REAL-TIME PREDICTION OF CASCADING FAILURES IN POWER SYSTEMS

A Thesis Submitted to the

College of Graduate and Postdoctoral Studies

In Partial Fulfillment of the Requirements

For the Degree of Doctor of Philosophy

In the Department of Electrical and Computer Engineering

University of Saskatchewan

Saskatoon

By

Mohamed Osama Mahgoub

PERMISSION TO USE

In presenting this thesis/dissertation in partial fulfillment of the requirements for a Postgraduate degree from the University of Saskatchewan, I agree that the Libraries of this University may make it freely available for inspection. I further agree that permission for copying of this thesis/dissertation in any manner, in whole or in part, for scholarly purposes may be granted by the professor or professors who supervised my thesis/dissertation work or, in their absence, by the Head of the Department or the Dean of the College in which my thesis work was done. It is understood that any copying or publication or use of this thesis/dissertation or parts thereof for financial gain shall not be allowed without my written permission. It is also understood that due recognition shall be given to me and to the University of Saskatchewan in any scholarly use which may be made of any material in my thesis/dissertation.

Requests for permission to copy or to make other uses of materials in this thesis/dissertation in whole or part should be addressed to:

Head of the Department of Electrical and Computer Engineering
57 Campus Drive
University of Saskatchewan
Saskatoon, Saskatchewan (S7N 5A9)
Canada

OR

Dean
College of Graduate and Postdoctoral Studies
University of Saskatchewan
116 Thorvaldson Building, 110 Science Place
Saskatoon, Saskatchewan (S7N 5C9)
Canada

DISCLAIMER

The use and reference of MATLAB, IEEE, TensorFlow, and COSMIC were exclusively created to meet the thesis and/or exhibition requirements for the degree of Doctor of Philosophy at the University of Saskatchewan. Reference in this dissertation to any specific commercial products, process, or service by trade name, trademark, manufacturer, or otherwise, does not constitute or imply its endorsement, recommendation, or favoring by the University of Saskatchewan. The views and opinions of the author expressed herein do not state or reflect those of the University of Saskatchewan and shall not be used for advertising or product endorsement purposes.

ABSTRACT

Blackouts in power systems cause major financial and societal losses, which necessitate devising better prediction techniques that are specifically tailored to detecting and preventing them. Since blackouts begin as a cascading failure (CF), an early detection of these CFs gives the operators ample time to stop the cascade from propagating into a large-scale blackout. In this thesis, a real-time load-based prediction model for CFs using phasor measurement units (PMUs) is proposed. The proposed model provides load-based predictions; therefore, it has the advantages of being applicable as a controller input and providing the operators with better information about the affected regions. In addition, it can aid in visualizing the effects of the CF on the grid. To extend the functionality and robustness of the proposed model, prediction intervals are incorporated based on the convergence width criterion (CWC) to allow the model to account for the uncertainties of the network, which was not available in previous works. Although this model addresses many issues in previous works, it has limitations in both scalability and capturing of transient behaviours. Hence, a second model based on recurrent neural network (RNN) long short-term memory (LSTM) ensemble is proposed. The RNN-LSTM is added to better capture the dynamics of the power system while also giving faster responses. To accommodate for the scalability of the model, a novel selection criterion for inputs is introduced to minimize the inputs while maintaining a high information entropy. The criteria include distance between buses as per graph theory, centrality of the buses with respect to fault location, and the information entropy of the bus. These criteria are merged using higher statistical moments to reflect the importance of each bus and generate indices that describe the grid with a smaller set of inputs. The results indicate that this model has the potential to provide more meaningful and accurate results than what is available in the previous literature and can be used as part of the integrated remedial action scheme (RAS) system either as a warning tool or a controller input as the accuracy of detecting affected regions reached 99.9% with a maximum delay of $400\ ms$. Finally, a validation loop extension is introduced to allow the model to self-update in real-time using importance sampling and case-based reasoning to extend the practicality of the model by allowing it to learn from historical data as time progresses.

ACKNOWLEDGEMENT

In the beginning, all praise is due to Allah for giving me life, sustenance, the mental capacity to pursue higher education, and above all Islam. Moreover, my deepest gratitude to everyone who helped me complete my PhD in any form, and even if your name is not in this acknowledgement due to length constraints you are definitely in my mind and soul.

I would like to thank my supervisors Prof. Tony C.Y. Chung and Sherif O. Faried for their support throughout my Ph.D. program. I have learned a deeper understanding of academia and research from their experience and guidance. Also, I would like to thank the University of Saskatchewan and more specifically the Electrical and Computer Engineering Department for providing the opportunity and resources required to complete my PhD. Special thanks go to the members of my Advisory Committee for their insightful comments and suggestions.

On a personal level, I owe my deepest gratitude to my parents, Osama and Faiza, as words can never express the amount of support that they provided me throughout the years. Also, the gratitude is extended to my loving wife, Ethar, for her continuous support in all aspects of my life and for her constant “you can do it” phrase. In addition, I thank all my siblings, Ahmed, Safa, Asma, Fatima Al-Zahraa, and Laila, each one of you helped me in a unique way and it is my honor to mention each one of you by name.

Finally, I would like to thank prof. Tony C.Y. Chung and Sherif O. Faried for financially supporting my PhD. The thanks are also extended to NSERC, SaskPower, and the University of Saskatchewan.

DEDICATION

This PhD is dedicated to my wife Ethar, and my parents Osama and Faiza.

TABLE OF CONTENTS

LIST OF TABLES

TABLE OF FIGURES

ABBREVIATIONS

AC	Alternating Current
BIP	Bus Isolation Probability
CAT	Cascade Analysis Tool
$CAMS$	Computing and Analytic Methods Subcommittee
CBR	Case-Based Reasoning
CF	Cascading Failure
CI	Confidence Interval
CFP	Cascading Failure Prediction
$COSMIC$	Cascading Outage Simulator with Multiprocess Integration Capabilities
CWC	Coverage Width Criterion
DC	Direct Current
$DDR4$	Double Data Rate (4^{th} Generation) [RAM model]
DF	Data Fusion
DNN	Deep Neural Network
EC	Eigenvector Centrality
$EENS$	Expected Energy Not Supplied
FRT	Fault Ride Through
FW	Floyd Warshall
GEI	Generalized Entropy Index
GPS	Global Positioning System
HDD	Hard Disk Drive
HF	Hidden Failure
$IEEE$	Institute of Electrical and Electronics Engineers
IMS	Importance Sampling
$LERIS$	LSTM ensemble with reduced input set
LG	Line to Ground Fault
LOL	Loss of Load
$LOLE$	Loss of Load Expectation

$LSTM$	Long Short-Term Memory
$LUBE$	Lower Power Bound Estimation
MC	Monte Carlo
$MPIW$	Mean Prediction Interval Width
$NMPIW$	Normalized Mean Prediction Interval Width
$NERC$	North America Electric Reliability Corporation
NN	Neural Network
OC	Overcurrent
OPA	ORNL PSERC Alaska Model
OPF	Optimal Power Flow
PES	Power & Energy Society
PI	Proportional Integral (controller)
PI	Prediction Interval
PM	Prediction Model
PMU	Phasor Measurement Unit
POL	Polish winter peak system
$POM - PCM$	Physical and Operational Margins - Potential Cascading Modes (commercial software suite)
PSA	Power System Analyzer (software)
PSO	Particle Swarm Optimization
QSS	Quasi Steady State
$R\&D$	Research and Development
RAM	Random Access Memory
RAS	Remedial Action Scheme
RNN	Recurrent Neural Network
$SCADA$	Supervisory Control and Data Acquisition System
SLD	Single Line Diagram
SNR	Signal to Noise Ratio
SVM	Support Vector Machine
TAM	Texas A&M CF modeling tool

TI	Theil Index
TVE	Total Vector Error
UF	Underfrequency
$UFLS$	Under Frequency Load Shedding
USD	United States Dollar
UV	Undervoltage
$UVLS$	Under Voltage Load Shedding
$WAMS$	Wide Area Monitoring System

NOMENCLATURE

M	Machine inertia constant
ω_i	Angular frequency of the i^{th} generator bus
$P_{m,i}$	Mechanical power input at i^{th} generator bus
$P_{g,i}$	i^{th} Generator real power output
D	Damping constant
δ_i	Rotor angle
f_o	Fundamental frequency
$\left\|E'_{a,i}\right\|$	Transient open circuit voltage magnitude
$\|V_i\|$	Voltage magnitude
$\delta_{m,i}$	Mechanical rotor angle
$X_{d,i}$	Direct axis synchronous reactance
$X'_{d,i}$	Direct axis transient reactance
$X_{q,i}$	Quadrature axis synchronous reactance
$\sigma(x)$	sigmoid activation function
T_t	PI controller time constant
P_{ref}	Reference power
R	Droop
T	Temperature
F	Power flow
$\varsigma(x)$	Line Heating tripping function
S	State
t	Time
$\xi(x)$	UVLS tripping function
v_n	Voltage at load n
v_{Th}	Threshold voltage
τ	Time delay
f_n	Frequency at bus n

ϱ	UFLS tripping function
px	Probability that x successes
qx	Probability that x fails
X	Input vector
PL	Real power vector
VP	Voltage magnitude vector
TP	Voltage angle vector
P_{L_n}	Real power at load bus n
v_m	Voltage magnitude at bus m
θ_m	Voltage angle at bus m
i	Generator index
j	Branch index
m	Bus index
n	Load bus index
I	Total number of generator buses
J	Total number of branches
M	Total number of buses
N	Total number of load buses
ω	System frequency
$\mathcal{F}_N$	NN model for load bus N
$Z^{\geq}$	Set of nonnegative integers
δ	Samples at different time delays
P_F	Probability of a wrong prediction
P_n	Probability of a correct prediction
$\mathbb{C}$	Mapping function
$\mathbb{R}$	Real number
$\mathbb{R}^n$	Real coordinate space of dimension n
α	The likelihood that the prediction lies outside the confidence interval
Y	Desired output
D	Set of input and output vectors pairs

U	Upper bound of prediction
L	Lower bound of prediction
Z	Total number of datapoints
z	Datapoints index
g	Number of features
$\mathcal{R}$	Output range
η	First hyperparameter of the CWC
μ	Second hyperparameter of the CWC
c	Boolean for correct predictions in CWC
N	Number of elements in a network
k	Number of failed components
x	Single value input
s	RNN's internal state
M	Number of NN inputs
N	Number of NN cells
σ	Sigmoid activation function
C	LSTM cell's state
H	LSTM cell's hidden state
W	LSTM weighting function
f	LSTM forget gate
k	LSTM auxiliary gate
i	LSTM input gate
$\tanh(x)$	Tanh activation function
$\otimes$	Bitwise multiplication
o	LSTM output gate
ℵ	Node
B	Observed bus
$w_{i,B,d}$	Weighting of bus i for observing bus B according to distance index
Ω	Set
$d_{i,B}$	Distance between node i and node B

d_{max}	Maximum distance (hyperparameter of the distance index)
α	Attenuation factor (hyperparameter of the distance index)
β	Difference coefficient of the GEI
n	Number of datapoints
s	Scenario index
S	Number of scenarios
$v_{i,s,t}$	Voltage of bus i in the s^{th} simulation at time t
$w_{i,e}$	Weighting of bus i according to the entropy index
S_o	Subgraph
S	Graph
$\mathrm{V}(S)$	Vertices of graph S
$\mathrm{E}(s)$	Edges of graph E
γ	Tunable parameter of the centrality index
A	Adjacency matrix
A^*	Modified adjacency matrix
$w_{i,B,F,c}$	Weighting of bus i for observing bus B for a fault at F according to distance index
c_i	Centrality of node i
k	Eigenvector Centrality constant
$\alpha(w)$	Weighting function
Ψ	Set of important buses
μ	mean
σ^2	Variance
sk	Skewness
kr	Kurtosis
s	Standard deviation
Ω^M	Set of buses
Ω^N	Set of load buses
Ω^J	Set of lines
SL	System's line state vector
FL	System's fault location vector

W	Subset vector of the input vector
T	Maximum prediction delay
t_{up}	Actual update time
t_{tr}	Training time
t_{de}	Desired update time
t_{mx}	Maximum allowed training time
d_t	Retraining dataset
d_s	Subset of the synthetic dataset
d_h	Subset of the historical dataset
d_n	Subset of the new dataset
d_S	Synthetic dataset
d_H	Historical dataset
d_N	New dataset
f_{rw}	Rejecting updates flag
M	Prediction model (for validation loop)
acc_{Th}	Threshold accuracy
f_{mw}	Mismatch warning flag
k	Training iteration
$\mathrm{E}(X)$	Expected value
g	Modified distribution function
p	Sampling probability

1 INTRODUCTION

1.1 General

Power systems are getting larger and more complex, and with the inclusion of renewable energy resources, they are also getting more stochastic. This complex and stochastic nature of power systems increases the probability of undesirable events, such as line trippings, forced outages…etc. [1]. Moreover, transmission lines are typically operated very close to their maximum capacities (thermal limits) to maximize profit, which significantly increases the probability of overheating the line conductors. Typically, system operators apply reliability analysis tests and try to maintain the N-1 criteria or the 0.1 days/year criteria to keep the system operational within acceptable operation levels even after a line trip or a forced outage [2]. However, operators' mistakes, hidden failures, extreme weather, or sabotage can induce multiple trippings and forced outages that can lead the system into a cascading failure (CF) which can then lead to a total or partial blackout [3]. Moreover, forecasting errors, e.g., when predicting loads or wind speed, increases the chances of a particular fault leading to a CF. In general, the increased uncertainties in the network increase the likelihood of failures. In addition, renewable energy resources can increase the probability of CFs and blackouts. For example, wind farms have lower inertia and lower fault ride through (FRT) capabilities than synchronous generators. Thus, a high penetration of wind energy means that the system is getting weaker in withstanding faults. Thus, increasing the probability of a CF. Additionally, due to their low FRT capabilities, it is possible that a fault within the wind farm or close to it could result in hundreds of wind turbines being tripped off in a cascading manner leading to major blackouts [1]. Hence, based on the previous scenarios and multiple points of failure, power systems have a multitude of ways by which CFs and blackouts can happen.

Major blackouts that occurred in the 21st century include the 2003 US-Canada Northeast blackout, the 2011 California blackout, the 2016 southeast Australia blackout, and the 2018 Brazilian blackout just to name a few [4]. Even more recently, there is the 2021 Texas blackout caused by winter storms and severe weather conditions during February, which caused damages in the range of billions of dollars [5]. A summary of major blackouts in the last twenty years and their estimated impacts is shown in Table 1.1. The recurrence of blackouts indicates that blackouts

are still happening albeit the numerous techniques used to suppress them and the multitude of reliability metrics that power grids adhere to. Additionally, all these blackouts result in huge supply interruptions and losses in the order of millions, and sometimes even billions, of dollars to both utilities and consumers. Besides, these blackouts have hidden costs consisting of their negative impact on other aspects of life, such as increased crime rate and increased mortality, which can never be quantified in terms of dollars [6]. This repetitive nature of blackouts in different regions and the huge losses associated with blackouts indicate that more analysis of blackouts and how to prevent them is still needed [7].

Table 1.1. Notable blackouts in the last twenty years (2001-2021), their estimated costs, and their main causes.

Ref	Location	Year	Loss of load (MW)	Time duration	Estimated cost (million $)[1]	Main causes for the initiation or propagation of the blackout
[5], [8]	Texas	2021	34,000	72-336 h	195,000	Severe windstorms, disruption in natural gas supply, and frigid winter contributed to the initiation and propagation of the blackout.
[1]	South Australia	2016	2,900	3-24 h	285	Severe thunderstorms caused the initial faults. Then, the fault ride through (FRT) mechanism in wind turbines caused CFs within the wind farms in the grid.
[9], [10]	Kenya	2016	180	<3 h	N.A.	A monkey caused a major transformer to trip resulting in a CF (the system was not N-1 secure at the time of failure).
[9]	Sri Lanka	2016	800	>4 h	N.A.	N.A.
[9], [11]	Turkey	2015	32.2	>7 h	700	The system was not N-1 secure due to maintenance and construction work. Then, a trip in a major tie-line caused islanding within the grid. Then, frequency instability initiated CFs in both islands.

[1] Cost estimates can vary greatly based on the approach used to estimate the costs, because there are direct costs (e.g., restoration cost), indirect costs (e.g., loss of revenue in utilities and other corporations), damage costs (e.g., sustained failure in equipment), indirect damage cost (e.g., loss of life), and future costs (e.g., loss of revenue due to bankruptcy). Hence, based on the considered factors the reported amounts significantly vary.

Table 1.1. Continued.

Ref	Location	Year	Loss of load (MW)	Time duration	Estimated cost ($)	Main causes for the initiation or propagation of the blackout
[9], [12]	India	2012	48	2-8 h	6,000	The system suffered from a weak transmission line system since multiple key tie-lines between northern and western regions were overloaded. Then, trippings in these tie-lines initiated the CF and the widescale blackout.
[9]	Brazil	2011	8.88	>3 h	N.A.	N.A.
[9]	Brazil and Paraguay	2009	24,436	4-6 h	N.A.	Direct cause undetermined, but the blackout was initiated by a fault causing the Itaipu dam to stop production, which initiated a CF due to the huge power imbalance.
[13]	Australia	2007	2,150	4-8 h	129	Bushfires and extreme weather conditions caused multiple failures and overheated transmission lines.
[9], [14]	Colombia	2007	6.64	>4 h	130	CF initiated by a maintenance mistake.
[9], [15]	Europe	2006	14.5	2 h	N.A.	Trippings of multiple high voltage lines formulated three islands that either collapsed or suffered major shedding because of power imbalances.
[9], [16]	Pakistan	2006	11.16	5-6 h	N.A.	Tripping of a large transformer that led to the tripping of its parallel counterpart. This initiated a CF causing multiple major tie-lines operating near their capacity to trip.
[13]	Sweden and Denmark	2003	6,550	2 h	N.A.	Loss of a 1.8 and a 1.2 GW plants due to two unrelated faults within five minutes of each other.
[9]	Italy	2003	24	5-9 h	1200	Tree flashover in a major tie line initiated a power deficit that led to a CF.

Table 1.1. Continued.

Ref	Location	Year	Loss of load (MW)	Time duration	Estimated cost (million $)	Main causes for the initiation or propagation of the blackout
[9]	London	2003	724	>30 m	N.A.	Hidden failures due to wrong relay settings and delayed maintenance work initiated the CF.
[9], [13]	North America	2003	63,000	5-72 h	Over 10,000	A heatwave that led to demand peaks and insufficient reactive power, incorrect state estimation of the grid, and bad vegetation management leading to multiple L-G faults.

The first step to prevent blackouts is to understand and detect how they start and expand. However, blackouts are difficult to analyze due to the numerous mechanisms involved, some of which are extremely hard to model, such as operators' mistakes [7]. Therefore, another approach for predicting blackouts is to analyze CFs, which are defined as the uncontrolled successive failure of elements in an interconnected system initiated by a starting failure [17]. Thus, by predicting and preventing CFs, blackouts will also be prevented by extension since they are a subset of CFs. However, modeling CF is also complex due to the involvement of numerous mechanisms and the complexity of power systems. Hence, there is no individual tool that captures all of the CF mechanisms, which can be a barrier that prevents building reliable and meaningful CF prediction models [18]. Nonetheless, it has been shown in the literature that the existing tools can provide results similar to historical data in terms of risk and probability distribution [19]. Also, they model the most significant parts that contribute to blackouts and CFs. Thus, these tools are sufficient for building models to analyze and predict CF [7].

1.2 Smart Grids and Wide Area Monitoring Systems

A major problem in power systems that contributed to the recurrence of blackouts and CFs in the past was the need to estimate the state of the grid due to a lack of the availability of the status of its major buses in real time. However, with the advent of Wide Area Monitoring Systems (WAMS), this issue is solved by introducing Phasor Measurement Units (PMUs) which provide

real-time measurements of voltage magnitude and phase angle. These measurements effectively eliminated the need for state estimation (assuming enough PMUs are used) [20]. In addition, PMUs are significantly better than the typical supervisory control and data acquisition system (SCADA), because PMUs allow for higher sampling rates and can provide synchronized measurements even between dispersed areas in the power network [21]. On the other hand, the added complexities of smart grids could lead to more ways by which a system fails due to the use of such technology [22]. Hence it is important to use PMUs to offset the negative impacts of smart grids technologies and improve the grid reliability and resiliency.

Since PMUs are essential to CF prediction, it is essential to summarize their structure and working principle. Initially, PMUs receive analog inputs from current transformers (CTs) and potential transformers (PTs) connected to the transmission line. Then, these measurements are passed through a low-pass filter and an analog to digital (A/D) converter to satisfy the Nyquist criteria and reduce noise. Afterwards, these measurements are time aligned and time stamped using a Phase Locked Loop (PLL) synchronized with a nanosecond accurate clock to have the synchronicity needed to generate synchronized voltage measurements. Typically, GPS's second of century (SOC) is used for synchronization. Afterwards, these measurements are passed to an internal microprocessor to calculate the magnitude and phase of both voltage and current, as well as, real power, reactive power, power factor, total harmonic distortion, frequency, frequency deviation, and other measurements. Finally, an internal modem generates the data packets for the PMU measurements, and the packets are sent to the control center via Phasor Data Concentrators (PDC). Figure 1.1 summarizes the block diagram of this working principle. Although PMUs appear to be similar to other conventional voltage and current measurement devices, the usage of GPS's SOC allows for the synchronicity that guarantees that all measurements can be precisely aligned together to form a snapshot of the system's state at a given time, which is the main advantage of PMUs.

1.3 Problem Statement

Supply interruption negatively affects the lives of millions and causes significant societal and economic impacts when it occurs. The largest and severest of supply interruptions are blackouts, which could affect millions or even hundreds of millions when they occur. Nonetheless, the rarity of blackouts and the complexities involved in how they occur makes it hard to predict

their occurrences. Thus, blackouts keep on reoccurring, such as the 2003 US-Canada blackout and the hundreds of other blackouts mentioned in the literature. However, in all these blackouts, a lack of information regarding whether the contingencies are going to lead to a blackout, or an operational system was among the main causes why these blackouts propagated and expanded. Thus, formulating a tool capable of detecting the possible occurrence of blackouts can help system operators and emergency controllers to react in a proper way that ensures the elimination of blackouts or at the very least the limiting of their devastating outcomes.

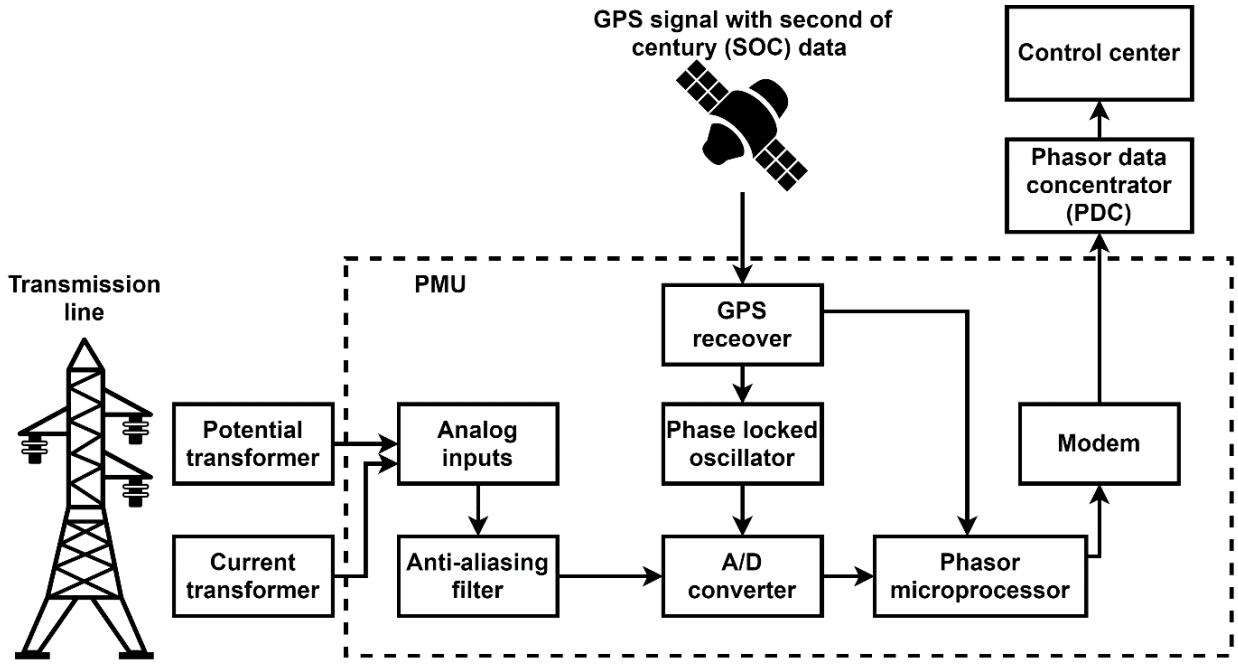

Figure 1.1. Phasor Measurement Unit (PMU) block diagram with GPS input, transmission line inputs and control center outputs.

1.4 Research Objectives

- Conduct a literature review on cascading failures and blackouts.
- Conduct a literature review on real-time prediction in power systems with a focus on real-time cascading failure prediction.
- Design a model to predict cascading failures, and by extension blackouts, in real-time with high accuracy and minimal delay.
- Predict the affected regions by the cascade in real-time using the aforementioned model.

- Improve the model to account for the different operating regions and the stochastic nature of the power network, as well as the uncertainties within the proposed model.
- Improve the scalability in the model to make it applicable to large interconnected systems.
- Improve the model by proposing a framework to update and improve the model in real-time as it receives more data from the power grid.

1.5 Significance

The main contribution of this thesis is improving the real-time cascading failure prediction and by extension blackout prediction. Although there are many tools for CF prediction (CFP), this research work is unique and needed due to three main additions compared to existing works. These will be elaborated further when discussing the literature review and the methodology, but a summary of these additions is listed here for the completeness of the introduction. The proposed model uses recurrent neural networks (RNN), which addresses some of the limitations of previous works like having a predictor that depends only on a small set of past values. In addition, the model uses information theory, graph theory, and probabilistic techniques to reduce the input count without affecting the accuracy. This aids in the expansion of the model to large interconnected systems. Moreover, the model divides the CF prediction problem into a load-point-based prediction, which makes the predictions more accurate and provides the operators with more information regarding the locations affected by the cascading failure or blackout.

1.6 Methods Overview

The relationship between each chapter's aims and methodology and the general flow of the thesis are summarized in Figure 1.2.

1.7 Research Outcome

The main outcome of this thesis is proposing a new framework for real-time prediction of cascading failures in power systems. Since the work is a framework, it is highly versatile and includes many components that can be changed according to the operators' requirements and specifications. Moreover, the framework using the setup described in Chapter 5 yielded results that are superior to other models in both accuracy and speed. Hence, it is expected that this framework could be used as an additional warning tool in the control center for detecting CFs and blackouts.

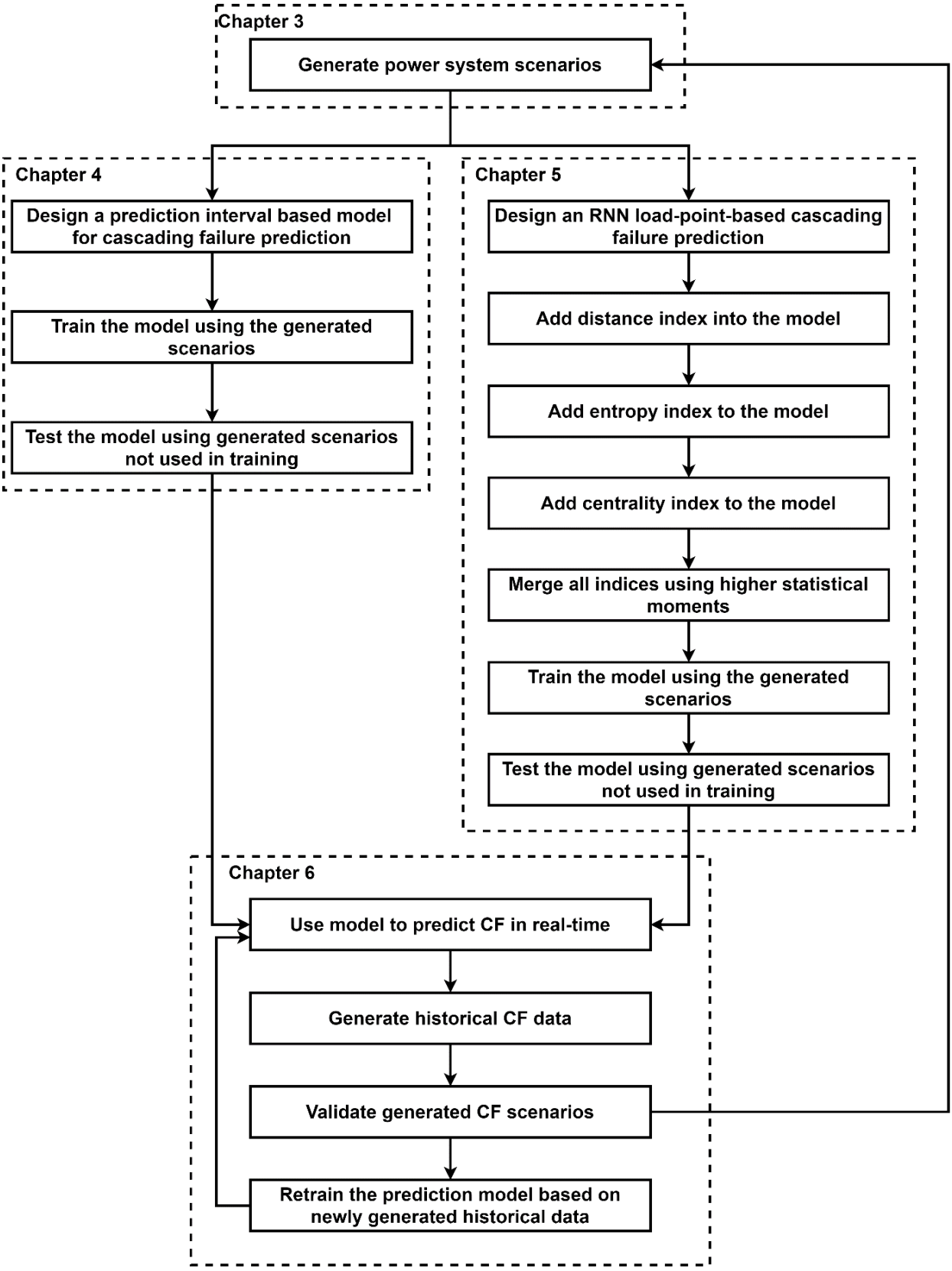

Figure 1.2. Methods overview and methodology flowchart.

1.8 Definitions

A complete list of definitions and clarifications regarding terminologies used in this thesis is given in Appendix A.

1.9 Thesis Organization

The thesis is organized into seven chapters, a list of references section, and five appendices. The main topics of each chapter are as follows:

- Chapter 1 introduces the nature of cascading failures and blackouts, and the benefits of PMU in CF prediction. The objective of the research is also presented in this chapter.
- Chapter 2 discusses the literature review of cascading failure in detail.
- Chapter 3 explains the dataset and how it is generated.
- Chapter 4 introduces the first method in this thesis, the prediction-interval-based model for cascading failure prediction, along with its results and conclusion.
- Chapter 5 expands on the initial method by incorporating RNN, graph theory, information theory, probabilistic techniques, and load point predictions into the formulation.
- Chapter 6 expands the model further by incorporating a validation and update loop to allow the model to improve in real-time.
- Chapter 7 summarizes the thesis, lists its conclusions, and provides areas of improvements where future work could be implemented.
- Appendix A lists the definitions of used terminologies in this thesis.
- Appendix B summarizes the dynamic equations included in the power system model used in this thesis.
- Appendix C includes the information for the IEEE 39-bus system.
- Appendix D provides a sample of the used data to generate the IEEE 39-bus system for reference and reproducibility.
- Appendix E includes the pseudocodes for all the algorithms used in this thesis.

2 CASCADING FAILURES IN POWER SYSTEMS LITERATURE REVIEW

2.1 Introduction

Cascading failures (CF) have an extremely low probability of occurrence but could lead to catastrophic outcomes, such as wide scale blackouts. Hence, operators and decision-makers need to know if the current grid (or dispatch) is aiding in reducing the probability and severity of CF, or the opposite. This applies to both the planning and operating of the gird. In planning, its importance stems from determining if a certain change to the gird or electricity dispatch would increase or decrease the probability (or severity) of blackouts, or CF in general. In operation, CF analysis is even more critical as CFs must be quickly mitigated and suppressed before they turn into blackouts.

Researching and analyzing CFs in power systems is a complex problem with many different, and sometimes contradicting, approaches and paradigms. However, since they are interconnected with each other, it is necessary to address all of them in the literature review to highlight the pros and cons of each paradigm and how they interact with each other. Moreover, it is important to understand the limitations of any CF prediction model, especially if it is a real-time model, by realizing the limitations of the different CF paradigms, to avoid making wrong decisions that could potentially expand the CF rather than suppressing it. With that in mind, the literature pertaining to CF could be classified into one or more of the following categories:

1. CF modeling
2. Component modeling
3. Angle stability
4. CF mitigation
5. Cybersecurity and malicious attacks
6. Interdependent systems
7. CF prediction
8. Reliability analysis and risk assessment
9. Literature review and benchmarking.

The relationship of these components with respect to the operation of the power system is shown in Figure 2.1. The figure divides the CF literature into 5 overlapping categories: research and development (R&D), dispatch, normal operation, contingency, and wide scale blackout. Most of the CF literature focuses on offline operation, i.e., R&D and dispatch, followed by online operation during contingencies. This focus on offline operation is mainly because CF modeling and analysis is still an ongoing research and there are not many industrial applications pertaining to CF in online operation. Based on the previous classification of CF categories, the remainder of this literature review addresses each one of these categories separately, but with emphasis on CF modeling and CF prediction as they are the focus of this work.

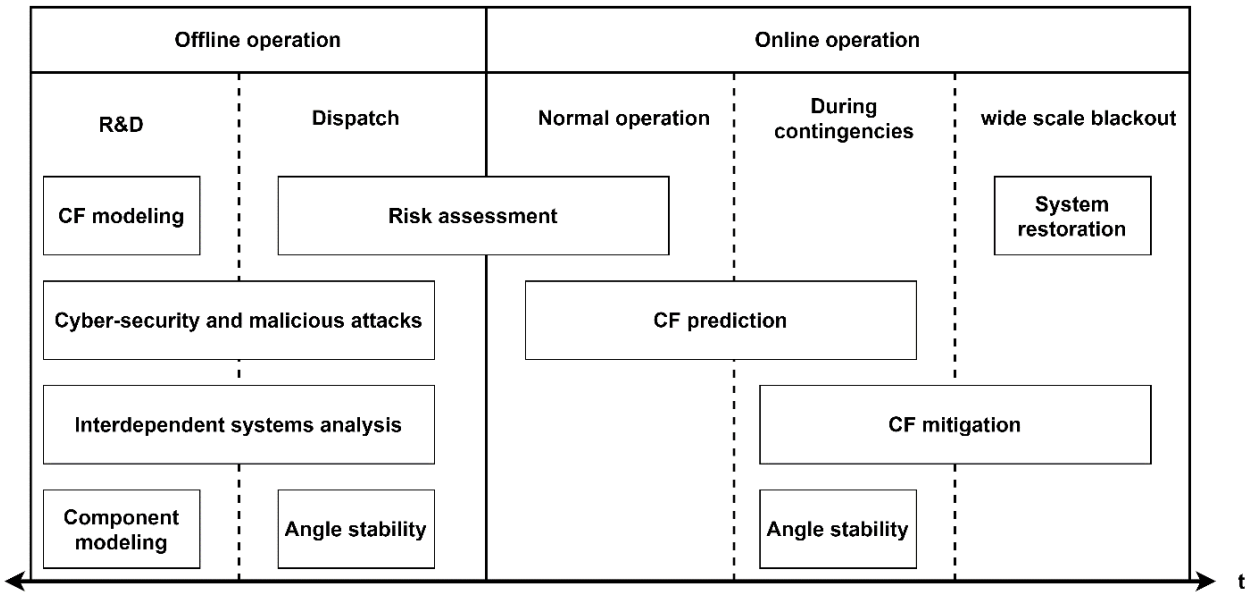

Figure 2.1. Classification of the CF literature according to the stages when they are relevant.

2.2 CF Modeling

2.2.1 Introduction

The purpose of CF modeling is to propose models capable of simulating or estimating CFs. These models could derive their concepts from areas such as network theory, probabilistic analysis, and time-domain simulations. Considering how detailed or abstract is the CF model, CF modeling can be classified into three types:

1. Topological models
2. Quasi steady state models
3. Dynamic models.

There are other types of models such as stochastic and interdependent system models, but they can also be classified using these three types. In topological models, power systems are assumed to be a graph network where buses and transmission lines are transformed into nodes and graphs. The advantage of such models is that they are quick to compute and can easily generate approximations for the likelihood of CFs given a specific network topology and the associated risk of these CFs. However, these models ignore multiple important power system aspects such as voltage violations and angle stability to name a few. On the other hand, quasi steady state (QSS) models aim to address the limitations of topological models by considering some power system aspects such as transmission lines' MW capacities, voltage violations, and the redistribution of power after the loss of a line or a generator. Nonetheless, QSS models ignore other power system dynamics such as angular stability and frequency violations to minimize the simulation time. Hence, they provide a trade-off between speed and accuracy. Finally, dynamical models, or time-domain models, simulate power system dynamics and mechanisms as accurately as possible to provide a realistic view about the power system behaviour following a CF. However, their models do not include all possible aspects as some are either impossible to model or computationally infeasible. Comparing all the three models, the trade-off is always between the simulation time and the accuracy of describing the power system behavior as shown in Table 2.1.

Table 2.1. Relative comparison between CF models in terms of speed and accuracy to calculate CF risk.

Model	**Simulation speed**	**Accuracy of the model**
Topological model	Extremely fast (milliseconds)	Low
Quasi steady state (QSS) model	Fast (seconds)	Medium
Dynamic model	Slow (hours)	High

Modeling of CFs is not only limited to the simulation speed and accuracy of the model, but it also depends on which causes of CFs are being modeled. Considering the different causes of CFs and how they propagate, CFs can generally be initiated, or propagated, by any of the following scenarios.

1. CF can happen due to a line tripping causing another line, or lines, to be overloaded. This leads to that other line, or lines, to trip as well. Then, this cascaded tripping continues until the entire system collapses or a new stable island, or islands, emerges. Alternatively, this propagation could be stopped by load shedding or other control schemes by relieving the congested lines before they trip due to overcurrent.
2. Another CF scenario is when a tripping causes a huge part of the generation to be lost, and then the demand becomes more than the supply. Afterwards, the system enters a CF, because the frequency of the system would drop leading to more generator tripping, to protect the generators from being damaged, until the entire system collapses. The opposite can also happen (i.e., the generation is far greater than the supply, because a huge part of the demand is disconnected) and this could also lead to a CF, but it is rarer and easier to deal with.
3. Finally, CFs can be initiated by the occurrence of a fault subsequent to another one that its effect has not yet been mitigated. This can result in a CF, either due to the system not being able to withstand that particular N-2 fault, or because the transient of these faults resulted in more tripping and failures, which initiated the cascade. According to NERC, 73.5% of CFs are due to or aggravated by hidden failures, which is one of the main reasons for having faults happening within a short time frame of another seemingly unrelated fault [23].

However, this list is not exhaustive, as there are many other ways by which CFs propagate. Regardless, it presents the most prominent ways by which blackouts and CFs happen and propagate. Moreover, the propagation of all the blackouts presented in Table 1.1 can be categorized using the three classifications stated above. Although the initiating events of these blackouts might not all fall into the three classifications, the important aspect of CF is the propagation, because this is the main difference between typical contingencies and CFs.

In all the different scenarios by which CFs propagate, cascading outages can be described as oscillating events between fast and slow CFs. The “slow cascade” is when the system appears

to be operating normally but is actually under stress due to one or more transmission lines operating very close or above their capacity (this also applies to other components, but it is more common in transmission lines). This slow cascade can then cause line trippings that may trigger electrical instability or move the system to another slow cascade. The time frame for the slow cascade is between a few seconds and hours. In case electrical instabilities occur, the system will then be in the "fast cascade". During the fast cascade, the system has successive trippings (mainly due to frequency and voltage violations). The time frame for fast cascade is between milliseconds to few seconds. After the "fast cascade", the system reaches one of four states: a partial blackout, a total blackout, another "slow cascade", or a new steady state with multiple islandings and major load sheddings [24]. This distinction between fast and slow cascades highlights one of the main reasons why topological and QSS models are inaccurate since both models only consider the slow cascade and ignore the fast cascade.

2.2.2 Existing CF models

For each of the three main CF models (topological, QSS, and time domain), as well as the other models (e.g., stochastic, and interdependent models), there are many existing tools both for industrial and research applications. Table 2.2 compares the various existing CF modeling tools both in the industry and research ([9], [24]–[27]). The comparison shows that each model has different advantages over the others. For example, probabilistic or risk-based models, like OPA, are more suited for general information about the risk of blackouts and CFs which is important during planning or when comparing the risks of different dispatches [26]. On the other hand, detailed deterministic models, like COSMIC, are more suited for simulating specific cases to determine their stability and exact behavior [27]. For example, this is useful when simulating a critical subset of N-2 contingencies.

Table 2.2. Comparison of existing CF analysis models.

Model	**Main advantages**	**Load flow**	**Additional information**
Hidden Failure	Modeling hidden failures in relays and relay systems. Uses fast heuristic approaches to randomly search for severe hidden failure scenarios and estimate their likelihood and severity to provide the expected risk.	AC	Research tool

Table 2.2. Continued.

Model	Main advantages	Load flow	Additional information
PSA	Includes isolation and connectivity analysis. Computes the least number of cuts for isolation. Can simulate grids with a large number of nodes.	AC/DC	Research tool
TAM	A reliability assessment tool that produces common reliability indices (e.g., LOLP and BIP). Uses MC simulation.	AC	Research tool
OPA	Considers tree contact, dispatching, and planning. Most useful in planning when comparing a base case with an updated grid layout and comparing their progression as years pass and demand increases.	AC/DC	Research tool. AC is only available in improved OPA. Ignores many dynamical processes and time elements in CF.
Manchester	Uses MC simulation to obtain Risk. Includes hidden failures and forced outage probabilities. Includes restoration modeling and estimation of restoration time.	AC	Research tool
COSMIC	Considers generators', turbines', and exciters' dynamics. Uses ZIPE load models. Open-source software.	AC	Research tool Uses dynamic timesteps.
ASSESS	Uses both analytical and Monte Carlo based techniques. Includes security-constrained AC OPF. Includes both time-domain simulation and QSS simulation.	AC/DC	Commercial tool Requires a specialist user since it covers a wide range of different simulation paradigms.
TRELSS/ TransCARE	Reliability assessment tool with CF analysis. Can model protection systems in detail. Has detailed control actions including redispatch, transformer tap, and reactor switching. Has fast decoupled power-flow algorithm.	AC/DC	Commercial tool

Table 2.2. Continued.

Model	Main advantages	Load flow	Additional information
CAT	Measures if a contingency leads to another one or not (e.g., CF initiation). Considers voltage and thermal overload violations. Uses a QSS AC approach to access losses due to load drops.	AC	Commercial tool
POM-PCM	Identifying all cascading paths and chains in one simulation. Ranking the CF chains based on vulnerability and likelihood. Has mitigation strategies incorporated.	AC (steady state and dynamic)	Commercial tool

2.2.3 Importance of time domain dynamic simulation in CFP

To obtain an accurate description of the CF propagation for the purpose of CF prediction (CFP), the detailed dynamical model of the system is needed to avoid the pitfalls of static and QSS simulations. In a dynamic model, the generators' dynamic following a system change (e.g., generation demand mismatch) is governed by the following equations:

$$M_i \frac{d\omega_i}{dt} = P_{m,i} - P_{g,i} - D_i(\omega_i - 1) \tag{2.1}$$

$$\frac{d\delta_i}{dt} = 2\pi f_o(\omega_i - 1) \tag{2.2}$$

$$P_{g,i} = \frac{\left|E'_{a,i}\right||V_i|}{X'_{d,i}} \sin\delta_i + \frac{|V_i|^2}{2}\left(\frac{1}{X_{q,i}} - \frac{1}{X'_{d,i}}\right)\sin 2\delta_i \tag{2.3}$$

$$\frac{dP_{m,i}}{dt} = \sigma\left(\frac{1}{T_{t_i}}\left(\sigma\left(P_{ref_i} - \frac{1}{R_i}\Delta\omega_i\right) - P_{m,i}\right)\right) \tag{2.4}$$

where:

M_i is generator i's inertia constant in seconds,

ω_i is generator i's speed in per unit,

$P_{m,i}$ is the mechanical power input of the i^{th} generator in per unit,

$P_{g,i}$ is the generator real power output of the i^{th} generator in per unit,

D_i is the damping coefficient of the i^{th} generator in per unit,

δ_i is generator i's load angle in radians,

f_o is the synchronous frequency in hertz,

$\left|E'_{a,i}\right|$ is the generator i's transient open circuit voltage magnitude in per unit,

$|V_i|$ is generator i's voltage magnitude in per unit,

$X'_{d,i}$ is the direct-axis transient reactance in per units,

$X_{q,i}$ is the quadrature-axis synchronous reactance in per units,

σ is the sigmoid activation function,

T_{t_i} is generator i's proportional integral (PI) controller time constant in seconds,

P_{ref_i} is generator i's reference power (e.g., desired power output) in per units, and

R_i is generator i's droop in per units.

However, in a static simulation (and some QSS simulations), generator outputs are assumed to be constant using the following assumptions:

$$\frac{d\omega_i}{dt} = \frac{d\delta_i}{dt} = 0, \qquad P_{m,i} = P_{g,i}, \qquad \omega_i = 1 \; in \; p.u. \tag{2.5}$$

This disparity between dynamic and static simulation can often be neglected when looking for the statistical behavior of power systems. To elaborate, the disparity mainly affects outcomes in the tail of the distribution, which has an extremely low probability ($p \ll 0.05$). Thus, the difference in the estimated risk between the dynamic and static simulation is negligible [27]. However, when predicting the outcome of a sole case unfolding in real-time, these differences become significant. For example, cascading events with multiple failures and massive demand loss are already in the tail of the distribution. Hence, they are far away from the region where the simplifications and linearizations of static simulations are valid [27]. Moreover, dynamic models allow for more realistic line heating characteristics, which is essential in large-scale CF propagation as evident from the 2003 Northeast blackout as the heating and tripping of lines was

a critical component in the propagation of the blackout [28]. Also, in [24] different cascading paths were investigated and the results indicate that similar events (assuming static simulation) could propagate differently considering dynamic simulation. Moreover, [24] shows that no mechanism of CF propagation can be ignored a priori, e.g., angular stability and frequency deviations. Hence, time domain simulations are necessary for the formulation of CFP models.

2.2.4 Literature review

Numerous studies have been made regarding the proposing of new CF modeling schemes and the improvement of existing models. In addition to the models listed in Table 2.2 more in-depth analyses of the different paradigms in CF modeling are summarized in Table 2.3 [26], [27], [29]–[36].

Table 2.3. CF literature review matrix (subtopic: CF modeling).

Ref	**Title**	**Year**	**Main aim**
[29]	A Study of Self-Organized Criticality of Power System Under Cascading Failures Based on AC-OPF With Voltage Stability Margin	2008	Designing a CF simulation model based on self organized criticality that includes AC OPF and lines upgrades.
[26]	An Improved OPA Model and Blackout Risk Assessment	2009	Extending the OPA model to account for power dispatch and relay protection among other elements that were not considered in the original OPA model.
[30]	A "Random Chemistry" Algorithm for Identifying Collections of Multiple Contingencies That Initiate Cascading Failure	2012	Using Random Chemistry algorithm to identify unique N-2 to N-5 contingencies that cause CFs with a minimal number of simulations that is orders of magnitude faster than Monte Carlo (MC) simulation.
[31]	Blackout Model Considering Slow Process	2013	Proposing an improved OPA model considering dispatching center actions and protective relays simulation.
[32]	Blackout Probabilistic Risk Assessment and Thermal Effects: Impacts of Changes in Generation	2013	Modeling and simulating the impact of thermal effects (wind changes and vegetation contact) in CF.

Table 2.3. Continued.

Ref	Title	Year	Main aim
[33]	Cascading Failure Analysis with DC Power Flow Model and Transient Stability Analysis	2015	proposing a DC power flow simulation for CF and performing a comparative analysis of its results with transient stability assessment tools and deriving insights about the discrepancies between them.
[34]	Revealing the Impact of Multiple Solutions in DCOPF on the Risk Assessment of Line Cascading Failure in OPA Model	2016	Extending the OPA Model to more realistic consideration of the OPF solution and dealing with outlier cases in the numerical solutions.
[35]	Cascading Failure Analysis for Indian Power Grid	2016	Modeling and simulating the Indian grid blackout and validating the findings of the simulation using the existing PMU data prior to the blackout.
[36]	Cascading failure model in power grids using the complex network theory	2016	Using complex network theory to analyze CF while also extending the model to include hidden failures and other electrical characteristics of the power network, instead of relying on topology alone.
[27]	Dynamic Modeling of Cascading Failure in Power Systems	2016	Proposing a new CF simulation algorithm (COSMIC) that uses dynamic simulations with variable time steps to enhance the prediction accuracy while minimizing the computer processing time.
[37]	Cascading Power Outages Propagate Locally in an Influence Graph That is Not the Actual Grid Topology	2017	Showing that the CF propagates non locally and constructing an influence graph of this propagation and proving it is Markovian in nature.
[38]	Impacts of Operators' Behavior on Reliability of Power Grids During Cascading Failures	2018	Attempting to model operators' behaviours (i.e., operators mistakes) in CF modeling based on historical data and interviews with power grid operators.

2.3 Component Modeling

In addition to the general approach of modeling CFs and their propagation, another method of studying CFs is to observe and model how certain components in the power grid behave under

CFs. For example, motors during stalling or the interaction between vegetation and transmission lines as time progresses. For example, regarding transmission lines and thermal line rating modeling, fixed thermal rating values are used in [39], [40]. While detailed modeling of vegetation growth and wind cooling effects are in [31], [32], [41]. Also, [42] expanded the transmission line simulation by estimating the mechanical state of the line. As another example of component modeling, protection devices' interaction with CF were modeled in [27], [43]–[45] considering aspects like zone 3 distance relay maloperation among other considerations. A summary of selected literature of component modeling in CF is presented in Table 2.4 [46]–[52].

2.4 Angle Stability

Angle stability refers to generators' ability to maintain synchronism following a contingency [53]. Angle stability is not directly related to CF. Nonetheless, it is important in modeling and predicting CFs, because the fast-cascading portion of the CF is usually an angle stability or a frequency stability issue. Hence, it is an important subtopic of CF. However, when analyzing angle stability from a CF perspective it has the major issue of being a cause and an outcome at the same time. To elaborate, an angle stability issue could initiate a CF, but a CF could also initiate an angle stability issue. Hence, dedicated CF analysis tools are still needed even with the existence of analysis methods for angle stability, and transient stability in general. A summary of the work related to using angle stability within CF is given in Table 2.5 [54]–[57].

2.5 CF Mitigation and Restoration

2.5.1 CF mitigation

Another important aspect of CF analysis and literature is the mitigation of CF. CF mitigation is concerned with the ways by which CF can be suppressed either by reducing its probability or reducing its severity. According to [58], one of the most effective ways to propel the system towards blackout is the cascaded failure of overloaded lines. Hence, an effective emergency remedial action against CFs is to perform load shedding to quickly alleviate the congestion. However, there are other ways to mitigate CFs, such as:

- Investing to upgrade the grid's infrastructure.
- Changing the operating point of the system to another point with lower risk.

