated voltages added in the PowerFactory model

Trans-former	From Bus	To Bus	Ur HV in kV	Ur LV in kV	r in p.u.	x in p.u.	Transformer final turns ratio
Trf_0004_0007	4	7	132.0	1.0	0.0	0.20912	0.978
Trf_0004_0009	4	9	132.0	33.0	0.0	0.55618	0.969
Trf_0005_0006	5	6	132.0	33.0	0.0	0.25202	0.932
Trf_0007_0008	7	8	11.0	1.0	0.0	0.17615	0.000
Trf_0007_0009	7	9	33.0	1.0	0.0	0.11001	0.000

A.2 IEEE 14 bus data (load flow results)

A.2.1 Results of buses

Name	Ul, Magnitude in kV	u, Magnitude in p.u.	U, Angle in deg
Bus_0001	139.92	1.060	0.00
Bus_0002	137.94	1.045	-4.98
Bus_0003	133.32	1.010	-12.72
Bus_0004	134.46	1.019	-10.32
Bus_0005	134.67	1.020	-8.78
Bus_0006	35.31	1.070	-14.22
Bus_0007	1.06	1.062	-13.37
Bus_0008	11.99	1.090	-13.37
Bus_0009	34.86	1.056	-14.95
Bus_0010	34.69	1.051	-15.10
Bus_0011	34.88	1.057	-14.80
Bus_0012	34.82	1.055	-15.08
Bus_0013	34.66	1.050	-15.16
Bus_0014	34.18	1.036	-16.04

A.2.2 Results of generators

Name	Active Power in MW	Reactive Power in Mvar
Gen_0001	232.39	-16.89
Gen_0002	40.00	42.40
Gen_0003	0.00	23.39
Gen_0006	0.00	12.24
Gen_0008	0.00	17.36

A.2.3 Results of lines

Name	Losses in MW	Reactive Losses in Mvar	Capacitive Loading in Mvar	Current in kA
Line_0001_0002/1	2.1474	3.6318	2.9246	0.3263
Line_0001_0002/2	2.1474	3.6318	2.9246	0.3263
Line_0001_0005	2.7638	6.0843	5.3248	0.3121
Line_0002_0003	2.3202	5.1494	4.6256	0.3067
Line_0002_0004	1.6770	1.1062	3.9824	0.2352
Line_0002_0005	0.9023	-0.8712	3.6260	0.1738
Line_0003_0004	0.3714	-2.6120	3.5598	0.1018
Line_0004_0005	0.5165	0.2990	1.3303	0.2727
Line_0006_0011	0.0547	0.1146	0.0000	0.1328
Line_0006_0012	0.0717	0.1492	0.0000	0.1336
Line_0006_0013	0.2115	0.4166	0.0000	0.3129
Line_0009_0010	0.0131	0.0348	0.0000	0.1123
Line_0009_0014	0.1168	0.2484	0.0000	0.1677
Line_0010_0011	0.0123	0.0288	0.0000	0.0678
Line_0012_0013	0.0062	0.0057	0.0000	0.0294
Line_0013_0014	0.0536	0.1091	0.0000	0.0980