- Relieving congestion before it propagates into CF.
- Improving the efficacy of restoration equipment and protocols.
- Regular maintenance checks and re-evaluation of tripping settings to prevent having devices with wrong settings or faulty components in the grid.

Table 2.4. CF literature review matrix (subtopic: component modeling and analysis in CF).

Ref	Title	Year	Main aim
[46]	Cascading Stall of Many Induction Motors in a Simple System	2012	Modeling the cascaded stall of motors following a failure in the grid.
[47]	Analysis of Induction Motor Cascading Stall in a Simple System Based on the CASCADE Model	2013	Proposing an analytical model to find the probability distribution of motor stalling CF in power systems.
[48]	On Self-Organized Criticality of the East China AC–DC Power System—The Role of DC Transmission	2013	Modeling CF for systems with DC transmission lines and calculating their blackout risk. Also, proposing a framework for choosing DC transmission operating points to reduce blackout risks.
[49]	Probabilistic Indicators for Assessing Age- and Loading-Based Criticality of Transformers to Cascading Failure Events	2014	Investigating the relation between transformers' hidden failures and CF initiation while also finding the important parameters for modeling transformer failures in CF.
[50]	Risk Assessment in Extreme Events Considering the Reliability of Protection Systems	2015	Risk assessment of power systems under extreme conditions (e.g., severe weather). The model also considers detailed circuit breakers (CB) and protective relay modeling.
[51]	A New Dynamic Performance Model of Motor Stalling and FIDVR for Smart Grid Monitoring/Planning	2016	Modeling induction motor stalling and delayed recovery within a CF model to propose an enhanced monitoring scheme for smart grids.
[52]	Real-Time Cascading Failures Prevention for Multiple Contingencies in Smart Grids Through a Multi-Agent System	2018	Simulating CF with hardware-in-loop (HIL) and introducing a multi-agent algorithm to prevent the propagation of CF.

Table 2.5. CF literature review matrix (subtopic: angle stability).

Ref	Title	Year	Main aim
[54]	A Novel Bio-Inspired Technique for Rapid Real-Time Generator Coherency Identification	2015	Rapidly identifying coherency groups in power systems using flocking behavior technique.
[55]	Angle Stability Analysis of Power System with Multiple Operating Conditions Considering Cascading Failure	2017	Analyzing angle stability while considering CF and reducing the computational burden of angle stability calculations.
[56]	Control strategy for relieving transient potential energy accumulation in power system	2018	Designing controllers for a power system based on relieving the accumulation of transient potential energy in order to prevent CFs.
[57]	Hybrid method for power system transient stability prediction based on two-stage computing resources	2018	Proposing a method for transient stability prediction in real-time based on trajectory fitting and extreme machine learning.

Since there are numerous methods to mitigate CFs, the literature pertaining to them is diverse as well. It includes improvements to relays, transmission lines, power dispatch, detecting hidden failures, estimating system information and missing data, and enhancing the power grid resiliency against cyber- and physical- attacks. The summary of these works is shown in Table 2.6.

Table 2.6. CF literature review matrix (subtopic: CF mitigation).

Ref	Title	Year	Main aim
[43]	Blocking of Zone 3 Relays to Prevent Cascaded Events	2008	Proposing a methodology to differentiate between line overloads and faults in zone 3 distance relay to reduce CF probability.
[59]	Long-Term Effect of the n-1 Criterion on Cascading Line Outages in an Evolving Power Transmission Grid	2008	Assessing long-term impacts of reliability policies on CF risks as the grid evolves, increasing both generation and demand.
[60]	Digital Grid: Communicative Electrical Grids of the Future	2011	Dividing the grids into cells based on a new concept called 'Digital Grid' to better utilize HVDC and increase renewable energy penetration while minimizing CFs probability.

Table 2.6. Continued.

Ref	Title	Year	Main aim
[61]	Sparse Overcomplete Representations for Efficient Identification of Power Line Outages	2012	Using limited voltage phasor data to identify multiple line outages in near real-time. It also aims to maximize the information gained from a limited set of PMUs or smart meters.
[62]	Determination of available transfer capability with implication of cascading collapse uncertainty	2014	Determination of the available transfer capacity in transmission lines considering the likelihood of CF using the bootstrap technique.
[63]	Efficient Location Identification of Multiple Line Outages with Limited PMUs in Smart Grids	2015	Using a location identification scheme to identify line outages with very limited PMU measurements.
[64]	Dynamic Detection of Transmission Line Outages Using Hidden Markov Models	2016	Detecting line outages (estimating the state of all lines) using limited information and under a noisy environment.
[65]	Boosting the Power Grid Resilience to Extreme Weather Events Using Defensive Islanding	2016	Improving the resiliency of the grid by introducing a risk assessment method that would pre-emptively separate a grid into stable, self-sustaining islands.
[66]	CCPA: Coordinated Cyber-Physical Attacks and Countermeasures in Smart Grid	2017	Analyzing cyber-physical attacks and devising methods to counteract them in smart grids using secure PMUs and online tracking of the power system impedance.
[45]	Real-time monitoring of zone 3 vulnerable distance relays to prevent maloperation under load encroachment condition	2017	Proposing a monitoring scheme to rank distance relay in real-time to detect the most vulnerable relays to maloperation.
[44]	A D-S Evidence Theory-Based Relay Protection System Hidden Failures Detection Method in Smart Grid	2018	Detecting hidden failures in relay protection system using D-S evidence theory.
[67]	Decentralized Implementation of Unit Commitment with Analytical Target Cascading: A Parallel Approach	2018	Proposing an implementation of the unit commitment optimization problem that is less prone to cyber attacks.

Table 2.6. Continued.

Ref	Title	Year	Main aim
[68]	Alleviation of post-contingency overloads by SOCP based corrective control considering TCSC and MTDC	2018	Controlling and optimizing the DC transmission lines power flow to reduce CF probabilities and blackout risk.
[69]	A Robustness-Oriented Power Grid Operation Strategy Considering Attacks	2018	Designing a constrained optimal power flow (COPF) framework that extends the existing OPF by considering CFs and malicious attacks.
[70]	Frequency Derivative-Based Inertia Enhancement by Grid-Connected Power Converters with a Frequency-Locked-Loop	2019	Introducing an inertia enhancement method for power grids with high penetration of renewable energy sources to reduce load sheddings and CFs.
[71]	Mitigating False Data Attacks Induced Overloads Using a Corrective Dispatch Scheme	2019	Implementing a corrective dispatch mechanism that is secure against cyber threats such as data attacks.

2.5.2 CF restoration

In case a CF event is not stopped or suppressed successfully then the only form of CF mitigation that can be applied is restoration. CF restoration refers to the steps and protocols that can aid in the quick restoration of normal operation. It includes elements such as black start protocols and guidelines for managing loads following a partial blackout. CF restoration is essential to CF, because it minimizes the expected energy not supplied (EENS) by reducing the total time required to return to normal operation. Hence, reducing the overall severity of the CF. Selected articles about grid restoration techniques following a CF are shown in Table 2.7.

2.6 Cybersecurity and Malicious Attacks

Following the malicious cyber attack on the Ukrainian grid network in 2015, the interest in cybersecurity significantly increased [72]. Moreover, since that cyber attack induced a CF, more research has been done in analyzing which malicious attacks can initiate CFs and proposing methods to counteract these attacks. Also, the study of the ‘worst possible attack’ was investigated in many articles, where the aim is to find the set of lines or bus stations that has the highest susceptibility to induce a large-scale blackout if a cyberattack targeted them and proposing

methods to protect such lines and bus stations. Table 2.8 provides a summary of the recent cybersecurity research with emphasis on CFs and blackouts.

Table 2.7. CF literature review matrix (subtopic: system restoration).

Ref	Title	Year	Main aim
[73]	A Hybrid Multiagent Framework With Q-Learning for Power Grid Systems Restoration	2011	Designing a Q-learning based framework for grid restoration control (loads and transmission lines prioritization) for both centralized and decentralized grids
[74]	Conceptual Design of a Multi-Agent System for Interconnected Power Systems Restoration	2012	Designing a multi-agent-based system for managing various aspects of grid restoration.
[75]	Incentivizing Energy Reduction for Emergency Demand Response in Multi-Tenant Mixed-Use Buildings	2018	Minimizing the energy shedding in emergency demand response cases and minimizing the losses by maximizing the benefits of the available emergency energy.

Table 2.8. CF literature review matrix (subtopic: cybersecurity).

Ref	Title	Year	Main aim
[76]	A Coordinated Multi-Switch Attack for Cascading Failures in Smart Grid	2014	Assessing multi switch attack cybersecurity threats and finding the worst-case scenarios of CFs.
[77]	Bilevel Model for Analyzing Coordinated Cyber-Physical Attacks on Power Systems	2016	Analyzing the effects of a cyber-physical coordinated attack to identify the most damaging and hard to detect attacks. Also, proposing methods to counteract them.
[78]	Price Modification Attack and Protection Scheme in Smart Grid	2017	Investigating price modifications attacks and their effects on CF. Then, devising approaches to analyze the problem and form a protection scheme against it.
[79]	Cyber cascades screening considering the impacts of false data injection attacks	2018	Finding the most severe false data injection attacks to a power grid in terms of their resulting CF.
[80]	Local Cyber-Physical Attack for Masking Line Outage and Topology Attack in Smart Grid	2019	Modeling a worst-case scenario malicious cyber-physical attack that can deceive the control center into assuming that the outage is at a different location.

2.7 Interdependent Systems

With the advent of smart grids and renewable energy sources, power grids became increasingly dependent on other networks and systems. Among the most notable interaction is the coupling between the power grid and the communication network due to smart sensors and WAMS. Another common interaction is the one between the transmission of power through transmission lines and the transmission of natural gas through pipelines as both are coupled through the generators' dependency on gas availability. In the case of communication networks, the failure in the power grid means a failure in the communication system (unless backup generators are used to power the communication system), but at the same time, a failure in the communication system could lead to a failure in the power grid due to incorrect or missing information regarding the state of the system. This dual interaction between power grids and communication systems leads to more ways by which a power grid can fail and enter a CF that could potentially lead to huge blackouts in both systems. A similar event can happen between gas and electricity, albeit with different interactions, and it is one of the main causes for the propagation of the 2021 Texas outage [8]. To understand and analyze these phenomena, multiple research works addressed the modeling of coupled systems and proposed new metrics to calculate the modified risks of having a CF or a blackout in a coupled system. Table 2.9 provides a summary of the recent interdependent systems research with emphasis on CFs and blackouts interactions with communication grids and vice versa.

Table 2.9. CF literature review matrix (subtopic interdependent system analysis).

Ref	Title	Year	Main aim
[81]	Detecting Critical Nodes in Interdependent Power Networks for Vulnerability Assessment	2013	Detection of the critical nodes that leads to the severest CF within an interdependent system.
[82]	Cascading Failures in Interdependent Infrastructures: An Interdependent Markov-Chain Approach	2016	Analyzing coupled systems and interdependent systems in terms of reliability and CF propagation using Markov chains.
[83]	Cascading Failure Analysis Considering Interaction Between Power Grids and Communication Networks	2016	Analyzing cyber physical systems in terms of CF interactions and proposing metrics and methods to improve their security.

Table 2.9. Continued.

Ref	Title	Year	Main aim
[84]	Framework for vulnerability assessment of communication systems for electric power grids	2016	Assessing the vulnerability of communication systems and proposing metrics to quantify the severity of the vulnerability.
[39]	Control of Communications-Dependent Cascading Failures in Power Grids	2019	Investigating the communication delay and communication failure effects on worsening the CF propagation and how to model control schemes that are more resilient to such issues.
[85]	Cost Efficient Data Aggregation Point Placement with Interdependent Communication and Power Networks in Smart Grid	2019	Proposing a method to optimally place data aggregators for sensors (PMUs or smart meters) in a power system while considering the CF in the communication and power network.

2.8 CF Prediction and Risk Assessment

Conventionally, power grid operators perform risk assessment during dispatch, in day ahead, or after a system change instead of predicting or assessing CFs. In risk assessment, the power grid is analyzed against a list of all possible N-1 contingencies and a selected set of N-2 contingencies using time-domain simulations with detailed system dynamics. This set of all possible N-1 and the selected N-2 contingencies can include up to 20,000 scenarios in some practical cases, which are depending on the application simulated each hour or even each couple of minutes [86]. In addition to applying detailed dynamic models, risk assessment also includes a DC reliability analysis that can even go up to N-5 when assessing the risks of blackouts (only the subset of the N-k cases that has high probability are considered). Typically, these analyses yield multiple vulnerability and criticality metrics that the industry uses to maintain an acceptable level of risk. The 0.1 days/year outage is often used, which means that the power system on average, and each bus therein, will only suffer a cumulative of 1 day worth of outages every 10 years. Another metric is the N-1, which indicates that the system can withstand any individual failure in any element in the grid (e.g., transmission line, generator, transformer…etc.). However, these metrics and similar metrics are no longer sufficient to prevent large-scale blackouts as evident from the blackouts that happened in the past. Moreover, these methods do not always satisfy the requirements to perform a real-time online assessment of large, interconnected power systems

[87]. Hence, CF prediction and analysis emerged, where the goal is to understand and suppress CF events regardless of their perceived rarity. For example, the 2003 blackout and the CF that initiated it happened when the system was undergoing an N-5 contingency, which is an exceedingly unlikely event from a probabilistic point of view. Nonetheless, operators are still reluctant to use new monitoring systems concerning CFs, because they are not rigorously tested to provide sufficient verification of their accuracy. However, the ongoing research in CF can lead to an improved understanding and testing methods of the newly formulated CF monitoring schemes, which should encourage future integration of these methods to the power grid.

Current CF prediction and risk assessment literature tackles the problem from a variety of approaches. One approach is the vulnerability and risk assessment with CF events consideration. These works aim to perform the typical severity metrics found in the industry but using CF models. A second approach is to assess the stochastic propagation of CFs considering random factors such as hidden failures and deviations between the calculated and actual thermal line rating to quantify the power grid sensitivity to these unexpected failures. Also, system assessment can be achieved by scanning the space of N-k (where $k \geq 2$) contingencies to quickly find all the CFs that induce large-scale blackouts and propose system modifications to mitigate them. Another approach is to propose models to predict the propagation and severity of CFs and blackout in real-time. Table 2.10 provides a summary of CF prediction literature, Table 2.11 summarizes the CF literature with focus on risk assessment and Table 2.12 provides a summary of the system assessment approach to CF.

2.9 Literature Review and Benchmarking

Considering the diverse nature of CF, multiple research works summarizing existing CF literature are available. Most significantly are the papers, proceedings, and presentations from the IEEE PES CAMS Working Group on Understanding, Prediction, Mitigation, and Restoration of Cascading Failures [88]. Among these works is [19], which aims to design a framework for the future testing and benchmarking of existing CF modeling and analysis tools, as well as, elaborating on the difficulties that prevent having a comprehensive benchmarking standard for CF. Furthermore, [9] provides overall information regarding blackouts, detailed CF modeling and analysis tools comparison, and a thorough review of existing CF literature. More research work on CF modeling and predictions is reported in the references given in Table 2.13.

Table 2.10. CF literature review matrix (subtopic: CF prediction).

Ref	Title	Year	Main aim
[89]	Estimating the Propagation and Extent of Cascading Line Outages from Utility Data with a Branching Process	2012	Predicting the distribution of line outages in a CF using branching process based on utility data.
[40]	Stochastic Analysis of Cascading-Failure Dynamics in Power Grids	2014	Proposing a tractable probabilistic model for CF dynamics based on the number of failures, maximum capacity of failed lines, and estimation errors.
[90]	An Interaction Model for Simulation and Mitigation of Cascading Failures	2015	Generating an interaction network from either historical or simulated data to quickly find the CF propagation paths alongside the critical lines and components in the network.
[91]	A Three Stages Decision Tree-Based Intelligent Blackout Predictor for Power Systems Using Brittleness Indices	2018	Using decision trees with brittleness indices based on PMU measurements from the grids to predict the blackout size following a disturbance in real-time.
[92]	Efficient Estimation of Component Interactions for Cascading Failure Analysis by EM Algorithm	2018	Estimating the CF propagation from an incomplete set of simulations to drastically reduce simulation time.

Table 2.11. CF literature review matrix (subtopic: risk assessment considering CF).

Ref	Title	Year	Main aim
[41]	Composite Power System Vulnerability Evaluation to Cascading Failures Using Importance Sampling and Antithetic Variates	2013	Quickly estimating the expected energy not served (EENS) using MC simulations with importance sampling. Moreover, the research considers overgrown vegetation, relay overtripping, and other practical considerations.
[93]	Towards Estimating the Statistics of Simulated Cascades of Outages with Branching Processes	2013	Estimating the blackout probability distribution using branching process.
[94]	Splitting Method for Speedy Simulation of Cascading Blackouts	2013	Using splitting method to quickly find CF with low probabilities but high impact.

Table 2.11. Continued.

Ref	Title	Year	Main aim
[95]	N-k Induced Cascading Contingency Screening	2015	Screening for CF inducing contingencies in N-k space to quickly estimate risk.
[96]	Efficient Splitting Simulation for Blackout Analysis	2015	Proposing a technique to quickly estimate the probability of rare events to efficiently estimate blackout risk and locate vulnerable lines.
[97]	Estimating Cascading Failure Risk with Random Chemistry	2015	Introducing a Random Chemistry based risk estimation that is multiple orders of magnitude faster than Monte Carlo simulations in estimating the risks of CF. Also, the method introduces a new metric to find strategies to reduce CF risk
[24]	A Two-Level Probabilistic Risk Assessment of Cascading Outages	2016	Introducing a two-level framework to simulate CFs as two phases (slow and fast cascade) and deriving risk and EENS from it.
[98]	Quantifying the Influence of Component Failure Probability on Cascading Blackout Risk	2018	Deriving the relationship between blackout risk and component failure probability.
[99]	Vulnerable transmission line identification considering depth of K-shell decomposition in complex grids	2018	Using a K-shell decomposition method to locate vulnerable lines in large power systems.

Table 2.12. CF literature review matrix (subtopic: system assessment).

Ref	Title	Year	Main aim
[42]	Mechanical State Estimation for Overhead Transmission Lines with Level Spans	2008	Monitoring and estimating the mechanical state (e.g., sag levels) of transmission lines to provide early warning to operators regarding lines vulnerable to tree contacts or dangerous sagging levels.
[100]	Vulnerability Assessment Scheme for Power System Transmission Networks Based on the Fault Chain Theory	2011	Using fault chain theory to assess the vulnerability of transmission lines efficiently.
[101]	Method for evaluating the importance of power grid nodes based on PageRank algorithm	2014	Modifying PageRank algorithm to suit the importance ranking of power grid buses problem to order the buses based on their importance in contributing to CFs and blackouts.

Table 2.12. Continued.

Ref	Title	Year	Main aim
[102]	Multi-attribute node importance evaluation method based on Gini-coefficient in complex power grids	2016	Proposing a node importance evaluation metric based on both topological and electrical characteristics of the power grid.
[103]	Impact of Topology on the Propagation of Cascading Failure in Power Grid	2016	Analyzing blackouts by reducing a grid to its topology and analyzing it through branching process.
[104]	Fast Screening of Vulnerable Transmission Lines in Power Grids: a PageRank-based Approach	2017	Modeling CF as a directed weighted graph and using PageRank approach to quickly screen the vulnerable lines.
[105]	Vulnerable transmission line identification using ISH theory in power grids	2018	Using improved structure hole theory to identify the most vulnerable transmission lines in power grids.
[106]	Quickest Localization of Anomalies in Power Grids: A Stochastic Graphical Framework	2018	Detecting the locations of anomalies (e.g., line outages) in a power grid quickly and accurately using stochastic approaches.

Table 2.13. CF literature review matrix (literature review and benchmarking).

Ref	Title	Year	Main aim
[107]	Causes of the 2003 major grid blackouts in North America and Europe, and recommended means to improve system dynamic performance	2005	Analyzing the 3 major blackouts that happened in 2003 and proposing recommendations to prevent future blackouts based on the causes of these blackouts.
[25]	Survey of Tools for Risk Assessment of Cascading Outages	2011	Reviewing the state of the art in cascading outages risk assessments both in the literature and industrial applications. Also, it reviews and compares the current industry practices for evaluating risks.
[7]	Risk Assessment of Cascading Outages: Methodologies and Challenges	2012	Reviewing the methods of estimating risks and modeling and analyzing CFs and blackouts.

Table 2.13. Continued.

Ref	Title	Year	Main aim
[19]	Benchmarking and Validation of Cascading Failure Analysis Tools	2016	Reviewing the existing models, tools, and algorithms to simulate and analyze CF and proposing methods and criteria to benchmark and validate both future and existing CF tools.
[9]	A Critical Review of Cascading Failure Analysis and Modeling of Power Systems	2017	A detailed comparison between the different CF modeling techniques and models indicating the main pros and cons of each model.
[108]	Impact Assessment of Hypothesized Cyberattacks on Interconnected Bulk Power Systems	2018	A literature review of cyberattacks, cyber-infrastructure protection, and mitigation strategies in power grids.

3 GENERATION OF CASCADING FAILURE CASES

3.1 Introduction

CF prediction requires a large amount of data (either simulated cases, historical data, or a mixture of both) to cover the many scenarios by which a system can enter a CF and the subsequent propagation of these CFs. However, CFs are exceedingly rare events in power systems due to their high reliability. Thus, historical data alone are not sufficient for building an accurate CF prediction model. Moreover, studies that analyze previous blackouts and historical data, such as [107], only provide general information regarding common trends in CFs and blackouts, as well as, comparing historical data with the available simulation tools. Hence, these studies cannot be extrapolated to other contingencies, other CF paths, or other power grids. Thus, it is important to synthetically generate data for CFs and blackouts to train CF prediction models and validate their efficacy.

To generate CF scenarios, a CF simulation tool is needed. As shown in Section 2.2, topological and QSS models do not provide an accurate description of CFs. Thus, it is necessary to use dynamic models (i.e., time domain simulation) to generate accurate CF data. In this thesis, the COSMIC package is used to generate the data since it is an open-source MATLAB-compatible dynamic model for CF simulation. COSMIC has multiple advantages over other dynamic CF models. For example, it uses dynamic time steps and includes ZIPE (impedance, current, power, and exponential) load model. COSMIC models the power system as a set of differential and algebraic equations (DAE) with additional discrete equations to model relays' action. Following this model, the power system state can be defined using three state vectors $x(t)$, $y(t)$, and $z(t)$. The three state vectors are described by the following equations [27]:

$$\frac{dx}{dt} = \mathrm{f}(t, x(t), y(t), z(t)) \tag{3.1}$$

$$\mathrm{g}\big(t, x(t), y(t), z(t)\big) = 0 \tag{3.2}$$

$$\mathrm{h}\big(t, x(t), y(t), z(t)\big) < 0 \tag{3.3a}$$

$$\mathrm{d}\big(t, x(t), y(t), z(t)\big) = 0 \tag{3.3b}$$

where:

$x(t)$ corresponds to the continuous state variables and they are described by the set of differential equations $\mathrm{f}(t, x(t), y(t), z(t))$, e.g., the angular speed in the swing equation.

$y(t)$ corresponds to the state variables which only have algebraic equations and are thus described by the set of algebraic equations $\mathrm{g}\big(t, x(t), y(t), z(t)\big)$, e.g., bus voltages in the load flow equations.

$z(t)$ corresponds to state variables that are binary, $z \in \{0,1\}$, and they are described as a function $\mathrm{h}(t, x(t), y(t), z(t))$ that activates a counter function $\mathrm{d}\big(t, x(t), y(t), z(t)\big)$ when the conditions of h are not satisfied, e.g., relay tripping action modeling where the h function is the line temperature equation and the counteraction function d is changing the bus impedance matrix.

Hence, from these three state vectors, all the equations and variables in the power system can be transformed into equations in the form of (3.1)-(3.3) where the variables are subsets of $x(t)$, $y(t)$, and $z(t)$.

3.2 Data Generation

Typically, to generate the different CF scenarios Monte Carlo (MC) simulations are used, but it was avoided in the studies conducted in this thesis because it has many downsides. MC underestimates the probabilities of N-3 and above contingencies whereas CFs are very sparse in N-2 and are almost non-existing in N-1 cases. For example, during simulating 31,000 N-2 scenarios of the US eastern interconnection system, only 953 cases had overloaded lines and only 38 initiated a cascading event [23]. Hence, only one case out of a thousand cases resulted in a CF. Thus, MC simulations would require a very long simulation time to generate sufficient CF scenarios. Moreover, the aim of MC simulations is to give a good estimate of the risk (i.e., severity multiplied by likelihood) of a power grid or a particular dispatch. However, this means that MC does not provide an equal portion of normal and abnormal cases since power systems are highly reliable by design. Hence, an alternative approach to avoid the downsides of MC simulations is to generate cases such that the number of normal events and the number of undesirable events (CFs, partial blackouts, or total blackouts) are close to each other [109]. This goal can be achieved either

by sampling from a different probability distribution that gives more weights to N-2+ contingencies (e.g., using importance sampling (IS) [110]) or by directly generating events only from a sample space that only has N-2+ contingencies. Alternatively, the training itself can be skewed to give more importance to abnormal cases accuracy to penalize the system for lumping the abnormal cases with the larger set of the normal cases, which can be achieved using a modified extreme learning machine (ELM) [111]. However, this approach is more suited if the abnormal cases are a good representation of the entire space of abnormal cases, which is not the case in CFs. Hence, the preferable approach is to make both the normal and abnormal cases have an almost equal number of cases, or at the very least are within the same order of magnitude, using either IS or sampling from N-2+.

For the data generation in this research, the following factors are considered for modification between each case:

- Loading level
- Fault location
- Lines out of service
- Generators out of service.

These factors are selected based on two criteria, improving the robustness of the trained model, and generating meaningful information. For each factor, a different method is used to generate the random data. For the loading level, a uniform distribution between 0.6 to 1.5 of the nominal loads is used to generate the random datapoints. The uniform distribution is used, rather than the normal distribution, to generate more data at the tails of the distribution, as shown in Figure 3.1 that compares the sampling of the load using normal distribution and uniform distribution. Generating data at the tail of the distribution is crucial because CFs typically happen at the tail of the distribution. For fault location, it is generated at a random transmission line giving equal probability for each transmission line (i.e., assuming all transmission lines have equal outage probability). As for the lines out of service, it is sampled such that 80% of the cases have no lines out of service, 16% of the cases have one line out of service, and 4% of the cases have two lines out of service. This selection does not reflect the actual power grid, but rather forces the data generation to cover multiple topologies to increase the robustness of the prediction model and to increase the likelihood of scenarios with CFs. Similarly, generators out of service are sampled as

80% of cases have no generator out of service, and 20% have one generator out of service. In both the lines out of service and generators out of service the location of components is randomly sampled with equal probabilities.

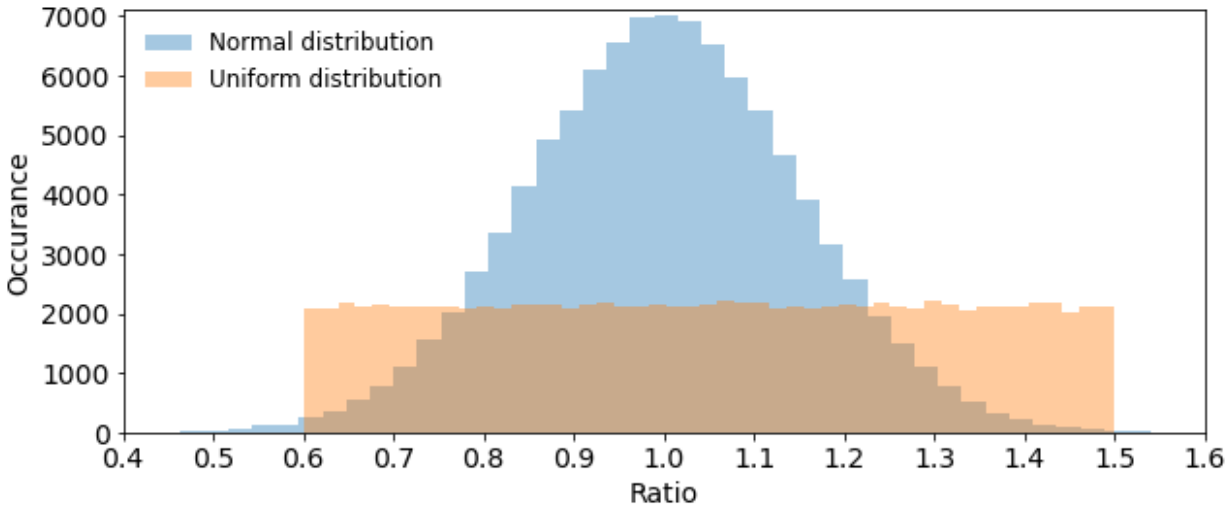

Figure 3.1. Comparison between sampling 100,000 load ratios using normal uniform distribution.

3.3 Constructing PMU Data

One of the COSMIC model advantages is that it uses variable time steps to have high accuracy during fast cascade (e.g., system transient behaviour immediately after a contingency) and high simulation speed during slow cascade (e.g., line heating dynamics). This means that the actual timestep between each measurement varies significantly, as shown in Figure 3.2. It can be seen from this figure that the gaps between the time steps are very narrow (around 25 ms) during the first half second of the simulation, while they reach 0.2 s when the system is experiencing slower dynamics (e.g., between $t = 3\text{–}4\ s$). In the dataset used in this thesis, the time step in COSMIC is set to be at a maximum of 1 s (during steady state) and a minimum of 1 μs (during fast transients), but it usually ranges between 10 ms and 100 ms. Hence, a 100 s simulation could take anywhere between 100 timesteps to >10,000 timesteps[1]. However, this means that the data cannot be directly used as PMU measurements. Hence, to construct the PMU measurements, linear

[1] Theoretically, the greatest number of timesteps in a 100s case is 100×10^6. However, in a dataset with 50,000 cases, the largest number of time steps was 20,000. This disparity between theoretical and actual maximum is because power systems oscillate between steady state and fast dynamics. Hence, there can never exist a case with fast transient events happening for the full duration of 100s.

interpolation is used between the datapoints to obtain the hypothetical PMU measurements (PMUs are assumed to be operating at a sampling rate of 60 samples per second). Figure 3.3 shows the constructed PMU measurements from a portion of the curve shown in Figure 3.2 where the number of PMU points between each time step varies between 4 to 9 points. These calculated points refer to the PMU measurements in the ideal scenario without considering PMU's internal delay, communication delay, noise, or the estimation error during transient. However, non-idealities will be considered in the next chapters by adding noise, and missing measurements to the constructed ideal PMU measurements.

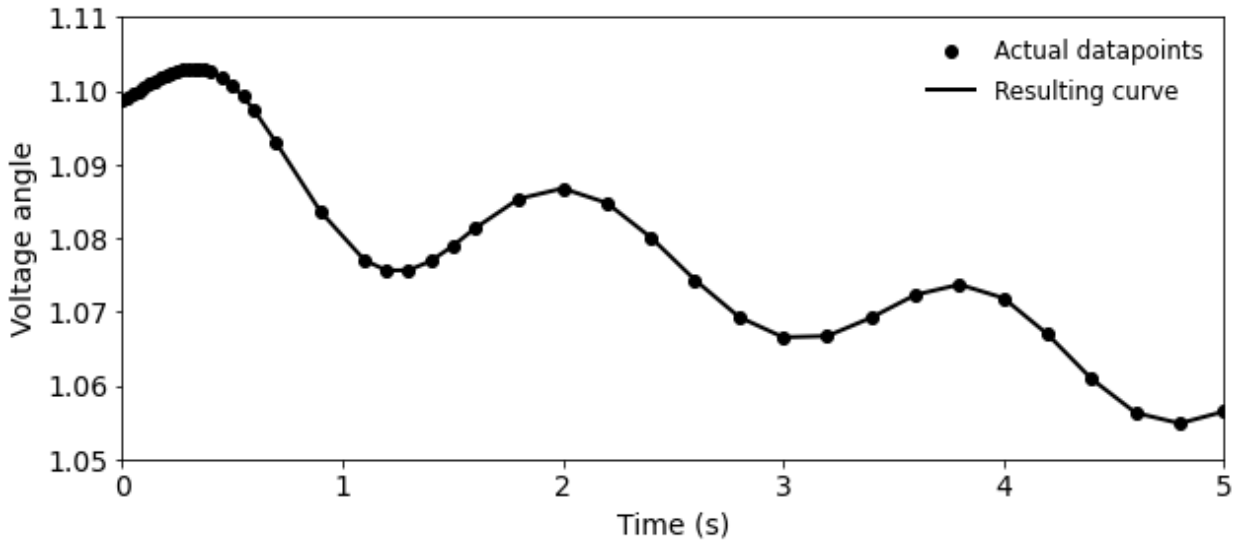

Figure 3.2. COSMIC variable time step. At the start of the contingency, the model uses more datapoints for high accuracy while larger time steps are used when there are slower dynamics.

3.4 IEEE 39-Bus System

3.4.1 System information and data summary

The IEEE 39-bus system is used in this thesis as an example of a small-scale power network. The single line diagram (SLD) of the system is shown in Figure 3.4, and additional information about it are provided in Appendix C. The number of generated cases is 54,756 and their description is provided in Table 3.1. These cases are generated according to the methodology mentioned in Section 3.2 and the construction of PMU data is done as per Section 3.3. In each case, the system is simulated until it collapses, stabilizes or the simulation time ends (100 s is selected as the simulation time).

To visualize the distribution of the data, the histogram of the LOL (in MW) in all cases is plotted in Figure 3.5 and the Pareto plot (ordered histogram, with the accumulated percentage of data) in Figure 3.6. It is observed from these figures that most of the LOL are in the intervals 0-1000 MW and 5000-7000 MW and only a negligible number of cases have LOL between 1000-5000 MW. This skewed distribution is due to multiple causes. First, most simulation tools, including COSMIC, overestimate the LOL by assuming that any nonconverging network (or nonconverging island) is suffering from a total blackout. Second, since the IEEE 39-bus system is a small system, multiple N-2+ contingencies can cause a significant load-generation imbalance that leads to instability issues leading to a total blackout (assuming no RAS intervention). Third, even with considering N-2+ contingencies, power systems are reliable. Hence, the most common outcome is having cases with 0% blackouts.

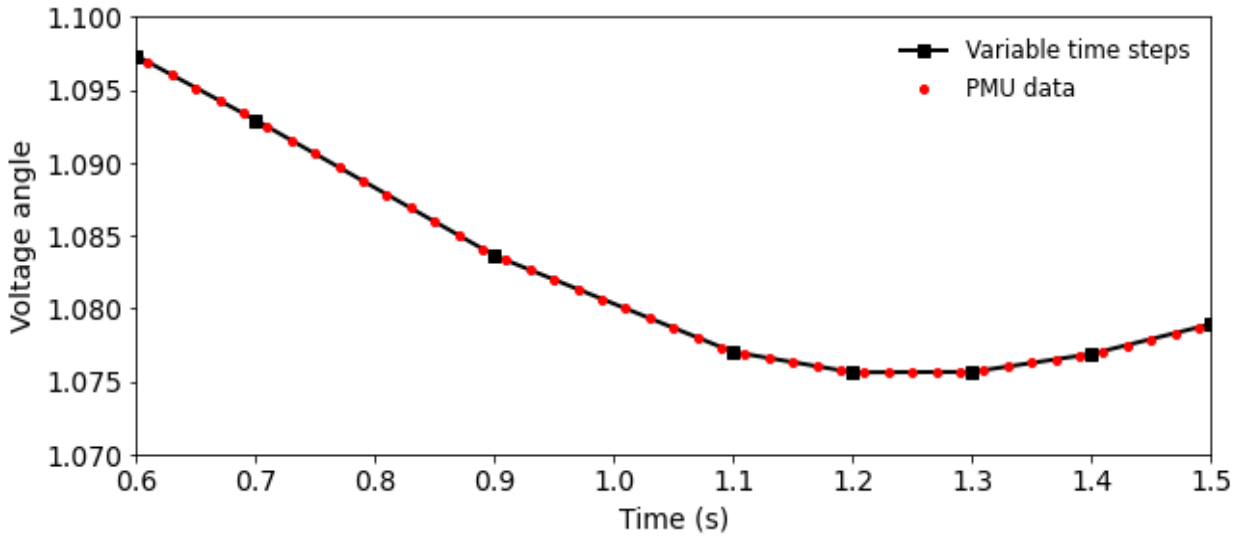

Figure 3.3. Linearly interpolating PMU measurements from the variable time steps generated by COSMIC.

Table 3.1. IEEE 39-bus system simulated cases summary.

Type	Number of cases
Simulated cases	54,756
Total blackout	24,364
Load shedding or partial blackout	8,329
No blackout	22,063

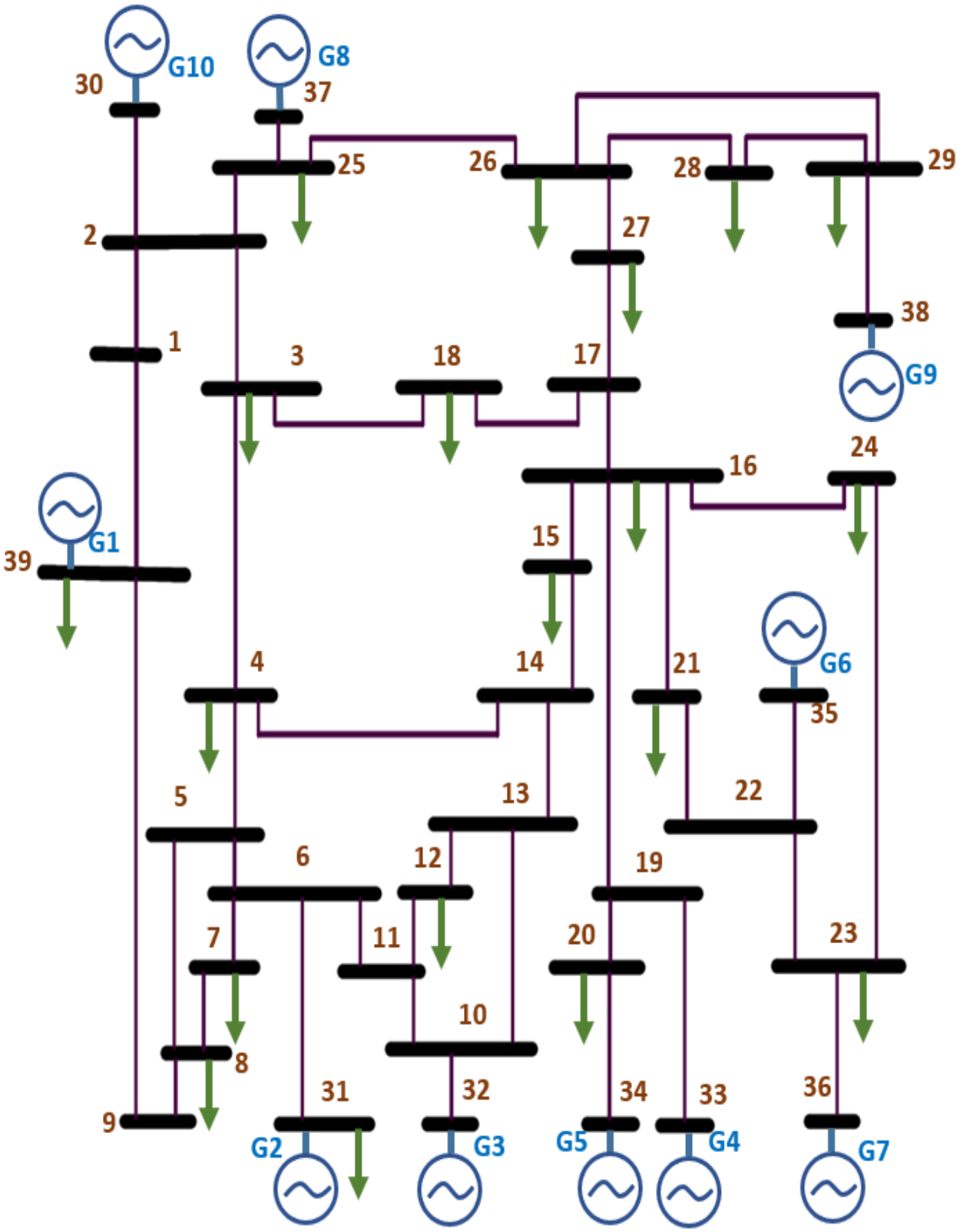

Figure 3.4. IEEE 39-bus system SLD.

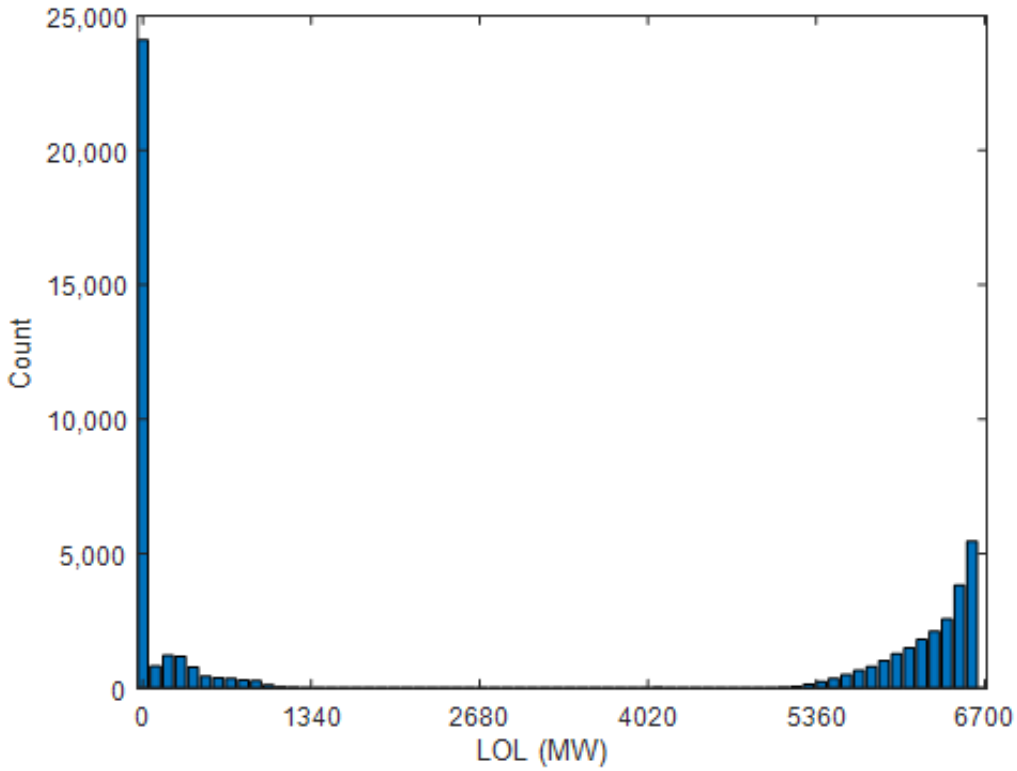

Figure 3.5. Distribution of the LOL (in MW) in the generated IEEE 39-bus system cases.

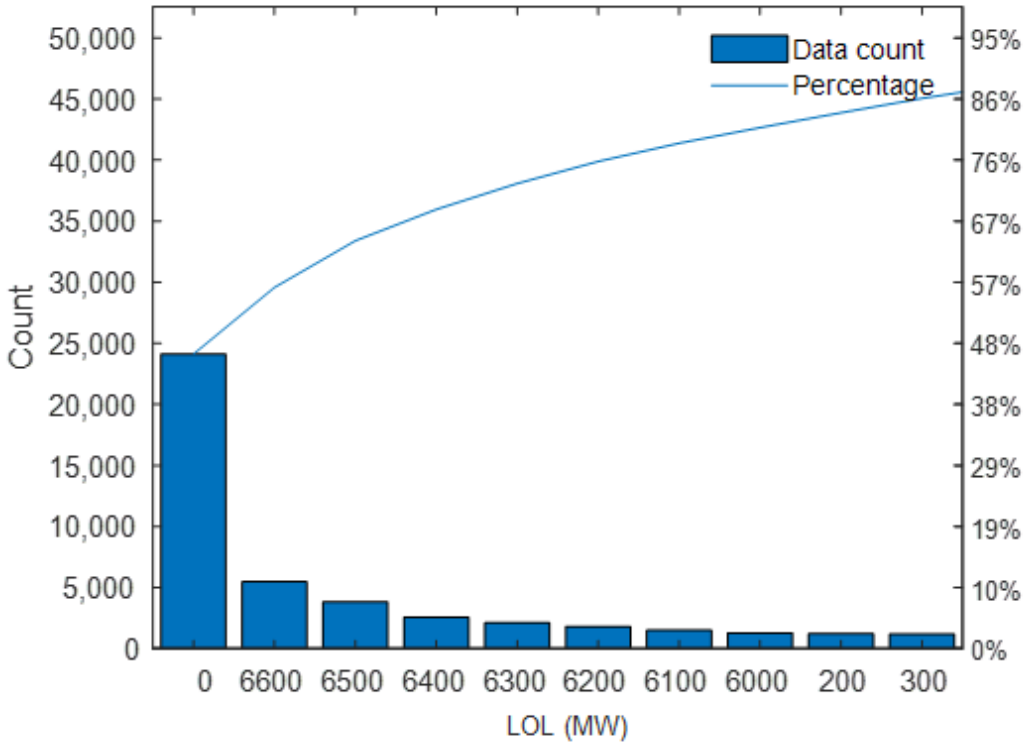

Figure 3.6. Pareto chart of the LOL (in MW) in the generated IEEE 39-bus system cases.

3.4.2 Sample scenarios

To demonstrate the content of each case, three case examples are selected from the simulated cases. The first case has a contingency that initiated a CF which then resulted in a total blackout, the second case has no outages (system stabilizes after the loss of lines due to the faults without any LOL), and the third case has a cascading event (a CF) that ends with a partial blackout due to improper islanding.

In the first example (Case number 5)[1], the system is stressed due to multiple loads having more than a 25% increase in their nominal values, then a fault happens at line 21 (between buses 16 and 17, Figure 3.4) causing that line to trip at $t = 10\ s$. The loading of the system at the time of the fault is described by Table 3.2 where the numbers refer to the load multiplier in comparison to the nominal loading of the bus (refer to Appendix C for load buses information and transmission line locations). The progression of the cascade is detailed in Figures 3.7 to 3.9. Figure 3.7 shows the voltages at each bus against time, Figure 3.8 shows the voltage angle (theta), and Figure 3.9 shows the rotor speed (omega). It can be seen from these figures that the first trip occurred at $t = 10\ s$ (exogenous event) and that the system exhibits unstable oscillations. At around $t = 20\ s$, the under-voltage load shedding (UVLS) scheme started tripping loads to reduce the dips in voltage. It tripped loads at buses 3, 4, 7, 8, 11, 13, 15 and 31. However, this increased the oscillations instead of stabilizing the system. Then, at about $t = 22.4\ s$ the system was declared to be in a total blackout state (assuming worst case scenario) since the dynamic simulation could not converge to any operating point. The summary of the events is shown in Figure 3.10 which displays the events log of the COSMIC simulation.

Table 3.2. Loading level at each load in Case number 5 (compared to nominal values).

Load number	1	2	3	4	5	6	7	8	9	10
Load multiplier	0.744	1.159	1.370	0.635	0.837	0.798	1.067	1.123	1.270	1.286
Load number	11	12	13	14	15	16	17	18	19	
Load multiplier	1.400	1.123	1.152	0.646	0.891	0.610	0.874	1.205	1.402	

[1] The numbering used in this Chapter refers to the internal numbering of cases in the 54,756 data. It is only provided for cross-referencing with the dataset.

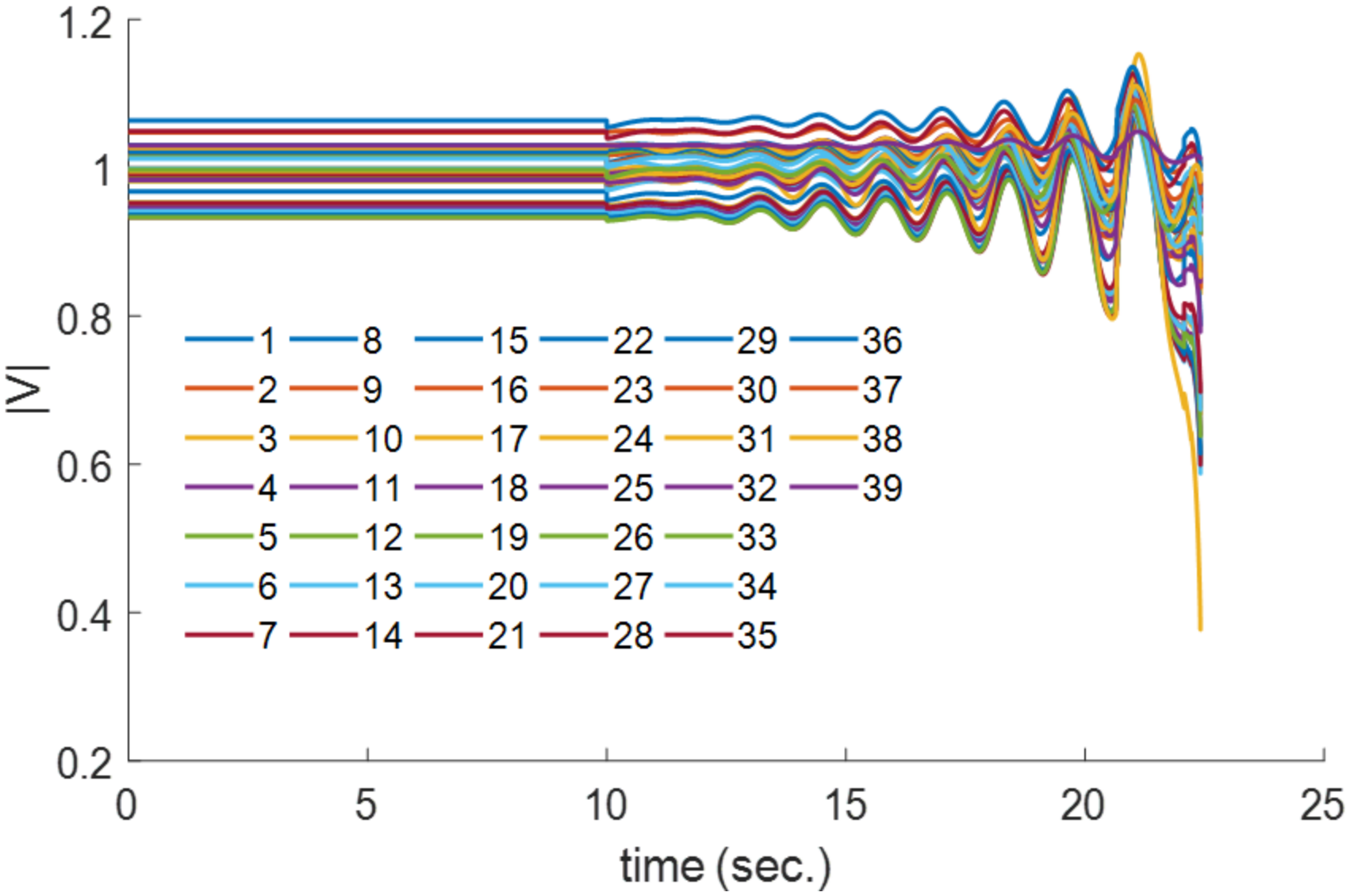

Figure 3.7. Voltage magnitude at each bus (Case number 5).

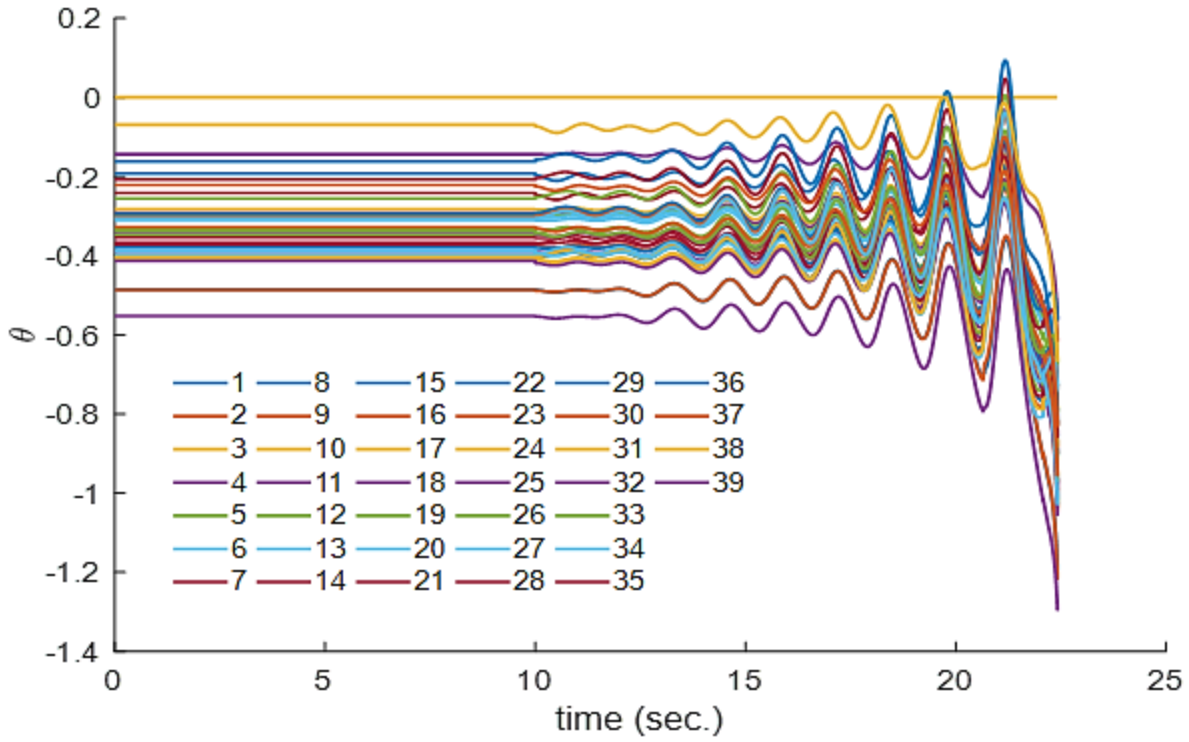

Figure 3.8. Voltage angle at each bus (Case number 5).

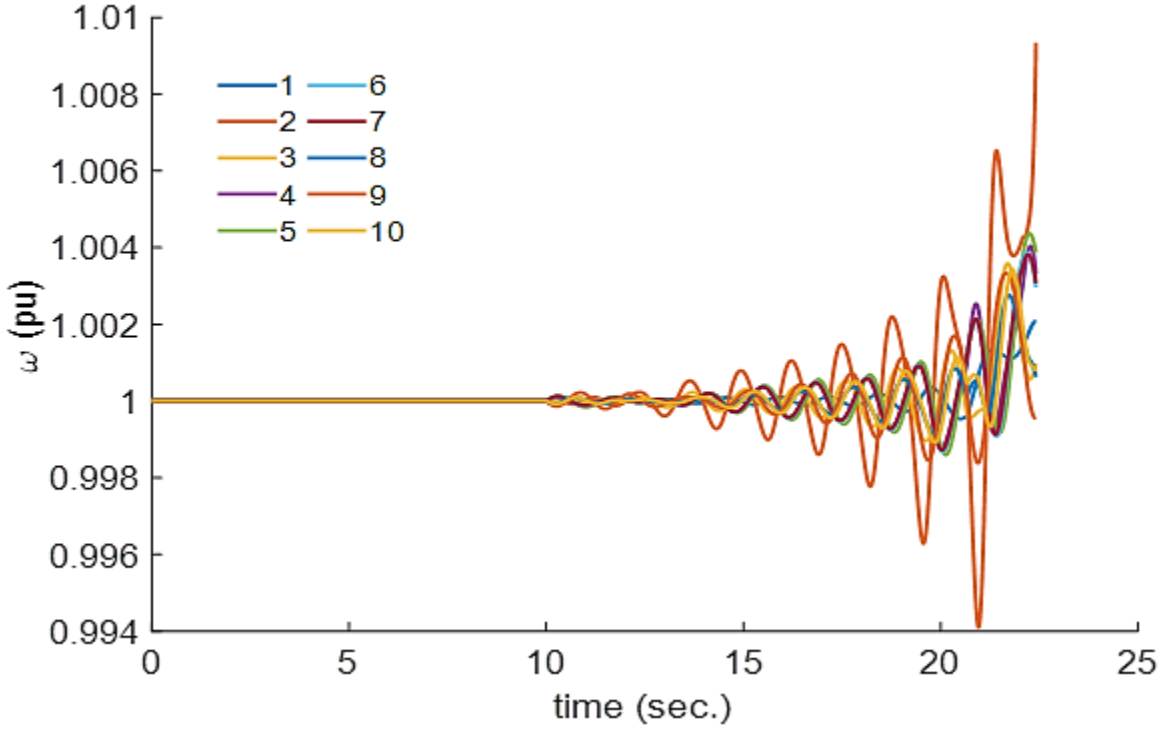

Figure 3.9. Rotor speed at each generator bus (Case number 5).

```
starting simulation at t = 0
writing results to IEEE_39_5_20210525T012401.csv
 simulation start event. simulation start event. simulation start event.
simulation start event.
 Simulating from t=0 s to t=10 s
 Completed simulation up until t=10
 Writing simulation results to IEEE_39_5_20210525T012401.csv
  t = 10.0000: Branch 21 tripped...

 Simulating from t=10 s to t=110 s
  t = 20.6510: UVLS relay trip at bus 12...
  t = 20.6510: 1.78 MW of load shedding at bus 12...
  t = 20.6550: UVLS relay trip at bus 7...
  t = 20.6550: 80.09 MW of load shedding at bus 7...
  t = 20.6650: UVLS relay trip at bus 8...
  t = 20.6650: 82.80 MW of load shedding at bus 8...
  t = 20.6800: UVLS relay trip at bus 4...
  t = 20.6800: 144.86 MW of load shedding at bus 4...
  t = 22.0610: UVLS relay trip at bus 31...
  t = 22.0610: 2.77 MW of load shedding at bus 31...
  t = 22.0750: UVLS relay trip at bus 15...
  t = 22.0750: 63.84 MW of load shedding at bus 15...
  t = 22.2200: UVLS relay trip at bus 3...
  t = 22.2200: 59.86 MW of load shedding at bus 3...
  t = 22.4150: UVLS relay trip at bus 18...
  t = 22.4150: 44.34 MW of load shedding at bus 18...
Reducing Newton step size.
Reducing Newton step size.
Reducing Newton step size.
Reducing Newton step size.
Reducing Newton step size.
Reducing Newton step size.
Reducing Newton step size.
Reducing Newton step size.
Reducing Newton step size.
Reducing Newton step size.
Reducing Newton step size.
Algorithm failure in the newton step.
 Completed simulation up until t=22.415
 Writing simulation results to IEEE_39_5_20210525T012401.csv
Completed simulation from 0 sec. to 110 sec.
```

Figure 3.10. Case number 5 events log.

In the second example, Case number 27444, that has the loading levels given in Table 3.3, the system has only one fault at line 33 (between buses 26 and 29) that causes the line to trip (at $t = 10\ s$). Figures 3.11 to 3.13 show the plot of voltage magnitude, voltage angle, and rotor speed against time, respectively. Also, Figure 3.14 displays the events log of the scenario. As seen from the plots and the events log, the system has an extended period of oscillations after the tripping of

the line, but the system then stabilizes at a new operating point and remains stable for the remainder of the simulation. Thus, this simulation indicates a case without any loss of load (i.e., no blackout) or CFs.

Table 3.3. Loading level at each load in Case number 27444 (compared to nominal values).

Load number	1	2	3	4	5	6	7	8	9	10
Load multiplier	1.402	1.258	1.111	1.139	0.765	1.169	0.880	1.286	0.999	0.873
Load number	11	12	13	14	15	16	17	18	19	
Load multiplier	1.440	0.856	0.694	1.393	1.281	0.697	0.906	0.667	0.824	

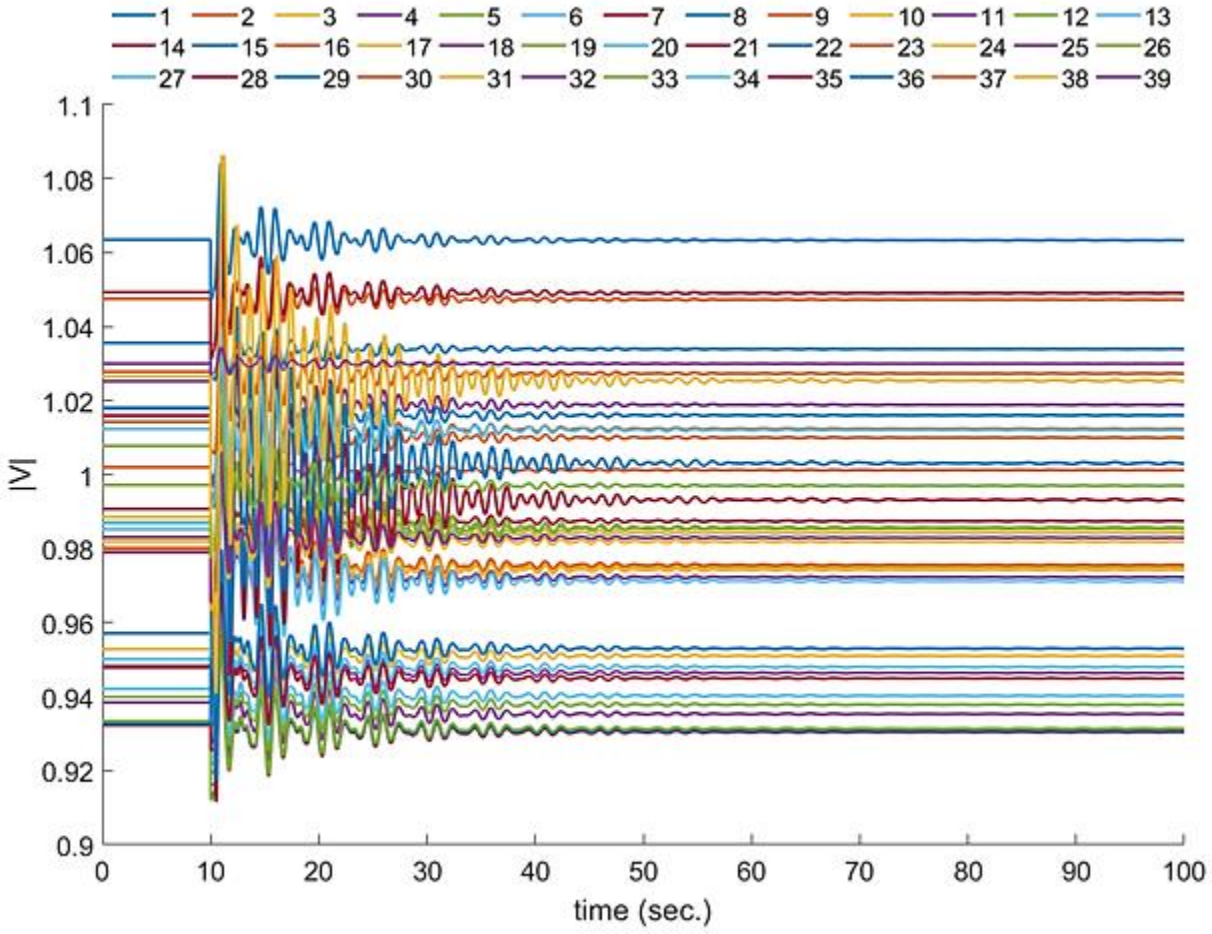

Figure 3.11. Voltage magnitude at each bus (Case number 27444).

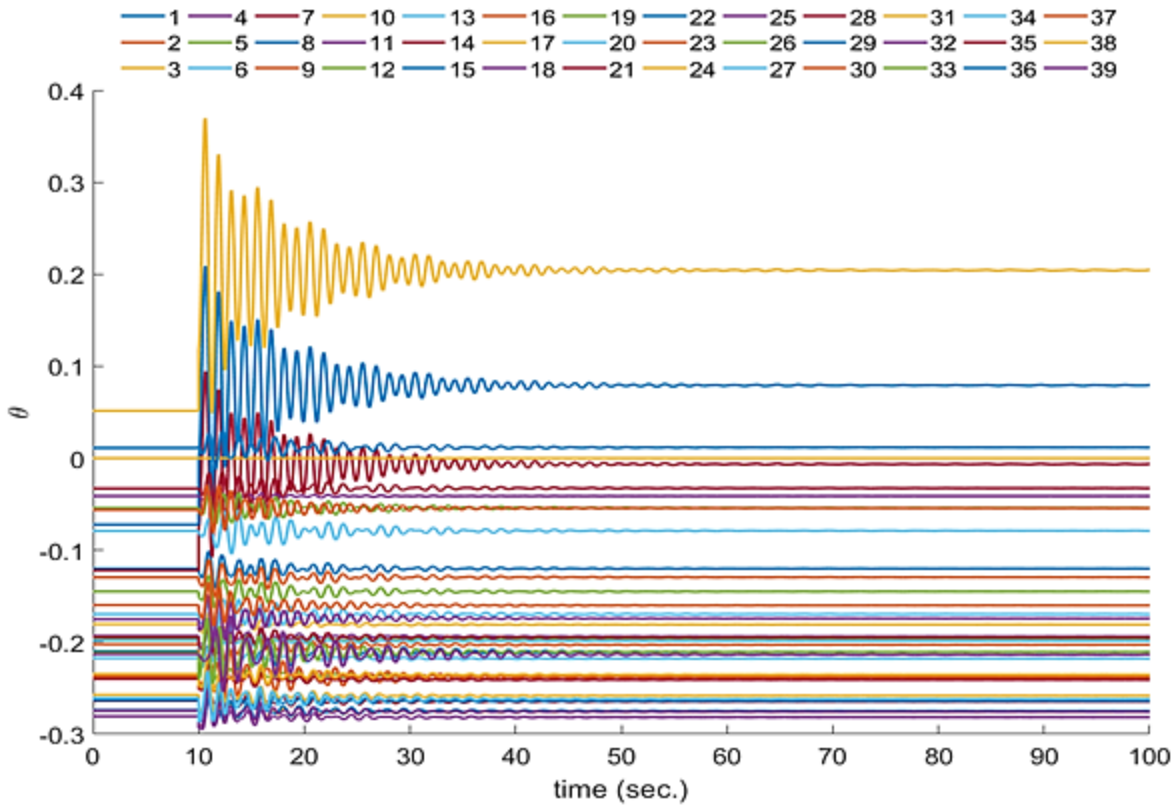

Figure 3.12. Voltage angle at each bus (Case number 27444).

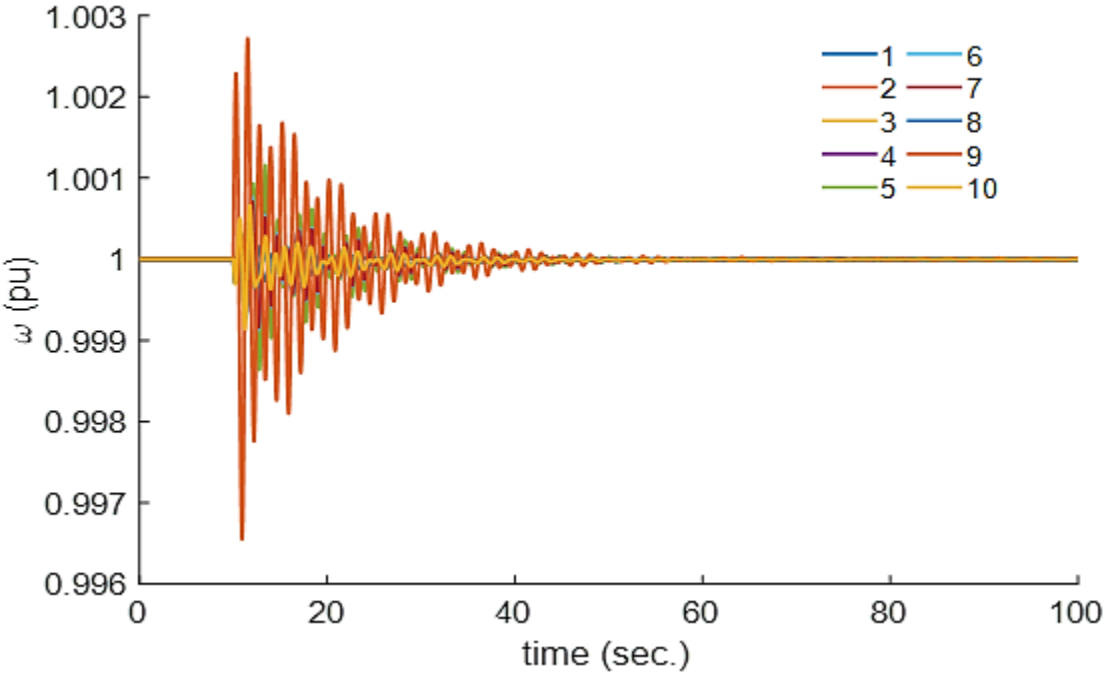

Figure 3.13. Rotor speed at each generator bus (Case number 27444).

```
starting simulation at t = 0
writing results to IEEE_39_27444_20210525T225442.csv
 simulation start event. simulation start event. simulation start event.
simulation start event.

 Simulating from t=0 s to t=10 s
 Completed simulation up until t=10
 Writing simulation results to IEEE_39_27444_20210525T225442.csv
  t = 10.0000: Branch 33 tripped...

 Simulating from t=10 s to t=110 s
 Completed simulation up until t=110
 Writing simulation results to IEEE_39_27444_20210525T225442.csv
Completed simulation from 0 sec. to 110 sec.
```

Figure 3.14. Case number 27444 events log.

In the third example, Case number 470, the loading levels are described in Table 3.4 and the system only has a fault at line 27 (between buses 21 and 22) that caused the line to trip (at $t = 10\ s$). As seen from Figure 3.16-Figure 3.18 and the events log in Figure 3.15, the system has a period of oscillations and almost stabilizes at around $t = 25\ s$, but at around $t = 28\ s$ the system has an endogenous trip (i.e., cascading event) which is the tripping of line 29 (between buses 23 and 24). Following that trip, four buses (22, 23, 35, and 36), including two generator buses (G6 and G7), were isolated from the grid, causing multiple load buses (e.g., buses 4, 7, 8, 12, 15, 21 and 24) to have partial load sheddings to restore the voltage to acceptable levels, while the smaller island collapsed due to over frequency at $t = 30\ s$ (assuming worst case scenario, since the smaller island did not converge). At the end of the event, the isolated loads, and the load sheddings amounted together to a 20% loss of load ($863\ MW$). Hence, this case is considered as a CF with a partial blackout.

Table 3.4. Loading level at each load in Case number 470 (compared to nominal values).

Load number	1	2	3	4	5	6	7	8	9	10
Load factor	1.402	1.258	1.111	1.139	0.765	1.169	0.880	1.286	0.999	0.873
Load number	11	12	13	14	15	16	17	18	19	
Load factor	1.440	0.856	0.694	1.393	1.281	0.697	0.906	0.667	0.824	

```
starting simulation at t = 0
writing results to IEEE_39_470_20210525T230307.csv
 simulation start event. simulation start event. simulation start event. simulation
start event.
 Simulating from t=0 s to t=10 s
 Completed simulation up until t=10
 Writing simulation results to IEEE_39_470_20210525T230307.csv
  t = 10.0000: Branch 27 tripped...

 Simulating from t=10 s to t=110 s
  t = 10.5000: UVLS relay trip at bus 4...
  t = 10.5000: 160.97 MW of load shedding at bus 4...
  t = 10.5000: UVLS relay trip at bus 7...
  t = 10.5000: 39.84 MW of load shedding at bus 7...
  t = 10.5000: UVLS relay trip at bus 8...
  t = 10.5000: 134.44 MW of load shedding at bus 8...
  t = 10.5000: UVLS relay trip at bus 12...
  t = 10.5000: 2.97 MW of load shedding at bus 12...
  t = 10.5000: UVLS relay trip at bus 15...
  t = 10.5000: 107.31 MW of load shedding at bus 15...
  t = 10.5000: UVLS relay trip at bus 16...
  t = 10.5000: 78.43 MW of load shedding at bus 16...
  t = 10.5000: UVLS relay trip at bus 21...
  t = 10.5000: 42.69 MW of load shedding at bus 21...
  t = 10.5000: UVLS relay trip at bus 24...
  t = 10.5000: 104.14 MW of load shedding at bus 24...
  t = 27.33: Over current relay trip at branch 29...
  t = 27.3300: Branch 29 tripped...
  t = 27.3300: The network partitioned into two islands...
Reducing Newton step size.
Reducing Newton step size.
Reducing Newton step size.
Reducing Newton step size.
Reducing Newton step size.
Reducing Newton step size.
Reducing Newton step size.
Reducing Newton step size.
Reducing Newton step size.
Reducing Newton step size.
Reducing Newton step size.
Reducing Newton step size.
Reducing Newton step size.
Reducing Newton step size.
Reducing Newton step size.
Reducing Newton step size.
Reducing Newton step size.
Reducing Newton step size.
Reducing Newton step size.
Reducing Newton step size.
Reducing Newton step size.
Reducing Newton step size.
Algorithm failure in the newton step.
 Completed simulation up until t=109.83
 Writing simulation results to IEEE_39_470_20210525T230307.csv
Completed simulation from 0 sec. to 110 sec.
```

Figure 3.15. Case number 470 events log.

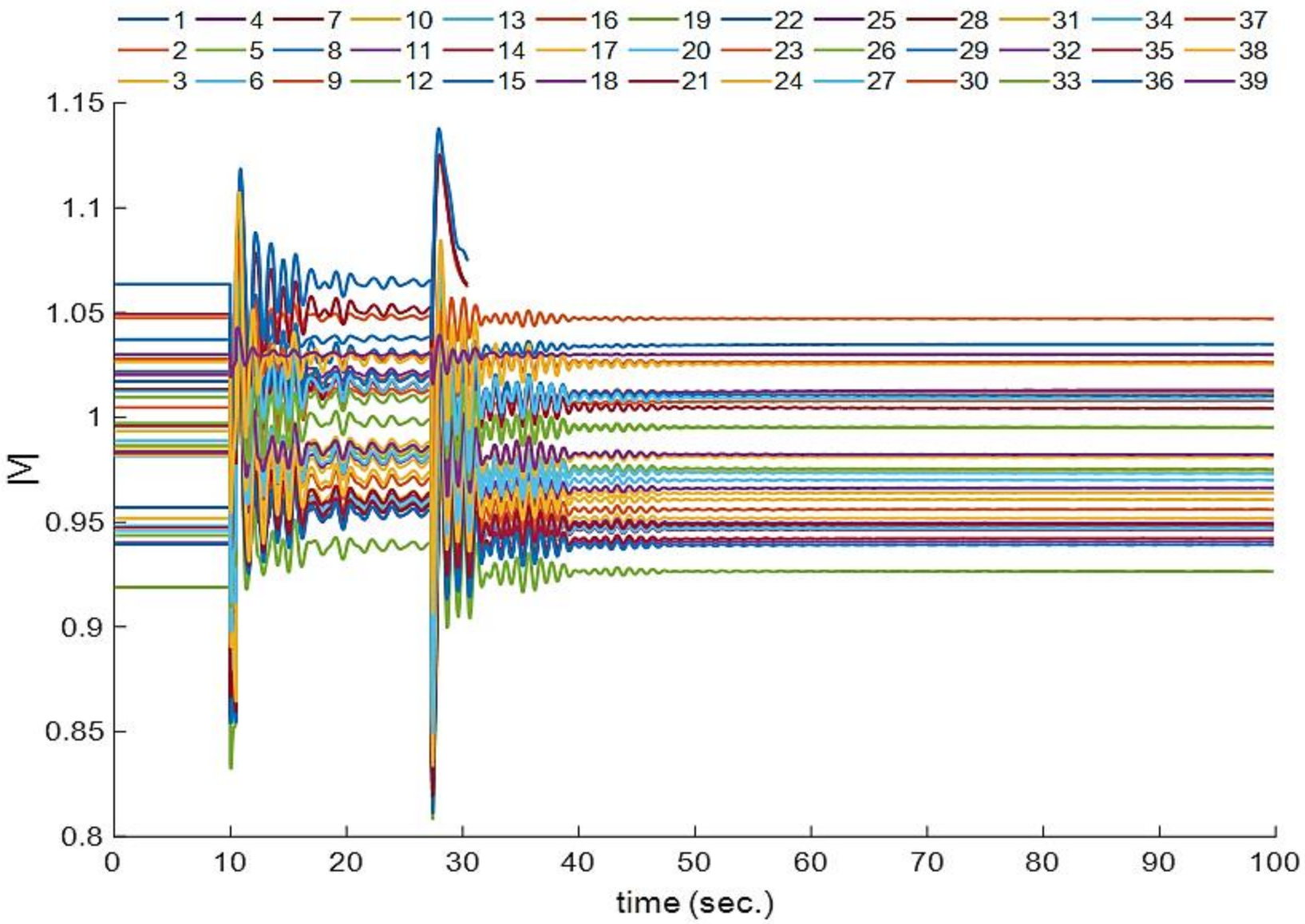

Figure 3.16. Voltage magnitude at each bus (Case number 470).

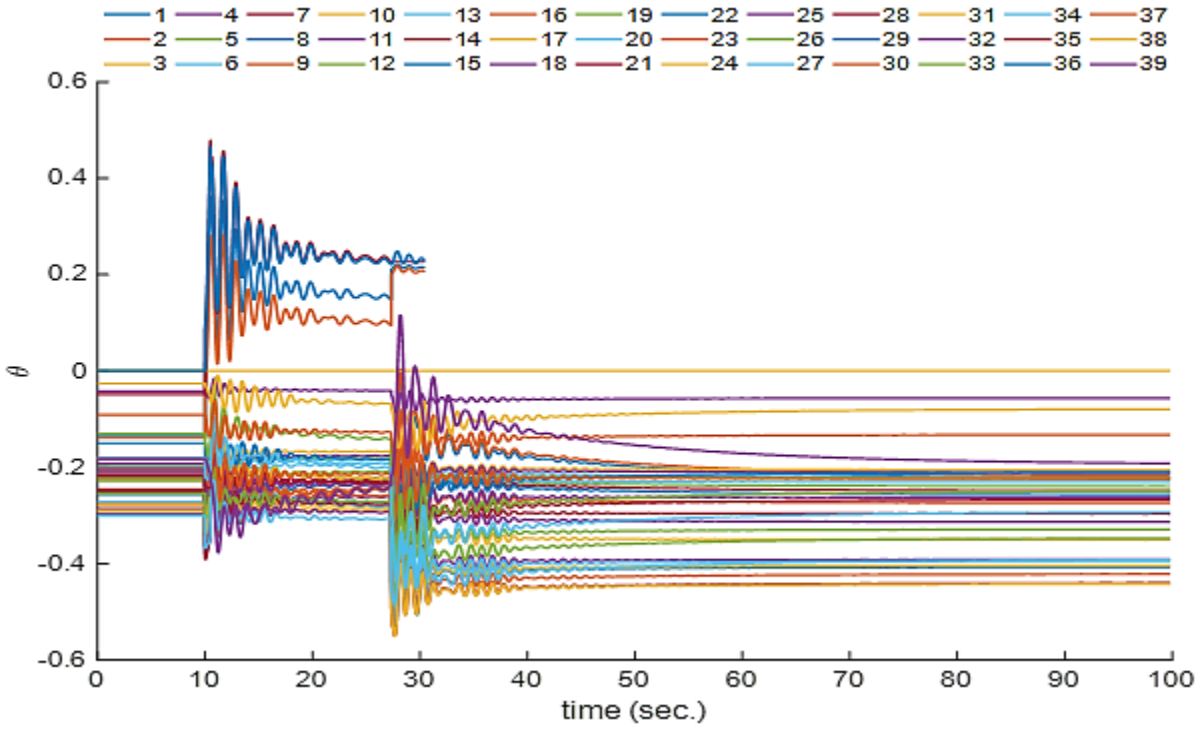

Figure 3.17. Voltage angle at each bus (Case number 470).

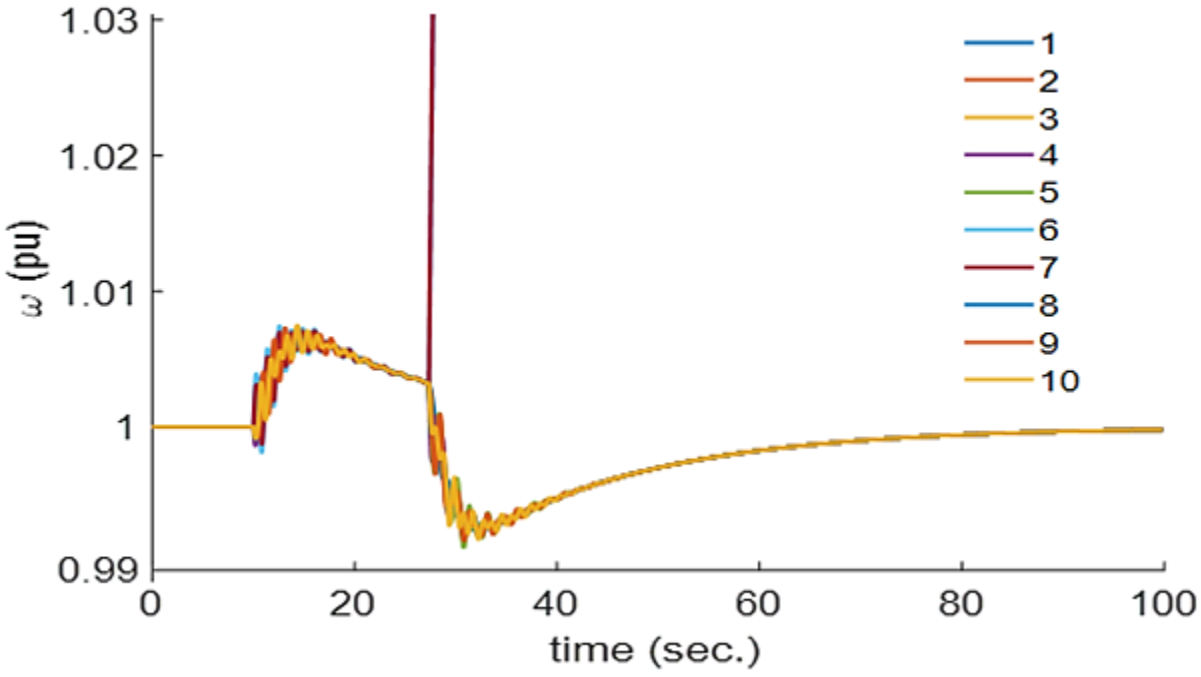

Figure 3.18. Rotor speed at each generator bus (Case number 470).

3.4.3 Prediction model input

To input the data from each scenario to the prediction models that are detailed in the next chapters, the data must be converted into a set of vectors or matrices. These vectors are divided into two categories:

- time-independent vectors
- time-dependent vectors.

If Case number 1 is used as an example, at $t = 0.01\ s$ (time is only relevant to for the time-dependent vectors), all the generated vectors are detailed in Tables 3.5 to 3.12 and are described as follows. The time-independent vectors represent the variables that do not change with time. These vectors include initial fault location (Table 3.5), generator outage status at the time of the fault (Table 3.6), line outage status at the time of the fault (Table 3.7), and load bus ratio prior to any contingency (Table 3.8). Each vector of the time-independent data, excluding load bus ratio, represents a binary vector whose values indicate whether the line or generator is having an outage (or fault) or not. The load bus ratio refers to the loading of the buses with respect to the nominal values (Appendix C). As for the time-dependent vectors, they represent vectors that typically change with respect to time. These vectors include: current timestamp (Table 3.9), generators' omega and delta (rotor speed and rotor angle) (Table 3.10), and the voltage magnitude and angle at each bus (Table 3.11 and Table 3.12). However, generators' deltas are not used for prediction, because they are internal measurements that are typically inaccessible.

Table 3.5. Fault location vector (fault at line 37).

Row 1-24	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Row 25-48	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0

Table 3.6. Generator outage vector (no generator outage.

Row 1-10	0	0	0	0	0	0	0	0	0	0

Table 3.7. Line outage location vector (line 37 is out of service).

Row 1-24	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Row 25-48	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0

Table 3.8. Load bus ratio vector.

Row 1-10	1.014	0.899	1.442	1.421	0.667	0.847	1.500	0.641	0.994	1.477
Row 11-19	0.747	0.763	0.824	1.467	1.318	0.620	0.759	1.338	0.635	

Table 3.9. Time vector for a single time step.

Time	0.01

Table 3.10. Generators' omega and delta vector for a single time step.

Row 1-10	-0.653	-0.813	-0.697	-0.670	-0.663	-0.678	-0.675	-0.664	-0.676	-0.667
Row 11-20	0.988	-1.000	0.922	0.948	0.949	0.941	0.935	0.936	0.9	0.987

Table 3.11. Voltage magnitude vector for a single time step.

Row 1-10	0.135	-0.126	-0.158	-0.201	0.339	0.157	0.458	0.436	0.233	-0.601
Row 11-20	-0.609	-0.525	-0.448	-0.186	-0.159	-0.150	-0.155	-0.157	-0.108	-0.102
Row 21-30	-0.136	-0.107	-0.102	-0.146	-0.120	-0.148	-0.162	-0.127	-0.111	-0.108
Row 31-39	-1.000	-0.543	-0.085	-0.087	-0.088	-0.074	-0.100	-0.113	0.298	

Table 3.12. Voltage angle vector for a single time step.

Row 1-10	0.732	0.631	0.521	0.488	0.481	0.476	0.485	0.731	0.512	0.499
Row 11-20	0.495	0.516	0.530	0.581	0.548	0.622	0.626	0.595	0.488	0.513
Row 21-30	0.566	0.440	0.345	0.637	0.346	0.564	0.384	0.405	0.820	-1.000
Row 31-39	0.564	0.639	0.515	0.638	0.557	0.683	0.461	0.926	0.717	

3.4.4 Data normalization

In artificial intelligence (AI) models, and especially in neural networks (NN), it is preferred to normalize the data to avoid coefficients that are close to infinity or close to zero for any of the internal parameters of the model. Ideally, this is achieved when the inputs are normalized to have a mean of zero and a maximum amplitude of one. Hence, the data for the IEEE 39-bus system has been normalized so that all inputs are between -1 and 1. For each input, the values are normalized using:

$$x_n = \frac{2(x - \min(X))}{\max(X) - \min(X)} - 1 \tag{3.4}$$

where:

x_n is the normalized input,

x is the original input,

X is the vector of all values of that input (for all scenarios and all timesteps),

$\min(X)$ and $\max(X)$ are the minimum and maximum of that input, respectively.

One main limitation of this normalization method is that it is dependent on the data to find the maximum and minimum. Thus, it is possible that the model could generate values outside the range of -1 to 1 while connected to the grid (i.e., the values are the local maxima and minima, instead of the global maxima and minima). However, in practical systems, this is not an issue, because each component in a power grid is expected to operate within a certain operational range before automatically tripping it (e.g., bus voltages between $0.9 - 1.1\ pu$).

3.5 2383-Bus Polish Winter Peak System

3.5.1 System information and data summary

The 2383-bus Polish winter peak system (henceforth POL 2383-bus system) is a publicly available model that represents the full Polish network during its winter peak of 1999-2000. The system has around 2800 lines, 1800 load buses and the total MW load of the system in it is nominal load level is $24\ GW$. The version of the model used here is based on its MATPOWER implementation [112]. This model is used as an example of a large interconnected system to show the efficacy of the prediction model presented in Chapter 5. The simulated cases include 15000 scenarios, described in Table 3.13. In the POL 2383-bus system cases, the data is normalized according to the same procedure described in Subsection 3.4.4. However, as for the lines out of service sampling, it was sampled such that 80% of the cases have one line out of service, 16% of the cases have two lines out of service, and 4% of the cases have three lines out of service. This discrepancy between the outage sampling in IEEE 39-bus and POL 2383-bus systems is to increase the number of cases with CFs, because the POL 2383-bus system is more reliable than the IEEE 39-bus system, due to its interconnectedness and larger size. The distribution of the cascading cases resulted loss of load (LOL) is shown as a histogram in Figure 3.19 and as a Pareto chart in

Figure 3.20, which both show that 99% of the cases have a LOL $< 70\ MW$ with more than half the cases having no LOL. Regarding the computer and simulation time, for each case, the simulation time was 60 seconds and the required computer time ranged from $180 - 240\ s$.

Table 3.13. POL 2383-bus system simulated cases summary.

Type	Number of cases
Simulated cases	15000
Total blackout	39
Load shedding or partial blackout	6505
No blackout	8456

3.5.2 Sample scenario

Similar to the IEEE 39-bus system, a sample of the POL 2383-bus system cases is presented here. Since normal operation cases do not present any new information compared to the previous cases they are not included here. Similarly, all the total blackouts in the POL 2383-bus system do not provide any information either because they are due to non-convergence. Hence, only a case of a partial blackout is presented here. However, due to the size of the system, only select buses are shown. Moreover, load ratios cannot be displayed as there are more than 1800 load buses. In Case number 163, the system suffers two exogenous faults that lead to line trippings at lines 2622 and 666. As seen from Figures 3.21-3.23, the system undergoes some minor oscillations before it stabilizes within 5 or 10 seconds depending on the buses. However, since the system is huge, the selected 50 buses do not reflect the entirety of the system even though these buses are selected from different regions of the system. To better understand the case, Figure 3.24 shows the events log of the case and it reveals that the network separated into two islands after the tripping of line 2622, and the newly formed small grid of 1 generator and 4 load buses collapsed due to nonconverging load flow caused by low frequency. The voltage of the affected load buses is plotted in Figure 3.24 and it shows that the collapse happened in less than two seconds after the grid separation. At the end of the event, the isolated loads amounted to a $50\ MW$ LOL, which is less than a 1% LOL since POL 2383-bus system has a total load of $24\ GW$, but it is still a significant blackout as it almost equates to 60,000 households losing electricity using the estimates in [113]. Hence, this case presents a partial blackout following a CF.

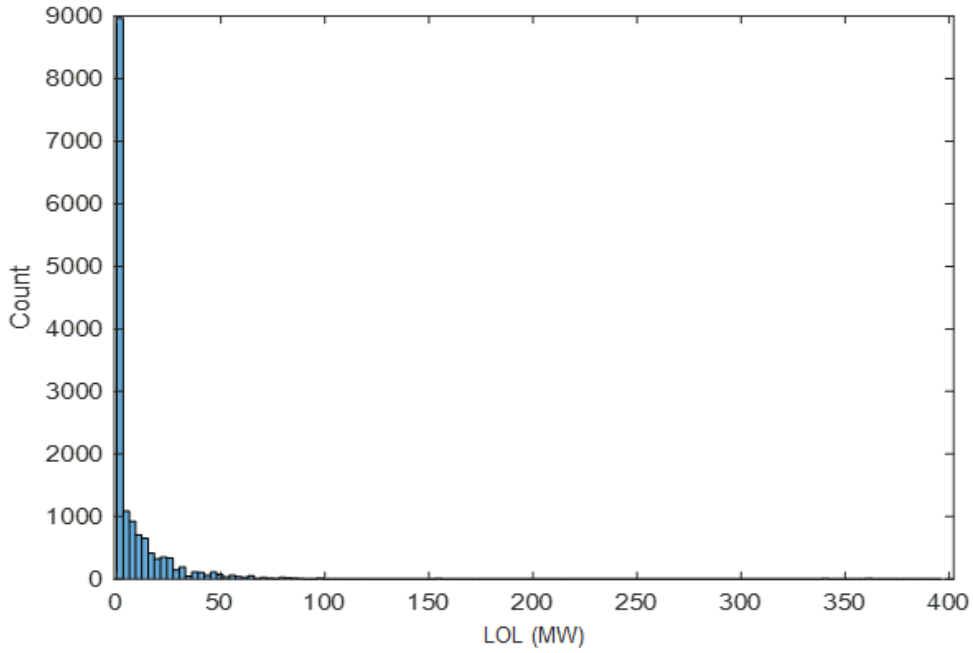

Figure 3.19. Distribution of the LOL in the generated POL 2383-bus system cases.

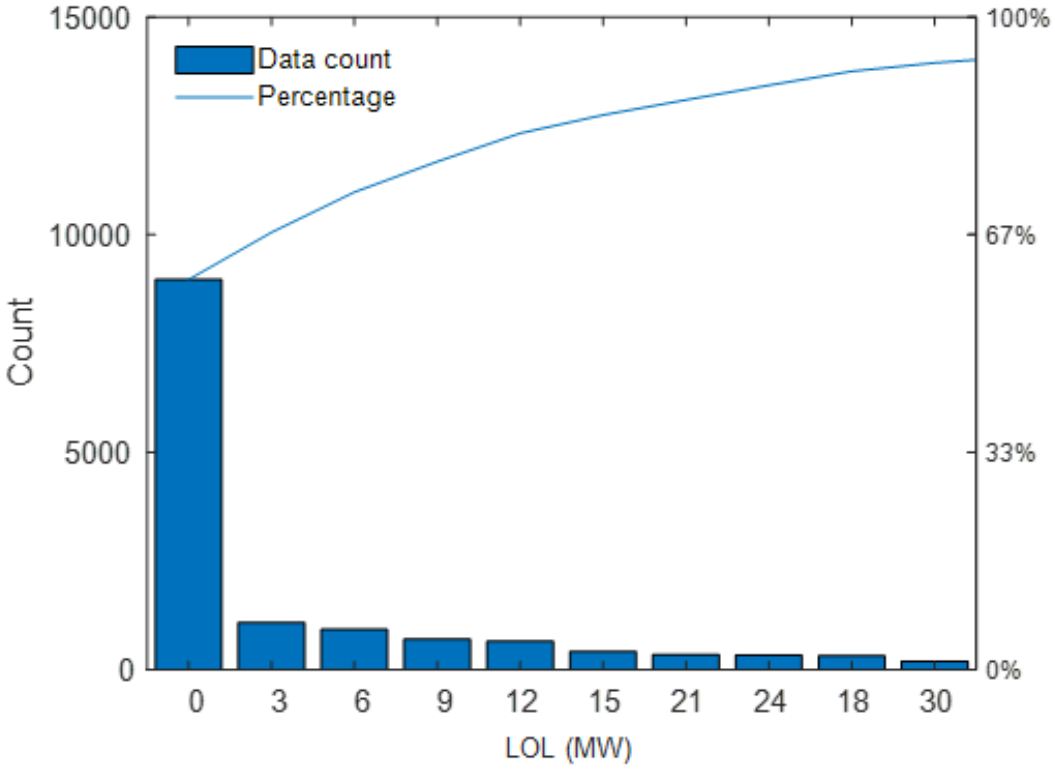

Figure 3.20. Pareto chart of the LOL in the generated POL 2383-bus system cases.

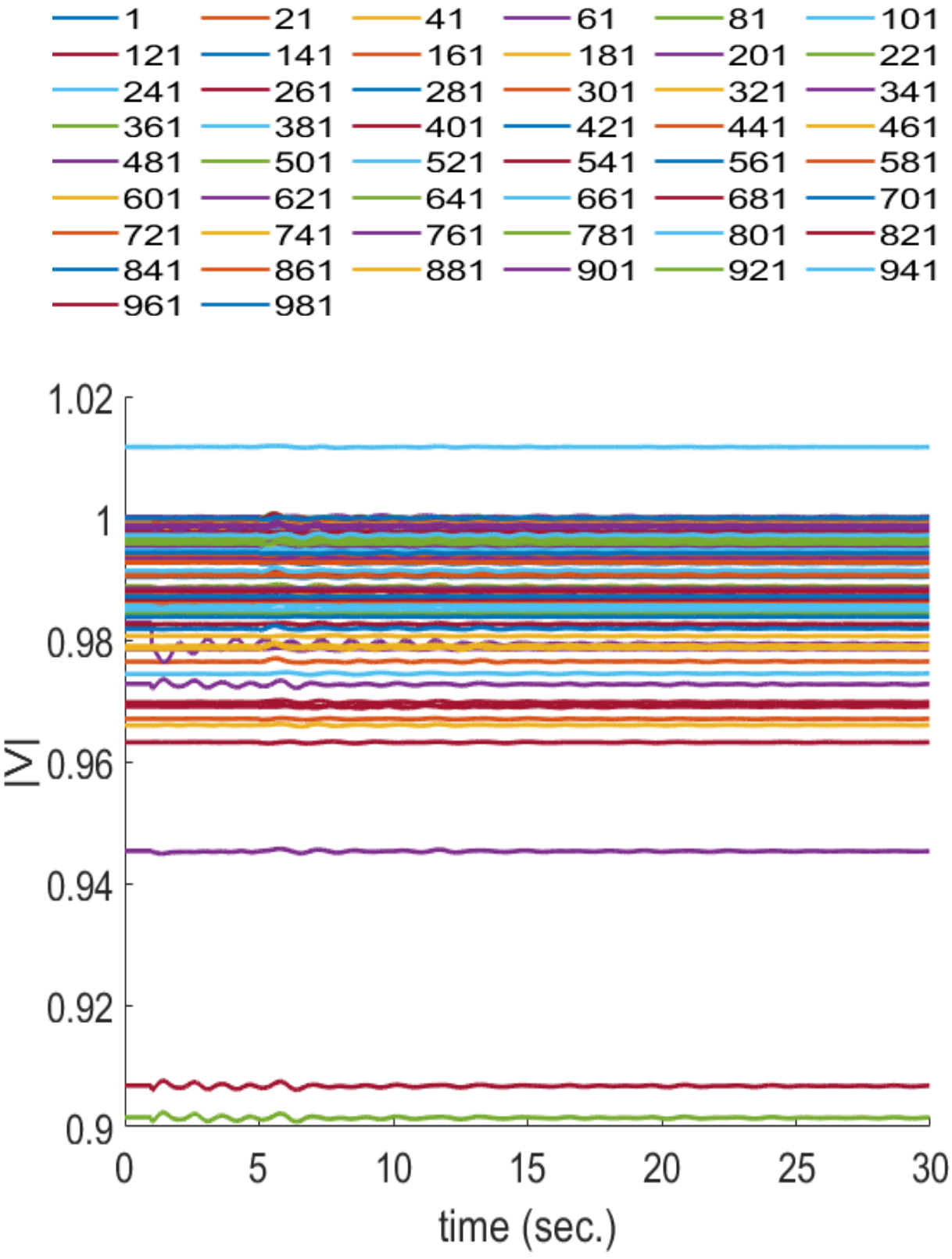

Figure 3.21. Voltage magnitude at selected buses (POL 2383-bus system Case number 163).

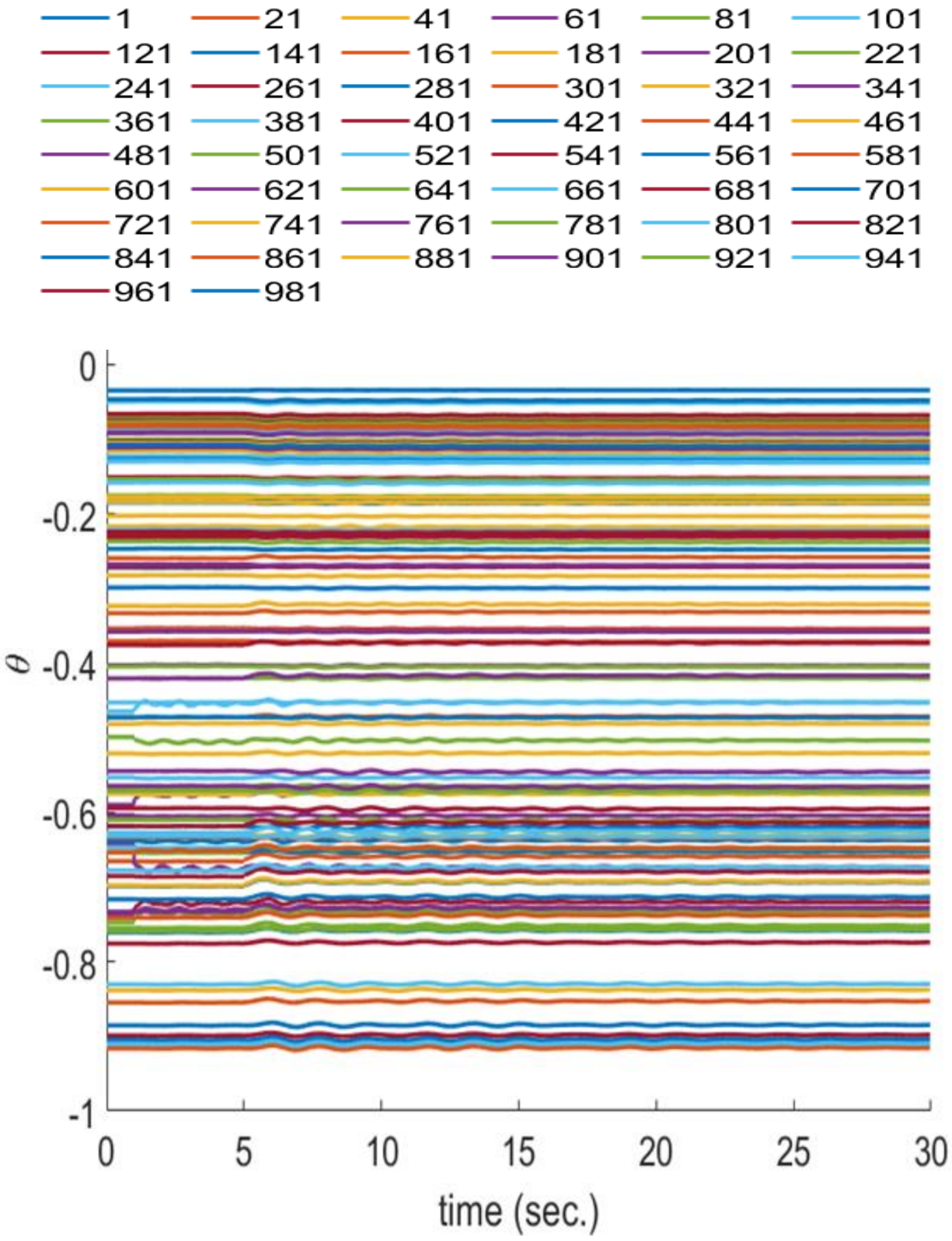

Figure 3.22. Voltage angle at selected buses (POL 2383-bus system Case number 163).