APPENDIX B

B.1 DPL (DIgSILENT) Code for performing PTS and Computing BCR

```
set sLines,sBuses, sGens;
object rLine,oLine,oBus,Cost,opf,oGen,oLine1,FaultL;
int n,d,o,r,m,z,a,f,p,samp,k,l,FaultT,i,j;
int Psl,Psl1,Gu,Gu1, Vec;
int w,x,W,X,P,N,D,O,st,B2,B5,t,g;
double q,y,s,b,R,RTUNA,RTUNA1,RA,RA1, Faultloc;
double c,V,h,RTUNV,RTUNV1,RV,H,RV1,HH;
double u,v,F,RTUNF,RTUNF1,RF,U,RF1,UU;
double G,A,ss,sv,aft,afl,afct,Z,B,B1,B3,B4,B6,C;
double ACost,VCost,FCost,Pg,Pg1,Pg2,Pload,Pload1;
double
Cr,C2,T1,T2,CA,CB,BCR,BB,CC,APCT,VPCT,FPCT,MV1,MV2,MA1,MA2,MF1,MF2
;
double APCT1,VPCT1,FPCT1;
SetRandomSeed(1);
Allservice.Execute();
ClearOutput();
sGens=Gens.Get();
sBuses=Bus.Get();
sLines=Lines.Get();
VecL.Clear();
VecFT.Clear();
VecFL.Clear();
VecFCT.Clear();
n=1;
d=1;
k=1;
j=1;
```

```
x=0;
X=0;
D=0;
N=0;
RA=0;
samp=2401;  !Choose Number of Monte Carlo Samples
MatRUNA.Init(0,0,0);
Mat.Init(0,0,0);
MatRUNA1.Init(0,0,0);
MatACost.Init(0,0,0);
MatBus.Init(0,0,0);
MatPg.Init(0,0,0);
MatF.Init(0,0,0);
MatAPCT.Init(0,0,0);
MatFinal.Init(0,0,0);
MatFinal1.Init(0,0,0);
MatA1.Init(0,0,0);
MatA2.Init(0,0,0);
for(z=1;z<=samp;z+=1)
{
c=0;
RTUNA=0;
RTUNA1=0;
oLine=sLines.First();
r=RndUnifInt(1,16);   !Select Fault Line Randomly
if(r=1)
{oLine=sLines.First();}
else
{
for(a=2;a<=r;a+=1)
```

```
{
oLine=sLines.Next();
}
 }
for(w=x;w<=x;w+=1)
{
VecL.Insert(w,oLine);          !Put faulted lines in vector
}
x+=1;
printf('\ccFaulted Line is %s',oLine:loc_name);
ShcEvent:p_target=oLine;
y=fRand(0,0,0.7);
printf('\cc--------------value of y is %f----------', y);
if(y>0.15.and.y<=0.7)
{f=2;}
if(y<=0.05)
{f=0;}
if(y>0.05.and.y<=0.1)
{f=3;}
if(y>0.1.and.y<=0.15)
{f=1;}
ShcEvent:i_shc=f; !Select Fault Type
printf('\ccvalue of f is %d', f);
for(W=X;W<=X;W+=1)              !Putting fault types in vector
{
VecFT.Insert(W,f);
}
X+=1;
oLine:fshcloc=fRand(0,0,100);  !Select Fault Location
A=oLine:fshcloc;
```

```
printf('\ccfault location is %f', A);
for(P=N;P<=N;P+=1)              !Putting fault locations in vector
{
VecFL.Insert(P,A);
}
N+=1;
ShcEventEnd:time=fRand(1,1.9,0.1); !Select FCT
ShcEventEnd:p_target=oLine;
printf('\ccfault clearinf time is %f', ShcEventEnd:time);
for(O=D;O<=D;O+=1)              !Putting clearing time in vector
{
VecFCT.Insert(O,ShcEventEnd:time);
}
D+=1;
ComLdf.Execute(); !Execute Load Flow
oBus=sBuses.First();
oGen=sGens.First();
g=1;
for(oBus=sBuses.First();oBus;oBus=sBuses.Next())
{
C=oBus:m:Pload;
for(t=g;t<=g;t+=1)
{
MatPload.Set(t,1,C);
}
 g+=1;
}
printf('\ccLoad on each bus is %o', MatPload);
g=1;
for(oGen=sGens.First();oGen;oGen=sGens.Next())
```

```
{
C=oGen:pgini;
for(t=g;t<=g;t+=1)
{
MatPg.Set(t,1,C);
}
g+=1;
}
printf('\ccActive power is %o', MatPg);
g=1;
for(oBus=sBuses.First();oBus;oBus=sBuses.Next())
{
C=oBus:m:u;
for(t=g;t<=g;t+=1)
{
MatBus.Set(t,1,C);
}
 g+=1;
}
ComI.Execute(); !Execute Time-domain Simulation
Coms.Execute();
oGen=sGens.First();
printf('\ccMax Angle Diff for Fault at %s: %.2f', oLine:loc_name,oGen:c:dfrotx);
s=oGen:c:dfrotx;
Pg=0;
g=1;
Pg2=0;
m=0;
for(oGen=sGens.First();oGen;oGen=sGens.Next())
{
```

```
if (oGen:s:outofstep=1)
{
 m+=1;
Pg2=MatPg.Get(g,1);
Pg+=Pg2;
}
g+=1;
}
Psl=m;
printf('\cc Number of pole slip in Gen value of Pg after fault %d ----- %f',Psl,Pg);
if(m>=1)
{
s+=360;
}
for(p=n;p<=n;p+=1)
{
Mat.Set(n,1,s);
}
n+=1;
b=((360-s)/(360+s));  !Compute Maximum rotor angle difference