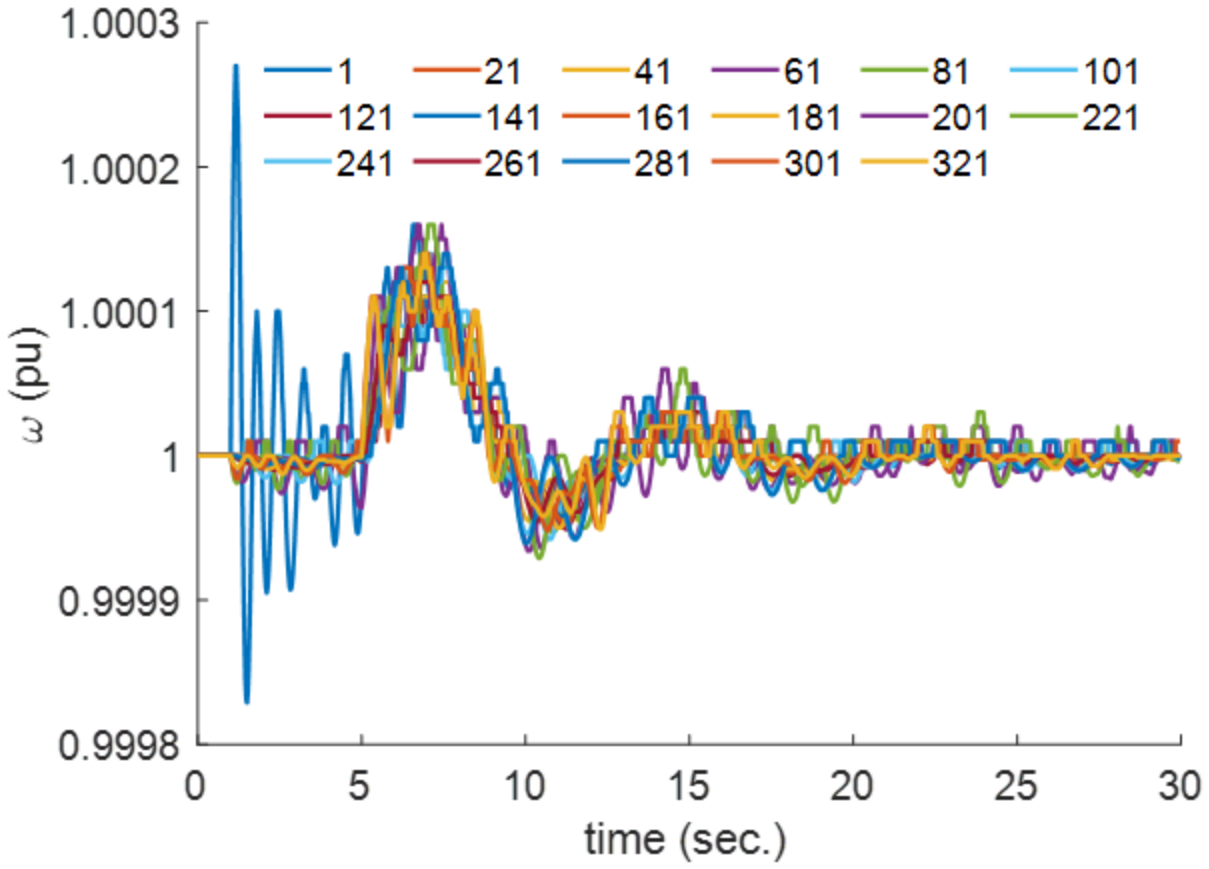

Figure 3.23. Rotor speed at selected generator buses (POL 2383-bus system Case number 163).

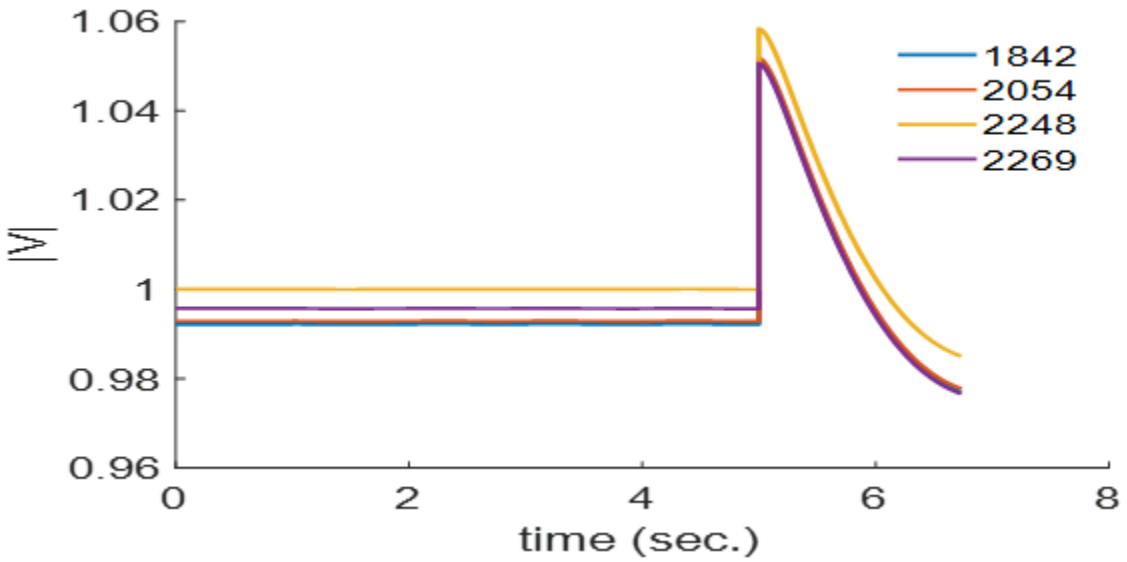

Figure 3.24. Voltage magnitude at isolated buses (POL 2383-bus system Case number 163).

```
starting simulation at t = 0
writing results to sim_case2383_20210616T155251.csv

 Simulating from t=0 s to t=1 s
 Completed simulation up until t=1
 Writing simulation results to sim_case2383_20210616T155251.csv
  t = 1.0000: Branch 666 tripped...

 Simulating from t=1 s to t=5 s
 Completed simulation up until t=5
 Writing simulation results to sim_case2383_20210616T155251.csv
  t = 5.0000: Branch 2622 tripped...

 Simulating from t=5 s to t=30 s
  t = 5.0000: The network partitioned into two islands...
  t = 6.7360: UFLS relay trip at bus 1842...
  t = 6.7360: 0.71 MW of load shedding at bus 1842...
  t = 6.7360: UFLS relay trip at bus 2054...
  t = 6.7360: 2.91 MW of load shedding at bus 2054...
  t = 6.7360: UFLS relay trip at bus 2248...
  t = 6.7360: 3.37 MW of load shedding at bus 2248...
  t = 6.7360: UFLS relay trip at bus 2269...
  t = 6.7360: 4.50 MW of load shedding at bus 2269...

Reducing Newton step size.
Reducing Newton step size.
Reducing Newton step size.
Reducing Newton step size.
Reducing Newton step size.
Reducing Newton step size.
Reducing Newton step size.
Reducing Newton step size.
Reducing Newton step size.
Reducing Newton step size.
Reducing Newton step size.
Reducing Newton step size.
Reducing Newton step size.
Reducing Newton step size.
Reducing Newton step size.
Reducing Newton step size.
Reducing Newton step size.
Reducing Newton step size.
Reducing Newton step size.
Reducing Newton step size.
Reducing Newton step size.
Reducing Newton step size.

Algorithm failure in the newton step.
 Completed simulation up until t=29.9667
 Writing simulation results to sim_case2383_20210616T155251.csv
Completed simulation from 0 sec. to 30 sec.
Elapsed time is 238.534823 seconds.
```

Figure 3.25. Events log of POL 2383-bus system Case number 163.

4 A CONFIDENCE INTERVAL BASED CASCADING FAILURE PREDICTION MODEL FOR POWER SYSTEMS

4.1 Overview

Several research works addressed the analyzing and predicting of CF [40], [91], [114]–[120]. In [40], continuous-time Markov chains are used to extract the blackout probability and severity, and feature selection is employed to allow for scalability without decreasing the accuracy. In [91], brittleness indices based on phasor measurement unit (PMU) measurements are used to predict the system's blackout size under different failures and maintenance topologies. In [114], a real-time probability function is proposed for the expected number of line outages using a branching process. In [115], machine learning and a Bayes network are used to predict the next likely propagation of the cascade (i.e., the chain of line trippings), whereas [116] uses a neural network (NN) to predict cascading events while also considering the effect of hidden failures. Consideration of weather factors (i.e., ice) on CF using Self Organized Criticality theory is carried out in [117]. In [118], the next line to trip is predicted using the transmission lines' power flow deviation from its nominal value. In [119], the probability distribution function of line outages for a given system is predicted using the Galton-Watson branching process theory. In [120], Support Vector Machine (SVM) and probabilistic models are used to build a prediction model for line outages. This approach has the advantage of using historical data, but it does not consider the system's transient behavior.

These prediction methods do provide insights for the operators that help in countering blackouts. However, most of them suffer from limitations such as assuming that the power grid is static, ignoring real-time operation, or assuming that generator dynamics can be ignored. Thus, it is essential to introduce a predictor based on dynamic data, which would require the use of wide-area measurement devices such as phasor measurement units (PMU), as well as constructing cases based on dynamical simulations of the power system. Examples of incorporating PMUs and detailed models in predictions are reported in [121], [122], where PMUs were used to predict the stability of power systems with renewable generation and to simulate the Arizona blackout. Moreover, another limitation of CF prediction (CFP) literature is that it only predicts the next line to trip, the expected number of line trips, or the expected size of the cascade [91], [114], [115].

Therefore, these algorithms do not inform the operator about the exact locations of the affected areas by the CF event. Hence, there is still a need for a CFP model for the locations affected by the CF that is based on dynamical simulations of the power grid.

Besides, CFP models also need to account for uncertainties because the errors are large in such models. The large uncertainty is due to considering multiple unknowns within the power system (e.g., load types, hidden failures…etc.). The issue of uncertainty becomes even more evident as more items are being predicted, which increases the likelihood of errors [115]. To counter this issue, prediction intervals (PI) were implemented in other smart grid fields such as generator coherency grouping and wind power forecast [123], [124]. Moreover, including prediction uncertainties to models based on deterministic predictions, for example decision trees or neural networks (NNs), increases the credibility and reliability of such predictions [125]. However, there is no such application of PI in CFP literature. Thus, the usage of PI in CFP to increase the reliability of the prediction model and account for the grid uncertainties requires further analysis.

In this chapter, to address the aforementioned limitations of CFP, a load-based PI is proposed for real-time CFP. The main contributions of this method are predicting the exact locations affected by the CF, as well as using PI instead of point predictions. The usage of PI allows the model to account for the uncertainties within its formulation. It thus provides the operators with more reliable information regarding the affected areas and the expected size of the cascade. Moreover, PI is needed to account for the added complexity of the proposed model in comparison to already existing methods such as predicting the next line to trip, or whether a CF will occur or not. Also, since the simulations used to build the model include dynamic simulations with multiple maintenance topologies, it addresses the static flow issue. Thus, the proposed model addresses multiple issues that exist in the current literature, and to the authors' knowledge, such a model has never been implemented before.

4.2 Prediction Intervals

Multiple methods to design PI in NN exist in the literature, such as the delta, Bayesian, MVE, Bootstrap, and LUBE methods. Each method has its unique advantages and limitations. Hence, this section provides a summary of each method.

The delta method assumes that the NN is a nonlinear regression model and, thus, it solves for the parameters that provide its best fit. Afterwards, the PI can be calculated by estimating the variance of the prediction error. Delta method's main advantage is that it provides an accurate PI if the NN model truly represents the best fit of the data. However, the main issue with the delta method is that it provides the same constant interval for all predictions. Hence, it does not consider when certain inputs correspond to a higher prediction accuracy [126].

On the other hand, the Bayesian method aims to minimize the sum of square errors by assuming that errors have a normal distribution and then finding the parameters that minimize the variance of the distribution. Also, its formulation makes it account for both the errors due to noise and incorrect or nonideal modeling. However, its main shortcoming is its huge computational demand [126].

Both the delta and Bayesian methods approach the PI problem as a single problem, but MVE addresses PI as two separate issues. In MVE, the PI is again assumed to be a normal distribution, but the formulation is divided into two predictions, the prediction of the mean and the prediction of the variance. Hence, each one is optimized separately. This formulation allows the model to assign different variances for each prediction, effectively solving the main issue of the delta method. Moreover, it is a simple model to implement since the first portion is just the typical point prediction model. However, the main limitation of this method is the assumption that the first model will produce an accurate estimate of the mean. Thus, the generated normal distribution parameters, and the intervals produced by them, can have large deviations from the actual distribution [126].

As for the Bootstrap method, it provides an easy to implement and accurate way to produced PIs. In Bootstrap, multiple NNs are trained, then the mean and variance are directly extracted from the mean and variance of the different models. Bootstrap's main pros are its simple formulation, its ability to account for error in both noise and misclassifications, and allowing for different variances (similar to MVE). However, it has multiple disadvantages. It is dependent on the number of NNs used. Therefore, a large number of NNs might be needed to get an accurate PI. Also, it is computationally expensive to get the initial training of all the models [126].

Finally, the LUBE method addresses the PI problem as an optimization function with two contradictory parameters, the interval width, and the coverage probability. The interval width refers to the distance between the upper and lower bounds of the PI, and the coverage probability refers to the probability that the true value lies within the PI. The optimization function aims to produce two points for each prediction, referring to the upper and lower bound of the estimation. Then, a fitness function is calculated based on the interval width and the coverage probability to find the fitness of the model and whether it provides better PIs compared to other models or not. The method is mathematically rigorous and can provide PIs that are both accurate and narrow. However, its main limitation is that the provided optimization function is highly nonlinear. Thus, the LUBE method is limited by the robustness of the used solver. Moreover, due to the high nonlinearity, heuristic algorithms need to be used, which are not mathematically tractable and can be trapped in local minima [125]. Overall, each PI method has different advantages and disadvantages. Hence, depending on the requirements of the model different methods should be used.

4.3 Cascading Failure Modeling

CF occurs when an initial failure in the network causes additional failures in the grid. The initial failure is referred to as the exogenous failure and typically it is not the focus of analysis in CF. However, the following failures are considered as endogenous failures and are the focus of CF. These additional failures are the main mechanism of how a CF propagates, and they can happen due to multiple causes. However, the most important causes are modeled as the following [3]:

4.3.1 Transmission lines overheat tripping (Overcurrent tripping)

Following the initial failure, the redistribution of transmission line power flows can cause some lines to overheat. The overheating of the lines can be modeled either by a constant limit, a variable limit based on transmission line heating, or a detailed model of circuit breaker behaviour (i.e., actual model). The comparison between the three types is demonstrated in Figure 4.1. The figure refers to the relation between time and current (normalized by the line’s current rating) in determining if a line will trip (if its state lies to the right or upper right of the curve) or not (if its state lies to the left or lower left of the curve). In static models, the dotted vertical line indicates that the model ignores the time factor and only assumes that the line will trip after exceeding a

constant value. Dynamic models extend the formulation by assuming an inverse relationship between time and allowed current. Finally, the actual model considers the detailed behaviour of the circuit breaker (CB) where the curve is made up of 3 or more subparts referring to the short time delay, long time delay, and instantaneous tripping. Moreover, it has an undetermined dead band, which could be modeled by sampling from a random distribution or ignored by only considering one of the curves or their averages.

In Figure 4.1, the three models can be compared by considering four different regions. In regions 1A and 1B, all models produce the same prediction (1A: normal operation, 1B: trip). In regions 2A and 2B, the dynamic and actual models behave the same (2B: normal operation, 2A: trip), but the static model differs (2A: normal operation, 2B: trip). In region 3, the static and actual models behave the same (trip), but the dynamic model differs (normal operation). In regions 4A, 4B, and 4C, the actual model is undetermined, but the dynamic model better approximates the behavior of the actual model than the static model. Since 1A and 1B regions are larger than the other regions, static models do provide good estimates for the overall risk. However, when considering real-time prediction or CF events with multiple line trippings, the difference between the static model and the more detailed models becomes more significant. Hence it is necessary to avoid static models when considering real-time prediction of CFs.

Assuming a dynamic model, the temperature change and line tripping are modeled in the power system as [27]:

$$\dot{T}_i = r_i F_i^2 - k_i T_i \tag{4.1}$$

$$S_{OC-relay,i} = \varsigma\left(\mathrm{T_i(t)}, \mathrm{F_i(t)}\right), \qquad S_{OC-relay,i} \in \{0,1\} \tag{4.2}$$

where:

T_i is the temperature of the line, in Celsius degrees, compared with the ambient value ($20\ ^\circ C$),

F_i is the current of line i in Amperes,

r_i is the heating constant,

k_i is the time constant,

$S_{OC-relay}$ is the state of the overcurrent relay,

ς is the tripping function, which depends on the temperature $T_i(t)$, line flow $F_i(t)$, and the relay's internal parameters, such as the delay setting τ and line rating.

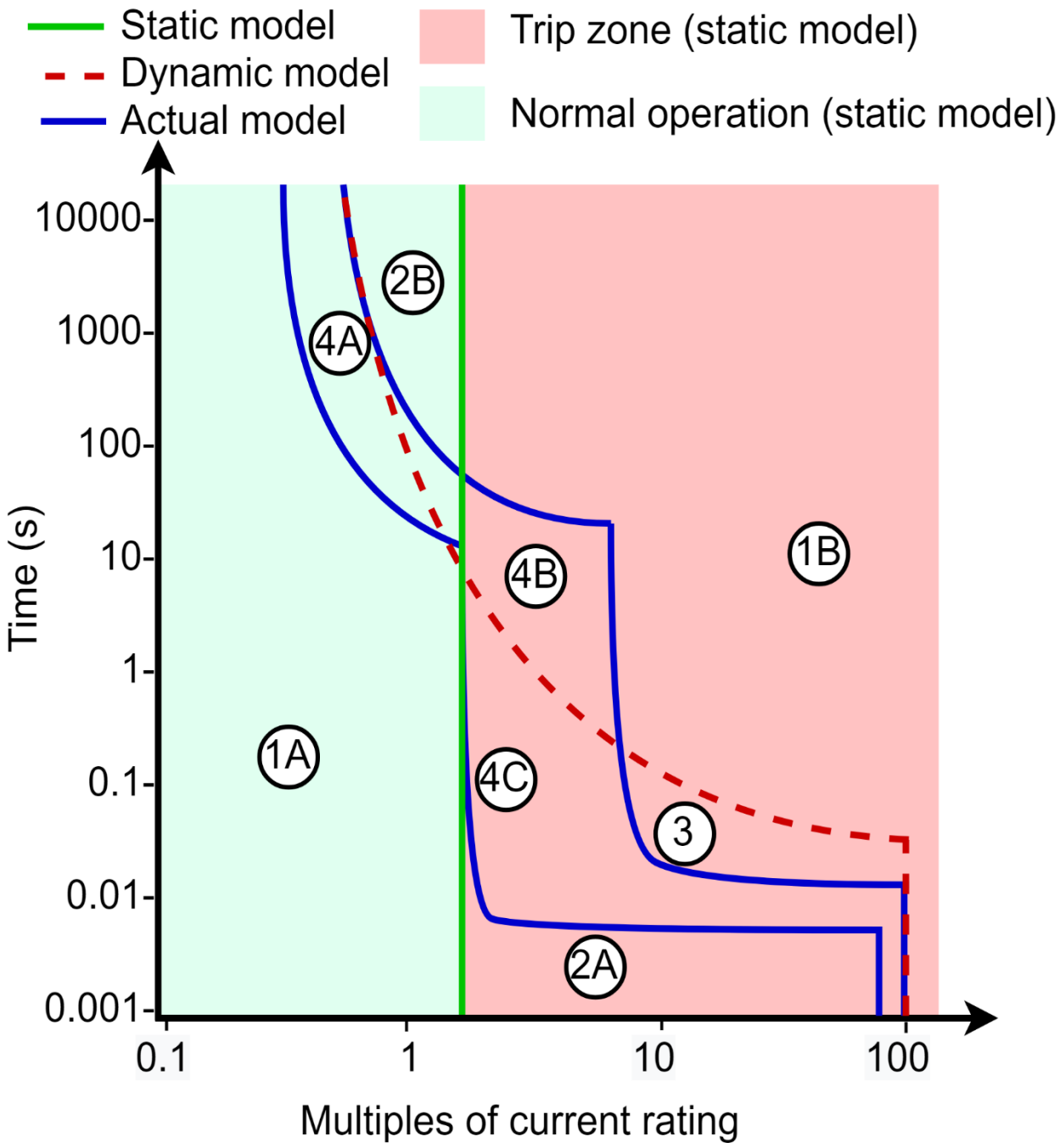

Figure 4.1. Comparison between the characteristics of static model, dynamic model, and actual line overheating model.

4.3.2 Under voltage load shedding (UVLS)

Following a large disturbance, if a large transmission line or generator becomes out of service some loads might sustain an extended duration of under-voltage. In this case, UVLS relay trips the load, or a portion of it, in an attempt to restore the voltage. Typically, this action is modeled as a binary operation [27]:

$$S_{UVLS-relay,n} = \xi(\mathrm{v_n(t)}), \qquad S_{UVLS-relay} \in \{0,1\} \tag{4.3}$$

where:

$S_{UVLS-relay}$ is the state of the UVLS relay,

ξ is the tripping function, which depends on the voltage at load bus n ($v_n(t)$), as well as the relay parameters such as the threshold voltage ($\mathrm{v_{Th}}$), delay (τ), and load shedding percentage.

Figure 4.2 represents a case where the IEEE 39-bus system initiated an UVLS to restore the voltage. The system suffered a contingency at $t = 129.6\ s$ (loss of a generator bus). Following the contingency, the system sustained critically low voltages for $0.5\ s$. Afterwards the UVLS tripped a portion of the affected loads, which effectively restored the voltage to normal levels ($0.9 - 1.1\ pu$).

4.3.3 Under frequency load shedding (UFLS)

Similar to UVLS, if the load sustained a long duration of under frequency state, the UFLS trips the load, or a portion of it, to restore the frequency. UFLS is also modeled as a binary operation [27]:

$$S_{\mathrm{UFLS}-relay.n} = \varrho(\mathrm{f_n(t)}), \qquad S_{\mathrm{UFLS}-relay} \in \{0,1\} \tag{4.4}$$

where:

$S_{\mathrm{UFLS}-relay}$ is the state of the UFLS relay,

ϱ is the tripping function, which depends on the measured frequency at load bus n terminal (f_n), and the settings of the relay such as frequency threshold (f_{Th}), delay (τ), and load shedding factor.

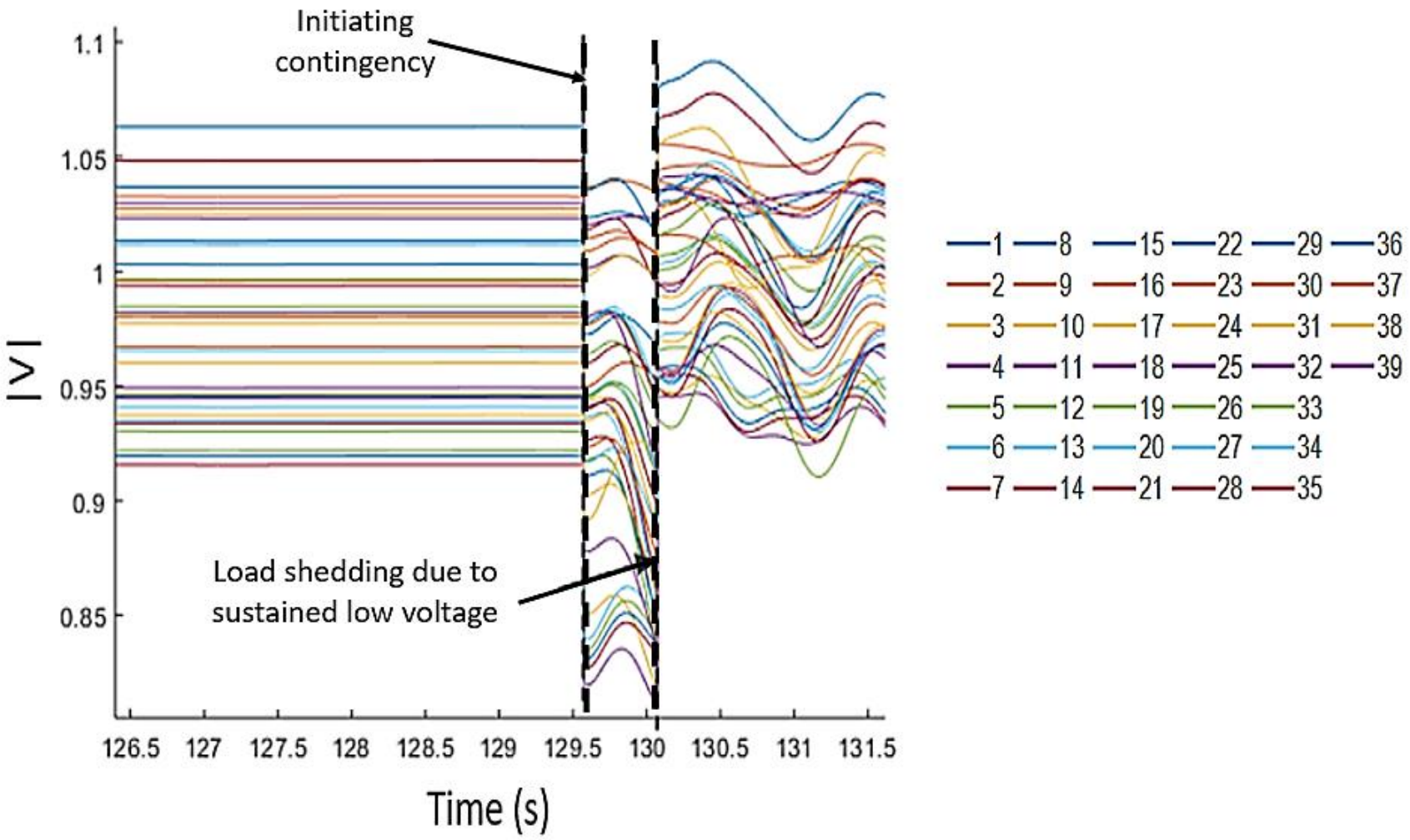

Figure 4.2. Restoration of the voltage following a loss of generation using load shedding in the IEEE 39-bus system.

4.3.4 Hidden failures

Hidden failures (HF) are mainly due to incorrect settings or malfunctioning equipment. HF refers to any failure that does not affect the system under normal operation but appears following an initial disturbance causing additional contingency. For example, if a circuit breaker (CB) is connected to a line with a nominal current of $1\ kA$ and is supposed to be set to trip at $10\ kA$, but is wrongly set to $5\ kA$ (or trips at $5\ kA$ due to a manufacturing error). Then, when the CB trips while the current is between $5\text{-}10\ kA$ it will be considered as a HF. Another common HF is tree contact failures when the current is lower than the expected tree contact threshold. As these failures are hard to predict, they are modeled as a fault with a random probability based on historical data. Generally, it is modeled as [127]:

$$px(t) + qx(t) = 1;\ px(t), qx(t) \in [0,1] \tag{4.5}$$

Where:

$px(t)$ is the probability that element x (relay, transformer…etc.) works as expected,

$qx(t)$ is the probability that element x has a hidden failure.

Finally, there are additional causes for CF outside of the causes mentioned here. However, they are either too difficult to model, such as operators' mistakes, a unique case of previous scenarios (e.g., multiple lines overheat trippings due to severe weather), or too rare (e.g., over-demand) [3], [19].

4.4 The Proposed Method

4.4.1 Neural network (NN)

In order to obtain an accurate prediction of the CF propagation, the detailed dynamical model of the system is needed to avoid the pitfalls of static simulations and to include the previously mentioned CF mechanism properly [23]. However, it is not feasible to run detailed dynamic simulations in real-time to predict CF. Thus, it is needed to transform the underlying complexity of the dynamic model as a relation between the state of the system and the expected loss of load (LOL) at a given load using methods such as neural networks (NN), Figure 4.3. The inputs for this model can be the voltages (magnitude and phase) and real power at each load as

obtained from the PMU data, while the output can be defined as the blackout size prediction for each load bus in the network. Thus, the input is formulated as:

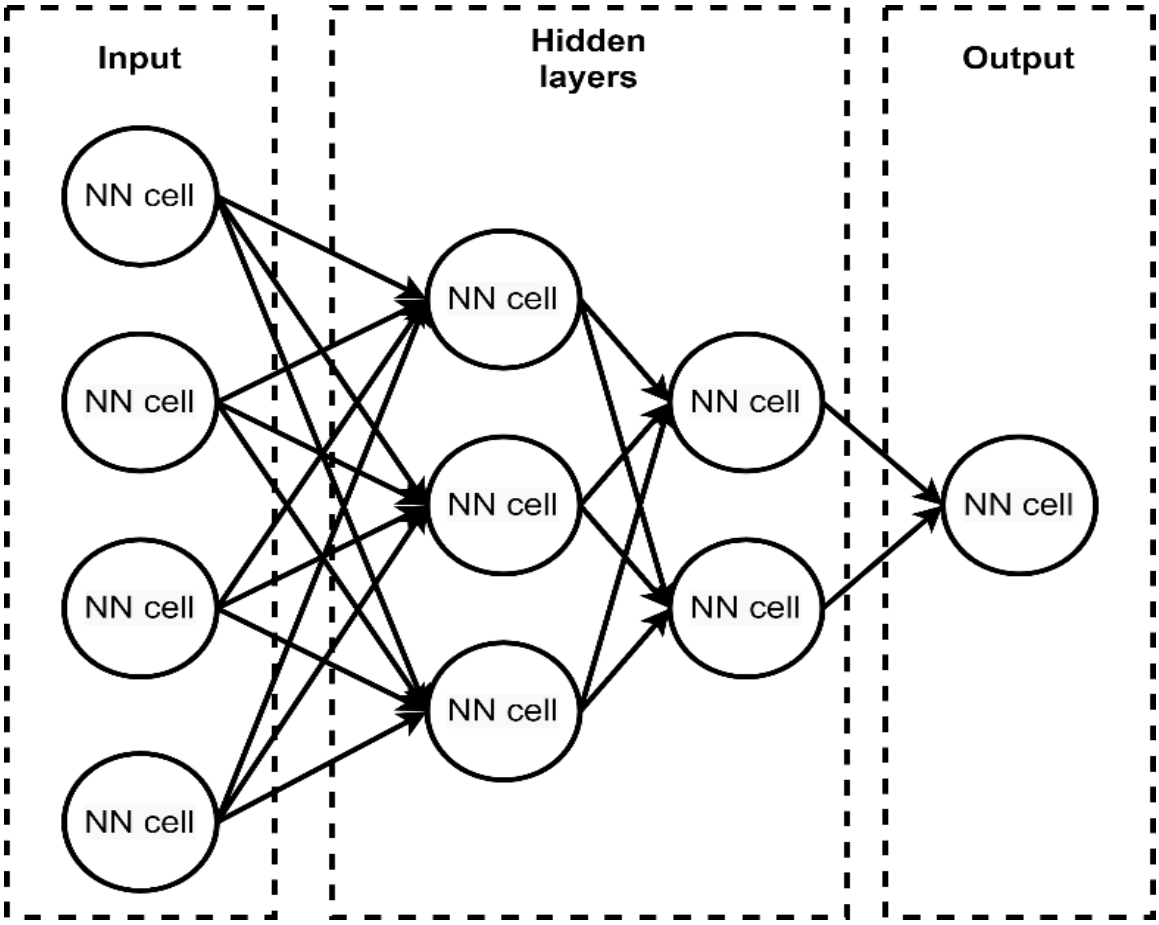

Figure 4.3. Structure of a neural network where hidden layers typically represent the approximation of the complexities of the original model.

$$X = [PL\ VP\ TP] \tag{4.6a}$$

$$PL = \left[p_{L_1} \ \dots p_{L_n} \dots p_{L_N}\right] \tag{4.6b}$$

$$VP = [v_1 \dots v_m \dots v_M] \tag{4.6c}$$

$$TP = [\theta_1 \dots \theta_m \dots \theta_M] \tag{4.6d}$$

$$m = 1,2, \dots M;\ n = 1,2, \dots N$$

Where:

PL is the real power vector,

VP and TP are the voltage magnitude and angle vectors,

P_{L_n} is the real power at load bus n,

v_m and θ_m are the voltage magnitude and angle at bus m, as obtained from the PMUs, respectively,

M is the number of buses,

N is the number of load buses in the network.

It is worth noting that under this formulation, the model does not receive any information regarding fault locations or line outages to avoid dependency on this data. This dependency is one of the primary causes of the propagation of the 2003 US-Canada Northeast blackout, because the system had an incorrect estimation of the state of line outages and fault locations [28].

Moreover, in order to provide a more accurate prediction and compensate for the missing information, X at multiple time steps are used for the prediction. Thus, the prediction of the LOL is obtained as:

$$LOL_n = \mathcal{F}_{\mathrm{n}}(X_t, X_{t-1} \dots, X_{t-\delta}), \qquad \delta \in \mathrm{Z}^{\geq} \tag{4.7}$$

where:

$\mathcal{F}_{\mathrm{n}}$ is the NN associated with predicting the LOL at load bus n and $\mathbf{Z}^{\geq}$ is the set of non-negative integers.

However, this formulation is highly likely to produce multiple errors as the error probability that at least one load prediction is wrong is:

$$P_F = 1 - \prod_{n=1}^{N} P_n \tag{4.8}$$

where:

P_F is the probability that at least one prediction is wrong,

P_n is the probability that $\mathcal{F}_n$ correctly predicted LOL_n.

It is evident that regardless of the individual accuracies P_n, the failure probability P_F quickly increases as the number of models increases. Thus, it is required to address this limitation of the

model. Hence, an additional sub-model is added to suppress this increasing error probability, which is data fusion.

4.4.2 Data fusion

Since power systems are interconnected systems, there is a high correlation between the state of a given load and the state of its surrounding loads. This correlation can be used to improve the accuracy of the prediction by correcting wrong predictions through data fusion (DF) techniques [128]. Although there are numerous DF techniques mentioned in the literature, in this CF prediction scenario the desired outputs are just the corrected inputs. Hence, both the inputs and outputs are from the same domain and range. Thus, that the DF can be reduced to the following equation:

$$LOL_{corrected} = \mathbb{C}(LOL_{old}), \qquad LOL \in \mathbb{R}^N \tag{4.9}$$

where:

LOL_{old} is the vector of the original LOL predictions,

$LOL_{corrected}$ is the vector of the corrected LOL predictions,

N is the number of load buses in the network,

$\mathbb{C}$ is the mapping or DF function.

Moreover, as an additional input, the original inputs can also be fed into the DF NN. This additional input helps in exploiting the correlations between the inputs and outputs, and in that case, the equation becomes:

$$LOL_{corrected} = \mathbb{C}(LOL_{old}, X_t, \dots, X_{t-\delta}) \tag{4.10}$$

To illustrate how this DF NN works, consider the following illustrative example. Given the simple 7-bus power system shown in Figure 4.4, and assuming that after a contingency the LOL prediction algorithms predicted the following. Loads 2,3,4,5 will suffer from a LOL, but loads 1 and 6 will not have a LOL. Then, under this scenario, the DF NN would either generate the same output if it considers it correct (i.e., 2,3,4, and 5 has LOL, 1 and 6 does not have a LOL) or modify the prediction to the most likely outcome. In this specific grid layout however, it is unlikely that load 6 (bus 7) does not have a LOL if load 3 (bus 4) has a LOL. Hence, the DF will

transform the prediction to the most likely outcome based on its training, which could be for example LOL in loads 2,3,4,5, and 6 while load 1 is the only load with no LOL as shown in Table 4.1. Similarly, in more complicated grids the DF NN compares a prediction result with its previous training dataset to detect and correct highly unlikely predictions. Finally, following the modified output from the DF NN, the LOL predictions are lump summed into a single predictor of the total blackout size (if any) using prediction intervals, which are added as the final sub-model of the prediction model.

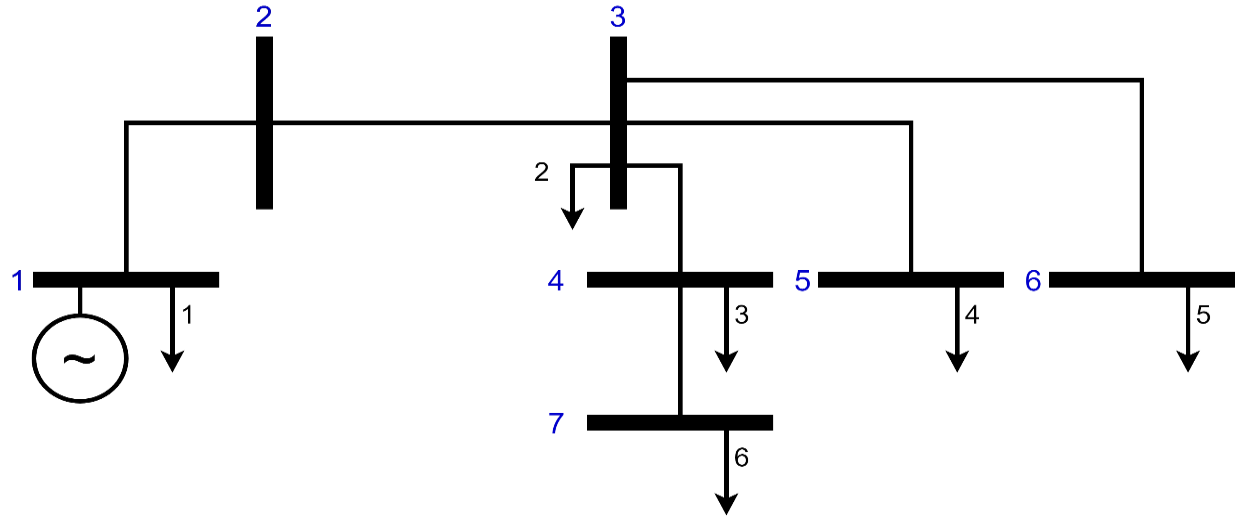

Figure 4.4. A simple 7-bus system with 1 generator and 6 loads.

Table 4.1. Prediction correction using data fusion for the 7-bus system example.

Load index	**1**	**2**	**3**	**4**	**5**	**6**
Original Prediction (LOL_{old})	0	1	1	1	1	0
Modified Prediction ($LOL_{corrected}$)	0	1	1	1	1	**1**

4.4.3 Prediction intervals

Typically, NNs are used to produce classifications or point predictions. However, these predictions lack crucial information, which is the probability that these predictions are correct. This uncertainty is even more critical as the number of predictions increases, which is the case in

this proposed model since each observed load has its LOL prediction. Thus, PIs are needed. Typically, PIs are the intervals at which there is a probability of $1-\alpha$ that the generated model will have its PI covering the actual value [125]. To achieve this, the Lower Upper Bound Estimation (LUBE) method, which is based on minimizing the Coverage Width-based Criterion (CWC) given a set D and is used. This method is selected over the other PI methods since it provides better PIs as it allows different outputs to have different PI widths and it does not require a huge computational burden. The LUBE method for PI is defined as follows [126]:

$$D = \{(X,Y)_{z=1}^{Z} | X_z \in \mathbb{R}^g, Y_z \in \mathbb{R}\} \tag{4.11}$$

$$CWC = NMPIW\left(1 + \nu(PICP)e^{-\eta(PICP-\mu)}\right) \tag{4.12}$$

$$NMPIW = \frac{MPIW}{\mathcal{R}} \tag{4.13}$$

$$MPIW = \frac{1}{Z}\sum_{z=1}^{Z}\left(U(X_z) - L(X_z)\right) \tag{4.14}$$

$$PICP = \frac{1}{Z}\sum_{z=1}^{Z} c_z \tag{4.15}$$

$$c_z = \begin{cases} 1, & y_z \in [L(X_z), U(X_z)] \\ 0, & otherwise \end{cases} \tag{4.16}$$

$$\nu = \begin{cases} 0, & PICP \geq \mu \\ 1, & PICP < \mu \end{cases} \tag{4.17}$$

Where:

Z is the number of data points,

g is the number of features,

MPIW is the mean prediction interval width,

NMPIW is the normalized MPIW,

PICP is the prediction interval coverage probability,

$\mathcal{R}$ is the output range y (the difference between the maximum and minimum value of y in the dataset),

η and μ are hyper-parameters of the LUBE method (μ can be set to the desired accuracy (i.e., $1-\alpha$) and η can just be set to an arbitrarily large exponent),

U and L are the upper and lower bounds of the prediction,

y_z is the desired output (i.e., the real value),

c_z indicates whether the desired output for a given input X_z is within the prediction interval or not and ν is a slack variable.

Thus, instead of minimizing the error between the predictions and desired outputs, and having a single point prediction for each input, by training the NN using CWC, each prediction becomes a set of an upper and lower limit with an associated probability $1-\alpha$ as shown in Figure 4.5. Hence, based on the previous formulation, the optimization of CWC aims to minimize two competing functions, the width of the prediction (i.e., NMPIW) and the accuracy of the prediction (i.e., PICP), to achieve a range that is both narrow and accurate.

A main limitation in using PI in a load point CF prediction framework is the scarcity of cases with partial load shedding. For example, in the given set of 20,000 cases, a bus could have around 8900 cases with total blackouts and 10900 cases with no blackouts, but only 200 cases with partial load shedding. This severe skewness in data makes the correct prediction of these cases irrelevant as it contributes to 1% of the cases. Moreover, even with considering negative sampling techniques, there are insufficient cases to validate the model [129]. Hence, the PIs will only be applied to the total blackout rather than individual buses, while individual buses will have the conventional point prediction NN. However, if sufficient cases with partial load shedding are generated, then a PI can be designed for each individual bus.

4.4.4 Full model

Incorporating both the PI and DF into the model, the final model becomes as shown in Figure 4.6. The full model is detailed as follows. Each LOL_n prediction model outputs a single point predictor of the expected LOL at bus n. Then, these predictions for the load losses vector (Y) along with the input vector (X) are fed into the DF model, which outputs the corrected load loss per load. Then, these outputs are inputted to the PI model, which is based on the LUBE method described in Subsection 4.4.3. Also, similar to data fusion, the original input could be added to the PI model as a feedforward input.

4.5 Results and Discussion

4.5.1 Case generation and simulation

As stated in Section 4.1, one of the main limitations of previous CFP methods is only using static simulation or, more generally, ignoring the dynamics of the power network. Thus, it is crucial to simulate cases that account for the dynamics of the power network, specifically the dynamics relating to CF. Thus, the COSMIC package was used for simulation since it is an open-source MATLAB-compatible detailed dynamic simulation [27]. It has multiple advantages over other approaches, such as variable time steps and more detailed line heating dynamics. Also, it can simulate all the cases mentioned in Section 4.3 by which CF occurs. In COSMIC formulations, OC relays directly trip the affected lines. However, UVLS and UFLS relays are implemented as a 4-phase tripping, where each phase trips 25% of the pre-contingency load and waits for 0.5 *s* before initiating another phase (if the UV or UF is still sustained). However, COSMIC does not directly simulate hidden failures. Thus, it was introduced by randomly sampling N-2 and N-3 failures.

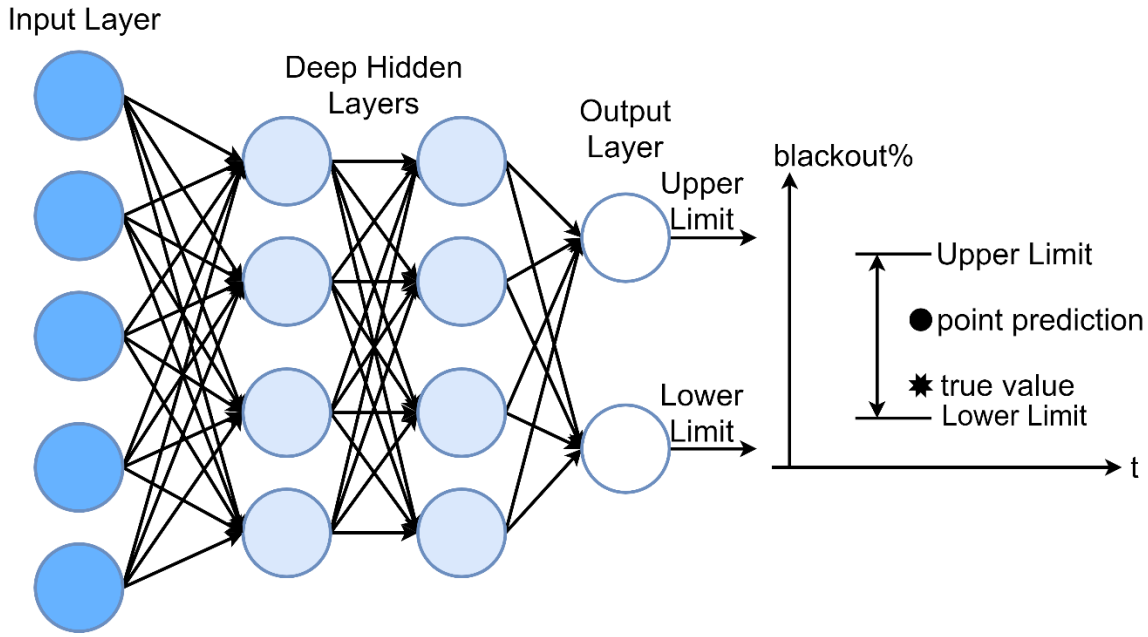

Figure 4.5. Deep neural network (DNN) model for generating prediction intervals (PI).

The IEEE 39-bus system (New England 10 machine power system) is used for evaluating the model. The simulated CF cases constitute 20,000 cases with different loading levels, repair

scenarios, and faults. Each case includes the full dynamical simulation of the case until either the system collapses, the cascading event saturates, or the simulation time ends (100 seconds). Table 4.2 summarizes the dataset (partial blackout includes both improper islanding and emergency load shedding). The large percentage of blackouts is attributed to two factors, the large number of N-3 contingencies and more importantly the fact that dynamic models tend to overestimate blackouts due to non-convergence [27].

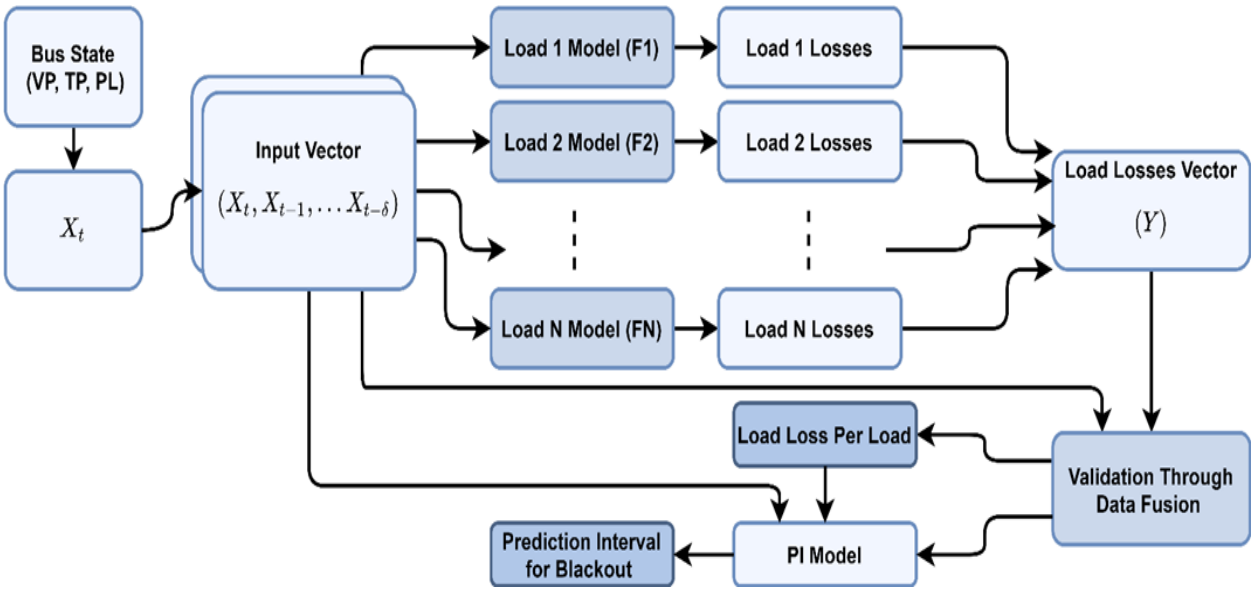

Figure 4.6. Load-point based CFP with prediction intervals and data fusion.

Table 4.2. IEEE 39-bus system simulated CF cases summary.

Type	Number of Cases
Simulated cases	20,000
- total blackout	8,933
- load shedding or partial blackout	3,015
- no blackout	8,052

4.5.2 NN specifications and data training

After simulating all the cases, the system's data recorded at each time-step are used as inputs for the training (voltage magnitude and angle). Also, the state of each load at the end of the simulation is used as the desired output. In addition, to avoid contaminating the training set with

the test set, the same train and test sets are used for all the components of the model. As for the ratios between training, validation, and testing, the data were split into 70% training, 15% validation, and 15% testing (both training and validation sets are used in training to avoid overfitting, while the results are based on the testing set only). All simulations are conducted in a PC with an i7-7700 processor (3.6 GHz quad-core processor), 16 GB DDR4 RAM, and no dedicated graphics card.

The implementation of the model is done as follows. The first component, the LOL prediction model, is built as 19 models since the IEEE 39-bus system has 19 load buses. As for the number of cycles, it is set to 3 samples (i.e., $50\ ms$ delay assuming 1 sample per cycle in a $60\ Hz$ system). Each model is built as a separate NN. The average training time for each model is $5\ s/epoch$ and $20\ epochs$ on average are needed to reach maximum validation accuracy. Moreover, since models can be trained in parallel, the number of models will not affect the total training time if enough processors are used. The second component, the DF model, is a single NN, with 19 outputs. Alternatively, this component can also be constructed as 19 models with a single output each. Each epoch requires an average of $2\ s$ and five epochs to reach maximum accuracy. The third component, the PI model, is built as a NN with two outputs (lower limit and upper limit). However, since the LUBE optimization function is nonlinear and non-differentiable, typical gradient descend algorithms cannot optimize it properly. As such, stochastic techniques are needed to optimize this component. The selection of the optimal stochastic algorithm is beyond the scope of this research. As such, the particle swarm optimization (PSO) technique is used since it is a general-purpose stochastic algorithm that can provide accurate global optima in a variety of cases [130]. The training time is $40\ s/iteration$ and the algorithm require around 30 iterations to reach maximum accuracy. The relatively long training time can be reduced by further tuning the PSO or using a more advanced algorithm. However, this time is still sufficient for CF prediction applications even for hour-ahead modeling.

4.5.3 Prediction model

Training the LOL prediction models (PM) on the IEEE 39-bus system dataset, the accuracy of detecting the LOL in the testing set is 90.9% on average and ranged from 89.5% to 92.2%, and it is detailed in Figure 4.7. The relatively low prediction accuracy is due to twofold limitations. First, the uncertainty of HF, as the model needs to predict the accuracy while accounting for the

cases where hidden failures exist. Thus, three scenarios can occur: no HF occurs, a HF occurs but does not significantly affect the final CF propagation, or a HF occurs and significantly changes the propagation of the CF. Hence, if the second scenario happens, the model must compromise between having a conservative estimate of the prediction (e.g., predict the output based on the worst possible HF) or not. Second, the model, by design, has no information regarding the locations of the faults and only uses three cycles after the fault, so there are cases where this information is not sufficient to produce a prediction.

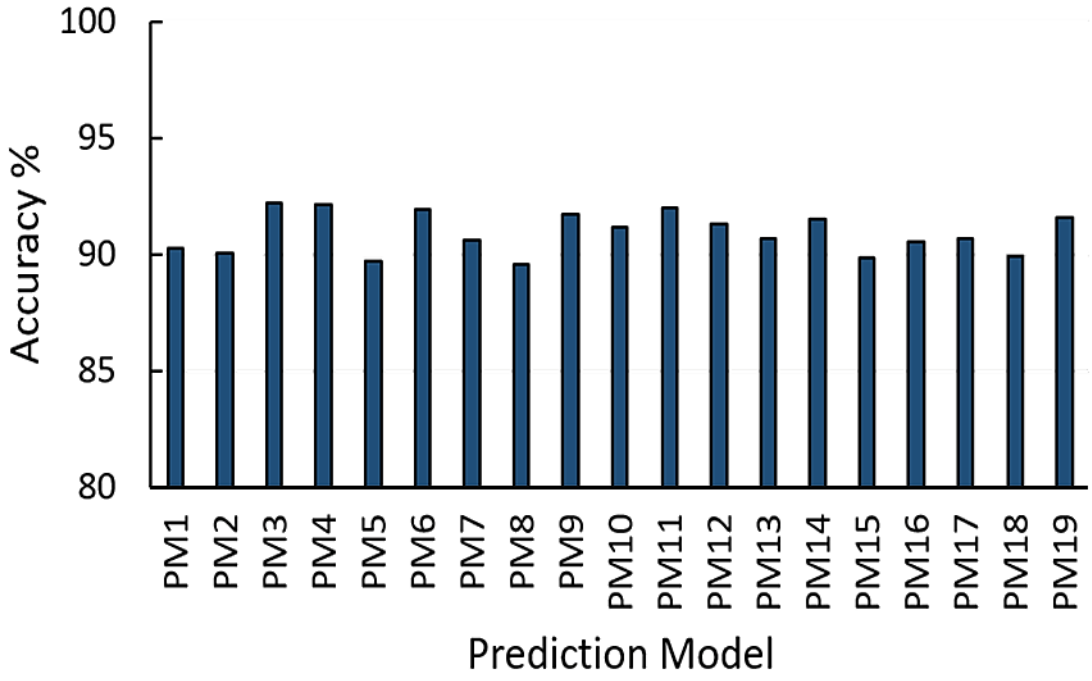

Figure 4.7. Testing set accuracy for LOL prediction models for the IEEE 39-bus system.

The next step in building the prediction model is to add the first additional component, the DF component. DF is applied as described in Subsection 4.4.2 (with feedforward from the input). The accuracy of the predictions increased for the training dataset by an average of 1.6%, as indicated in Figure 4.8. The accuracy for the testing set for each model now ranges from 91.5 to 93.3, with an average of 92.5. Thus, DF manages to correct 17.5% of the errors in the predictions on average (i.e., the average of the errors decreased from 9.1% to 7.5%). The reason for using the input feedforward model is that it significantly increases the accuracy without significantly increase the training time (the accuracy improvement when using the model without feedforward

is 0.8% on average). Also, a second observation is that the DF significantly reduced the errors in the models with the least accuracy. Thus, the deviation in accuracy between models is reduced.

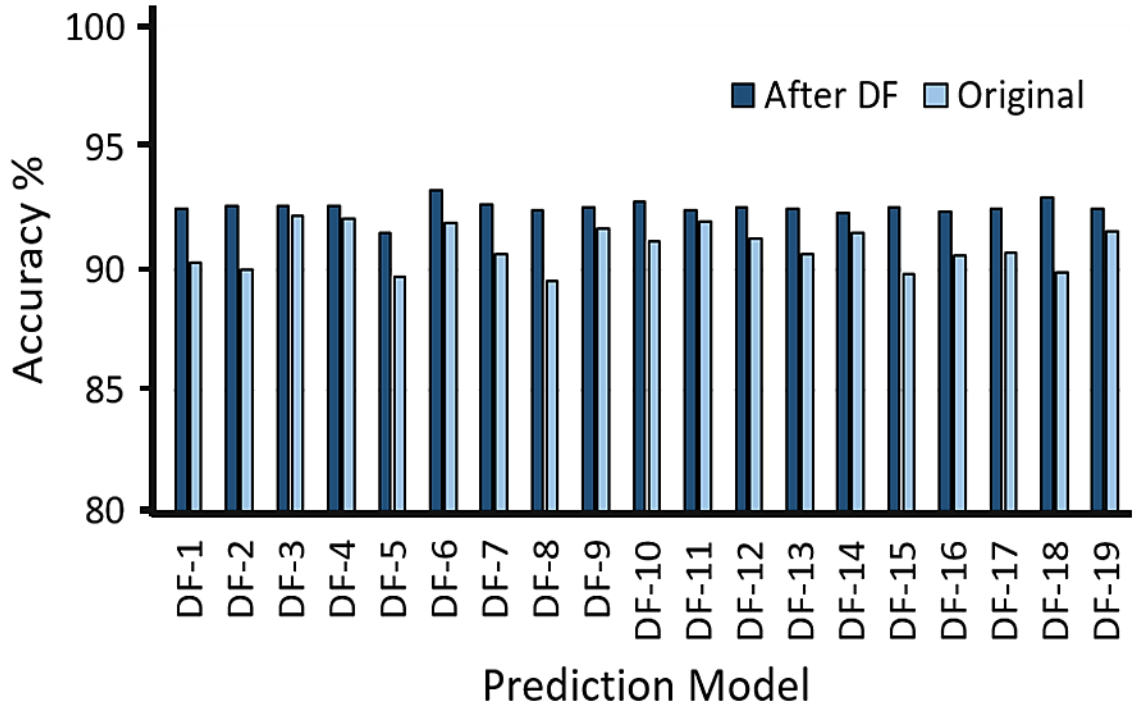

Figure 4.8. Testing set accuracy for LOL prediction models for the IEEE 39-bus system after applying data fusion (DF).

Finally, adding the PI component (without feedforward from the input) to output a single range describing the expected LOL of the system, the model achieved a 90% confidence interval (CI) with intervals having a width ranging from 0.1% to 10% with an average width of 5% of the total nominal load. The feedforward was not added in the PI as the stochastic training time significantly increased without notable improvement in the accuracy. Figure 4.9 shows 35 randomly selected cases, with correct predictions and various blackout percentages. Since the IEEE 39-bus system is relatively small and has very few redundancies in transmission lines, almost all blackouts larger than 20% results in the system collapsing with 90% or above blackout. Thus, it is rare to find cases with blackouts ranging between 20% and 90%.

To show the PI operation in real-time, the output of the PI following a disturbance in one of the simulated IEEE 39-bus system cases is presented in Figure 4.10. From the figure, the output of the model indicated that both the upper and lower limits are close to 0. Hence, the expected

blackout size is 0%. However, after the endogenous fault, the prediction increased its upper limit to 10% indicating that a partial blackout could happen, but it is unlikely since the lower limit is still 0%. Then, after the system stabilized, the PI returned to upper and lower limits of 0%, which indicates that the predictor is predicting normal operation (i.e., 0% blackout).

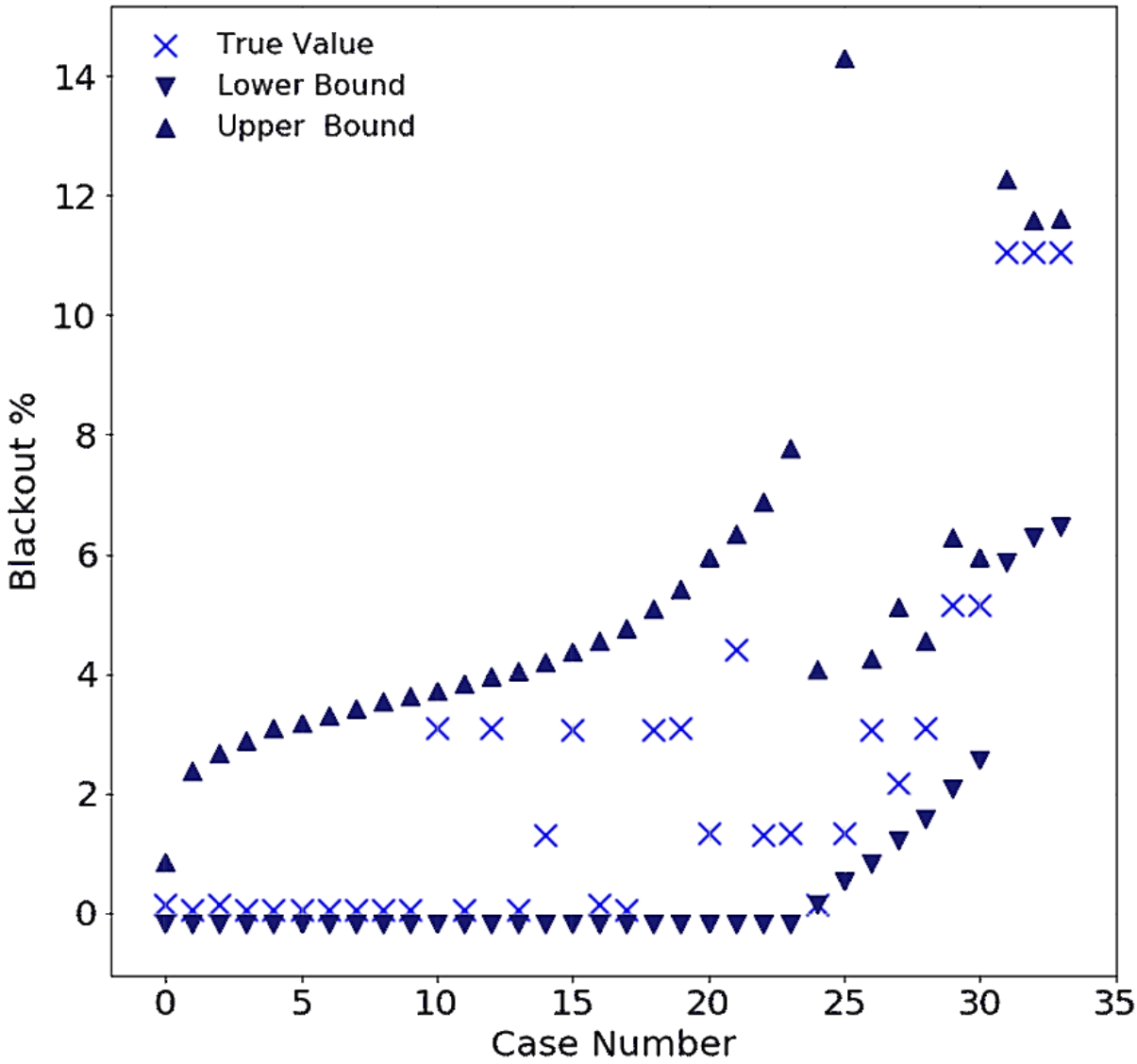

Figure 4.9. PI for a sample of the simulated cases (ordered by lower bound prediction).

Finally, comparing the obtained results with existing works, the width of the proposed PI (5%) is significantly narrower than the method proposed in [91] since its proposed brackets are 10-30% of the total nominal load, with an average width of 24%. However, this model uses a 90%

CI with a 50 ms delay, while the other model uses a 98.5% CI with 500 ms delay. As such, each model provides a different compromise between accuracy, speed, and interval width. In addition, this model includes the effect of hidden failures, which is ignored in [91]. Hence, the reported accuracy is an overestimate. The comparison between the proposed model and the existing model for PI-based CFP is summarized in Table II. Also, an additional PI model with 95% CI is added for comparison.

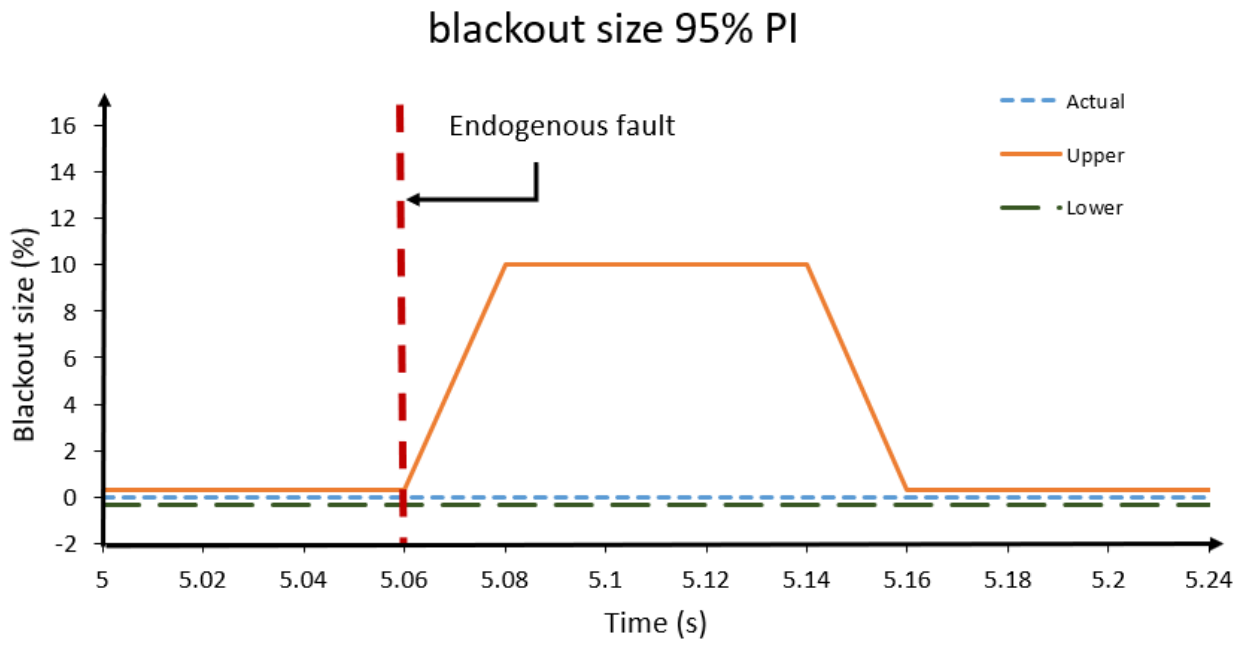

Figure 4.10. PI output in real-time following a disturbance.

Table 4.3 Models' comparison of different prediction interval schemes.

Model	Delay	Average Interval Width	Accuracy	Hidden Failures
LOL-DF-PI	50 ms	5%	90.1%	Yes
LOL-DF-PI-95%	50 ms	57%	95.4%	Yes
3-Stage-DT [91]	500 ms	25%	98.5%	No

4.6 Summary

In this chapter, typical CF propagation methods are explained to show the importance of dynamic models in comparison with static models. In addition, a new model is proposed for CF prediction in real-time applications. The proposed method reduces the problem of CF prediction to a set of parallel LOL predictors that further combine into a single predictor using DF and PI techniques. DF is used to exploit the correlation between the LOL predictors and the inputs to improve model accuracy. Moreover, PI is used to address a common limitation in conventional NN predictors, which is the lack of a reliability indicator. The results indicate that the model can detect the location of the buses affected by the cascade and provide narrow PI for the expected LOL in the grid using limited information and minimal delay, even while HF and multiple failures are considered. In conclusion, this model is expected to be applicable as a visualization or warning tool for the operators to react quickly and correctly to suppress contingencies that could propagate into a CF or a large-scale blackout.

5 A LOAD-POINT-BASED ONLINE CASCADING FAILURE PREDICTION FRAMEWORK FOR LARGE INTERCONNECTED POWER SYSTEMS

5.1 Introduction

Several published works address the analysis and prediction of CF in the literature [40], [91], [114]–[120]. Despite the variety of approaches and techniques, the majority of the current CF prediction (CFP) methods have still some limitations. First and foremost, these techniques do not precisely specify the areas and buses affected by a CF event. For example, although some of these techniques predict the next line to trip, reveal the expected number of line trips, and predict the overall expected size of a cascade, they do not precisely pinpoint the effects of CF on the individual system buses [40], [91], [118]–[120]. As such predictions are supposed to feed and trigger emergency control modules, it seems beneficial to share more detailed information about the impacts of CFs on the buses of interest.

Techniques for remedial action scheme (RAS) emergency control include actions such as load shedding and controlled islanding, and sufficient information about the state of the power system is needed to apply these emergency actions. Considering controlled islanding, for instance, an emergency islanding based on coherency groups of generators and minimization of power mismatch is discussed in [131]. Similarly, [132] introduced an islanding technique in case controlled islanding is needed following a fault. However, both algorithms require knowledge of whether a CF would cause uncontrolled islanding at a given region or not. As such, existing CFP methods that predict the next line trip or blackout percentages provide incomplete information for such emergency control techniques as they do not necessarily predict if a CF would lead to uncontrolled islanding or not. On the other hand, if a prediction algorithm can predict the buses affected by a CF, it is consequently capable of detecting the areas affected, i.e., islanded regions. Thus, several islanding methods reported in the literature, such as [131], [132], can be triggered with faster reaction time and better accuracy because more detailed information regarding the state of the network is revealed in the first place.

Another limitation in CFP is that most existing works are not ideal for real-time operation and are limited to planning-stage studies. For example, generator dynamics are ignored in [114] and [118]. Moreover, some methods capture system transients, but require a long time-series of data before making the prediction [91]. Additionally, most of these methods are slow to detect the final blackout size due to negligence of the transient state of the network and its dynamics [133]. Thus, it seems worthwhile to introduce a predictor that can utilize dynamic measurements obtained from PMUs, i.e., capable of updating the prediction based on each single data sample received from the measurement device. Such an approach can substantially increase both the speed and accuracy of CFP. Recurrent neural network (RNN) long short-term memory (LSTM) cells can accept temporal sequences of data and hence be used for this purpose. Successful applications of RNN-LSTM to predict dynamic events using PMU data are reported in [21]. In addition, LSTM allows for predictions to be produced at each time step, instead of waiting for a long delay period as in [91]. Hence, the application of LSTM for CFP can be further investigated based on PMU data to facilitate the incorporation of transient behavior in this area.

Last but not least, one can note that scaling is a major issue in CFP models when applied to large interconnected systems. In such cases, the number of inputs and parameters can grow exponentially depending on the model used, substantially increasing the computational complexity of the resulting model. This issue is partially addressed in [40], [91], but the proposed techniques do not consider the importance and ranking of transmission lines and buses [134]. However, this limits the efficacy of the model as the criticality of lines and buses is an important aspect of CF propagation [100]. Nonetheless, depending solely on critical lines and buses could result in a distorted view of the network, as a large portion thereof could be unobservable by the set of critical buses. Thus, it seems beneficial to use information theory and probabilistic models to find the ideal candidate features to input to the prediction models and minimize the loss of information.

In this chapter, a load-based real-time CFP framework is introduced to address the limitations of existing CFP literature. The main contribution of this method is the load perspective approach to CFP, which provides operators with more reliable information regarding the affected areas and the expected size of the blackout. Also, by incorporating the Theil index and eigenvector centrality for data selection, the model can scale to large interconnected systems without a significant loss of information. Moreover, higher order probabilistic moments are used to also

maximize the amount of retained information from a small set of inputs further aiding in the scaling of the model. As a secondary contribution of this work, RNN-LSTM cells are used to train the model to address the static flow issue and capture transient data from PMUs. Finally, several maintenance topologies, multiple failures, different loading levels, and various power dispatches are considered to make the model more robust to system changes. All in all, the model addresses multiple issues existing in the current CFP literature, and to the best of the authors' knowledge, such implementation has never been applied before.

5.2 Load-Point-Based Ensemble RNN

5.2.1 Overview

The main goal of a CFP model is to predict the propagation and extent of a potential CF using available data, e.g., PMU measurements, with the shortest delay and acceptable accuracy. Ideally, dynamic simulations are needed for such predictions to avoid the pitfalls of static simulations [27]. Running real-time dynamic simulations to identify and control the grid for a possible CF may not be feasible, so it might be helpful to transform the model into a machine learning prediction problem, which facilities estimation of the underlying complexities using methods such as neural networks, decision trees, or similar algorithms [27], [91]. However, this may lead to a major challenge with respect to capturing transient dynamic behaviors, as a constant delay d is needed to feed the model the vector of delayed inputs ($x = x_t, x_{t-1}, \dots, x_{t-d}$) to output the prediction ($y = y_t$) [91]. For example, Figure 5.1 (a) represents a model with a two time-step delay ($x = x_t, x_{t-1}, x_{t-2}, y = y_t$). To address this shortcoming, RNNs accept temporal sequence inputs and can hence produce output y_t based on both the inputs x_t and the state of the system s, rather than depending on the inputs alone [21]. Moreover, as shown in Figure 5.1(b), the RNN calculates its predictions using all inputs in $t \in (-\infty, 0]$. This interaction is elaborated by Figure 5.1(c) which shows the unfolded structure of RNN cells. Hence, models based on RNN cells can produce outputs for each time t, so building separate models for different delays is not required.

Although the use of an RNN solves a major issue in CFP, the curse of dimensionality may deteriorate model performance. This is because RNN training time rapidly increases as the training set size and number of RNN inputs increase [135]. To partially mitigate this effect, a load-based ensemble of RNNs is used. Specifically, instead of building a single RNN model to represent the

relationship between system state and loss of load (LOL), an ensemble of RNNs is employed as shown in Figure 5.2 where each essential load[1] has a separate RNN to predict its LOL. This load-based approach to divide RNNs achieves two benefits in addition to minimizing the negative impacts of the curse of dimensionality; it provides more meaningful information because the areas affected by the CF can be visualized, and it also allows for parallel processing, which aids in the scalability of the model to large-scale interconnected networks.

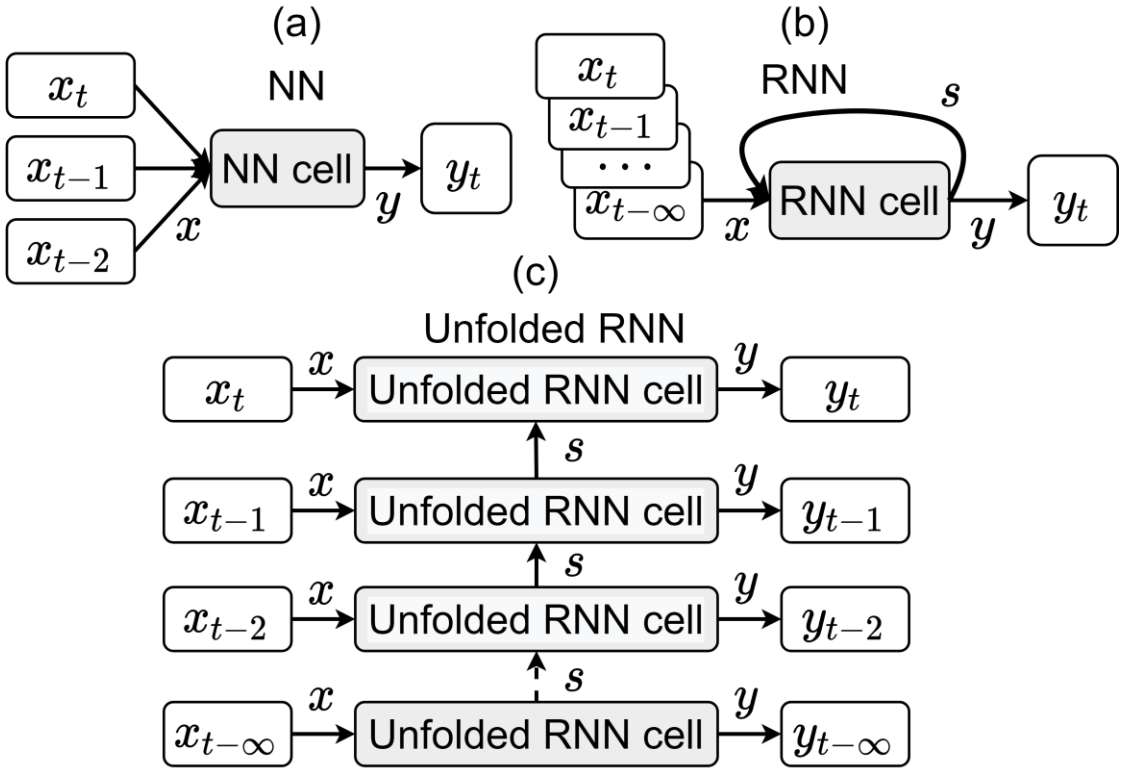

Figure 5.1. (a) NN structure (for a two time-step delay), (b) RNN structure, and (c) RNN structure after unfolding the internal state feedback loop.

However, having an ensemble of RNNs by itself is not sufficient to guarantee scalability. This is because the ensemble can reduce the size and training time of individual RNNs but cannot reduce the number of RNN inputs (i.e., the ensemble would turn a model with M inputs, C cells, and N outputs into N models with M inputs, $C - N + 1$ cells, and 1 output. Hence the number of

[1] Essential load refers to any load(s) of interest to the operator or decision maker. If enough resources are available, then all buses in the network could be considered as essential load. Hence, the ensemble would be applied to all buses.

inputs is the same albeit the size of the model is reduced). To address this limitation, ranking indices and weighted average probability indices are introduced for input selection, which guarantees the scalability of the model by limiting the number of inputs even in large networks.

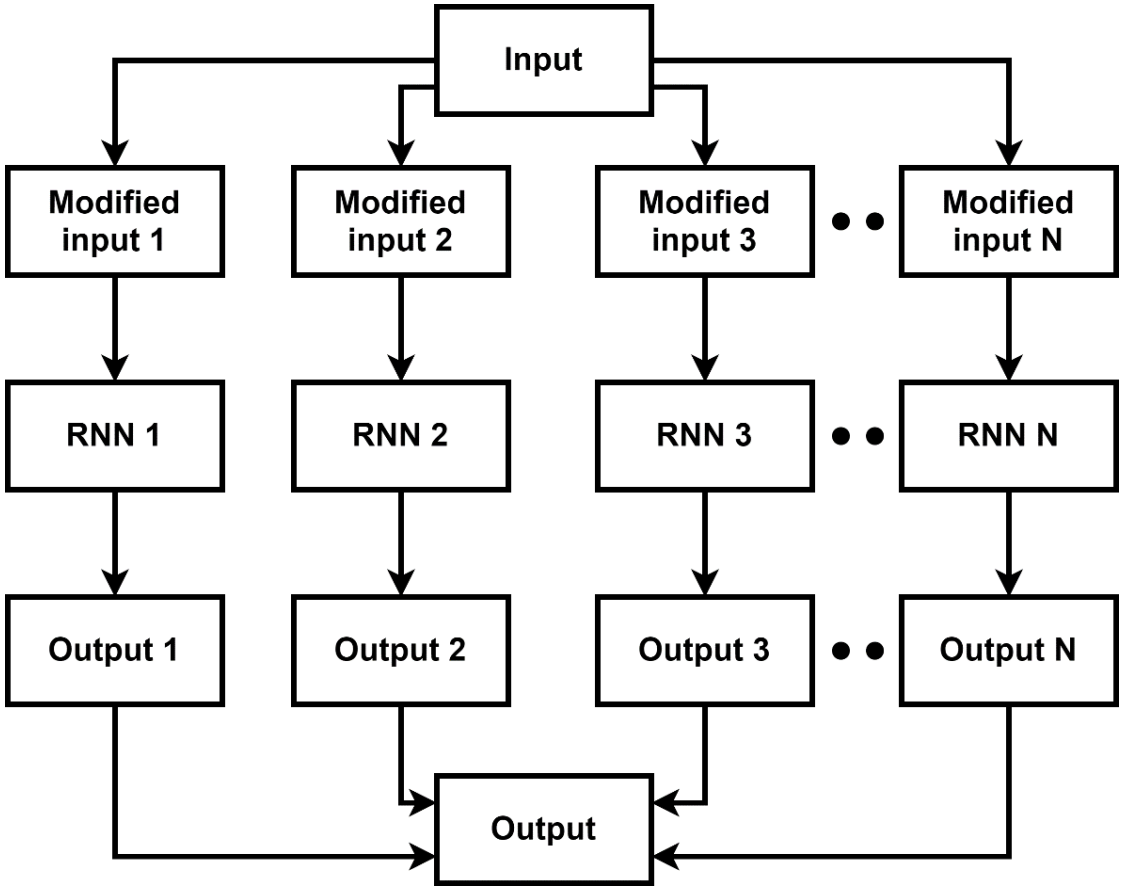

Figure 5.2. Structure of the proposed RNN ensemble.

5.2.2 LSTM structure

To build the RNN ensemble, each essential load requires a separate RNN that is comprised of one or multiple LSTM cells shown in Figure 5.3. The behavior of each LSTM cell can be described as follows [136]. First, the LSTM has three inputs (current input X_t, previous cell state C_{t-1}, and previous hidden state H_{t-1}) and two outputs (cell state C_t, and hidden state H_t). The inputs X_t and H_{t-1} are passed through the forget gate f_t:

$$f_t = \sigma(\overline{W}_f.\left[\overline{H}_{t-1}, \overline{X}_t\right] + b_f) \tag{5.1}$$

where:

W and b are the weightings and biases for each gate within the LSTM (the subscript indicates the corresponding gate),

σ is the sigmoid activation function[1].

Then, the same inputs are passed again through an auxiliary gate k_t and an input gate i_t:

$$k_t = \tanh(\overline{W}_k.\left[\overline{H}_{t-1}, \overline{X}_t\right] + b_k) \tag{5.2}$$

$$i_t = \sigma(\overline{W}_i.\left[\overline{H}_{t-1}, \overline{X}_t\right] + b_i) \tag{5.3}$$

where:

$tanh$ is the tanh activation function[2].

Subsequently, the outputs from these three gates merge with the previous cell state to produce the new state of the cell:

$$\overline{C}_t = \overline{f}_t \otimes \overline{C}_{t-1} + \overline{i}_t \otimes \overline{k}_t \tag{5.4}$$

where:

$\otimes$ is bitwise multiplication.

Finally, to obtain the new hidden state H_t, X_t and H_{t-1} are passed through the output gate o and merged with C_{t-1} through a $tanh$ activation function:

$$o_t = \sigma(\overline{W}_o.\left[\overline{H}_{t-1}, \overline{X}_t\right] + b_o) \tag{5.5}$$

$$\overline{H}_t = \overline{o}_t \otimes \tanh\left(\overline{C}_t\right) \tag{5.6}$$

This structure of gates makes LSTM suitable for capturing power system dynamics as the output of the LSTM depends on two cell states (the output state and the hidden state). This dual state structure of LSTM allows it to simultaneously capture the slow and fast dynamics of the power system as LSTMs can detect both long- and short-term dependencies [137]. For instance, if

[1] Sigmoid activation function is defined as: $\sigma(x) = \frac{1}{1-e^{-x}}$

[2] Tanh activation function is defined as: $\tanh(x) = \frac{e^{x}-e^{-x}}{e^{x}+e^{-x}}$

a disturbance happens while a line is also overheating, a single LSTM can track the effects of both the disturbance (short-term event) and the overheating (long-term event) without suffering from the vanishing gradient problem.

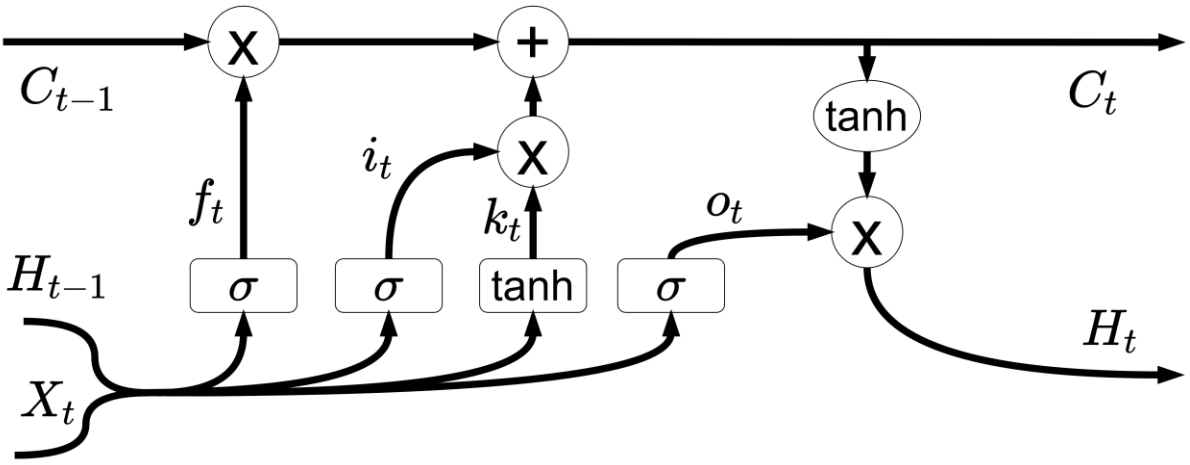

Figure 5.3. Structure of a single LSTM cell.

5.3 Ranking Indices and Reduced Input Set

Selecting only the most critical information as inputs is important for limiting the number of inputs to each RNN model within the ensemble. However, no single universal ranking index can measure the importance of buses and lines in power networks [134]. Thus, it is important to consider multiple indices to reduce biases towards any specific ranking metric. To achieve this goal, three ranking indices are introduced based on node distance, entropy change, and bus centrality. Mathematically, this ranking of buses to select the top buses represents a projection of the higher dimension point of all state variables (e.g., all voltages magnitude and angle) to a dimension with only the states of the important buses in the network (e.g., the voltages magnitude and angle of five buses). To illustrate this projection, Figure 5.4 represents the projection of classifications in an $\mathbb{R}^3$ space into an $\mathbb{R}^2$ space. In the figure, any information relating to the z-axis is not retained and a new space is formed for the prediction. However, it is still possible to build a classifier to differentiate between the two sets of points using only $\mathbb{R}^2$ space as the z axis information was not essential. A similar concept applies to the prediction model proposed here, but the dimension reduction cannot be visualized as the number of states even in a small power grid (e.g., IEEE 39-bus system) is too large to visualize.

After selecting the buses based on node distance, entropy change, and bus centrality, average weighting and statistical moments are used to merge all indices together. Finally, the inputs are selected based on available PMU data and passed to the RNN-LSTM ensemble model.

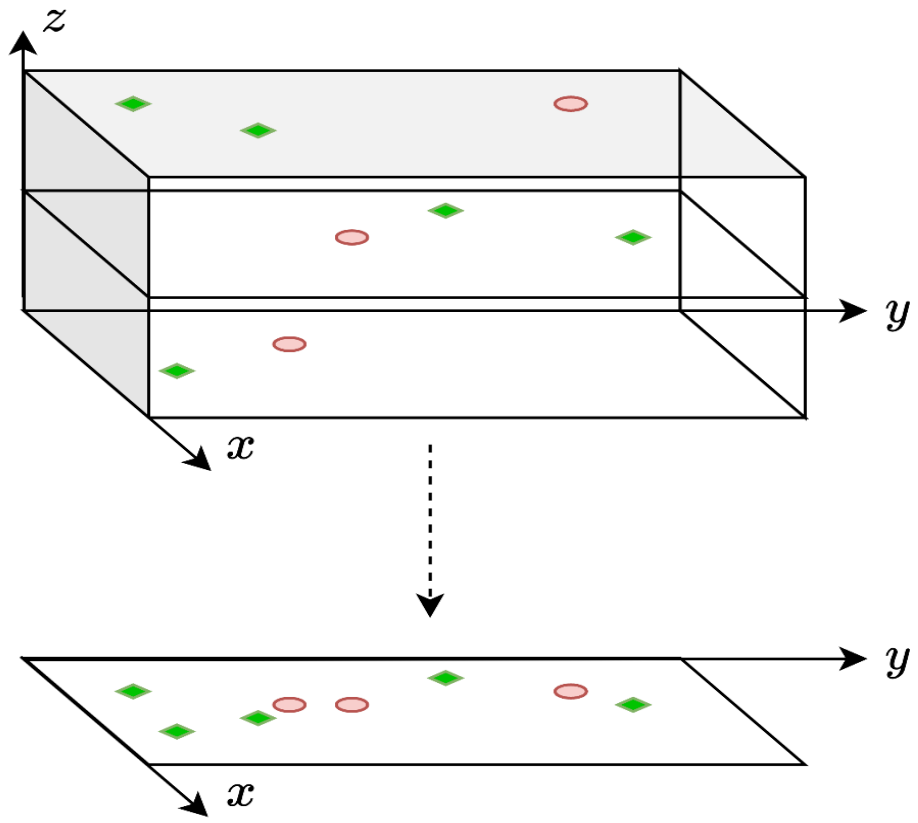

Figure 5.4. Projection of classifier points in an $\mathbb{R}^3$ space into an $\mathbb{R}^2$ space

5.3.1 Node distance

Considering graph theory, if a power network is assumed to be static, then it can be represented as a graph where buses are nodes and transmission lines are edges [138]. Under this assumption, the distance (number of edges) becomes the most important correlating factor between buses (nodes $\aleph$). Using the graph shown in Figure 5.5 which represents an 18-bus system as an example, if a fault happens at Node 0 ($\aleph_0$), then only $\aleph_1$ and $\aleph_2$ will be affected as they are in A_1 area, and only when these nodes fail, the fault will propagate to other nodes via A_2 edges (e.g., A_2 area). Subsequently, $\aleph_5$, which is connected to the faulted bus via an A_4 edge, would only fail as the 4th failure even when considering the shortest path ($t_0: \aleph_0 \rightarrow t_1: \aleph_1 \rightarrow t_2: \aleph_3 \rightarrow t_3: \aleph_4 \rightarrow$

t_4: $\aleph_5$). Similarly, if the fault is at $\aleph_5$, then it would only propagate to $\aleph_0$ as the 4th failure. Hence, as the number of edges between nodes increases the effect of their faults decreases. Thus, a ranking index for the importance of a bus can be formulated as follows:

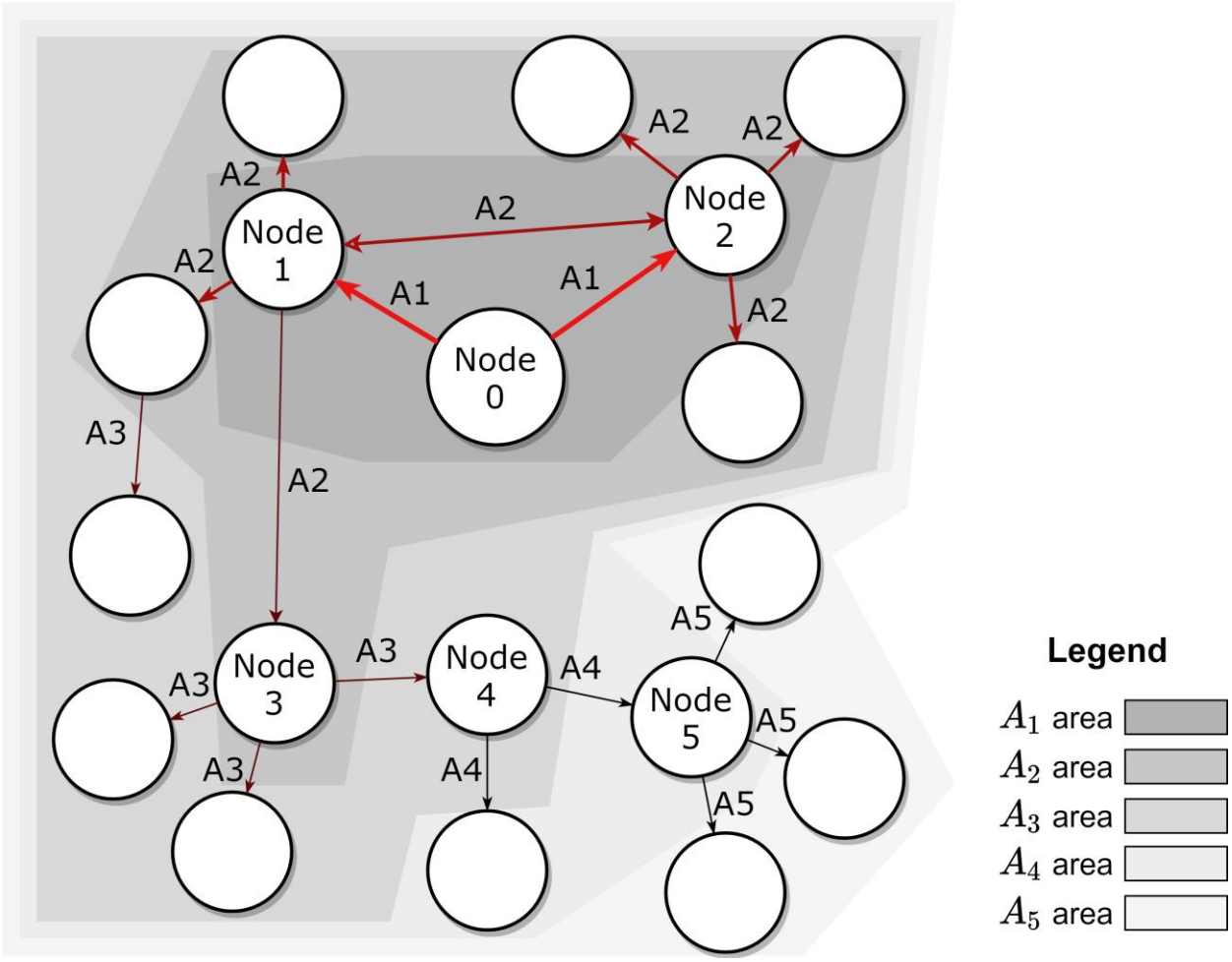

Figure 5.5. Network graph representation and fault propagation for an 18-bus power system (only nodes of interest are numbered).

Assigning bus B as the observed bus[1], the ranking factor of each bus i in the network from the perspective of B is calculated as:

[1] The observed bus B refers to the bus that is currently being observed using a single RNN in the RNN ensemble. Hence, for each essential bus (or all buses if considering that all buses as essential buses), there will be an RNN model where that bus is the observed bus B.

$$w_{i,B,d} = \begin{cases} \alpha^{-d_{i,B}}, \forall i \in \Omega_{d_{max}} \\ 0, \quad \forall i \notin \Omega_{d_{max}} \end{cases} \tag{5.7}$$

$$\Omega_{dmax} = \{i | d_{i,B} < d_{max}\} \tag{5.8}$$

Where:

$w_{i,B,d}$ is the weighting of bus i with respect to bus B according to the distance index,

α is the attenuation factor,

d_{max} is the largest allowable distance between bus B and other buses to consider the bus as significant (both α and d_{max} are tunable parameters of the node distance index),

$d_{i,B}$ is the minimum number of edges (lines) between bus i and bus B, which can be calculated for the entire system using the Floyd Warshall (FW) algorithm, which guarantees the global minimum distance for all buses [139].

Alternatively, the node distance index can be replaced by any graph, and it is not restricted to only the topological graph of the power system. For instance, [37] describes a modified graph that is not the grid topology, but the actual propagation of the CF. That modified graph, and similar graphs, could be used in generating the node distance indices using the same equations (5.7)-(5.8) but d_i and d_{max} will be calculated based on the modified graph instead of the grid topology.

5.3.2 Information entropy

Following disturbances in the power system, certain buses experience more voltage fluctuations than others. These buses are more important to observe during cascading events because they provide more information compared to buses that their voltages are almost constant even under different contingencies (i.e., provide less information). For example, in Figure 5.6, the voltage at bus 5, in a particular set of cases, is having a higher entropy compared to the voltage at bus 27, because the voltage has more variability. Whereas the voltage at bus 27 has most of its values close to 1.035, so it does not provide as much information as bus 5.

To quantify the amount of information entropy (e.g., the disparity between values), the generalized entropy index (GEI) can be used. The GEI can distinguish between redundant variables and crucial factors as it measures the degree of variation that a variable is exhibiting. Moreover, GEI can

differentiate between redundant variables regardless of normalization since its formulation includes both the magnitude and distribution of the change in variables. The GEI is calculated by [134]:

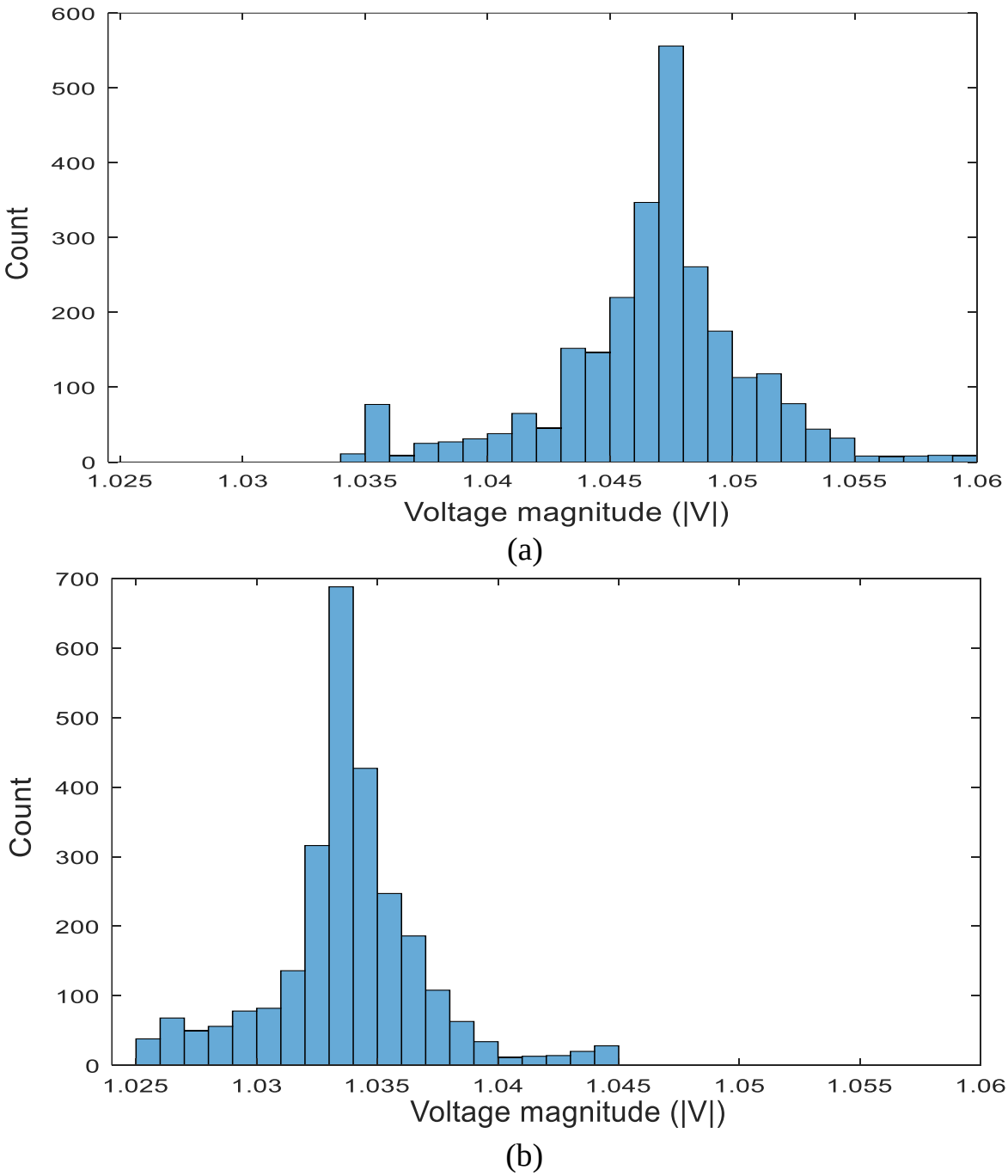

Figure 5.6. Distribution of the voltage for two buses in the IEEE 39-bus system: (a) voltage at bus 5 (high entropy case) and (b) voltage at bus 27 (low entropy case).

$$G(\beta) = \begin{cases} \frac{1}{\beta(\beta-1)}\left[\frac{1}{n}\cdot\sum_{i=1}^{n}\left(\frac{x_i}{\bar{x}}\right)^{\beta}-1\right], & \beta \neq 0,1 \\ \frac{1}{n}\cdot\sum_{i=1}^{n}\log\left(\frac{x_i}{\bar{x}}\right), & \beta = 0 \\ \frac{1}{n}\cdot\sum_{i=1}^{n}\frac{x_i}{\bar{x}}\cdot\log\left(\frac{x_i}{\bar{x}}\right), & \beta = 1 \end{cases} \quad (5.9)$$

$$\bar{x} = \frac{1}{n}\cdot\sum_{i=1}^{n} h_i \,.\, x_i \quad (5.10)$$

where:

n is the number of data points in a sample,

x_i is the i^{th} datapoint,

h_i is the weighting factor for x_i,

$\bar{x}$ is the mean of the sample,

β is the difference coefficient, which is a parameter of the GEI.

Within the GEI, when ($\beta = 1$) the index is referred to as the Theil Index (TI), which is preferred when measuring differences within the middle of the sample's distribution, and that is often the main focus of analyses. Extending the TI to CF simulation and given a set of S CF scenarios, the sensitivity (or information entropy) of bus i to the CF can be measured using the voltages (or phase angles) of the bus in these S scenarios. For brevity, a single time sample of voltage is used in the formulation, but the equations can be expanded to multiple time samples per scenario. Using a single time sample v_{i,s,t_1} from each CF simulation, the TI is calculated as:

$$\bar{v} = \frac{1}{S}\cdot\sum_{s=1}^{S} v_{i,s,t_1} \quad (5.11)$$

$$w_{i,e} = TI = G(1) = \frac{1}{S}\cdot\sum_{s=1}^{S}\frac{v_{i,s,t_1}}{\bar{v}}\cdot\log\left(\frac{v_{i,s,t_1}}{\bar{v}}\right) \quad (5.12)$$

where:

$w_{i,e}$ is the weighting of bus i according to the entropy index (note: this weight is independent of the observed bus B, i.e., $w_{i,B,e} = w_{i,e}$),

v_{i,s,t_1} is the voltage of bus i following fault scenario s at time t_1,

$\bar{v}$ is the mean of the voltage at bus i, and assuming all datapoints have equal weighting ($h_i = 1$) in (5.10).

5.3.3 Bus centrality

This index measures the importance of a bus by evaluating its centrality, which can indicate the influence of a node (bus) on the rest of the network. Centrality is an indicator of the connectedness of a node to the rest of the nodes. As such, buses connecting multiple regions (e.g., bus stations between long tie lines) would have higher centrality, while buses that are only connected to fewer buses (e.g., radial buses) would have lower centrality. However, in large networks, a large set of buses would have large centrality values. To address this concern, the formulation of centrality needs to take the observed bus and fault locations into account. This can be achieved by defining a new subgraph S of the original power network graph (S_o) such that:

$$S = \{(V(S), E(S)) \mid d_{iB} < \gamma \text{ and } d_{iF} < \gamma \quad \forall\, i \in V(S)\} \tag{5.13}$$

where:

$V(S)$ and $E(S)$ are the vertices and edges of S, respectively,

d_{iB} is the distance between bus i and the observed bus B,

d_{iF} is the distance between bus i and the fault location (the fault is assumed to be at the ending side of the transmission line that is closest to the fault),

γ is a tunable parameter for the subgraph size.

Then, using eigenvector centrality (EC), the centrality of each bus according to the new subgraph can be measured as follows [140]:

$$A_{ij}^{*} = \begin{cases} 1 & buses\ i\ and\ j\ are\ connected\ in\ S \\ 0 & otherwise \end{cases} \tag{5.14}$$

$$w_{i,B,F,c} = \begin{cases} c_i = k^{-1} \cdot \sum_{j \in S} A^*_{ij} \cdot c_j & i \in S \\ 0 & i \notin S \end{cases} \tag{5.15}$$

where:

A^* is the adjacency matrix of S,

$w_{i,B,F,c}$ is the weighting factor of bus i according to the centrality index when observing bus B while a fault happens at F,

c_i is the centrality of bus i,

k is a constant that should be larger than or equal to the maximum eigenvalue of matrix A^*, to ensure EC values are all non-negative.

5.3.4 Reduced set weighted average probability indices

To merge all the ranking indices, an average weighting is used to normalize the indices and adjust their contributions:

$$w_{i,B} = \alpha_{\mathrm{d}}(w_{i,B,d}) + \alpha_{\mathrm{e}}(w_{i,e}) + \alpha_{\mathrm{c}}(w_{i,B,F,c}) \tag{5.16}$$

where:

$w_{i,B}$ is the final weight of importance of bus i with respect to bus B,

$\alpha_{\mathrm{d}}, \alpha_{\mathrm{e}}$, and α_{c} are functions to normalize and adjust the weightings of distance, entropy, and centrality indices, respectively.