ACost=((200*Pg)+(60000*Psl)); !Compute transient instability cost
printf('\cc value of ACost  %f',ACost);
if(b<0)
{
if(f=0)
{
 R=0.05*0.000625*abs(b);
 }
if(f=1)
{
```

```
 R=0.15*0.000625*abs(b);
 }
if(f=2)
{
 R=0.70*0.000625*abs(b);
 }
if(f=3)
{
 R=0.10*0.000625*abs(b);
 }
RTUNA+=R;
RA+=RTUNA;
for(o=d;o<=d;o+=1)
{
MatRUNA.Set(d,1,RTUNA);
}
d+=1;
}
MatACost.Set(z,1,ACost);
}
printf('\ccValue of Maximum Angle Diff for each sample is %o', Mat);
printf('\ccvalue of Risk of angles for unstable samples is %o',MatRUNA);
printf('\ccvalue of RA is %f', RA);
G=RA;
if(G=RA)
{
d=0;
for(Z=1;Z<=samp;Z+=1)
{
ss=0;
```

```
st=Z-1;
oLine1=VecL.Get(st);
aft=VecFT.Get(st);
afl=VecFL.Get(st);
afct=VecFCT.Get(st);
ShcEvent:p_target=oLine1;     !same faulted line as above
ShcEvent:i_shc=aft;           !same fault type as above
oLine1:fshcloc=afl;           !same fault location  as above
ShcEventEnd:time=afct;        !ending fault at CCT
ShcEventEnd:p_target=oLine1;
ComI.Execute();
Coms.Execute();
oGen=sGens.First();
printf('\ccMax Angle Diff for Fault at %s: %.2f', oLine1:loc_name,oGen:c:dfrotx);
s=oGen:c:dfrotx;
m=0;
for(oGen=sGens.First();oGen;oGen=sGens.Next())
{
if (oGen:s:outofstep=1)
{
 m+=1;
}
}
if(m>=1)
{
s+=360;
}
b=0;
b=((360-s)/(360+s));
B=0;
```

```
Pg=0;
Pg1=0;
Psl=0;
Psl1=0;
ss=0;
APCT1=0;
if(b<0)
{
while(b<0)
{
m=0;
Pg=Pg1;
Psl=Psl1;
ss+=0.01;
ShcEvent:p_target=oLine1;     !same line as above
ShcEvent:i_shc=aft;           !same type of fault
oLine1:fshcloc=afl;           !same fault location
sv=afct-ss;
ShcEventEnd:time=sv;          !ending fault at CCT
APCT1=sv;
ShcEventEnd:p_target=oLine1;
B=b;
ComI.Execute();
Coms.Execute();
oGen=sGens.First();
s=oGen:c:dfrotx;
g=1;
for(oGen=sGens.First();oGen;oGen=sGens.Next())
{
if (oGen:s:outofstep=1)
```

```
{
 m+=1;
 Pg1=MatPg.Get(g,1);
}
g+=1;
}
Psl1=m;
if(m>=1)
{
s+=360;
}
b=0;
b=((360-s)/(360+s));
}
}
APCT=APCT1;
MatAPCT.Set(Z,1,APCT);
ACost=((200*Pg)+(60000*Psl));
R=0;
if(B<0)
{
if(aft=0)
{
 R=0.05*0.000625*abs(B);
 }
if(aft=1)
{
 R=0.15*0.000625*abs(B);
 }
if(aft=2)
```

```
{
 R=0.70*0.000625*abs(B);
 }
if(aft=3)
{
 R=0.10*0.000625*abs(B);
 }
 }
RTUNA1=R;
RA1+=RTUNA1;
d+=1;
for(o=d;o<=d;o+=1)
{
MatRUNA1.Set(d,1,RTUNA1);
}
MatACost.Set(Z,2,ACost);
}
printf('\cc--------value of Risk of angles for unstable samples is %o',MatRUNA1);
printf('\cc--------value of RA is %f', RA1);
}
if(G=RA)
{
printf('\cc--Matrix MatAcost %o',MatACost);
}
for(z=1;z<=samp;z+=1)
{
if(G=RA)
{
MA1=MatACost.Get(z,1);
MA2=MatACost.Get(z,2);
```

```
MatF.Set(z,1,MA1);
MatF.Set(z,2,MA2);
}
}
for(z=1;z<=samp;z+=1)
{
Vec=(z-1);
FaultL=VecL.Get(Vec);
Cr=MatF.Get(z,1);
T1=VecFCT.Get(Vec);
FaultT=VecFT.Get(Vec);
Faultloc=VecFL.Get(Vec);
if(G=RA)
{
T2=MatAPCT.Get(z,1);
}
CB=20000+3700/(T2-1);
BCR=Cr/2CB;
MatFinal.Set(z,1,Cr);    !Column Matrix for Cost of Risk
MatFinal.Set(z,2,T1);    !Column Matrix for FCT
MatFinal.Set(z,3,T2);    !Column Matrix for CCT
MatFinal.Set(z,4,CB);    !Column Matrix for single circuit breaker cost
MatFinal.Set(z,5,BCR);   !Column Matrix for BCR
MatA1.Set(z,1,FaultT);
MatA2.Set(z,1,Faultloc);
printf('\cc--Matrix Fault_type %o',MatA1); !Column Matrix for fault type
printf('\cc--Matrix Fault_location %o',MatA2); !Column Matrix for fault location
printf('\cc%o',FaultL); !Column Matrix for faulted line
}
printf('\cc--Matrix MatFinal %o',MatFinal); ! Final Matrix
```

APPENDIX C

C.1 MATLAB Codes

C.1.1 ANN Regression

```
% Solve an Input-Output Fitting problem with a Neural Network
% Script generated by Neural Fitting app
% This script assumes these variables are defined:
%   Reg - input data.
%   BCR - target data.
 x = Reg';
t = BCR';
 % Choose a Training Function
% For a list of all training functions type: help nntrain
% 'trainlm' is usually fastest.
% 'trainbr' takes longer but may be better for challenging problems.
% 'trainscg' uses less memory. Suitable in low memory situations.
trainFcn = 'trainlm';  % Levenberg-Marquardt backpropagation.
 % Create a Fitting Network
hiddenLayerSize = 20;
net = fitnet(hiddenLayerSize,trainFcn);
 % Setup Division of Data for Training, Validation, Testing
net.divideParam.trainRatio = 70/100;
net.divideParam.valRatio = 15/100;
net.divideParam.testRatio = 15/100;
 % Train the Network
[net,tr] = train(net,x,t);
 % Test the Network
y = net(x);
e = gsubtract(t,y);
performance = perform(net,t,y)
 % View the Network
```

```
view(net)
% Plots
% Uncomment these lines to enable various plots.
%figure, plotperform(tr)
%figure, plottrainstate(tr)
%figure, ploterrhist(e)
%figure, plotregression(t,y)
%figure, plotfit(net,x,t)
```

C.1.2 ANN Classification

```
% Solve a Pattern Recognition Problem with a Neural Network
% Script generated by Neural Pattern Recognition app
% This script assumes these variables are defined:
%   Class - input data.
%   Status - target data.
 x = Class';
t = Status';
% Choose a Training Function
% For a list of all training functions type: help nntrain
% 'trainlm' is usually fastest.
% 'trainbr' takes longer but may be better for challenging problems.
% 'trainscg' uses less memory. Suitable in low memory situations.
trainFcn = 'trainlm'; % Levenberg-Marquardt backpropagation.
 % Create a Pattern Recognition Network
hiddenLayerSize = 20;
net = patternnet(hiddenLayerSize, trainFcn);
 % Setup Division of Data for Training, Validation, Testing
net.divideParam.trainRatio = 70/100;
net.divideParam.valRatio = 15/100;
net.divideParam.testRatio = 15/100;
```

```
 % Train the Network
[net,tr] = train(net,x,t);
 % Test the Network
y = net(x);
e = gsubtract(t,y);
performance = perform(net,t,y)
tind = vec2ind(t);
yind = vec2ind(y);
percentErrors = sum(tind ~= yind)/numel(tind);
 % View the Network
view(net)
% Plots
% Uncomment these lines to enable various plots.
%figure, plotperform(tr)
%figure, plottrainstate(tr)
%figure, ploterrhist(e)
%figure, plotconfusion(t,y)
%figure, plotroc(t,y)
```