Selecting the top ψ buses according to their $w_{i,B}$ values, the set Ψ can be defined as the set of the most important buses to observe bus B. This Ψ set is the main input to the RNN and the information about the remaining buses could be discarded. However, to retain more information about the state of the grid, these buses are merged into a weighted average instead of discarding them:

$$\mu = \frac{1}{\sum_{i \notin \Psi} w_{i,B}} \cdot \sum_{i \notin \Psi} w_{i,B}\, x_i \tag{5.17}$$

Moreover, higher moments such as variance, skewness, and kurtosis are used to capture more information about the distribution of these values, as these higher moments change significantly during CF [120]. Thus, the weighted values of the higher moments are calculated as:

$$\sigma^2 = \frac{1}{\sum_{i\notin\psi} w_{i,B}} \cdot \sum_{i\notin\Psi} \left(w_{i,B}\, x_i - \mu\right)^2 \tag{5.18}$$

$$sk = \frac{1}{s^3 \cdot \sum_{i\notin\psi} w_{i,B}} \cdot \sum_{i\notin\Psi} \left(w_{i,B}\, x_i - \mu\right)^3 \tag{5.19}$$

$$kr = \frac{1}{s^4 \cdot \sum_{i\notin\psi} w_{i,B}} \cdot \sum_{i\notin\psi} \left(w_{i,B}\, x_i - \mu\right)^4 \tag{5.20}$$

where:

s is the standard deviation,

σ^2 is the variance,

sk is the skewness,

kr is the kurtosis.

Figure 5.7 illustrates the importance of the higher moments where each graph represents a distribution with the same mean and variance, but with a different skewness and kurtosis (according to the Gram-Charlier expansion of the normal distribution) [141]. Since completely different distributions (e.g., voltage distribution at the different buses of the power grid) could have similar mean and variance, having more moments reduces the probability of such similarity. However, there are infinite distributions with the same mean, variance, skewness, and kurtosis. Hence, these four measurements are not always sufficient to distinguish between the different distributions. Nonetheless, having four moments instead of two provides more information regarding the shape of the distribution as illustrated in the figure. As for the interpretation of the higher moments, it can change depending on the distribution, but in general they represent the following concepts:

- The mean represents the middle point (location) of the distribution.
- The variance is its scatter (spread).

- The skewness represents the shift in the distribution between its mean and median (i.e., lack of symmetry).
- The kurtosis is the shortness of the tail of the distribution (peakedness) [142].

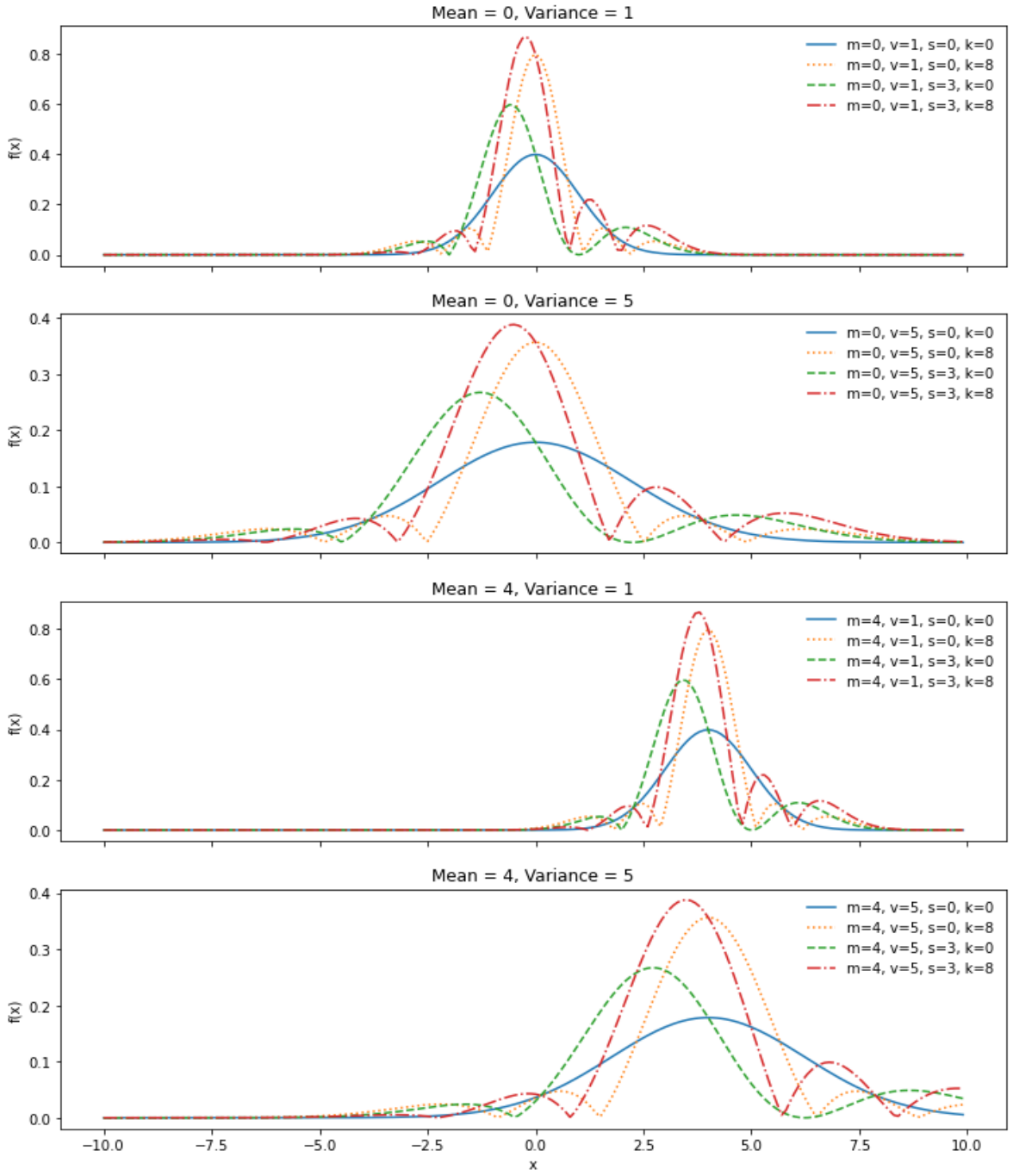

Figure 5.7: Visualization of the skewness and kurtosis effect on a normal distribution using the Gram-Charlier expansion, m is the mean, v is the variance, s is the skewness and k is the kurtosis of the distribution.

Based on the proposed Ψ set and the additional higher moment inputs, the total number of inputs to the RNN will only be $(|\Psi| + 4) * f$, where f is the number of features, regardless of the size of the network. Hence, the curse of dimensionality is significantly suppressed because the input count is independent of the total buses count in the network.

5.3.5 Model implementation

Given a power network with $|\Omega^M|$ buses, $|\Omega^N|$ load buses, and $|\Omega^J|$ lines, the set of possible input features for the proposed load-point prediction model can be defined as:

$$X = [SL\ FL\ PL\ VP\ TP] \tag{5.21}$$

where:

SL is the network's line state,

FL is the fault location,

PL is the real power loading,

VP, and TP are the voltage, and theta measurements obtained from the PMUs, respectively.

Also, each element in X is referred to as a set W (e.g., $W_1 = SL$), where each W set is a row vector defined as follows:

$$\begin{aligned}
&SL_j = \begin{cases} 1 & line\ j\ is\ out \\ 0 & otherwise \end{cases} \\
&FL_j = \begin{cases} 1 & j = fault\ location \\ 0 & otherwise \end{cases} \\
&PL = \{p_{L_n} | \forall\ n \in \Omega^N\}\ , \qquad VP = \{v_m | \forall\ m \in \Omega^M\}, \\
&TP = \{\theta_m | \forall\ m \in \Omega^M\} \\
&j = 1,2, \dots |\Omega^J|;\ m = 1,2, \dots |\Omega^M|;\ n = 1,2, \dots |\Omega^N|
\end{aligned} \tag{5.22}$$

where:

P_{L_n} is the real power at load bus n,

v_m and θ_m are the voltage magnitude and angle at bus m, respectively.

Ω^J, Ω^M, and Ω^N are the sets of transmission lines, load buses, and buses, respectively.

Then, for each set in X, only the values referring to the top ψ buses are selected while the remaining buses are reduced to probability indices according to (5.17)-(5.20). Hence, the set of inputs to the RNN are:

$$X' = \{\{W_{k\Psi}, \mu_{W_k}, \sigma_{W_k}, sk_{W_k}, kr_{W_k}\} | \forall W_k \in X\} \tag{5.23}$$

where:

X' is the updated input features,

$W_{k\Psi}$ is the W_k values for buses $i \in \Psi$. $\mu_{W_k}, \sigma_{W_k}, sk_{W_k}$, and kr_{W_k} are the mean, variance, skewness, and kurtosis of $W_k \backslash \Psi$ (the set of W_k excluding elements in Ψ), respectively.

Moreover, to address the trade-off between speed and accuracy, the model is set up to predict the LOL within a maximum of T cycles after the initiating fault. Thus, the prediction of the LOL is obtained as:

$$LOL_n = \mathcal{F}_{\mathrm{n}}(X'_T, X'_{T-1}, \dots, X'_0) \tag{5.24}$$

where:

$\mathcal{F}_{\mathrm{n}}$ is the RNN associated with predicting the LOL at load bus n and X'_T is the X' at cycle T.

However, note that RNN produces output for each set of inputs X'_t. Thus, T refers to the maximum delay, rather than a constant delay (i.e., the algorithm still produces predictions at all $0 \leq t \leq$ T). Afterward, if all load buses are considered as loads of interest (i.e., $n = 1,2, \dots |\Omega^N|$), the total size of the predicted cascade ($blackout\%$) is calculated as:

$$blackout\% = \frac{\sum_{n=1}^{|\Omega^N|} LOL_n}{\sum_{n=1}^{|\Omega^N|} P_{Ln}} \tag{5.25}$$

To sum it up, the offline portion of the model commences by generating the CF scenarios and simulating them while calculating the three indices (w_c, w_d, and w_e) using (5.7)-(5.15) in parallel. Thereafter, the values needed for X (5.21) and (5.22) are extracted from the data and passed to the weighting and selection phase according to (5.16) to find the top ranking buses ψ. Furthermore, for each load bus of interest (or all load buses if all buses are considered buses of interest), an RNN-LSTM prediction model is trained in parallel, where the inputs are the top buses

in addition to the weighted average of the remaining buses' probability indices using (5.23). Finally, for the online implementation, the data in X are already measured or calculated from PMUs. Hence, the data is directly passed to the weighting and selection phase which is already calculated offline. Therefore, only (5.17)-(5.20) are calculated online before inputting the data to the trained RNN-LSTM ensemble. Finally, the implementation of the model, with respect to online implementation and offline training, and all the relevant calculations are summarized in Figure 5.8.

5.4 Case Studies

5.4.1 Dynamic CF simulations and NN specifications

The CF simulations are conducted using the COSMIC package, an open-source MATLAB-compatible detailed dynamic simulation designed explicitly for CF. It has multiple advantages over other approaches, including variable time steps, generator ramping, and detailed line heating and generator dynamics [27]. The IEEE 39-bus system and the 2383-bus Polish winter peak system according to [27] (henceforth POL 2383-bus system) are used to test the model and demonstrate its applicability to both small- and large-scale networks. Each case has different loading levels and repair scenarios, and both single and multiple initiating faults are considered (faults range from N-1 to N-5). Moreover, each case includes the full dynamical simulation until either the system collapses, the cascading event saturates, or the simulation time ends ($60\,s$ is selected as the simulation time). Furthermore, the dynamic simulation has a variable time step with a minimum step of $1 \times 10^{-6}\,s$ (during fast cascade) and a maximum step of $1\,s$ (during slow cascade) to maximize simulation efficiency, while the PMUs are assumed to record the measurements once each cycle $\approx 16.7 \times 10^{-3}\,s$.

After simulating all cases, the system's data at each time step are recorded as the raw inputs. Similarly, the state of each load at the end of the simulation is assigned as the desired output. Moreover, since each case is a time-series, it can be divided into multiple cases, e.g., Case1 A: $t = 0\,s - 0.05\,s$, Case1 B: $t = 0.01\,s - 0.06\,s$ …etc. Thus, the final data size can be an integer multiple of the simulated case count. The implementation of the RNN-LSTM model was done in Python using TensorFlow [143] and all the simulations were carried out on a PC with an i7-7700

processor (3.6 GHz quad-core processor), 16 GB DDR4 RAM, 4 TB HDD, and no dedicated graphics card.

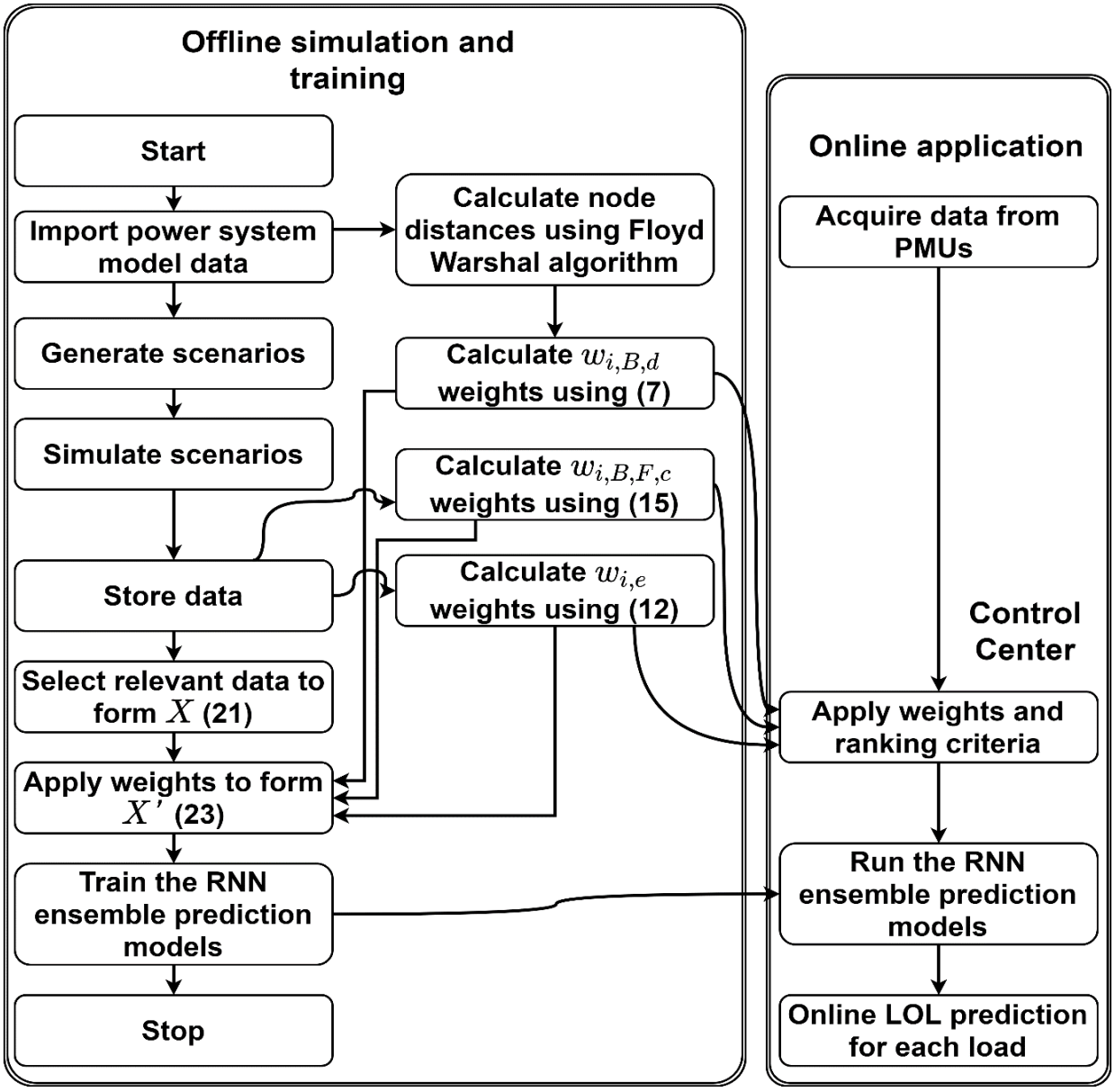

Figure 5.8. Implementation of the proposed method for predicting CF in power systems in real-time.

5.4.2 Rank index parameter tuning

The first case study aims to tune the model parameters, validate its efficacy, and observe its sensitivity against parameter tuning. Yet, tuning all parameters of the model to find the global

optimum point requires a dedicated optimizer with both RNN and big data in a loop, which is beyond the main scope of this research. An alternative and more realistic tuning approach is to separately tune the parameters of each part of the model and investigate the model's sensitivity to those parameters [144]. Furthermore, this approach can validate the efficacy of the model by showcasing how each portion of the model provides significant, but incomplete insights regarding the state of the power system under cascading events.

The tuning of the parameters is done on the IEEE 39-bus system. The case study includes 54,756 CF scenarios. The average computer time to process each scenario is around 14 s, while parallel processing was used to reduce the total time for simulating all the cases and training the RNNs by 75% (more significant reductions could be achieved using more logical processors or more PCs). Table 5.1 is a summary of all simulated cases and their outcomes; in all cases, the number of variables per time sample is 196.

Table 5.1. Summary of simulated cases for the IEEE 39-bus system.

Type	Number of Cases
Simulated cases	54,756
Total blackout	24,364
Load shedding or partial blackout	8,329
No blackout	22,063

5.4.2.1 Node distance index

Starting with the node distance index, finding the optimal values of the parameters α (attenuation factor) and d_{max} (max distance), is conducted by arbitrarily selecting bus 18 (the 8th load bus) in the IEEE 39-bus system, Figure 5.9, as the observed bus while varying α and d_{max}. Whilst forcing all other ranking indices to zero (i.e., $w_{i,B} = w_{i,B,d}$), limiting X in (4.6) to voltage only, and removing direct inputs (i.e., $|\Psi| = 0$). The RNN-LSTM model is built with six layers: input layer, three fully connected NN, LSTM RNN, and then a fully connected NN (Table 5.2). The accuracies of all the models ranged between 50 and 85% and a heat map of the obtained accuracies is depicted in Figure 5.10. The optimal values were found at $\alpha = 1.1$ and $d_{max} = [4,5]$.

However, since increasing d_{max} increases the complexity of the formulation, $d_{max} = 4$ is selected. Moreover, the accuracy of the model peaked at 85%, which indicates the distance index with voltage alone is not a sufficient indicator for cascading failures.

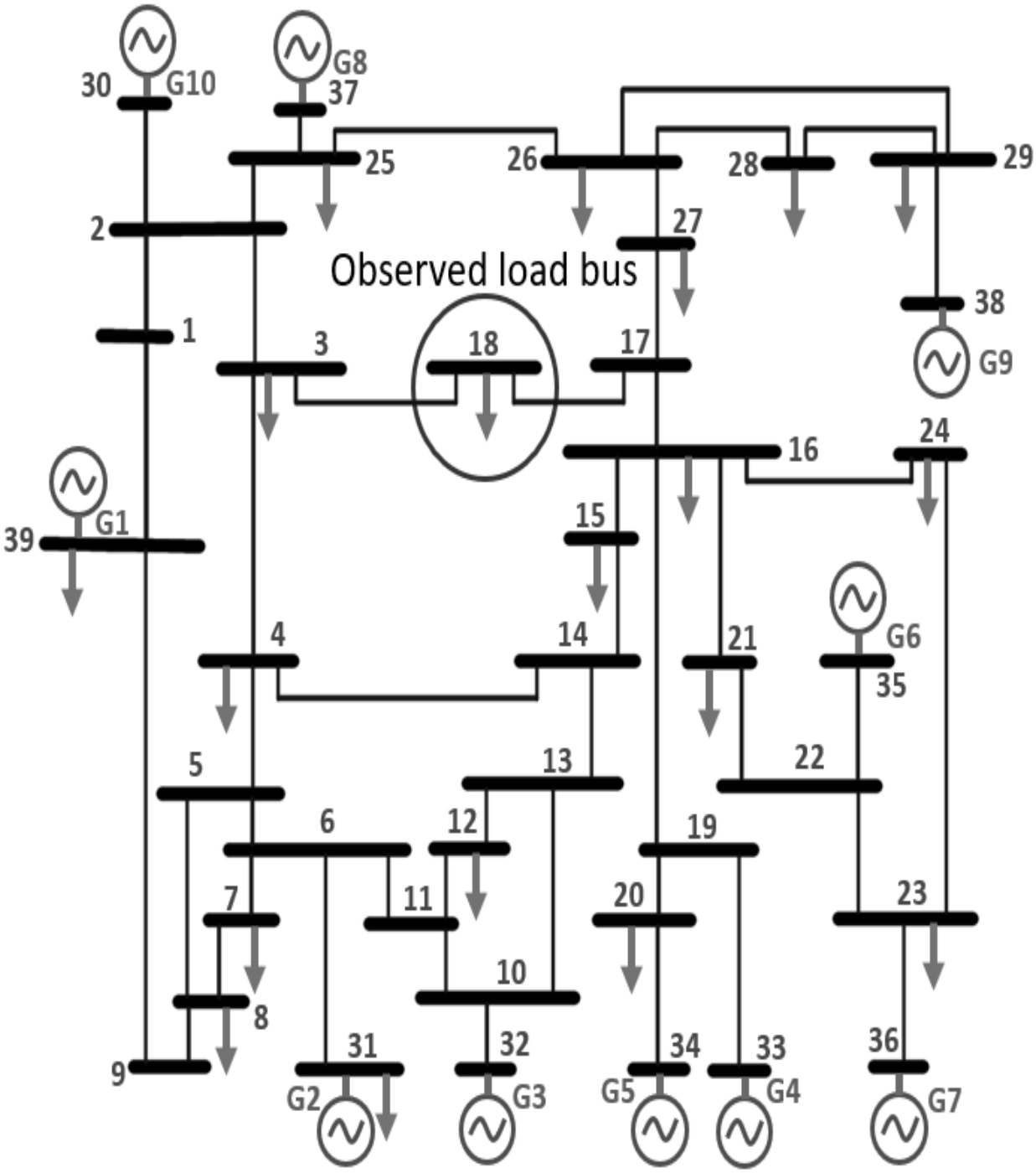

Figure 5.9. Single line diagram of IEEE 39-bus system

Table 5.2. RNN-LSTM network layer types and specifications.

Layer name and type	**Layer shape**	**Number of parameters**
Main input (Input layer)	(, 25, 36)	0
Reduction layer (Dense layer)	(, 25, 128)	4736
Reduction layer 2 (Dense layer)	(, 25, 64)	8256
Reduction layer 3 (Dense layer)	(, 25, 32)	2080
Main LSTM (LSTM layer)	(, 16)	3136
Main output (Dense layer)	(, 2)	34

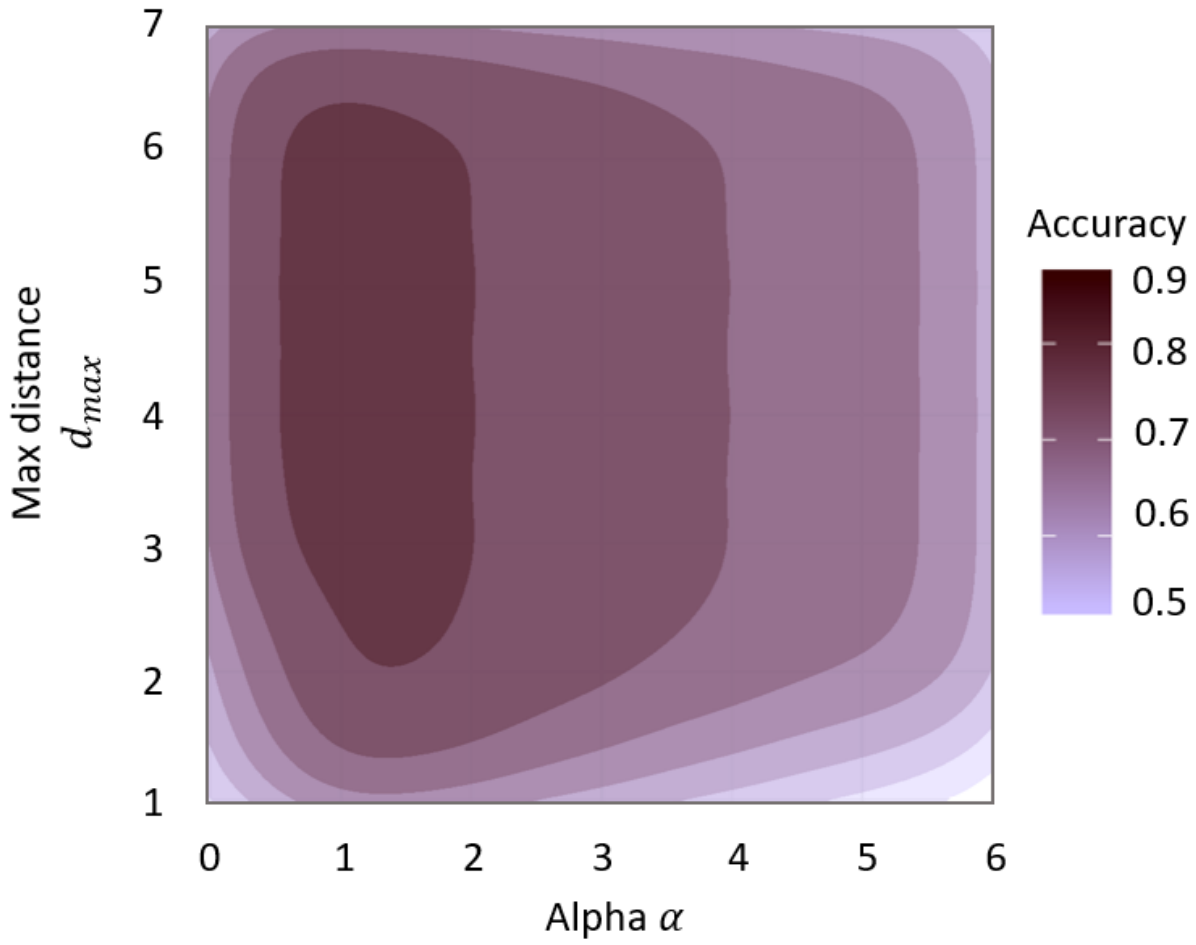

Figure 5.10. Effects of variation in α and d_{max} on the accuracy of the testing model (figure smoothed to account for RNN training variations).

5.4.2.2 Entropy index

For the entropy index, the only tunable parameter is the number of samples per CF scenario. A comparison between one and two samples per scenario is presented in Figure 5.11, which shows a negligible difference in the entropy index as the order of the buses (highest to lowest entropy) remains the same in both scenarios. Subsequently, the entropy index is used to train the RNN and predict CF with the rankings obtained from a single sample. However, unlike the distance index, $|\Psi| = 1$ was selected while eliminating the mean, variance, and higher moments. This is applied to measure the efficacy of the eliminated inputs as will be shown later. The prediction accuracy of the model under these specifications was 80%. Moreover, the accuracy of the entropy index is lower than the distance index which is expected because the distance index provides local information regarding the effect of CF whereas the entropy index uses the same bus ranking for all of its predictions.

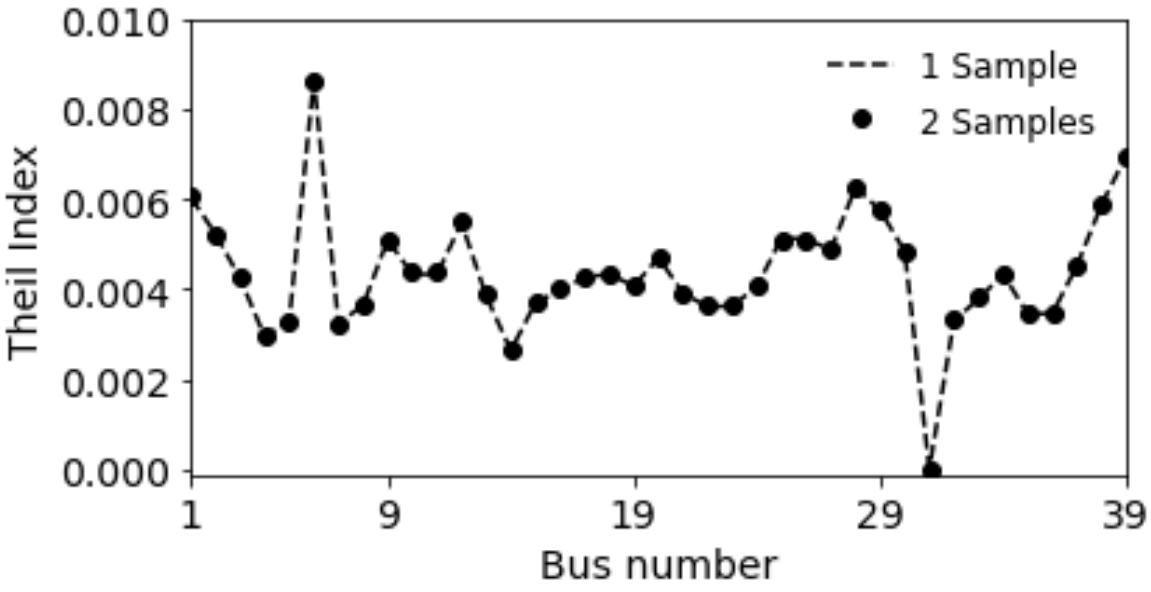

Figure 5.11. Theil index for one and two samples of the voltage.

5.4.2.3 Centrality index

The main issue with the centrality index is its huge computational burden compared to other indices. To elucidate, both the entropy and distance index are independent of fault location, and the entropy index is also independent of the observed bus. However, the centrality index is

dependent on both fault location and the observed bus. Moreover, it requires building a subgraph and calculating the centrality of each of its nodes. To mitigate this issue, the γ value can be tuned to limit the number of generated subgraphs as $d_{iB} > \gamma$ or $d_{iF} > \gamma$ would result in the trivial solution of an empty subgraph (i.e., the fault is far from the observed bus, hence $w_c = 0$ for all buses) or subgraphs with one node (i.e., only one bus, or a group of isolated buses are satisfying the criteria of being close to both the fault and the observed bus). The number of trivial solutions should be tuned properly as it is not desired to acquire this solution if the buses of interest are close to the fault location, but at the same time, trivial solutions are much faster to compute. Hence, γ is capped as follows:

$$\gamma = \frac{\max(d)}{r_f} \tag{5.26}$$

where:

$\max(d)$ is the maximum of the shortest distance between any two nodes, i.e., the maximum value from the Floyd Warshall algorithm, and r_f is a reduction factor.

Investigating the effects of r_f on the accuracy and trivial solutions, an RNN-LSTM based predictor with a similar setting as the distance index, but with $w_{i,B} = w_{i,B,F,c}$, is used to predict CF while varying the value of r_f. Figure 5.12 showcases the resulting comparison between the prediction accuracy and the percentage of trivial solutions in the dataset. The figure illustrates the value of r_f can significantly increase the trivial solutions without substantially affecting the accuracy as the accuracy is only marginally decreasing between r_f =3-11. However, an unexpected result is that the accuracy is lowest when there are no trivial solutions. This could be due to selecting the global central bus rather than the central bus according to the fault and observed bus locations. Finally, the maximum accuracy of the prediction model is also only 82% which indicates the centrality index is not a sufficient predictor by itself.

5.4.2.4 Correlation between parameters

To further showcase the importance of each element, the correlation between the parameters along with their correlation with the expected result is presented in Figure 5.13. The results indicate a very low correlation between the indices, meaning all of them can be helpful in the prediction. Because the low correlation indicates that each model is having a different set of

correct predictions. Moreover, the highest correlation with the desired solution was the distance index prediction that used the mean, variance, and higher moments, thus indicating the importance of statistical higher moments in the formulation.

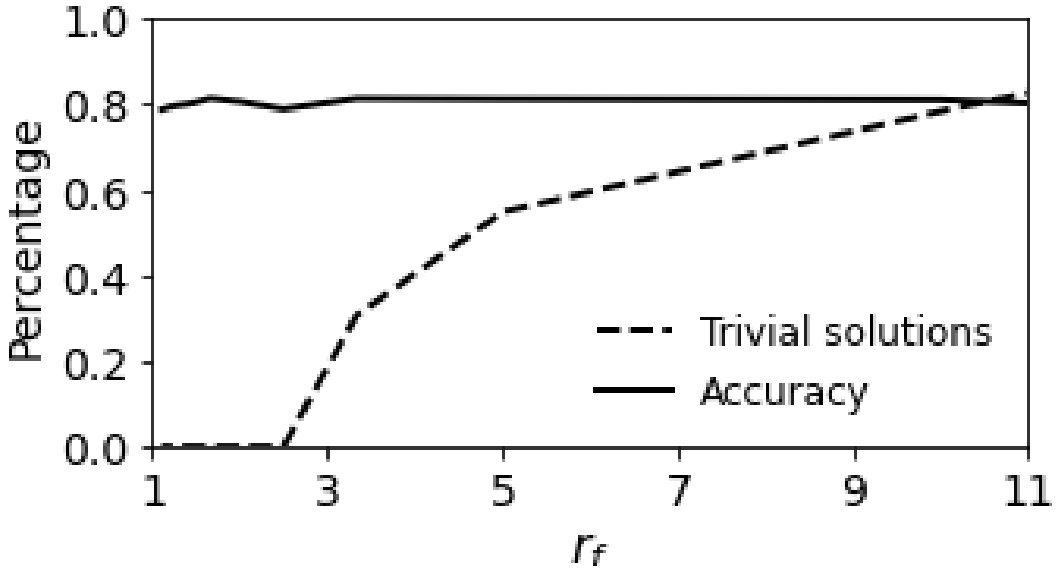

Figure 5.12. Prediction accuracy and trivial solutions as a function of r_f.

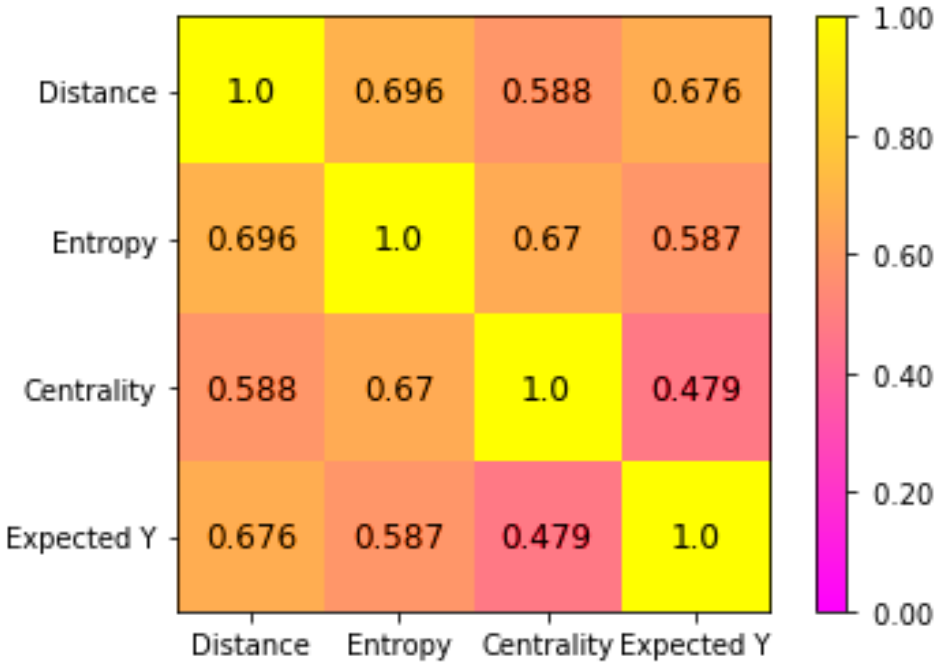

Figure 5.13. Correlation between parameters.

5.4.3 Input buses size $|\Psi|$

To merge all indices according to (5.16), they need to be adjusted according to a weighting function $\alpha(w)$. For brevity, $\alpha(w)$ is only used for normalization but can be used to give different weighting and importance to each index. Hence, the weighting is defined as:

$$w_{i_{new}} = \alpha(w_i) = \frac{w_i - \underline{w}}{\overline{w} - \underline{w}} \tag{5.27}$$

where:

$\overline{w}$ and $\underline{w}$ are the maximum and minimum values of a given index, respectively.

Finally, to evaluate the sensitivity and importance of the input size $|\Psi|$, the full model is tested on the IEEE 39-bus system and the results are depicted in Figure 5.14. In this figure, the number of buses $|\Psi|$ is varied from 0 to 25, which corresponds to varying the inputs from 8 to 58. Considering two features per bus (voltage magnitude and phase), and four probability indices per feature). As seen in the figure, the model accuracy increases from 83% to 91% when $|\Psi|$ is increased to 25. However, considering the range from 15 to 25, the accuracy remains almost constant (with variations due to model uncertainty). Hence, the additional inputs did not result in a significant increase in accuracy, which validates the claim that the most important buses are selected by the algorithm.

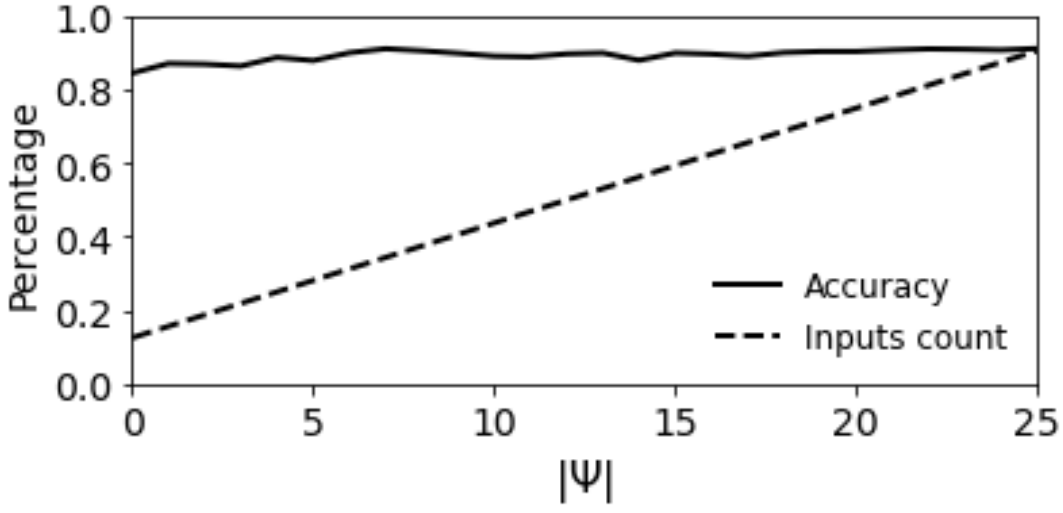

Figure 5.14. The number of inputs and model accuracy as a function of $|\Psi|$. (Input count is normalized by dividing it by its maximum value).

5.4.4 Time sensitivity analysis

One assumption in previous sections was that the number of time-samples remained constant to reduce the number of unknowns in the formulation. However, after acquiring the parameters for the indices and the appropriate values for α, d_{max}, r_f, and $|\Psi|$, the next step is to observe the sensitivity of the model to an increase in time samples. To evaluate the sensitivity, time-samples are assumed to be one cycle apart (i.e., $1/60\ s$ interval) where the first time-sample is taken at fault occurrence of the first exogenous tripping. As shown in Figure 5.15, the accuracy increases as the number of time-samples increases, but this will significantly lower the efficacy of the model as it introduces more delay. For instance, 25 samples correspond to a $0.417\ s$ delay with 90.5% accuracy, whereas three samples give an accuracy of 89.7% with only a $0.05\ s$ delay. Moreover, increasing the time does not always improve the accuracy as multiple high peaks are acquired at different time intervals. This behaviour of multiple peaks and drops is caused by multiple reasons. The variability within the RNN, because each model has a component of randomness, and the locations of peaks and drops within the data (i.e., whether the prediction model's last data is at a crest, trough, or equilibrium point). Hence, the ideal number of time samples can only be determined after training the model, which is 9 time-samples in this scenario considering the best trade-off between speed and accuracy, or 25 time-samples considering the maximum accuracy.

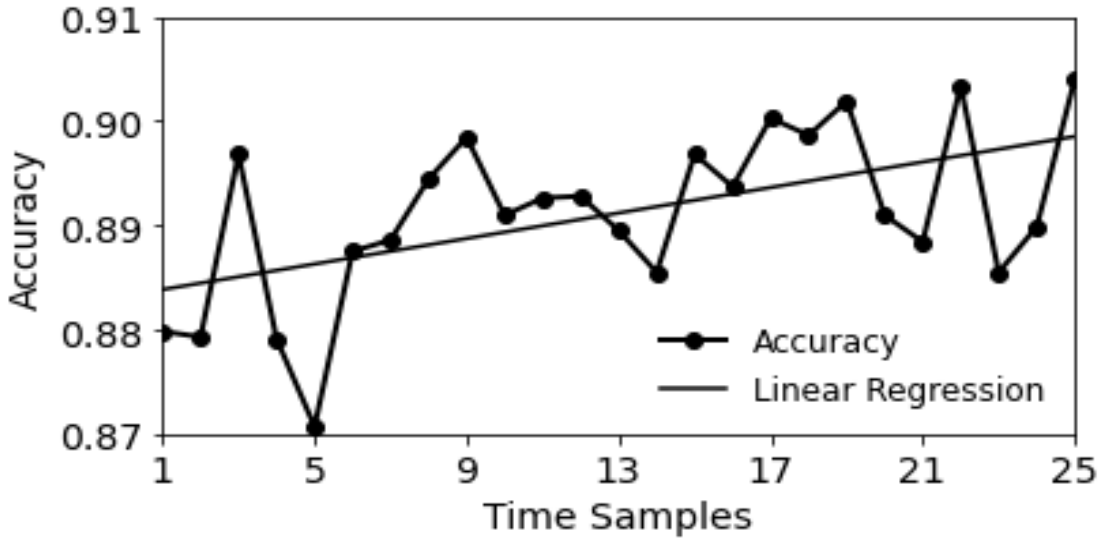

Figure 5.15. Model accuracy as a function of time samples.

5.4.5 IEEE 39-bus system

To fully test the proposed model, the RNN-LSTM ensemble is built as 19 RNN-LSTM models because the IEEE 39-bus system has 19 load buses. The average training time for each model is $50\,s/epoch$ and $10\,epochs$ on average are needed to reach maximum validation accuracy. Because models can be trained in parallel, the number of models will not affect the total training time if enough processors are used. The reduction in time in our setup was 75% due to the use of four processors, but further reductions could be reached with a maximum of 95% reduction in computing time using 19 or more processors. The LOL prediction models are built using the parameters described in Table II, which are the optimal values according to Subsections 5.4.2 to 5.4.4. Thus, this setup corresponds to the maximum accuracy of the corresponding indices. The accuracy of detecting the LOL for each bus for the testing set using these parameters was 94.5% on average and ranged from 94.1 to 94.9%.

Table 5.3. RNN-LSTM ensemble parameters.

Parameter	Value	Parameter	Value
α	1.1	r_f	10
d_{max}	4	$\|\Psi\|$	12
$t\ (for\ entropy)$	1	$time\ samples$	12

Furthermore, to showcase the efficacy of the RNN-LSTM framework in terms of improving its prediction accuracy as time progresses, the testing set accuracy of the model for bus 18 as it receives more data is shown in Figure 5.16. Only one bus is selected for brevity, but the other models also exhibit the same trend. The figure shows the model starts with an 87% prediction accuracy at fault occurrence, and as it receives more data the accuracy increases until it saturates at 94.2% after 22 cycles ($0.366\,s$). In other words, the model informs the operators whether a particular fault, regardless of how close or far it is from bus 18, is going to cause any load shedding or LOL at bus 18 or not with an initial accuracy of 87% at fault occurrence, but as time progresses it improves until it peaks at 94.2% within $0.360\,s$.

To illustrate this improvement of prediction, in one of the cases, the system is assumed to be at stress by setting the loads 20% higher than the nominal value (e.g., an unexpected heatwave or cold snap causing increased electrical consumption). Then, a fault causes the line between buses 4 and 5 to trip. For this scenario, the system stabilizes within $10s$ after the trip. However, after the initial fault, the line connecting buses 13 and 14 endogenously trips due to overcurrent, which causes the system to enter a transient instability that leads to a series of instabilities ending in a total blackout. Under this case, the proposed model could detect that bus 18 will suffer a loss of load within $0.15\ s$ after the first (exogenous) fault as shown in Figure 5.17. Hence, the model predicted the LOL even before line 13-14 trips. Moreover, it is essential to note that this prediction is for any curtailment in general. Thus, if a contingency would lead to a partial blackout that includes bus 18, or load shedding of bus 18 was needed as part of undervoltage or underfrequency load shedding schemes the prediction model will also issue a LOL prediction. Hence, the operators (or emergency controllers) have ample time to take corrective actions to ensure that bus 18, or the whole grid, if possible, does not suffer a LOL.

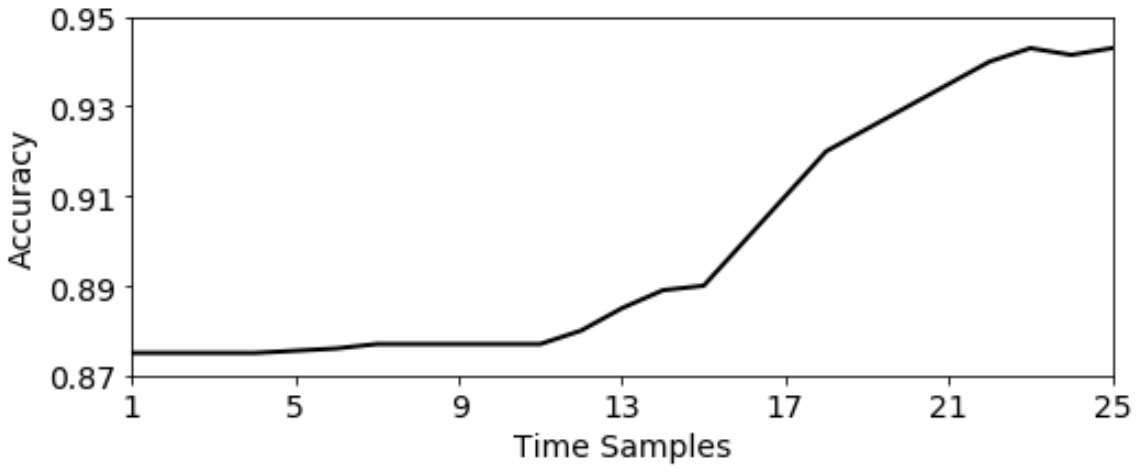

Figure 5.16. Model accuracy for the IEEE 39-bus system for bus 18 (based on the testing set) as the model receives more time samples.

5.4.6 Missing PMU data

In previous cases, PMUs are assumed to be ideal. However, in practice, many PMU non-idealities occur. A major non-ideality during CFs is missing PMU measurements because buses become unobservable due to line outages. To test the robustness of the proposed model to missing

PMU data, the PMU data is modified such that 1% of measurements are missing. Figure 5.18 compares the case with missing PMUs with the normal case and shows that the model is resilient to missing measurements especially as more time samples are obtained since both cases have similar accuracy trends especially from 14 time-samples and beyond. The main reason for this high resiliency is due to the dependency of the model on a smaller set of direct measurements compared to using all available measurements. Hence, there is a lower probability that a missing PMU measurement is needed.

5.4.7 Noisy PMU data

The second type of non-ideality considered in this model is noisy measurements. The noise is first modeled as a 2% total vector error (TVE) and then as a 10% TVE and the prediction accuracy of the model under these cases is shown in Figure 5.19. With a low TVE, the model is resilient and shows similar trends as the missing PMUs case. However, with a 10% TVE the prediction model performs poorly with almost a 9% reduction in accuracy at the first time-sample that shrinks to 5% with 24 time-samples. Although PMUs are typically designed to have a TVE

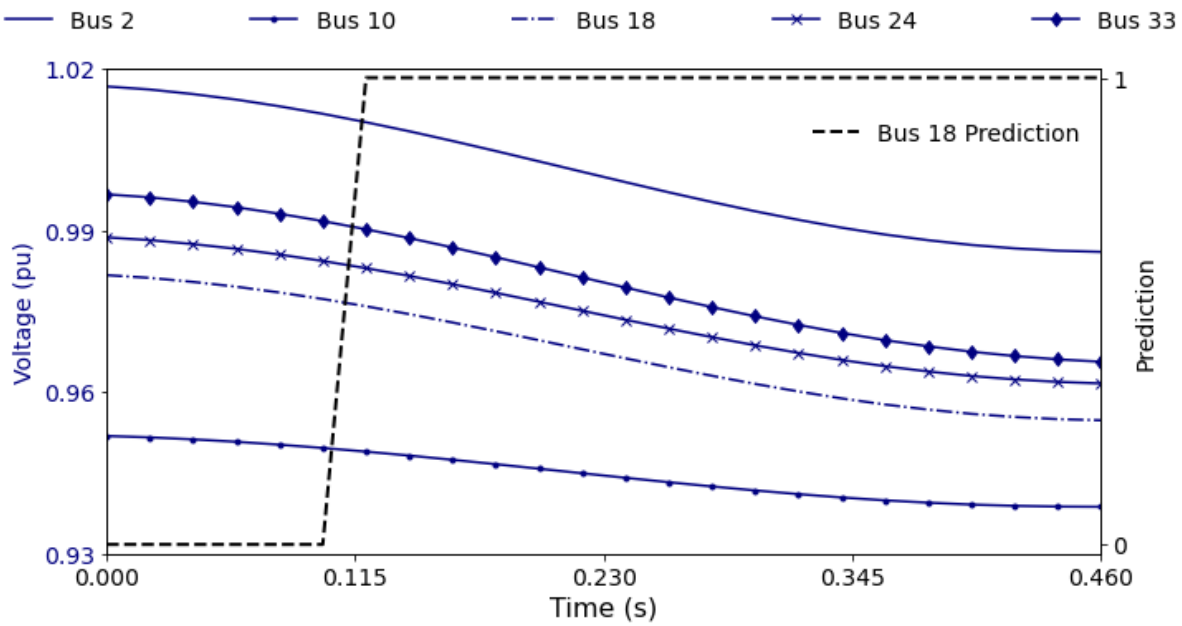

Figure 5.17. Voltages of select buses of the IEEE 39-bus system (voltages at other buses show a similar trend) and the prediction model for bus 18 following a trip in line 4-5 (Logic 1 indicates that the bus would suffer load curtailment, 0 otherwise).

$\leq 1\%$, which the model handles well, this small TVE only applies to steady state TVE. On the other hand, TVE during transient (i.e., first $100\ ms$ after a tripping or a failure) can reach up to 10% in some major disturbances [145]. Hence, depending on the PMUs' transient TVE characteristics the current prediction model might have a significant drop in accuracy between transient and steady states.

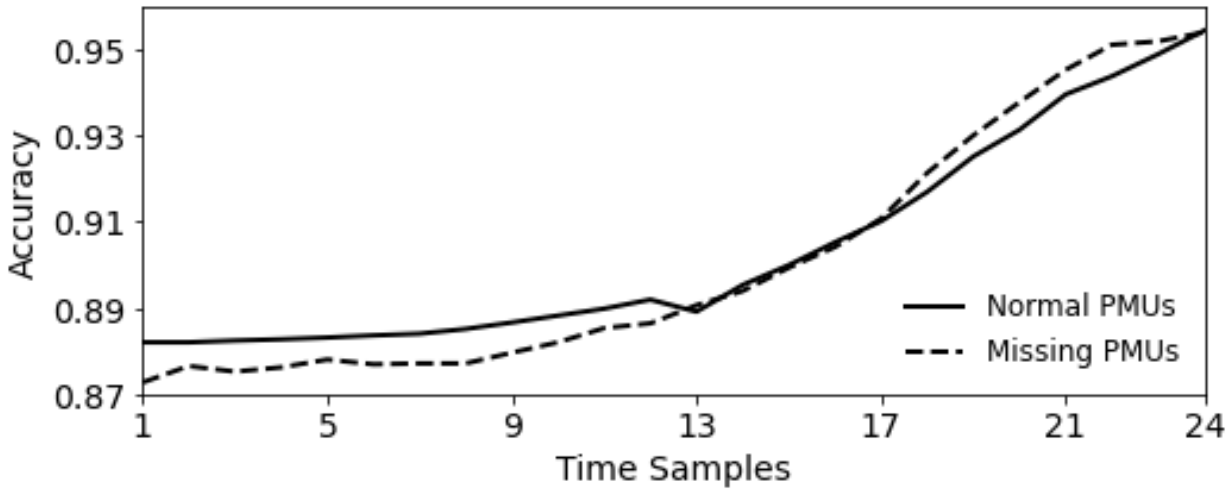

Figure 5.18. Comparison between the accuracy under normal conditions and with missing PMU data.