C.1.3 SVM Classification

```
function [trainedClassifier, validationAccuracy] = trainClassifier(trainingData)
% [trainedClassifier, validationAccuracy] = trainClassifier(trainingData)
% Returns a trained classifier and its accuracy. This code recreates the
% classification model trained in Classification Learner app. Use the
% generated code to automate training the same model with new data, or to
% learn how to programmatically train models.
%  Input
%
%      trainingData: A table containing the same predictor and response
%       columns as those imported into the app.
```

```
%
%  Output:
%      trainedClassifier: A struct containing the trained classifier. The
%       struct contains various fields with information about the trained
%       classifier.
%
%      trainedClassifier.predictFcn: A function to make predictions on new
%       data.
%
%      validationAccuracy: A double containing the accuracy in percent. In
%       the app, the History list displays this overall accuracy score for
%       each model.
%
% Use the code to train the model with new data. To retrain your
% classifier, call the function from the command line with your original
% data or new data as the input argument trainingData.
%
% For example, to retrain a classifier trained with the original data set
% T, enter:
%   [trainedClassifier, validationAccuracy] = trainClassifier(T)
%
% To make predictions with the returned 'trainedClassifier' on new data T2,
% use
%   yfit = trainedClassifier.predictFcn(T2)
%
% T2 must be a table containing at least the same predictor columns as used
% during training. For details, enter:
%   trainedClassifier.HowToPredict
  % Extract predictors and response
% This code processes the data into the right shape for training the
```

```
% model.
inputTable = trainingData;
predictorNames = {'type', 'loc', 'T1', 'load', 'LineNum'};
predictors = inputTable(:, predictorNames);
response = inputTable.Status;
isCategoricalPredictor = [false, false, false, false, false];
 % Train a classifier
% This code specifies all the classifier options and trains the classifier.
classificationSVM = fitcsvm(...
    predictors, ...
    response, ...
    'KernelFunction', 'gaussian', ...
    'PolynomialOrder', [], ...
    'KernelScale', 2.46297857551967, ...
    'BoxConstraint', 479.1629420174838, ...
    'Standardize', false, ...
    'ClassNames', [0; 1]);
 % Create the result struct with predict function
predictorExtractionFcn = @(t) t(:, predictorNames);
svmPredictFcn = @(x) predict(classificationSVM, x);
trainedClassifier.predictFcn = @(x) svmPredictFcn(predictorExtractionFcn(x));
 % Add additional fields to the result struct
trainedClassifier.RequiredVariables = {'LineNum', 'load', 'T1', 'loc', 'type'};
trainedClassifier.ClassificationSVM = classificationSVM;
trainedClassifier.About = 'This struct is a trained model exported from Classification
Learner R2020a.';
trainedClassifier.HowToPredict = sprintf('To make predictions on a new table, T, use: \n
yfit = c.predictFcn(T) \nreplacing "c" with the name of the variable that is this struct, e.g.
"trainedModel". \n \nThe table, T, must contain the variables returned by: \n
c.RequiredVariables \nVariable formats (e.g. matrix/vector, datatype) must match the
```

```
original training data. \nAdditional variables are ignored. \n \nFor more information, see
<a href="matlab:helpview(fullfile(docroot, "stats", "stats.map"),
"appclassification_exportmodeltoworkspace")">How to predict using an exported
model</a>.');
 % Extract predictors and response
% This code processes the data into the right shape for training the
% model.
inputTable = trainingData;
predictorNames = {'type', 'loc', 'T1', 'load', 'LineNum'};
predictors = inputTable(:, predictorNames);
response = inputTable.Status;
isCategoricalPredictor = [false, false, false, false, false];
 % Perform cross-validation
partitionedModel = crossval(trainedClassifier.ClassificationSVM, 'KFold', 5);
 % Compute validation predictions
[validationPredictions, validationScores] = kfoldPredict(partitionedModel);
 % Compute validation accuracy
validationAccuracy = 1 - kfoldLoss(partitionedModel, 'LossFun', 'ClassifError');
```

C.2 Reference

[1] L. L. Freris, A. M. Sasson: “Investigation of the load-flow problem,” *Proceedings of the Institution of Electrical Engineers*, vol. 115, no. 10, pp. 1459-1470, Oct. 1968.