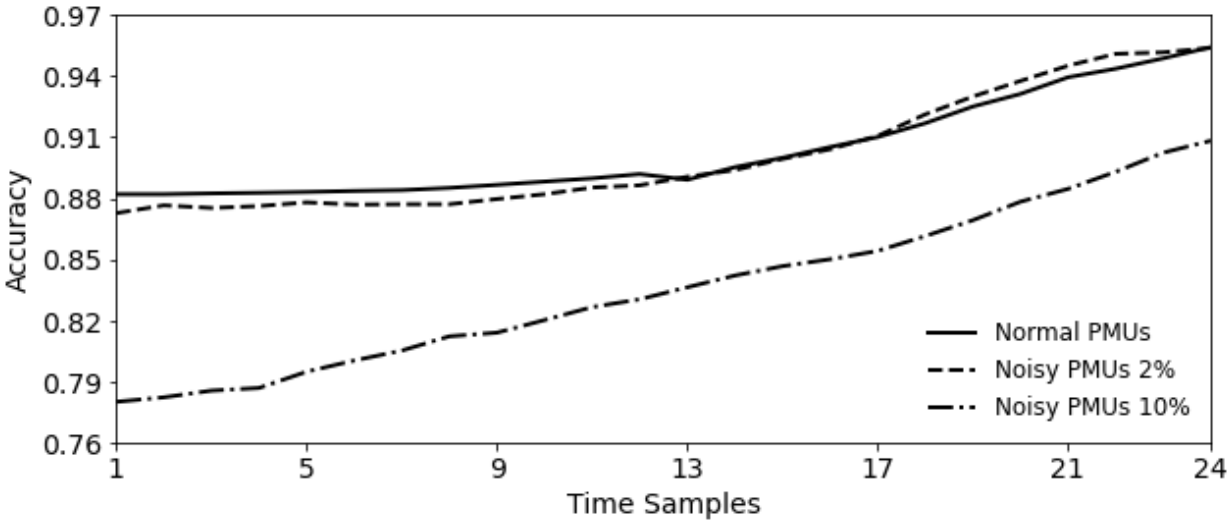

Figure 5.19. Comparison between the prediction accuracy under normal conditions and with noisy PMU data.

5.4.8 POL 2383-bus system

The POL 2383-bus system is used to validate the model's efficacy in a large, interconnected network. The simulated CF scenarios include 15,000 cases with an average computer time of 170 s per case. The number of variables per time sample for all cases is 9951. Table 5.4 shows a summary describing all the simulated cases. Note that the number of POL 2383-bus total blackout cases is negligible compared to the IEEE 39-bus system because the number of applied contingencies (ranging from N-1 to N-5) are not severe enough to cause many total blackouts in this large interconnected system. The RNN-LSTM ensemble is built as 23 models of the selected load buses (among more than 1800 load points). These buses were selected so the distance between them is maximized to cover different areas of the POL 2383-bus system. Also, each model is built using the same specifications as the IEEE 39-bus system.

Table 5.4. Summary of simulated cases for the POL 2383-bus system.

Type	Number of cases
Simulated cases	15000
Total blackout	39
Load shedding or partial blackout	6505
No blackout	8456

The LOL prediction model for these selected load buses had a 99.8% accuracy of detecting the loss of load on average and ranged from 99.5 to 100% as shown in Figure 5.20. The error is very narrow since POL 2383-bus is more decoupled than the IEEE 39-bus system. Thus, a localized predictor can accurately describe the state of the bus. Besides, since most blackouts are only partial, then from a load perspective, the probability of having a LOL is significantly lower. However, a valid concern would be that the predictor has high accuracy because the probability of LOL is much lower. To address this concern, the average of the extended confusion matrix for all the buses of interest is shown in Table 5.5, where 0 indicates no LOL (negative) and 1 indicates that there is a LOL (positive). From the table, it is observed that the false positives are orders of magnitude lower than the true negatives. Similarly, the false negatives with the true positives show a similar trend. Hence, the model is indeed accurate, and it is not biased towards the negative

result. As for the recall, it is only 89.1%, but with all CFP literature ignoring recall it is not possible to compare the results of this model with other models. Hence, this work can be the base for comparison for future CFP works.

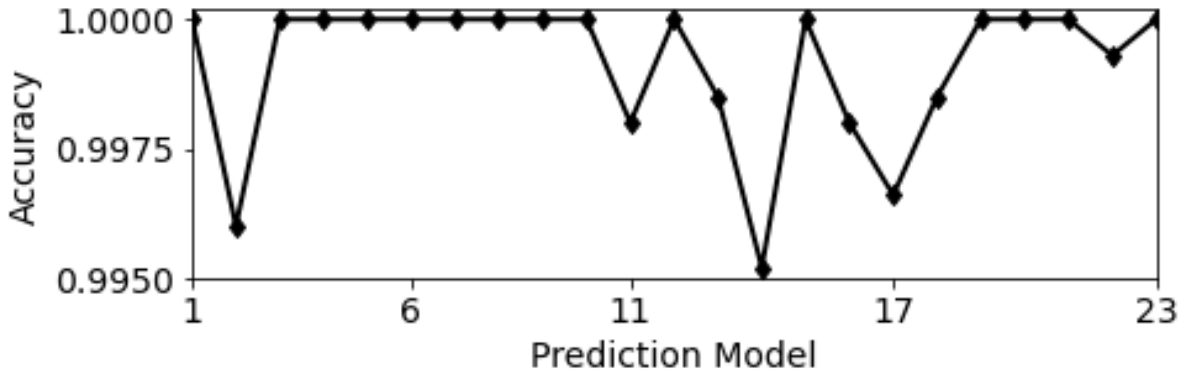

Figure 5.20. Prediction model accuracy for selected buses of the POL 2383-bus system.

Table 5.5. Average of the extended confusion matrices of the POL 2383-bus LOL Prediction models.

Type	Actual 0	Actual 1	Sum	Extended metrics
Predicted 0	99.464%	0.055%	99.52%	False omission rate = 0.0005%
Predicted 1	0.008%	0.472%	0.48%	Recall (Sensitivity) = 89.1%
Sum	99.47%	0.53%	100%	F1-score = 93.47%
Extended metrics	False positive rate = 0.0001%	Precision = 98.3%	Accuracy = 99.93%	

5.4.9 Comparison with existing work

One of the main contributions of the proposed method is its single bus prediction accuracy using limited inputs. As such, it is difficult to compare it with existing methods, because no other published work considered this aspect. However, one way to provide a meaningful comparison is to expand the model to the maximum number of buses and predict the expected blackout percentage in a similar fashion to existing work. Table 5.6 shows this comparison, where the proposed RNN-LSTM ensemble model (LSTM-e) is modified in two ways. Firstly, the model outputs ‘true’ if any single bus is having a LOL (LSTM-e Boolean), to compare it with [120] and [116]. Secondly, the model is set to output the expected blackout size out of four intervals (LSTM-

e Brackets) to compare it with [91]. It is evident from the table that the proposed method is more versatile and more accurate than existing CFP methods as it has dynamics simulation with higher or equal accuracy and a shorter delay.

Table 5.6. Models Comparison Summary.

Method	Notes	Predictor	ref	Acc.
BPDT	Dynamic, 500 ms delay, multiple topologies	Blackout% (4 brackets)	[91]	97.97%
SVM-PF	Static line trip, single topology	Blackout Boolean	[120]	100.0%*
FFNN	Static line trip, single topology	CF Boolean	[116]	99.95%
LSTM-e Boolean	Dynamic, 100-400 ms delay, multiple topologies	Blackout Boolean	-	99.99%
LSTM-e Brackets	Dynamic, 100-400 ms delay, multiple topologies	Blackout% (4 brackets)	-	99.99%

* With a wide range of 'indecisive region'

5.5 Summary

In this chapter, a new model is proposed to predict cascading failure in real-time. The method reduces the problem of cascading failure to a set of parallel loss of load predictors using the correlation between data points and their influence on the CF. The correlation and influence of the buses are extracted through graph theory, Theil index, eigenvector centrality, and skewed probability distributions to maximize information entropy while minimizing the number of inputs. Moreover, RNN LSTM cells are used to allow the model to produce self-improving results with each additional information as time progresses. The results indicate that the model successfully addresses multiple concerns in the literature, including versatility, faster response, fewer inputs, scalability to large systems, and the ability to detect the location of the buses affected by the cascade. In conclusion, this model is expected to be applicable in practical applications, either as a remedial action scheme input or a visualization tool for the operator.

6 FRAMEWORK FOR A REAL-TIME AUTONOMOUS CASCADING FAILURE PREDICTION MODEL

6.1 Introduction

To predict CFs, and by extension blackouts, multiple CF prediction algorithms and frameworks were proposed in the literature [40], [91], [114]–[120], [146]. However, in all these approaches, the modeling is divided into two distinct phases: the offline training and the online predicting as illustrated in Figure 6.1. However, this approach limits the potential of these models since the grid may change significantly as time progresses. Also, synthetic data might not be as accurate as real data. Hence, a better approach is to use adaptive or autonomous models [147]. In this case, the model constantly updates its prediction accuracy and adapts to the changes in the grid. This updated structure is demonstrated in Figure 6.2 where prediction results along with the PMU data are aggregated into a new prediction dataset that is then used to retrain the model along with the existing simulated scenarios.

Sampling from this new set of mixed synthetic and real datapoints can introduce multiple issues. It can generate a biased dataset, because power grids are by design highly reliable which leads to real datapoints having very few positive cases (i.e., cases with contingencies). Moreover, the mixed set can include contradictory datapoints if a synthetic result mismatches the outcome of a real result. To address the first issue, importance sampling (IMS) can be used to balance the sets [110]. In IMS, instead of sampling from the original distribution of the data, the datapoints are sampled from a new distribution according to their importance, which can be set so that the bias is minimized. Alternatively, the model training can be modified to give more weighting to the sparse classes using weighted regularization [111]. However, IMS has the advantage of keeping the original model unchanged. Hence it is less invasive to existing works. As for the second issue, a validation loop based on case-based reasoning (CBR) can be added to detect and remove any contradictory synthetic data [148]. Moreover, it can be used to generate synthetic data to validate the model by comparing the generated data with existing real data. Thus, verifying the efficacy of the simulation tool used to generate the data. Furthermore, CBR's 4R cycle (Retrieve, Reuse, Revise, Retain) can aid in the formulation of the autonomous prediction model [149]. Hence, both IMS and CBR provide important additions and fixes to the proposed autonomous model.

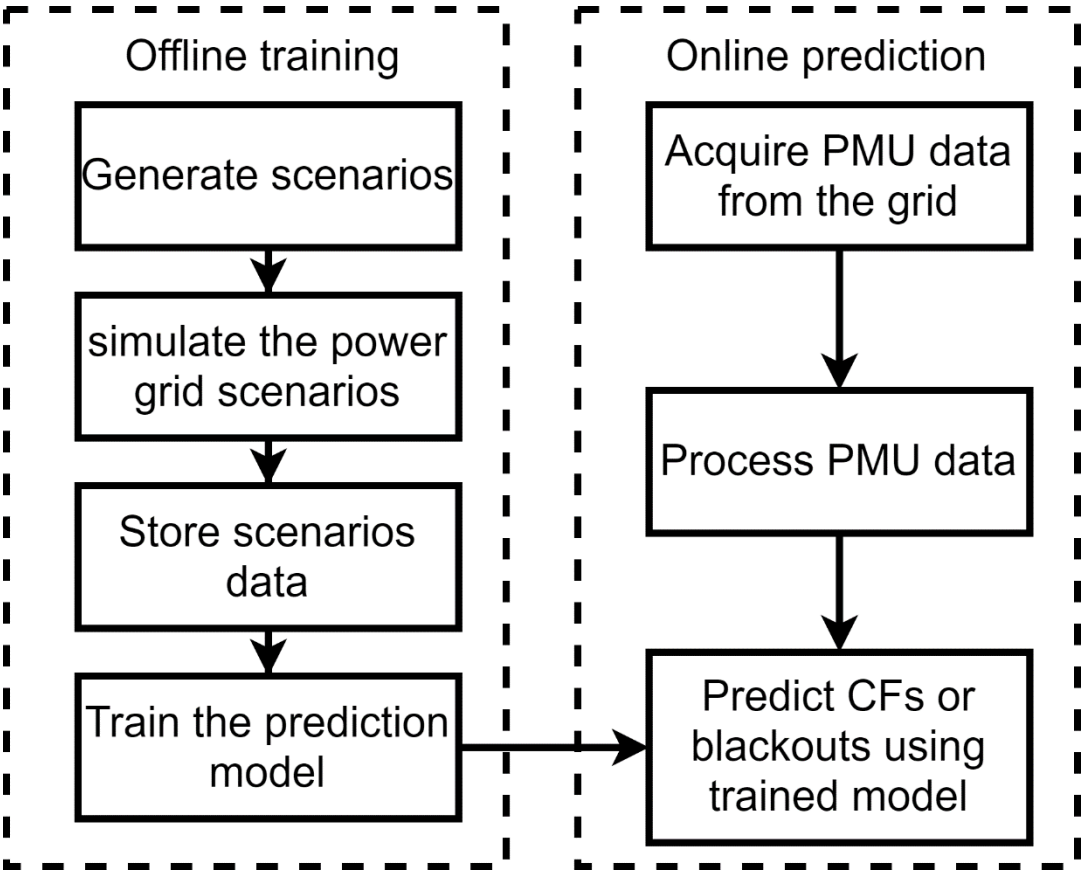

Figure 6.1. Typical flowchart of a real-time CF prediction model.

Based on the previous limitations and insights, a framework for an adaptive real-time CF prediction model is introduced in this chapter. The framework uses IMS to address the biased samples issues and to select the most important samples for the training of the model. Moreover, it includes a validation loop based on CBR that corrects contradictions in the dataset between real data and existing synthetic data. In addition, it allows for the validation of the simulation tool used to generate the synthetic data. Therefore, this framework significantly enhances existing CF prediction models and can easily be incorporated in prediction models in other power engineering areas.

The remainder of the chapter is organized as follows. Section 6.2 describes the adaptive prediction paradigm and how to implement it in a CF prediction framework. Section 6.3 introduces IMS and its application to address the skewness in the data. The results and discussions are in Section 6.4. Finally, Section 6.5 concludes the chapter.

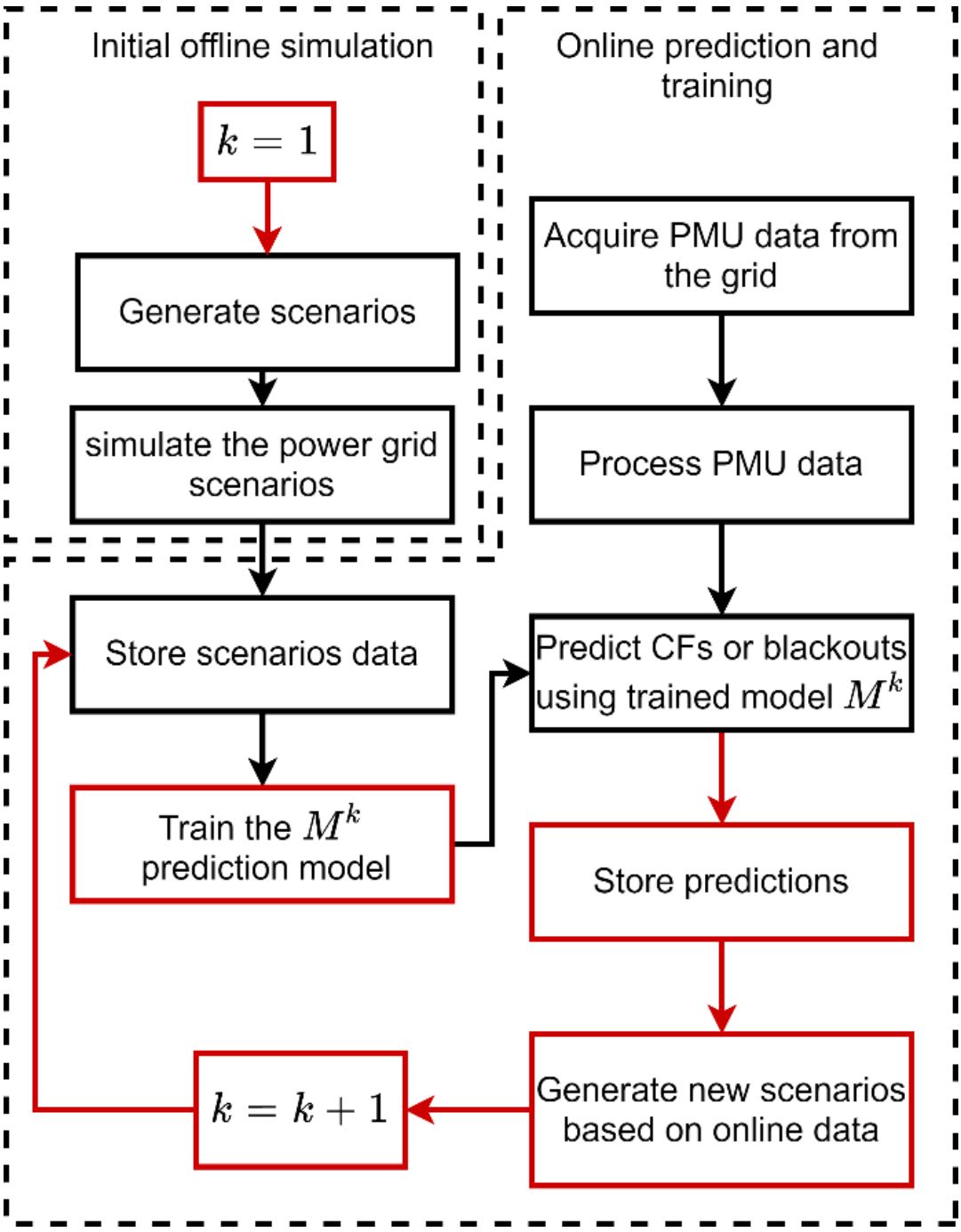

Figure 6.2. Proposed flowchart for the adaptive real-time CF prediction model.

6.2 Adaptive Cascading Failure Prediction

Power grids produce a large amount of data through PMUs, smart meters, and other devices at each second. This data should be utilized in prediction models to enhance their prediction. However, in conventional prediction models, models are only trained when considering a different

dispatch or when a different grid topology is introduced (depending on the purpose of the model). An alternative and more efficient way to retrain the model and utilize existing data is to have the model dynamically adapt to newly acquired data by incorporating the data within a learning loop.

To build this autonomous learning loop the following components need to be defined: update scheduler, dataset selector, update acceptor, and data validator components. Collectively, these four components formulate the 4R cycle of the CBR as update scheduler handles retaining, dataset selector handles retrieving, update acceptor handles reusing, and data validator handles case revising. The structure of CBR and its relation to the proposed autonomous learning loop is shown in Figure 6.3.

The update scheduler handles the amount of time between each iteration k of the model's retraining. Ideally, this time should be the same as the rate of the measurement devices (e.g., $1/60\ s$ in the case of PMUs). However, practically this is impossible, because the update time is restricted by the training speed of the model. Hence, the update time would be the maximum between the desired update time and the training time. Alternatively, the training can be interrupted if it exceeds a predefined maximum allowed time. Therefore, the training time for the update scheduler can be defined as:

$$t_{up} = \min\left(\max(t_{tr}, t_{de}), t_{mx}\right) \tag{6.1}$$

where:

t_{up} is the actual update time,

t_{tr} is the training time,

t_{de} is the desired update time,

t_{mx} is the maximum allowed time for training.

The second component is the dataset selector, and it is responsible for defining the subset of data used to retrain the model. Considering Figure 6.2, the available data to construct the retraining dataset can be classified into three sets: synthetic data, historical data that were used to train previous prediction models, and newly stored historical data that were not used to train any model yet. Depending on the application of the prediction model, a different ratio between these three sets can be selected. Hence, a general formula for the retraining dataset is:

$$d_t = \{d_s, d_h, d_n\}, d_s \subseteq d_S, d_h \subseteq d_H, d_n \subseteq d_N \quad (6.2)$$

where:

d_t is the retraining dataset,

d_s, d_h, and d_n are the subsets used in the retraining from synthetic data, historical data, and new data, respectively.

d_S, d_H, and d_N are the full set of synthetic data, historical data, and new data, respectively.

Since tuning the sizes of the subsets in d_t is a nontrivial problem, IMS can be used to select the optimal set of datapoints, which is discussed in Section 6.3.

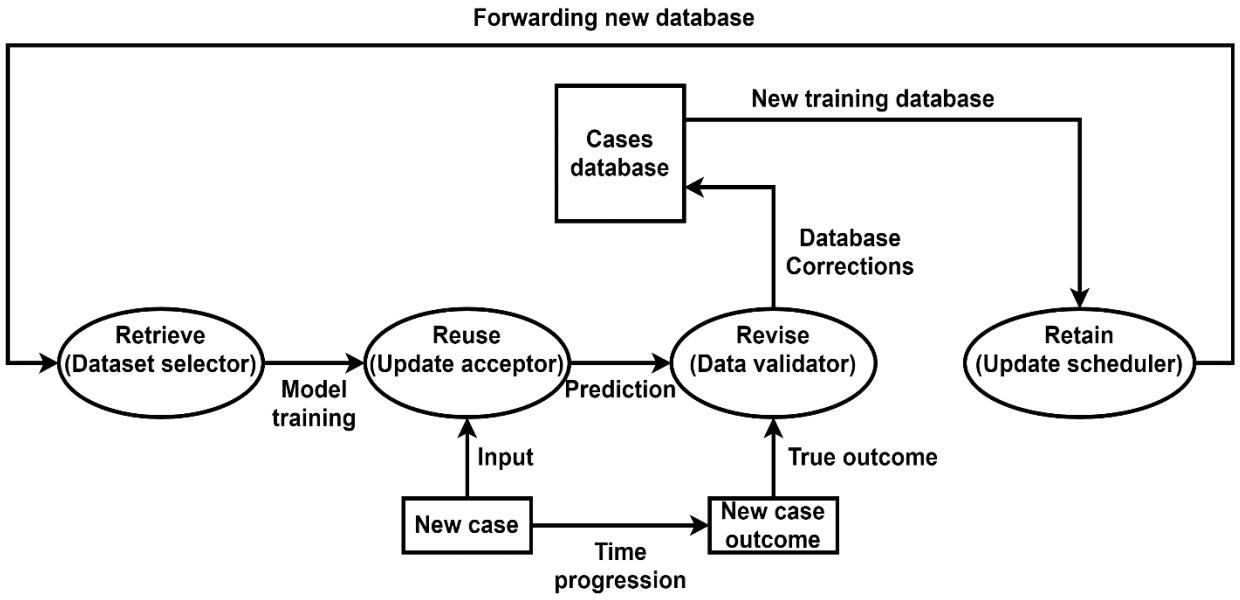

Figure 6.3. Case-based reasoning (CBR) 4R cycle (Retrieve, Reuse, Revise, and Retain) according to the proposed learning loop components.

The third component is the update acceptor, which is used for accepting or rejecting updates. It also maintains the accuracy of the model by either rejecting an update, repeating it, or changing its parameters if the resulting trained model is below a certain acceptance criterion. Typically, this acceptance criterion is the model accuracy, but precision, F1 score, or other criteria can be used as well. Alternatively, this component could just issue a warning to the operator following a violation. Since there are multiple ways to formulate this component and each

approach requires different parameters, it is harder to define a general update acceptor component. Nonetheless, a general accuracy-based acceptor with warning can be represented as:

$$f_{rw} = \text{acc}(M^k) < acc_{Th} \tag{6.3}$$

$$M = \begin{cases} M^k, & f_{rw} = 0 \\ M^{k-1}, & f_{rw} = 1 \end{cases} \tag{6.4}$$

where:

f_{rw} is the internal Boolean flag for rejecting updates and issuing warnings,

k is the model update counter,

$\text{acc}(M^k)$ is the accuracy of model M^k,

acc_{Th} is the threshold accuracy to accept the model.

Finally, the last component, the validator, compares the newly acquired data at each loop with the existing ones and deletes any contradictory datapoints (i.e., two points with the same inputs but different outputs). In addition, for increased robustness of the model, the validator can also generate new synthetic data to compare it with existing historical data that have the same input to validate the correctness of the simulation tool. Thus, the validator is formulated as a case generation and comparison loop that produces a Boolean flag f_{mw} to denote a mismatch warning, where $f_{mw} = 1$ if any contradictory predictions are detected.

Based on (6.1)-(6.4), the adaptive autonomous real-time CF prediction model shown in Figure 6.2 can be extended as shown in Figure 6.4. Initially, the model counter k is initialized to 1 and an initial set of synthetic data is generated using the desired simulation tool depending on the required specifications [25]–[27]. This ends the initial offline phase of the model. Then, these simulated scenarios are used to create a new subset according to (6.2). However, since there are no other datasets yet, $d_t = d_S$ can be used. After that, the model is trained, but if the maximum time in (6.1) is violated the training will be abruptly stopped. Regardless of whether the training ended prematurely or not, the current iteration of the training will be considered as the M^k model. Hence, the next step is to measure f_{rw} to decide if the model is accepted or rejected. Next, the model is updated ($M = M^k$) if its accuracy exceeds the required threshold or kept as it is if the accuracy of the new model is low ($M = M^{k-1}$). The first iteration cannot fail since there are no

models before it. The chosen model is then deployed as the prediction model, and the PMU data are processed according to the prediction model requirements and then passed to produce the prediction. The prediction and the inputs are then stored. After the actual outcome of the grid becomes known, the input data with the true outcome form a new scenario for the model. These new scenarios are then matched with the existing scenarios to find and delete any contradictory datapoints (issuing a warning if any are discovered). Then, the model starts a new iteration of training with the new scenarios added to the existing ones. Additionally, since the update time is much slower than the prediction time, the portion of the online model labeled as prediction loop (the dotted area) will run multiple times before the model is trained and k is incremented (i.e., the other portion of the online model). Finally, for the compactness of the flowchart, the validator's synthetic cases generation loop is omitted from the figure.

6.3 Sampling Method

In a prediction model with significantly skewed classifiers, e.g., failure prediction in a highly reliable system, the goal of modifying the sampling method is to maintain a higher ratio of failure cases to normal operation cases. This can be achieved by either reducing the number of normal operation cases or sampling more times from the pool of failure cases. However, before applying any of these techniques, the data needs to be analyzed to check if there is any biasing in the dataset.

Considering the datasets used in the proposed CF prediction framework, the following observations regarding synthetic data, historical data, and new data can be made. For the synthetic data d_S, the size of the dataset is $|d_S|$, with P_{d_s} as the ratio of normal operation cases, and $1 - P_{d_S}$ for the failure cases ratio. However, $P_{d_s} \cong 1 - P_{d_S} \cong 0.5$ in well-designed synthetic data because equal portions between positive and negative cases is typically a design requirement. As for historical data d_H, the dataset size is $|d_H|$, the ratio of normal operation cases is P_{d_H}, and the ratio of failure cases is $1 - P_{d_H}$. However, $P_{d_H} \gg 1 - P_{d_H}$ since power systems are highly reliable, then $P_{d_H} \cong 1$. As for new data, they follow the same ratio as historical data, but with a much smaller dataset since $|d_N| \ll |d_H|$ and $|d_N| \ll |d_S|$. Thus, $|d_t| = |d_S| + |d_H| + |d_N| \cong |d_S| + |d_H|$. Hence, if a sample X is randomly drawn from the pool of datapoints, it will have the following expectation:

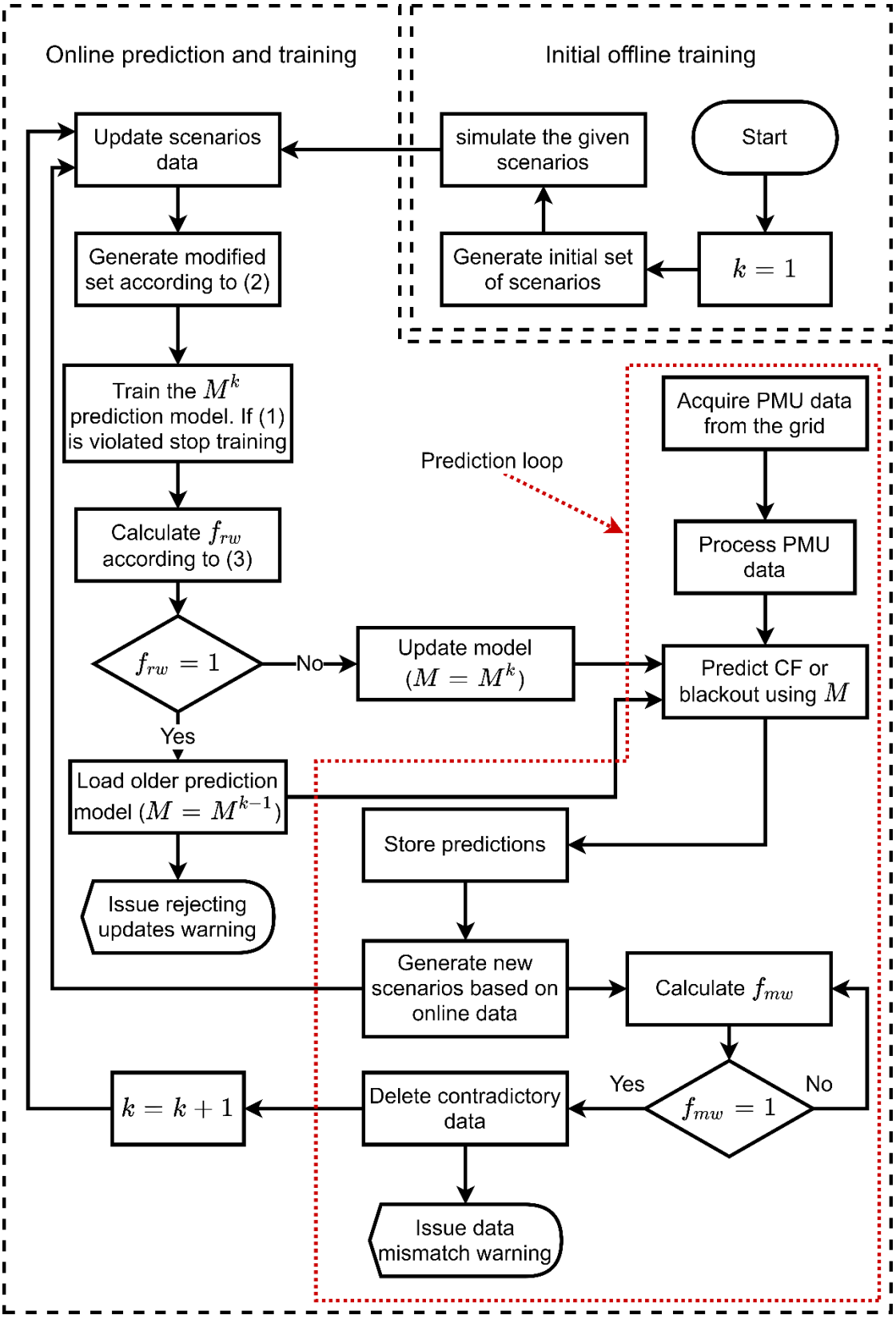

Figure 6.4. Expanded adaptive CF prediction model.

$$\mathrm{E}(X) = \frac{|d_S| P_{d_s} + (|d_H| + |d_N|) P_{d_H}}{|d_S| + |d_H| + |d_N|} \tag{6.5}$$

Simplifying the terms using the approximations provided:

$$\mathrm{E}(X) \cong \frac{|d_S| * 0.5 \ + |d_H|}{|d_S| + |d_H|} = 0.5 \frac{|d_S|}{|d_S| + |d_H|} + \frac{|d_H|}{|d_S| + |d_H|} \tag{6.6}$$

Hence, as the simulation begins, $|d_S| \gg |d_H|$ and thus the expectation is close to 0.5, which is desired as prediction models typically provide better results with unbiased data. However, as the simulation continues and $|d_H|$ gets bigger, the ratio approaches 1. Thus, a very skewed dataset is formulated where most of the cases are normal operation points. Therefore, there is bias in the dataset, but it happens as a function of time.

To solve the issue of the skewed dataset, IMS can be used to ensure that the sampled data samples more from the "important samples", which in the case of the proposed model is the sampling that makes the expectation closer to 0.5. IMS achieves this goal by sampling from a new distribution function g. Ideally, the aim of g is to minimize the variance in the scaled sample [150]:

$$\mathrm{g}^*(X) = \min_{\mathrm{g}} \mathrm{var}_{\mathrm{g}}(X \frac{\mathrm{f}(X)}{\mathrm{g}(X)}) \tag{6.7}$$

which can be achieved by:

$$\mathrm{g}^*(X) = \frac{|X| \mathrm{f}(X)}{\int |x| \mathrm{f}(x) dx} \tag{6.8}$$

Considering the given datapoints and the target of $\mathrm{E}(X) = 0.5$, g can be defined as the following. Given the original sampling method, the drawn sample can be defined as:

$$x = \begin{cases} x_S, & p = \dfrac{|d_S|}{|d_N| + |d_S| + |d_H|} \\ x_H, & p = \dfrac{|d_H|}{|d_N| + |d_S| + |d_H|} \\ x_N, & p = \dfrac{|d_N|}{|d_N| + |d_S| + |d_H|} \end{cases} \tag{6.9}$$

where:

x is the drawn sample,

x_s, x_H, and x_n refer to a sample drawn from synthetic, historical, and new data, respectively,

p is the probability of drawing from that particular sample.

Then, the modified $g(x)$ to maintain $E(x) = 0.5$ would modify the sampling into:

$$x = \begin{cases} x_S, & p = 1 \\ x_H, & p = 0 \\ x_N, & p = 0 \end{cases} \tag{6.10}$$

However, this function does not consider the "importance" of real data (both historical and new). Hence, one solution to solve this issue would be to set $E(X) = \alpha$ where $\alpha > 0.5$. Additionally, since d_N is small and contains data not previously used in training, it should be included in the training dataset without sampling (i.e., $d_n = d_N$). Thus, the new sampling function to maintain the new expectation ($E(X) = \alpha$) is:

$$x = \begin{cases} x_S, & p = \dfrac{\alpha - P_{d_H}}{P_{d_S} - P_{d_H}} \\ x_H, & p = 1 - \dfrac{\alpha - P_{d_H}}{P_{d_S} - P_{d_H}} \end{cases} \tag{6.11}$$

Hence, the CF prediction dataset can be formulated by sampling from the original datasets using the modified sampling function (6.11) where the number of samples drawn is $|d_t| - |d_N|$. Then, concatenating the d_N datapoints to formulate d_t.

6.4 Results and Discussion

To test the validity of the proposed model, a set of COSMIC simulations of the IEEE 39-bus system are conducted and separated into two categories: synthetic data, and real-time data. For the synthetic dataset, the dataset has 50% of the cases with CF events, while the real-time dataset has only 1% of the cases with CF events. The detail of each dataset is summarized in Table 6.1. For the CF prediction, the model in [146] is used, but with RNN instead of NN to improve prediction accuracy, while load bus 9 of the IEEE 39-bus system is arbitrarily selected as the bus under analysis (other buses show similar trends). The simulation starts with training the model using only synthetic data. Then, for each time step, new data is added from the pool of real-time data. Afterwards, the simulation finishes when all the data from the real-time category are

included. As for the parameters of the model, they are listed in Table 6.2. Finally, since both datasets are based on COSMIC, the validator will not detect any contradictions. Hence, it is omitted.

Initiating the simulation according to Figure 6.4 and the specified settings in Table 6.2, the accuracy of the prediction model as k progresses is shown in Figure 6.5. Initially, the model started with only 89.2% accuracy, but as more historical data is fed into the model its accuracy improved until it peaked at $k = 5$ with 94.7% accuracy. However, since the models did not achieve a better prediction accuracy afterwards, the M^5 model remained as the selected prediction model until the end of the simulation at $k = 9$. This result shows that the model is capable of learning from the data and subsequently updating the prediction model as more information about the grid is known. Moreover, the maximum training time (t_{tr}) was $300\ s$; thus, the model under this setup can produce updated predictors each 5 minutes.

Table 6.1. Synthetic and real data description.

Case type	Synthetic data Count	Real data count
Normal Operation	2000	15840
Blackout or CF	2000	160
Total	4000	16000

Table 6.2. CF prediction model parameters.

Parameter	Value	Parameter	Value
t_{mx}	600 s	$\|d_t\|$	10000
t_{de}	1 s	$\|d_N\|$	2000
t_{tr}	30-300 s	acc_{Th}	$acc(M^{k-1})$

To further test the model, two additional case studies are conducted. In the first case study, the effect of modifying $|d_N|$ is investigated by reducing it to 1000. The new CF prediction model parameters are shown in Table 6.3 where the t_{tr} also decreased significantly, indicating that this setup can provide quicker updates (around one new update every 2 minutes). The accuracy of this new configuration is plotted in Figure 6.6 and it shows double the number of iterations since the model needs half the amount of new data to formulate a new model. The accuracy shows a similar

trend as the original model but has more variations between the models and deeper dips in accuracy. Also, the best accuracy is slightly lower.

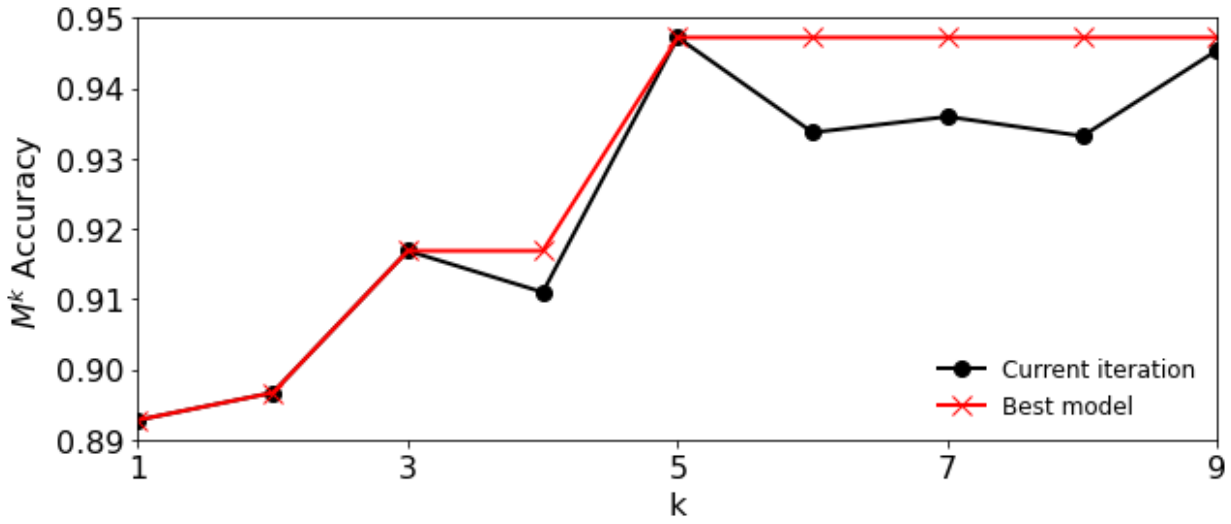

Figure 6.5. Accuracy of the model M^k as k is incremented.

Table 6.3. CF prediction model parameters (with reduced $|d_N|$).

Parameter	Value	Parameter	Value
t_{mx}	600 s	$\|d_t\|$	10000
t_{de}	1 s	$\|d_N\|$	1000
t_{tr}	30-150 s	acc_{Th}	$acc(M^{k-1})$

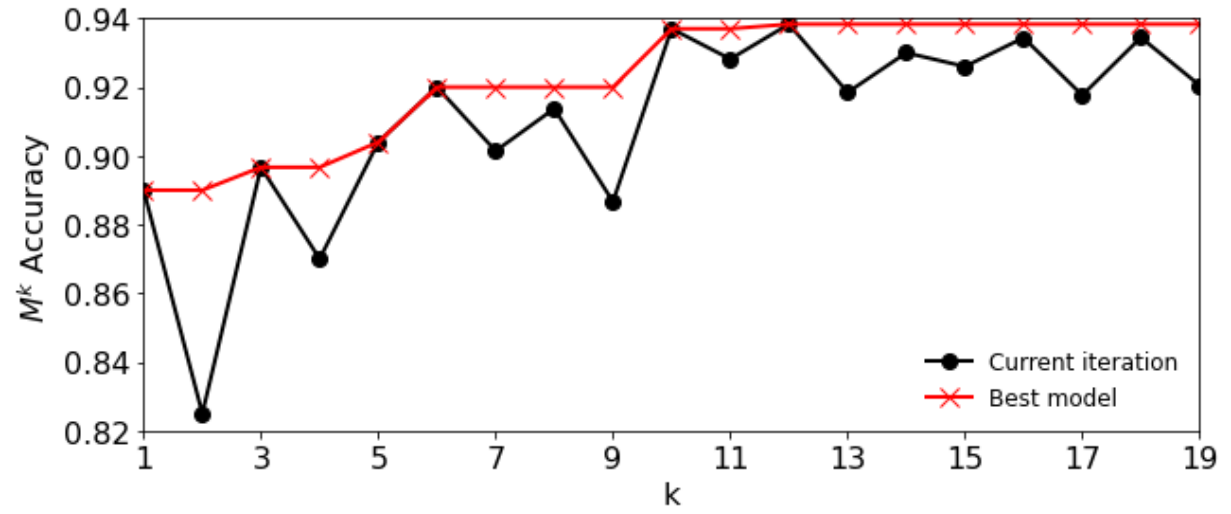

Figure 6.6. Accuracy of the model M^k as k is incremented with reduced $|d_N|$).

The second case study is modifying the data size used for training so that only new data is used ($|d_t| = |d_N|$). However, the model is modified so that it allows for incremental change (i.e., new models use the older model as its based model before training the data). The new parameters are listed in Table 6.4. With the new configuration, the training time is significantly less than one minute, indicating that the model is suitable even for five-minute dispatching. As shown in Figure 6.7, the model under this case study has fewer variations than the reduced $|d_N|$, but more than the original model. However, the incremental nature of the model provided accuracies higher than the previous two models as the maximum accuracy was 95.5%.

Table 6.4. CF prediction model parameters (with $|d_t| = |d_N|$).

Parameter	Value	Parameter	Value
t_{mx}	600 s	$\|d_t\|$	1000
t_{de}	1 s	$\|d_N\|$	1000
t_{tr}	5-25 s	acc_{Th}	$acc(M^{k-1})$

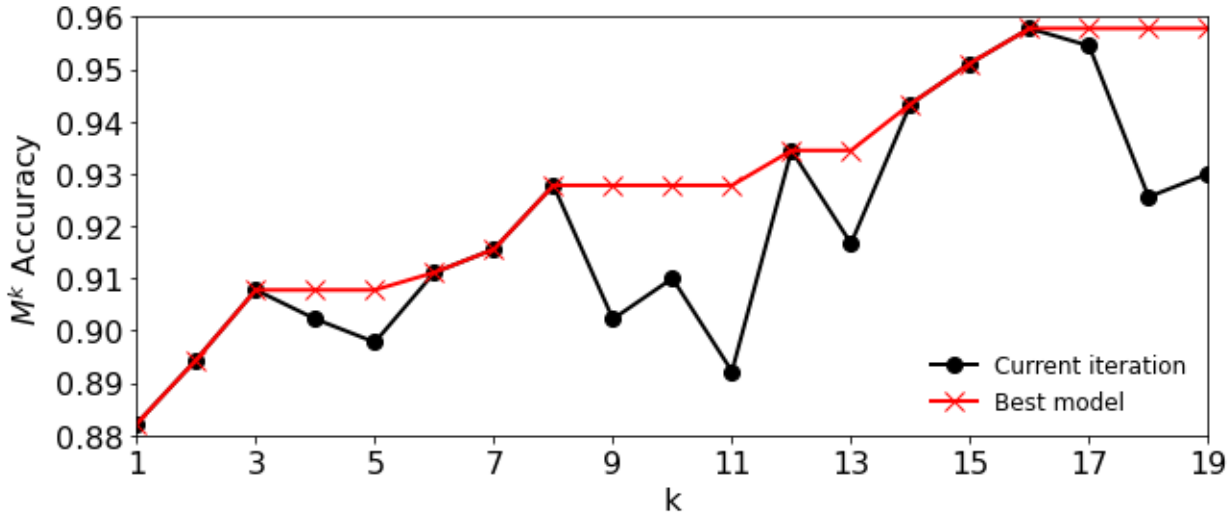

Figure 6.7. Accuracy of the model M^k as k is incremented with $|d_t| = |d_N|$).

6.5 Summary

In this chapter, an autonomous CFP framework is proposed. The framework addresses multiple issues in the current CFP models by allowing the model to automatically learn from new datasets as time progresses. Moreover, IMS and CBR extend the functionality of the autonomous framework and increase its convergence speed and its efficacy in extracting information from the

newly acquired datasets. Moreover, this framework is expected to be applied as an auxiliary warning system to detect discrepancies between synthetic and real-time data to help aid in the improvement of the CF simulation models.

7 SUMMARY AND CONCLUSIONS

7.1 Summary

Blackout losses cost billions of dollars to both utilities and consumers. To reduce the risk of blackouts, system operators use criteria such as the N-1 outage and the 0.1 days/year outage to maintain acceptable operation levels even after contingencies. However, the recurrence of blackouts necessitates better assessment and analysis methods. Hence, cascading failure (CF) analysis emerged because it better describes the propagation of blackouts in power systems since all blackouts propagate as a CF. Then, the recent developments in CF analysis, most notably CF modeling, increased the interest in designing CF prediction models.

This thesis presented multiple CF prediction frameworks to address the different design requirements of real-time CF prediction. A confidence-interval-based prediction framework was designed to predict CFs in real-time without the need for fault information. However, it had the limitation of requiring PMU data from all buses, which limits its scalability. Then, a CF prediction framework based on RNN-LSTM ensemble to capture the dynamics of the power networks was proposed. It used information theory and other reduction techniques to significantly reduce the number of inputs while maintaining high accuracy. Hence, it addressed the main limitation of the first model. Finally, a validation loop was introduced to allow the proposed models to update in real-time while being connected to the grid to reduce the dependency on synthetic data and to adapt to system changes.

The following paragraphs provide a chapter-wise summary of this thesis.

Chapter 1 introduces the importance of CF prediction and blackouts prevention. Also, the impact of smart grids and WAMS on the used paradigms to analyze and suppress CFs and blackouts is discussed.

A detailed literature review on CF is conducted in Chapter 2. The reviewed topics included CF modeling, reliability analysis and its relation to CF, CF prediction, and the benchmarking attempts to both CF modeling and prediction. The chapter also addresses the limitations of static modeling which is often used in CF analysis. Moreover, the limitations of angle stability and risk assessment from a CF perspective are discussed.

Chapter 3 presents the methodology used to generate the CF data in this thesis and the downsides of using normal distributions and Monte Carlo simulations to create CF scenarios. The analyzed generated CF and blackout data have shown the significant difference between the LOL distribution in small and large systems.

Chapter 4 presents the first CF prediction model proposed in this thesis. The model uses prediction intervals (PI) to deliver more meaningful information to the operator regarding the likelihood and size of the blackout. Also, it uses data fusion to improve the prediction accuracy. Moreover, the proposed formulation does not depend on the failure location and line outage data to produce its prediction. Hence, it is less prone to estimation errors. However, it requires inputs from all PMUs which might not be practical in large systems.

Chapter 5 presents a second CF prediction model, namely the real-time load-point-based CF prediction for large interconnected systems. RNN-LSTM cells are used in this model for better modeling of the dynamic grid behaviour so that the predictor can provide quicker and more accurate predictions at each time step. Moreover, a load-based RNN ensemble is used to provide more information to the operators and RAS controllers regarding the locations affected by the CF. Also, the ensemble partially decreases the size of the RNN to reduce the curse of dimensionality. In addition, the usage of network theory, eigenvector centrality, information entropy, and higher moments to obtain the most relevant information to CF prediction drastically decreases input count without affecting the accuracy. Hence, it allows the model to be applicable to large interconnected power systems, which was the main limitation of the first model. Moreover, PMU non-idealities are considered in the formulation of the CF prediction model, and it is shown that the model is resilient to both missing PMU data and noisy PMU measurements.

Chapter 6 proposes an autonomous learning framework for CF prediction. This framework aims at improving the efficacy of CF prediction algorithms by training and updating them in real-time while they are connected to the grid. The framework uses importance sampling and case-based reasoning to optimize the data size and the training time to make the framework suitable for real-time CF prediction. The results indicate that this framework can significantly improve the efficacy of existing CF prediction models by allowing them to improve in real-time without notably altering their original formulation.

7.2 Contributions

Following are the main contributions of this thesis:

1. **Designing a real-time CF prediction framework** that can be scaled to large interconnected power systems.
2. **Tailoring prediction interval techniques to the CF prediction problem** and adjusting the problem formulation accordingly.
3. **Proposing methodologies to significantly reduce the number of inputs used in CF prediction** using higher moments, bus centrality, information entropy, and graph theory.
4. **Designing a load-based RNN-LSTM ensemble,** which provides more meaningful information to the operators regarding the state of the grid while also simplifying the formulation and training of the model.
5. **Formulating an autonomous framework for CF prediction training in real-time** based on importance sampling and case-based reasoning.

7.3 Conclusions

The following conclusions are drawn from the studies conducted in this thesis:

1. Typical MC simulations are not suited for CF analysis. Hence, modified data generation schemes are needed to efficiently generate CF data in order to test and train CF prediction models.
2. Data fusion substantially improves the prediction accuracy of load-based predictions by exploiting the correlation between the buses (it corrects around 20% of the prediction errors).
3. Applying a load-based approach to CF prediction improves the accuracy of the predictions and allows the operators to have meaningful information about the affected regions.
4. RNN-LSTM cells are suitable for modeling both the fast and slow cascades of the power network. Hence, they provide higher prediction accuracy (99.9%) with less delays (the maximum time delay is $400\ ms$).

5. The inclusion of hidden failures (HFs) significantly reduces the accuracy of a prediction model (maximum accuracy is below 95%). However, it more accurately represents the uncertainties of the network. Thus, there is a trade-off between efficacy and accuracy when including HFs within CF prediction.
6. RNN-LSTM ensemble with input reduction techniques can significantly suppress the curse of dimensionality in CF prediction schemes and provides very high accuracy in large interconnected systems (inputs are reduced by 98%, but the accuracy is still higher than 99%).
7. The validation loop allows the CF prediction model to improve naturally as time progresses. Hence, the validation loop eliminates the need for recurrent offline training. Moreover, it shows the applicability of importance sampling and case-based reasoning on CF prediction frameworks.

7.4 Suggestions for Future Work

The following studies are recommended for future investigations based on the research carried out in this thesis.

1. **EENS Prediction:** Currently, all models proposed in this thesis, and the CFP literature in general, output the LOL expectation (LOLE). However, a more desired output is the expected energy not supplied (EENS). This can be achieved by modeling the restoration process to estimate the duration of the LOLE. Hence, future work could incorporate such a process.
2. **Bus-Branch-based CF prediction**: The CF analysis in this thesis uses the Bus-Branch modeling of the power system. However, CF analysis can be extended to include the Node-Breaker model, when considering failures near or within substations.
3. **Detailed PMU dynamics:** Only some PMUs non-idealities are considered in this research (total vector error (TVE) as a Gaussian noise and missing PMU data). However, a more holistic system could incorporate more PMUs non-idealities. For example, the delay due to communication networks, which can range between $100\ ms$ to $3000\ ms$ depending on the communication network infrastructure. Also, the TVE due to transient response and internal PMU delays. Hence, future work should address these limitations and build a more robust CF prediction framework.

REFERENCES

[1] R. Yan, N. Al -Masood, T. Kumar Saha, F. Bai, and H. Gu, "The Anatomy of the 2016 South Australia Blackout: A Catastrophic Event in a High Renewable Network," *IEEE Trans. Power Syst.*, vol. 33, no. 5, pp. 5374–5388, Sep. 2018, doi: 10.1109/TPWRS.2018.2820150.

[2] M. S. E. Sharif, M. Chowdhury, M. Moniruzzaman, and A. Mohit, "A Probabilistic LOLE Based Method to Determine Optimal Reserve and Control Power System Frequency," in *2017 3rd International Conference on Electrical Information and Communication Technology (EICT)*, 2017, vol. 2018-Janua, pp. 1–6, doi: 10.1109/EICT.2017.8275201.

[3] H. Haes Alhelou, M. Hamedani-Golshan, T. Njenda, and P. Siano, "A Survey on Power System Blackout and Cascading Events: Research Motivations and Challenges," *Energies*, vol. 12, no. 4, p. 682, Feb. 2019, doi: 10.3390/en12040682.

[4] O. P. Veloza and F. Santamaria, "Analysis of Major Blackouts from 2003 to 2015: Classification of Incidents and Review of Main Causes," *Electr. J.*, vol. 29, no. 7, pp. 42–49, Sep. 2016, doi: 10.1016/j.tej.2016.08.006.

[5] D. Wu *et al.*, "An Open-source Model for Simulation and Corrective Measure Assessment of the 2021 Texas Power Outage," pp. 1–13, Apr. 2021, [Online]. Available: http://arxiv.org/abs/2104.04146.

[6] G. B. Anderson and M. L. Bell, "Lights Out," *Epidemiology*, vol. 23, no. 2, pp. 189–193, Mar. 2012, doi: 10.1097/EDE.0b013e318245c61c.

[7] Vaiman *et al.*, "Risk Assessment of Cascading Outages: Methodologies and Challenges," *IEEE Trans. Power Syst.*, vol. 27, no. 2, pp. 631–641, May 2012, doi: 10.1109/TPWRS.2011.2177868.

[8] "ERCOT Calls for Rotating Outages as Extreme Winter Weather Forces Generating Units Offline," 2021. http://www.ercot.com/news/releases/show/225210 (accessed Apr. 26, 2021).

[9] H. Guo, C. Zheng, H. H. C. Iu, and T. Fernando, “A Critical Review of Cascading Failure Analysis and Modeling of Power System,” *Renew. Sustain. Energy Rev.*, vol. 80, no. April, pp. 9–22, 2017, doi: 10.1016/j.rser.2017.05.206.

[10] J. Gettleman, “Monkey in Kenya Survives After Setting Off Nationwide Blackout,” 2016. https://www.nytimes.com/2016/06/09/world/africa/monkey-kenya-survives-blackout-internet-vervet.html (accessed Apr. 26, 2021).

[11] “Report on Blackout in Turkey on 31st March 2015,” 2015. [Online]. Available: https://eepublicdownloads.entsoe.eu/clean-documents/SOC documents/Regional_Groups_Continental_Europe/20150921_Black_Out_Report_v10_w.pdf.

[12] L. L. Lai, H. T. Zhang, C. S. Lai, F. Y. Xu, and S. Mishra, “Investigation on July 2012 Indian blackout,” in *2013 International Conference on Machine Learning and Cybernetics*, 2013, vol. 1, no. July 2012, pp. 92–97, doi: 10.1109/ICMLC.2013.6890450.

[13] A. Atputharajah and T. K. Saha, “Power System Blackouts - Literature Review,” *ICIIS 2009 - 4th Int. Conf. Ind. Inf. Syst. 2009, Conf. Proc.*, no. December, pp. 460–465, 2009, doi: 10.1109/ICIINFS.2009.5429818.

[14] C. A. Ruiz, N. J. Orrego, and J. F. Gutierrez, “The Colombian 2007 Black Out,” in *2008 IEEE/PES Transmission and Distribution Conference and Exposition: Latin America*, 2008, pp. 1–5, doi: 10.1109/TDC-LA.2008.4641845.

[15] UCTE, “Final Report System Disturbance on 4 November 2006 union for the co-ordination of transmission of electricity,” 2006.

[16] M. W. Younas and S. A. Qureshi, “Analysis of Blackout of National Grid System of Pakistan in 2006 and the Application of PSS and FACTS Controllers as Remedial Measures,” in *2007 International Conference on Electrical Engineering*, 2007, pp. 1–6, doi: 10.1109/ICEE.2007.4287339.

[17] Z. Huang, C. Wang, M. Stojmenovic, and A. Nayak, “Characterization of Cascading Failures in Interdependent Cyber-Physical Systems,” *IEEE Trans. Comput.*, vol. 64, no. 8,

pp. 2158–2168, Aug. 2015, doi: 10.1109/TC.2014.2360537.

[18] Y. Xiaohui, Z. Wuzhi, S. Xinli, W. Guoyang, L. Tao, and S. Zhida, "Review on Power System Cascading Failure Thoeries and Studies," in *2016 International Conference on Probabilistic Methods Applied to Power Systems (PMAPS)*, 2016, pp. 1–6, doi: 10.1109/PMAPS.2016.7764167.

[19] J. Bialek *et al.*, "Benchmarking and Validation of Cascading Failure Analysis Tools," *IEEE Trans. Power Syst.*, vol. 31, no. 6, pp. 4887–4900, Nov. 2016, doi: 10.1109/TPWRS.2016.2518660.

[20] G. C. Patil and A. G. Thosar, "Application of Synchrophasor Measurements using PMU for Modern Power Systems Monitoring and Control," in *2017 International Conference on Computation of Power, Energy Information and Commuincation (ICCPEIC)*, 2017, vol. 2018-Janua, pp. 754–760, doi: 10.1109/ICCPEIC.2017.8290464.

[21] J. J. Q. Yu, D. J. Hill, A. Y. S. Lam, J. Gu, and V. O. K. Li, "Intelligent Time-Adaptive Transient Stability Assessment System," *IEEE Trans. Power Syst.*, vol. 33, no. 1, pp. 1049–1058, Jan. 2018, doi: 10.1109/TPWRS.2017.2707501.

[22] P. Rezaei, "Cascading Failure Risk Estimation and Mitigation in Power Systems," Ph.D. dissertation, University of Vermont, Burlington, VT, 2016.

[23] H. M. Merrill, M. A. Hossain, and M. Bodson, "Nipping Blackouts in the Bud: Introducing a Novel Cascading Failure Network," *IEEE Power Energy Mag.*, vol. 18, no. 4, pp. 64–75, Jul. 2020, doi: 10.1109/MPE.2020.2986085.

[24] P. Henneaux, P. E. Labeau, J. C. Maun, and L. Haarla, "A Two-Level Probabilistic Risk Assessment of Cascading Outages," *IEEE Trans. Power Syst.*, vol. 31, no. 3, pp. 2393–2403, 2016, doi: 10.1109/TPWRS.2015.2439214.

[25] M. Papic *et al.*, "Survey of Tools for Risk Assessment of Cascading Outages," in *2011 IEEE Power and Energy Society General Meeting*, 2011, pp. 1–9, doi: 10.1109/PES.2011.6039371.

[26] S. Mei, F. He, X. Zhang, S. Wu, and G. Wang, "An Improved OPA Model and Blackout

Risk Assessment," *IEEE Trans. Power Syst.*, vol. 24, no. 2, pp. 814–823, 2009, doi: 10.1109/TPWRS.2009.2016521.

[27] J. Song, E. Cotilla-Sanchez, G. Ghanavati, and P. D. H. Hines, "Dynamic Modeling of Cascading Failure in Power Systems," *IEEE Trans. Power Syst.*, vol. 31, no. 3, pp. 2085–2095, May 2016, doi: 10.1109/TPWRS.2015.2439237.

[28] B. Liscouski and W. Elliot, "U.S.-Canada Power System Outage Task Force," 2004. doi: 10.1021/ac203394r.A.

[29] S. Mei, Y. Ni, G. Wang, and S. Wu, "A Study of Self-Organized Criticality of Power System under Cascading Failures Based on AC-OPF with Voltage Stability Margin," *IEEE Trans. Power Syst.*, vol. 23, no. 4, pp. 1719–1726, 2008, doi: 10.1109/TPWRS.2008.2002295.

[30] M. J. Eppstein and P. D. H. Hines, "A 'Random Chemistry' Algorithm for Identifying Collections of Multiple Contingencies that Initiate Cascading Failure," *IEEE Trans. Power Syst.*, vol. 27, no. 3, pp. 1698–1705, 2012, doi: 10.1109/TPWRS.2012.2183624.

[31] J. Qi, S. Mei, and F. Liu, "Blackout Model Considering Slow Process," *IEEE Trans. Power Syst.*, vol. 28, no. 3, pp. 3274–3282, 2013, doi: 10.1109/TPWRS.2012.2230196.

[32] P. Henneaux, P. E. Labeau, and J. C. Maun, "Blackout Probabilistic Risk Assessment and Thermal Effects: Impacts of Changes in Generation," *IEEE Trans. Power Syst.*, vol. 28, no. 4, pp. 4722–4731, 2013, doi: 10.1109/TPWRS.2013.2263851.

[33] J. Yan, Y. Tang, H. He, and Y. Sun, "Cascading Failure Analysis With DC Power Flow Model And Transient Stability Analysis," *IEEE Trans. Power Syst.*, vol. 30, no. 1, pp. 285–297, 2015, doi: 10.1109/TPWRS.2014.2322082.

[34] X. Liu and Z. Li, "Revealing the Impact of Multiple Solutions in DCOPF on the Risk Assessment of Line Cascading Failure in OPA Model," *IEEE Trans. Power Syst.*, vol. 31, no. 5, pp. 4159–4160, 2016, doi: 10.1109/TPWRS.2015.2508642.

[35] V. Rampurkar, P. Pentayya, H. A. Mangalvedekar, and F. Kazi, "Cascading Failure Analysis for Indian Power Grid," *IEEE Trans. Smart Grid*, vol. 7, no. 4, pp. 1951–1960, 2016, doi: 10.1109/TSG.2016.2530679.

[36] F. Wenli, L. Zhigang, H. Ping, and M. Shengwei, “Cascading Failure Model in Power Grids using the Complex Network Theory,” *IET Gener. Transm. Distrib.*, vol. 10, no. 15, pp. 3940–3949, 2016, doi: 10.1049/iet-gtd.2016.0692.

[37] P. D. H. Hines, I. Dobson, and P. Rezaei, “Cascading Power Outages Propagate Locally in an Influence Graph That is Not the Actual Grid Topology,” *IEEE Trans. Power Syst.*, vol. 32, no. 2, pp. 958–967, 2017, doi: 10.1109/TPWRS.2016.2578259.

[38] Z. Wang *et al.*, “Impacts of Operators’ Behavior on Reliability of Power Grids During Cascading Failures,” *IEEE Trans. Power Syst.*, vol. 33, no. 6, pp. 6013–6024, Nov. 2018, doi: 10.1109/TPWRS.2018.2825348.

[39] J. Cordova-Garcia, X. Wang, D. Xie, Y. Zhao, and L. Zuo, “Control of Communications-Dependent Cascading Failures in Power Grids,” *IEEE Trans. Smart Grid*, vol. 10, no. 5, pp. 5021–5031, Sep. 2019, doi: 10.1109/TSG.2018.2873217.

[40] M. Rahnamay-Naeini, Z. Wang, N. Ghani, A. Mammoli, and M. M. Hayat, “Stochastic Analysis of Cascading-Failure Dynamics in Power Grids,” *IEEE Trans. Power Syst.*, vol. 29, no. 4, pp. 1767–1779, Jul. 2014, doi: 10.1109/TPWRS.2013.2297276.

[41] Q. Chen and L. Mili, “Composite Power System Vulnerability Evaluation to Cascading Failures using Importance Sampling and Antithetic Variates,” *IEEE Trans. Power Syst.*, vol. 28, no. 3, pp. 2321–2330, 2013, doi: 10.1109/TPWRS.2013.2238258.

[42] P. Ramachandran, V. Vittal, and G. T. Heydt, “Mechanical State Estimation for Overhead Transmission Lines with Level Spans,” *IEEE Trans. Power Syst.*, vol. 23, no. 3, pp. 908–915, 2008, doi: 10.1109/TPWRS.2008.926093.

[43] S. Il Lim, C. C. Liu, S. J. Lee, M. S. Choi, and S. J. Rim, “Blocking of Zone 3 Relays to Prevent Cascaded Events,” *IEEE Trans. Power Syst.*, vol. 23, no. 2, pp. 747–754, 2008, doi: 10.1109/TPWRS.2008.919417.

[44] Z. Jiao, H. Gong, and Y. Wang, “A D-S Evidence Theory-Based Relay Protection System Hidden Failures Detection Method in Smart Grid,” *IEEE Trans. Smart Grid*, vol. 9, no. 3, pp. 2118–2126, 2018, doi: 10.1109/TSG.2016.2607318.

[45] T. Ghanizadeh Bolandi, M. R. Haghifam, and M. Khederzadeh, “Real-Time Monitoring of Zone 3 Vulnerable Distance Relays to Prevent Maloperation Under Load Encroachment Condition,” *IET Gener. Transm. Distrib.*, vol. 11, no. 8, pp. 1878–1888, 2017, doi: 10.1049/iet-gtd.2016.0486.

[46] H. Wu and I. Dobson, “Cascading Stall of Many Induction Motors in a Simple System,” *IEEE Trans. Power Syst.*, vol. 27, no. 4, pp. 2116–2126, 2012, doi: 10.1109/TPWRS.2012.2189420.

[47] H. Wu and I. Dobson, “Analysis of Induction Motor Cascading Stall in a Simple System Based on The CASCADE Model,” *IEEE Trans. Power Syst.*, vol. 28, no. 3, pp. 3184–3193, 2013, doi: 10.1109/TPWRS.2013.2244924.

[48] J. Tu, H. Xin, Z. Wang, D. Gan, and Z. Huang, “On Self-Organized Criticality of the East China AC-DC Power System - The Role of DC Transmission,” *IEEE Trans. Power Syst.*, vol. 28, no. 3, pp. 3204–3214, 2013, doi: 10.1109/TPWRS.2013.2251913.

[49] S. K. E. Awadallah, J. V. Milanović, P. N. Jarman, and Z. Wang, “Probabilistic Indicators for Assessing Age- and Loading-Based Criticality of Transformers to Cascading Failure Events,” *IEEE Trans. Power Syst.*, vol. 29, no. 5, pp. 2558–2566, 2014, doi: 10.1109/TPWRS.2014.2308581.

[50] X. Liu, M. Shahidehpour, Y. Cao, Z. Li, and W. Tian, “Risk Assessment in Extreme Events Considering the Reliability of Protection Systems,” *IEEE Trans. Smart Grid*, vol. 6, no. 2, pp. 1073–1081, 2015, doi: 10.1109/TSG.2015.2393254.

[51] H. Zheng and C. L. DeMarco, “A New Dynamic Performance Model of Motor Stalling and FIDVR for Smart Grid Monitoring/Planning,” *IEEE Trans. Smart Grid*, vol. 7, no. 4, pp. 1989–1996, 2016, doi: 10.1109/TSG.2016.2548082.

[52] A. A. Babalola, R. Belkacemi, and S. Zarrabian, “Real-Time Cascading Failures Prevention for Multiple Contingencies in Smart Grids Through a Multi-Agent System,” *IEEE Trans. Smart Grid*, vol. 9, no. 1, pp. 373–385, 2018, doi: 10.1109/TSG.2016.2553146.

[53] J. D. Glover, M. S. Sarma, and T. J. Overbye, *Power system analysis and design*. Stamford,

CT: Cengage Learning, 2008.

[54] J. Wei, D. Kundur, and K. L. Butler-Purry, “A Novel Bio-Inspired Technique For Rapid Real-Time Generator Coherency Identification,” *IEEE Trans. Smart Grid*, vol. 6, no. 1, pp. 178–188, 2015, doi: 10.1109/TSG.2014.2341213.

[55] J. Ma *et al.*, “Angle Stability Analysis of Power System With Cascading Failure,” *IEEE Trans. Power Syst.*, vol. 32, no. 2, pp. 873–882, 2017, doi: 10.1109/TPWRS.2016.2566672.

[56] X. Meng, M. Liu, and Q. Zhang, “Control Strategy for Relieving Transient Potential Energy Accumulation in Power System,” *IET Gener. Transm. Distrib.*, vol. 12, no. 3, pp. 704–712, 2018, doi: 10.1049/iet-gtd.2017.0077.

[57] Y. Tang, F. Li, Q. Wang, and Y. Xu, “Hybrid Method for Power System Transient Stability Prediction Based on Two-Stage Computing Resources,” *IET Gener. Transm. Distrib.*, vol. 12, no. 8, pp. 1697–1703, Apr. 2018, doi: 10.1049/iet-gtd.2017.1168.

[58] M. R. Aghamohammadi, S. Hashemi, and A. Hasanzadeh, “A New Approach for Mitigating Blackout Risk by Blocking Minimum Critical Distance Relays,” *Int. J. Electr. Power Energy Syst.*, vol. 75, pp. 162–172, Feb. 2016, doi: 10.1016/j.ijepes.2015.08.019.

[59] Hui Ren, I. Dobson, and B. A. Carreras, “Long-Term Effect of the n-1 Criterion on Cascading Line Outages in an Evolving Power Transmission Grid,” *IEEE Trans. Power Syst.*, vol. 23, no. 3, pp. 1217–1225, Aug. 2008, doi: 10.1109/TPWRS.2008.926417.

[60] R. Abe, H. Taoka, and D. McQuilkin, “Digital Grid: Communicative Electrical Grids of the Future,” *IEEE Trans. Smart Grid*, vol. 2, no. 2, pp. 399–410, Jun. 2011, doi: 10.1109/TSG.2011.2132744.

[61] H. Zhu and G. B. Giannakis, “Sparse Overcomplete Representations for Efficient Identification of Power Line Outages,” *IEEE Trans. Power Syst.*, vol. 27, no. 4, pp. 2215–2224, 2012, doi: 10.1109/TPWRS.2012.2192142.

[62] A. Safdarian, M. Fotuhi-Firuzabad, M. M. Othman, I. Musirin, N. A. Salim, and M. S. Serwan, “Determination of Available Transfer Capability with Implication of Cascading Collapse Uncertainty,” *IET Gener. Transm. Distrib.*, vol. 8, no. 4, pp. 705–715, 2014, doi:

10.1049/iet-gtd.2013.0395.

[63] J. Wu, J. Xiong, and Y. Shi, “Efficient Location Identification of Multiple Line Outages With Limited PMUs in Smart Grids,” *IEEE Trans. Power Syst.*, vol. 30, no. 4, pp. 1659–1668, 2015, doi: 10.1109/TPWRS.2014.2357751.

[64] Q. Huang, L. Shao, and N. Li, “Dynamic Detection of Transmission Line Outages using Hidden Markov Models,” *Proc. Am. Control Conf.*, vol. 2015-July, no. 3, pp. 5050–5055, 2015, doi: 10.1109/ACC.2015.7172125.

[65] M. Panteli, D. N. Trakas, P. Mancarella, and N. D. Hatziargyriou, “Boosting the Power Grid Resilience to Extreme Weather Events Using Defensive Islanding,” *IEEE Trans. Smart Grid*, vol. 7, no. 6, pp. 2913–2922, 2016, doi: 10.1109/TSG.2016.2535228.

[66] R. Deng, P. Zhuang, and H. Liang, “CCPA: Coordinated Cyber-Physical Attacks and Countermeasures in Smart Grid,” *IEEE Trans. Smart Grid*, vol. 8, no. 5, pp. 2420–2430, 2017, doi: 10.1109/TSG.2017.2702125.

[67] A. Kargarian, M. Mehrtash, and B. Falahati, “Decentralized Implementation of Unit Commitment with Analytical Target Cascading: A Parallel Approach,” *IEEE Trans. Power Syst.*, vol. 33, no. 4, pp. 3981–3993, 2018, doi: 10.1109/TPWRS.2017.2787645.

[68] R. Bi, T. Lin, R. Chen, J. Ye, X. Zhou, and X. Xu, “Alleviation of Post-Contingency Overloads by SOCP Based Corrective Control Considering TCSC and MTDC,” *IET Gener. Transm. Distrib.*, vol. 12, no. 9, pp. 2155–2164, 2018, doi: 10.1049/iet-gtd.2017.1393.

[69] Y. Xiang, L. Wang, and N. Liu, “A Robustness-Oriented Power Grid Operation Strategy Considering Attacks,” *IEEE Trans. Smart Grid*, vol. 9, no. 5, pp. 4248–4261, 2018, doi: 10.1109/TSG.2017.2653219.

[70] J. Fang, R. Zhang, H. Li, and Y. Tang, “Frequency Derivative-Based Inertia Enhancement by Grid-Connected Power Converters With a Frequency-Locked-Loop,” *IEEE Trans. Smart Grid*, vol. 10, no. 5, pp. 4918–4927, Sep. 2019, doi: 10.1109/TSG.2018.2871085.

[71] L. Che, X. Liu, and Z. Li, “Mitigating False Data Attacks Induced Overloads Using a Corrective Dispatch Scheme,” *IEEE Trans. Smart Grid*, vol. 10, no. 3, pp. 3081–3091, May

2019, doi: 10.1109/TSG.2018.2817515.

[72] G. Liang, S. R. Weller, J. Zhao, F. Luo, and Z. Y. Dong, "The 2015 Ukraine Blackout: Implications for False Data Injection Attacks," *IEEE Trans. Power Syst.*, vol. 32, no. 4, pp. 3317–3318, Jul. 2017, doi: 10.1109/TPWRS.2016.2631891.

[73] D. Ye, M. Zhang, and D. Sutanto, "A Hybrid Multiagent Framework with Q-Learning for Power Grid Systems Restoration," *IEEE Trans. Power Syst.*, vol. 26, no. 4, pp. 2434–2441, 2011, doi: 10.1109/TPWRS.2011.2157180.

[74] F. Ren, M. Zhang, D. Soetanto, and X. Su, "Conceptual Design of a Multi-Agent System for Interconnected Power Systems Restoration," *IEEE Trans. Power Syst.*, vol. 27, no. 2, pp. 732–740, 2012, doi: 10.1109/TPWRS.2011.2177866.

[75] N. H. Tran, C. Pham, M. N. H. Nguyen, S. Ren, and C. S. Hong, "Incentivizing Energy Reduction for Emergency Demand Response in Multi-Tenant Mixed-Use Buildings," *IEEE Trans. Smart Grid*, vol. 9, no. 4, pp. 3701–3715, 2018, doi: 10.1109/TSG.2016.2639033.

[76] S. Liu, B. Chen, T. Zourntos, D. Kundur, and K. Butler-Purry, "A Coordinated Multi-Switch Attack for Cascading Failures in Smart Grid," *IEEE Trans. Smart Grid*, vol. 5, no. 3, pp. 1183–1195, 2014, doi: 10.1109/TSG.2014.2302476.

[77] Z. Li, M. Shahidehpour, A. Alabdulwahab, and A. Abusorrah, "Bilevel Model for Analyzing Coordinated Cyber-Physical Attacks on Power Systems," *IEEE Trans. Smart Grid*, vol. 7, no. 5, pp. 2260–2272, 2016, doi: 10.1109/TSG.2015.2456107.

[78] S. Mishra, X. Li, T. Pan, A. Kuhnle, M. T. Thai, and J. Seo, "Price Modification Attack and Protection Scheme in Smart Grid," *IEEE Trans. Smart Grid*, vol. 8, no. 4, pp. 1864–1875, 2017, doi: 10.1109/TSG.2015.2509945.

[79] L. Che, X. Liu, Z. Shuai, Z. Li, and Y. Wen, "Cyber Cascades Screening Considering the Impacts of False Data Injection Attacks," *IEEE Trans. Power Syst.*, vol. 33, no. 6, pp. 6545–6556, 2018, doi: 10.1109/TPWRS.2018.2827060.

[80] H.-M. Chung, W.-T. Li, C. Yuen, W.-H. Chung, Y. Zhang, and C.-K. Wen, "Local Cyber-Physical Attack for Masking Line Outage and Topology Attack in Smart Grid," *IEEE Trans.*

Smart Grid, vol. 10, no. 4, pp. 4577–4588, Jul. 2019, doi: 10.1109/TSG.2018.2865316.

[81] D. T. Nguyen, Y. Shen, and M. T. Thai, "Detecting Critical Nodes in Interdependent Power Networks for Vulnerability Assessment," *IEEE Trans. Smart Grid*, vol. 4, no. 1, pp. 151–159, 2013, doi: 10.1109/TSG.2012.2229398.

[82] M. Rahnamay-naeini and M. M. Hayat, "Cascading Failures in Interdependent Infrastructures :," *IEEE Trans. Smart Grid*, vol. 3053, no. c, pp. 1997–2006, 2016, doi: 10.1109/TSG.2016.2539823.

[83] Y. Cai, Y. Cao, Y. Li, T. Huang, and B. Zhou, "Cascading Failure Analysis Considering Interaction Between Power Grids and Communication Networks," *IEEE Trans. Smart Grid*, vol. 7, no. 1, pp. 530–538, 2016, doi: 10.1109/TSG.2015.2478888.

[84] Q. Wang, M. Pipattanasomporn, M. Kuzlu, Y. Tang, Y. Li, and S. Rahman, "Framework for Vulnerability Assessment of Communication Systems for Electric Power Grids," *IET Gener. Transm. Distrib.*, vol. 10, no. 2, pp. 477–486, 2016, doi: 10.1049/iet-gtd.2015.0857.

[85] P. Y. Kong, "Cost Efficient Data Aggregation Point Placement with Interdependent Communication and Power Networks in Smart Grid," *IEEE Trans. Smart Grid*, vol. 10, no. 1, pp. 74–83, 2019, doi: 10.1109/TSG.2017.2731988.

[86] M. Vaiman, "Cascading Analysis in a Control Room: Real-time Applications," in *Challenges of Cascading Failure Analysis: Modeling, Data and Interdependence*, 2018, [Online]. Available: http://site.ieee.org/pes-cascading/files/2018/08/3-Vaiman.pdf.

[87] Z. Pannell, B. Ramachandran, and D. Snider, "Machine Learning Approach to Solving the Transient Stability Assessment Problem," in *2018 IEEE Texas Power and Energy Conference (TPEC)*, 2018, vol. 2018-Febru, pp. 1–6, doi: 10.1109/TPEC.2018.8312089.

[88] Cascading Failure Working Group, "Working Group Publications," 2021. https://site.ieee.org/pes-cascading/publications/ (accessed May 04, 2021).

[89] I. Dobson, "Estimating the Propagation and Extent of Cascading Line Outages from Utility Data with a Branching Process," *IEEE Trans. Power Syst.*, vol. 27, no. 4, pp. 2146–2155, 2012, doi: 10.1109/TPWRS.2012.2190112.

[90] J. Qi, K. Sun, and S. Mei, “An Interaction Model for Simulation and Mitigation of Cascading Failures,” *IEEE Trans. Power Syst.*, vol. 30, no. 2, pp. 804–819, 2015, doi: 10.1109/TPWRS.2014.2337284.

[91] M. R. Salimian and M. R. Aghamohammadi, “A Three Stages Decision Tree-Based Intelligent Blackout Predictor for Power Systems Using Brittleness Indices,” *IEEE Trans. Smart Grid*, vol. 9, no. 5, pp. 5123–5131, Sep. 2018, doi: 10.1109/TSG.2017.2680600.

[92] J. Qi, J. Wang, and K. Sun, “Efficient Estimation of Component Interactions for Cascading Failure Analysis by em Algorithm,” *IEEE Trans. Power Syst.*, vol. 33, no. 3, pp. 3153–3161, 2018, doi: 10.1109/TPWRS.2017.2764041.

[93] J. Qi, I. Dobson, and S. Mei, “Towards Estimating the Statistics of Simulated Cascades of Outages with Branching Processes,” *IEEE Trans. Power Syst.*, vol. 28, no. 3, pp. 3410–3419, 2013, doi: 10.1109/TPWRS.2013.2243479.

[94] J. Kim, J. A. Bucklew, and I. Dobson, “Splitting Method for Speedy Simulation of Cascading Blackouts,” *IEEE Trans. Power Syst.*, vol. 28, no. 3, pp. 3010–3017, 2013, doi: 10.1109/TPWRS.2012.2231887.

[95] Y. Jia, K. Meng, and Z. Xu, “N-k Induced Cascading Contingency Screening,” *IEEE Trans. Power Syst.*, vol. 30, no. 5, pp. 2824–2825, 2015, doi: 10.1109/TPWRS.2014.2361723.

[96] S. P. Wang, A. Chen, C. W. Liu, C. H. Chen, J. Shortle, and J. Y. Wu, “Efficient Splitting Simulation for Blackout Analysis,” *IEEE Trans. Power Syst.*, vol. 30, no. 4, pp. 1775–1783, 2015, doi: 10.1109/TPWRS.2014.2359920.

[97] P. Rezaei, P. D. H. Hines, and M. J. Eppstein, “Estimating Cascading Failure Risk With Random Chemistry,” *IEEE Trans. Power Syst.*, vol. 30, no. 5, pp. 2726–2735, 2015, doi: 10.1109/TPWRS.2014.2361735.

[98] J. Guo, F. Liu, J. Wang, M. Cao, and S. Mei, “Quantifying the Influence of Component Failure Probability on Cascading Blackout Risk,” *IEEE Trans. Power Syst.*, vol. 33, no. 5, pp. 5671–5681, 2018, doi: 10.1109/TPWRS.2018.2809793.

[99] W. L. Fan, X. M. Zhang, S. W. Mei, and S. W. Huang, “Vulnerable Transmission Line

Identification Considering Depth of K-Shell Decomposition in Complex Grids," *IET Gener. Transm. Distrib.*, vol. 12, no. 5, pp. 1137–1144, 2017, doi: 10.1049/iet-gtd.2017.0767.

[100] A. Wang, Y. Luo, G. Tu, and P. Liu, "Vulnerability Assessment Scheme for Power System Transmission Networks Based on the Fault Chain Theory," *IEEE Trans. Power Syst.*, vol. 26, no. 1, pp. 442–450, Feb. 2011, doi: 10.1109/TPWRS.2010.2052291.

[101] C. Li *et al.*, "Method for Evaluating the Importance of Power Grid Nodes Based on Pagerank Algorithm," *IET Gener. Transm. Distrib.*, vol. 8, no. 11, pp. 1843–1847, 2014, doi: 10.1049/iet-gtd.2014.0051.

[102] F. Wenli, H. Ping, and L. Zhigang, "Multi-Attribute Node Importance Evaluation Method Based on Gini-Coefficient in Complex Power Grids," *IET Gener. Transm. Distrib.*, vol. 10, no. 9, pp. 2027–2034, 2016, doi: 10.1049/iet-gtd.2015.0803.

[103] P. Dey, R. Mehra, F. Kazi, S. Wagh, and N. M. Singh, "Impact of Topology on the Propagation of Cascading Failure in Power Grid," *IEEE Trans. Smart Grid*, vol. 7, no. 4, pp. 1970–1978, 2016, doi: 10.1109/TSG.2016.2558465.

[104] Z. Ma, C. Shen, F. Liu, and S. Mei, "Fast Screening of Vulnerable Transmission Lines in Power Grids: A PageRank-Based Approach," *IEEE Trans. Smart Grid*, vol. 10, no. 2, pp. 1982–1991, Mar. 2019, doi: 10.1109/TSG.2017.2785267.

[105] F. Wenli, Z. Xuemin, M. Shengwei, H. Shaowei, W. Wei, and D. Lijie, "Vulnerable Transmission Line Identification using ISH Theory in Power Grids," *IET Gener. Transm. Distrib.*, vol. 12, no. 4, pp. 1014–1020, 2018, doi: 10.1049/iet-gtd.2017.0571.

[106] J. Heydari and A. Tajer, "Quickest Localization of Anomalies in Power Grids: A Stochastic Graphical Framework," *IEEE Trans. Smart Grid*, vol. 9, no. 5, pp. 4679–4688, 2018, doi: 10.1109/TSG.2017.2666543.

[107] N. D. Hatziargyriou *et al.*, "Causes of the 2003 Major Grid Blackouts in North America and Europe , and Recommended Means to Improve System Dynamic Performance Causes of the 2003 Major Grid Blackouts in North America and Europe , and Recommended Means to Improve System Dynamic Perform," *IEEE Trans. Power Syst.*, vol. 20, no. 4, pp. 1922–

1928, 2005.

[108] C. W. Ten, K. Yamashita, Z. Yang, A. V. Vasilakos, and A. Ginter, “Impact Assessment Of Hypothesized Cyberattacks on Interconnected Bulk Power Systems,” *IEEE Trans. Smart Grid*, vol. 9, no. 5, pp. 4405–4425, 2018, doi: 10.1109/TSG.2017.2656068.

[109] P. Shukla and K. Bhowmick, “To Improve Classification of Imbalanced Datasets,” in *2017 International Conference on Innovations in Information, Embedded and Communication Systems (ICIIECS)*, 2017, vol. 2018-Janua, pp. 1–5, doi: 10.1109/ICIIECS.2017.8276044.

[110] O. A. Ansari and C. Y. Chung, “A Hybrid Framework for Short-Term Risk Assessment of Wind-Integrated Composite Power Systems,” *IEEE Trans. Power Syst.*, vol. 34, no. 3, pp. 2334–2344, May 2019, doi: 10.1109/TPWRS.2018.2881250.

[111] P. R. Bal and S. Kumar, “WR-ELM: Weighted Regularization Extreme Learning Machine for Imbalance Learning in Software Fault Prediction,” *IEEE Trans. Reliab.*, vol. 69, no. 4, pp. 1355–1375, Dec. 2020, doi: 10.1109/TR.2020.2996261.

[112] R. D. Zimmerman, C. E. Murillo-sánchez, and R. J. Thomas, “MATPOWER : Steady-State Operations , Systems Research and Education,” *IEEE Trans. Power Syst.*, vol. 26, no. 1, pp. 12–19, 2011.

[113] A. Parekh, P. Wang, and T. Strack, “Survey Results of User-Dependent Electricity Loads in Canadian Homes,” in *ACEEE Summer Study on Energy Efficiency in Buildings*, 2012, vol. 9, pp. 240–251.

[114] P. Dey, M. Parimi, A. Yerudkar, and S. R. Wagh, “Real-Time Estimation of Propagation of Cascade Failure using Branching Process,” in *2015 IEEE 5th International Conference on Power Engineering, Energy and Electrical Drives (POWERENG)*, 2015, vol. 2015-Septe, pp. 629–634, doi: 10.1109/PowerEng.2015.7266390.

[115] R. Pi, Y. Cai, Y. Li, and Y. Cao, “Machine Learning Based on Bayes Networks to Predict the Cascading Failure Propagation,” *IEEE Access*, vol. 6, pp. 44815–44823, 2018, doi: 10.1109/ACCESS.2018.2858838.

[116] N. H. Idris, N. A. Salim, M. M. Othman, and Z. M. Yasin, “Prediction of Cascading

Collapse Occurrence Due to the Effect of Hidden Failure Protection System using Different Training Algorithms Feed-Forward Neural Network," *Int. J. Electr. Electron. Syst. Res.*, vol. 13, no. 2, pp. 366–375, 2017.

[117] L. Xinyao, L. Huaqiang, W. Yimiao, L. Peiqing, and H. Qiang, "Cascading Failures Forecasting Research to Power Grid Based on Self-Organized Criticality," in *2014 International Conference on Power System Technology*, 2014, pp. 820–825, doi: 10.1109/POWERCON.2014.6993880.

[118] D. Dagur, M. Parimi, and S. R. Wagh, "Prediction of Cascade Failure using Probabilistic Approach with AC Load Flow," in *2014 IEEE Innovative Smart Grid Technologies - Asia (ISGT ASIA)*, 2014, pp. 542–547, doi: 10.1109/ISGT-Asia.2014.6873850.

[119] I. Dobson, J. Kim, and K. R. Wierzbicki, "Testing Branching Process Estimators of Cascading Failure with Data from a Simulation of Transmission Line Outages," *Risk Anal.*, vol. 30, no. 4, pp. 650–662, Apr. 2010, doi: 10.1111/j.1539-6924.2010.01369.x.

[120] S. Gupta, R. Kambli, S. Wagh, and F. Kazi, "Support-Vector-Machine-Based Proactive Cascade Prediction in Smart Grid Using Probabilistic Framework," *IEEE Trans. Ind. Electron.*, vol. 62, no. 4, pp. 2478–2486, Apr. 2015, doi: 10.1109/TIE.2014.2361493.

[121] T. Wang, T. Bi, H. Wang, and J. Liu, "Decision Tree Based Online Stability Assessment Scheme for Power Systems with Renewable Generations," *CSEE J. Power Energy Syst.*, vol. 1, no. 2, pp. 53–61, 2015, doi: 10.17775/cseejpes.2015.00019.

[122] M. Bin Gani and S. M. Brahma, "Dynamic Simulation of the Arizona-Southern California Blackout to Develop a Wide Area Testbed," *2018 North Am. Power Symp. NAPS 2018*, pp. 1–6, 2019, doi: 10.1109/NAPS.2018.8600682.

[123] S. M. Mazhari, N. Safari, C. Y. Chung, and I. Kamwa, "A Quantile Regression-Based Approach for Online Probabilistic Prediction of Unstable Groups of Coherent Generators in Power Systems," *IEEE Trans. Power Syst.*, vol. 34, no. 3, pp. 2240–2250, 2019, doi: 10.1109/TPWRS.2018.2888831.

[124] N. Safari, S. M. Mazhari, and C. Y. Chung, "Very Short-Term Wind Power Prediction

Interval Framework via Bi-Level Optimization and Novel Convex Cost Function," *IEEE Trans. Power Syst.*, vol. 34, no. 2, pp. 1289–1300, 2019, doi: 10.1109/TPWRS.2018.2872822.

[125] A. Khosravi, S. Nahavandi, D. Creighton, and A. F. Atiya, "Lower Upper Bound Estimation Method for Construction Of Neural Network-Based Prediction Intervals," *IEEE Trans. Neural Networks*, vol. 22, no. 3, pp. 337–346, 2011, doi: 10.1109/TNN.2010.2096824.

[126] A. Khosravi, S. Nahavandi, D. Creighton, and A. F. Atiya, "Comprehensive Review of Neural Network-Based Prediction Intervals and New Advances," *IEEE Trans. Neural Networks*, vol. 22, no. 9, pp. 1341–1356, Sep. 2011, doi: 10.1109/TNN.2011.2162110.

[127] F. Yang, A. P. S. Meliopoulos, G. J. Cokkinides, and Q. B. Dam, "Effects of Protection System Hidden Failures on Bulk Power System Reliability," *2006 38th Annu. North Am. Power Symp. NAPS-2006 Proc.*, pp. 517–523, 2006, doi: 10.1109/NAPS.2006.359621.

[128] Y. Zheng, "Methodologies for Cross-Domain Data Fusion: An Overview," *IEEE Trans. Big Data*, vol. 1, no. 1, pp. 16–34, 2015, doi: 10.1109/tbdata.2015.2465959.

[129] J. Robinson, C. Y. Chuang, S. Sra, and S. Jegelka, "Contrastive Learning with Hard Negative Samples," pp. 1–21, Oct. 2020, [Online]. Available: http://arxiv.org/abs/1911.12247.

[130] A. Deb, J. S. Roy, and B. Gupta, "Performance Comparison of Differential Evolution, Particle Swarm Optimization and Genetic Algorithm in The Design of Circularly Polarized Microstrip Antennas," *IEEE Trans. Antennas Propag.*, vol. 62, no. 8, pp. 3920–3928, 2014, doi: 10.1109/TAP.2014.2322880.

[131] Y. Li and S. Wu, "Controlled Islanding for a Hybrid AC/DC Grid with VSC-HVDC using Semi-Supervised Spectral Clustering," *IEEE Access*, vol. 7, pp. 10478–10490, 2019, doi: 10.1109/ACCESS.2018.2886533.

[132] O. Gomez and M. A. Rios, "Real Time Identification of Coherent Groups for Controlled Islanding Based on Graph Theory," *IET Gener. Transm. Distrib.*, vol. 9, no. 8, pp. 748–758, May 2015, doi: 10.1049/iet-gtd.2014.0865.

[133] H. Li, A. Bose, and V. M. Venkatasubramanian, “Wide-Area Voltage Monitoring and Optimization,” *IEEE Trans. Smart Grid*, vol. 7, no. 2, pp. 785–793, Mar. 2016, doi: 10.1109/TSG.2015.2467215.

[134] L. Changchao, K. Zhongjian, Y. Hongguo, R. Kihong, and Z. Shichao, “Identifying the Key Transmission Lines Considering the Power Flow Impact Force,” *IEEE Access*, vol. 7, pp. 96297–96308, 2019, doi: 10.1109/ACCESS.2019.2928702.

[135] Z. Huang, G. Zweig, M. Levit, B. Dumoulin, B. Oguz, and S. Chang, “Accelerating Recurrent Neural Network Training Via Two Stage Classes and Parallelization,” in *2013 IEEE Workshop on Automatic Speech Recognition and Understanding*, 2013, no. 2, pp. 326–331, doi: 10.1109/ASRU.2013.6707751.

[136] M. Dong and L. Grumbach, “A Hybrid Distribution Feeder Long-Term Load Forecasting Method Based on Sequence Prediction,” *IEEE Trans. Smart Grid*, vol. 11, no. 1, pp. 470–482, Jan. 2020, doi: 10.1109/TSG.2019.2924183.

[137] Y. Yuan *et al.*, “Using An Attention-Based LSTM Encoder–Decoder Network for Near Real-Time Disturbance Detection,” *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.*, vol. 13, pp. 1819–1832, 2020, doi: 10.1109/JSTARS.2020.2988324.

[138] D. Lv, A. Eslami, and S. Cui, “Load-Dependent Cascading Failures in Finite-Size Erdös-Rényi Random Networks,” *IEEE Trans. Netw. Sci. Eng.*, vol. 4, no. 2, pp. 129–139, Apr. 2017, doi: 10.1109/TNSE.2017.2685582.

[139] H. Azis, R. D. Mallongi, D. Lantara, and Y. Salim, “Comparison of Floyd-Warshall Algorithm and Greedy Algorithm in Determining the Shortest Route,” in *2018 2nd East Indonesia Conference on Computer and Information Technology (EIConCIT)*, 2018, pp. 294–298, doi: 10.1109/EIConCIT.2018.8878582.

[140] M. Newman, *Networks*, vol. 1. Oxford University Press, 2018.

[141] R. Freedman, “On Gram-Charlier Approximations,” *IEEE Trans. Commun.*, vol. 29, no. 2, pp. 122–125, Feb. 1981, doi: 10.1109/TCOM.1981.1094977.

[142] W. H. Press, S. A. Teukolsky, W. T. Vetterling, and B. P. Flannery, *Numerical Recipes:*

The Art of Scientific Computing. Cambridge University Press, 2007.

[143] M. Abadi *et al.*, "TensorFlow: Large-Scale Machine Learning on Heterogeneous Distributed Systems," *OSDI'16 Proc. 12th USENIX Conf. Oper. Syst. Des. Implement.*, pp. 265–283, Mar. 2016, [Online]. Available: http://arxiv.org/abs/1603.04467.

[144] D. Fernández Slezak, C. Suárez, G. A. Cecchi, G. Marshall, and G. Stolovitzky, "When the Optimal Is Not the Best: Parameter Estimation in Complex Biological Models," *PLoS One*, vol. 5, no. 10, p. e13283, Oct. 2010, doi: 10.1371/journal.pone.0013283.

[145] W. Meng, X. Wang, Z. Wang, and I. Kamwa, "Impact of Causality on Performance of Phasor Measurement Unit Algorithms," *IEEE Trans. Power Syst.*, vol. 33, no. 2, pp. 1555–1565, 2018, doi: 10.1109/TPWRS.2017.2734662.

[146] M. O. Mahgoub, S. M. Mazhari, C. Y. Chung, and S. O. Faried, "A Prediction Interval Based Cascading Failure Prediction Model for Power Systems," in *2020 IEEE Electric Power and Energy Conference (EPEC)*, 2020, vol. 3, pp. 1–6, doi: 10.1109/EPEC48502.2020.9320125.

[147] T.-H. Teng, A.-H. Tan, and J. M. Zurada, "Self-Organizing Neural Networks Integrating Domain Knowledge and Reinforcement Learning," *IEEE Trans. Neural Networks Learn. Syst.*, vol. 26, no. 5, pp. 889–902, May 2015, doi: 10.1109/TNNLS.2014.2327636.

[148] R. K. Akhmadulin, I. N. Gluhih, and I. Y. Karyakin, "An Object-Oriented Model of Case-Based Reasoning System using Situations Tree," in *2016 IEEE 10th International Conference on Application of Information and Communication Technologies (AICT)*, 2016, pp. 1–4, doi: 10.1109/ICAICT.2016.7991668.

[149] S. Guo, W. Zhou, and K. Li, "Multi-layer Case-based Reasoning Approach of Complex Product System," in *2012 Third World Congress on Software Engineering*, 2012, pp. 107–110, doi: 10.1109/WCSE.2012.27.

[150] R. Y. Rubinstein and D. P. Kroese, *Simulation and the Monte Carlo Method*. Wiley, 2016.

APPENDIX A: DEFINITIONS

Table A.1. Definitions and clarifications.

Terminology	Definition and clarifications
Cascading failure	The uncontrolled successive loss of elements in an interconnected system that is initiated by the loss or maloperation of an element or elements. Moreover, all blackouts propagate as a cascading failure, so the term refers to both cascading failures and blackouts.
(Total) Blackout	The loss of power in the entire power system network.
(X% blackout)	The partial loss of power either due to load shedding, disconnection between generation and demand, or generation-demand mismatch. Sometimes this term is referred to as brownout in the literature, but X% blackout is used in this thesis.
Reliability	The system's ability to maintain normal operation following a contingency.
Load shedding	Tripping of the load by the operator (or automatic controllers) to help stop CF propagation or restore the grid to an acceptable operating point. Mainly done to have generation-demand balance, relieving congested transmission lines, or restoring frequency to an appropriate level.
Load curtailment	The loss of load (partially or completely) due to load shedding, improper islanding, disconnection between generation and demand, or not having transmission lines with sufficient capacity.
(Uncontrolled) Islanding	The separation of the grid into one or more islands following a contingency. This is also called improper islanding in some literature.
Controlled islanding	The separation of the grid into one or more islands during a contingency or a CF to isolate the affected regions from the normally operating regions.
Resilience	The system's ability to quickly recover to normal operation following a contingency.
Robustness	The model's ability to operate normally in different operating conditions. More generally, it is the model's ability to maintain its reliability and resilience under different initial conditions and contingencies.

Slow cascade	The slow change in the transmission line temperature that could extend the cascading event by tree contact, overcurrent relays, or other slow dynamics.
Fast cascade	The electrical instability that extends the cascading event by triggering frequency and voltage violations causing generator or load trippings.

Additional information regarding CF terminologies are in [7], [9], [19].

APPENDIX B: DYNAMIC EQUATIONS OF COSMIC

This appendix presents a summary of the dynamic equations used in the open-source MATLAB-based simulation tool used to simulate the CF scenarios in this thesis. The detailed description of COSMIC[1] is in [27]. This summary is only provided for reference.

List of included dynamics and their equations:

Swing equation and rotor angle for each generator in the network:

$$M\frac{d\omega}{dt} = P_m - P_g - D(\omega - 1) \tag{B.1}$$

$$\frac{d\delta}{dt} = 2\pi f_o(\omega - 1) \tag{B.2}$$

Generator model:

$$P_g = \frac{|E'_a||V|}{X'_d}\sin\delta_{m,i} + \frac{|V|^2}{2}\left(\frac{1}{X_q} - \frac{1}{X'_d}\right)sin2\delta_m \tag{B.3}$$

$$Q_g = \frac{|E'_a||V|}{X'_d}\cos\delta_m + |V|^2\left(\frac{\cos^2\delta_m}{X'_d} + \frac{\sin^2\delta_m}{X_q}\right) \tag{B.4}$$

$$\frac{d|E'_a|}{dt} = -|E'_a|\frac{X_d}{T'_{do}} + \left(\frac{X_d}{X'_d} - 1\right)\frac{|V|}{T'_{do}}\cos\delta_m + \frac{E_{fd}}{T'_{do}} \tag{B.5}$$

Exciter model:

$$\frac{d|E_{fd}|}{dt} = \frac{1}{T_E}(K_E.\sigma\left(\left(1 - \frac{T_A}{T_B}\right)E_1 + \frac{T_A}{T_B}(V_{ref} - V_t)\right) - E_{fd}) \tag{B.6}$$

$$\frac{d|E_1|}{dt} = \frac{1}{T_B}(V_{ref} - V_t - E_1) \tag{B.7}$$

[1] Cosmic is a dynamic cascading outage simulator in power systems, developed by Jiajia Song and Eduardo Cotilla-Sanchez (Oregon State University), Paul Hines and Goodarz Ghanavati (University of Vermont). Its code is free to use for research purposes (GNU General Public License V3.0) and is available at: https://github.com/ecotillasanchez/cosmic.

Governor model:

$$\frac{dP_m}{dt} = \sigma\left(\frac{1}{T_t}\left[\sigma\left(P_{ref} - (\frac{1}{R}\Delta\omega + P_3)\right) - P_m\right]\right) \quad \text{(B.8)}$$

$$\frac{dP_3}{dt} = \frac{1}{R.T_i}\Delta\omega \quad \text{(B.9)}$$

where each variable and constant are described according to the Table below.

Table B.1. COSMIC dynamic equation variables.

Variable/Constant	Definition	Variable/Constant	Definition
M	Machine inertia constant (s)	E_{fd}	Excitation voltage (p.u.)
ω	Rotor speed (p.u.)	T_E	Exciter time constant E (s)
P_m	Mechanical power input (p.u.)	K_E	Exciter gain
P_g	Electrical power output (p.u.)	$\sigma(x)$	Sigmoid limiter function
D	Damping constant	T_A	Exciter time constant A (s)
δ	Rotor angle (radian)	T_B	Exciter time constant B (s)
f_o	Synchronous speed (Hz)	E_1	Internal variable of the exciter
E'_a	Transient open-circuit voltage (p.u.)	V_{ref}	Desired reference voltage (p.u.)
V	Bus voltage (p.u.)	V_t	Terminal voltage (p.u.)
δ_m	Mechanical rotor angle (radians)	T_t	Servomotor time constant (t)
X'_d	Direct-axis transient reactance (p.u.)	P_{ref}	Reference Power (p.u.)
X_q	Quadrature-axis reactance (p.u.)	R	Droop
X_d	Direct-axis reactance (p.u.)	P_3	Internal variable of the PI controller
T'_{do}	Direct-axis transient time constant	T_i	PI time constant

APPENDIX C: IEEE 39-BUS SYSTEM INFORMATION

The single line diagram (SLD) of the IEEE 39-bus system is shown in Figure C.. The figure also includes the numbering of the generator buses. Moreover, the numbering of transmission lines and load bus locations are shown in Tables A. 2 and A. 3.

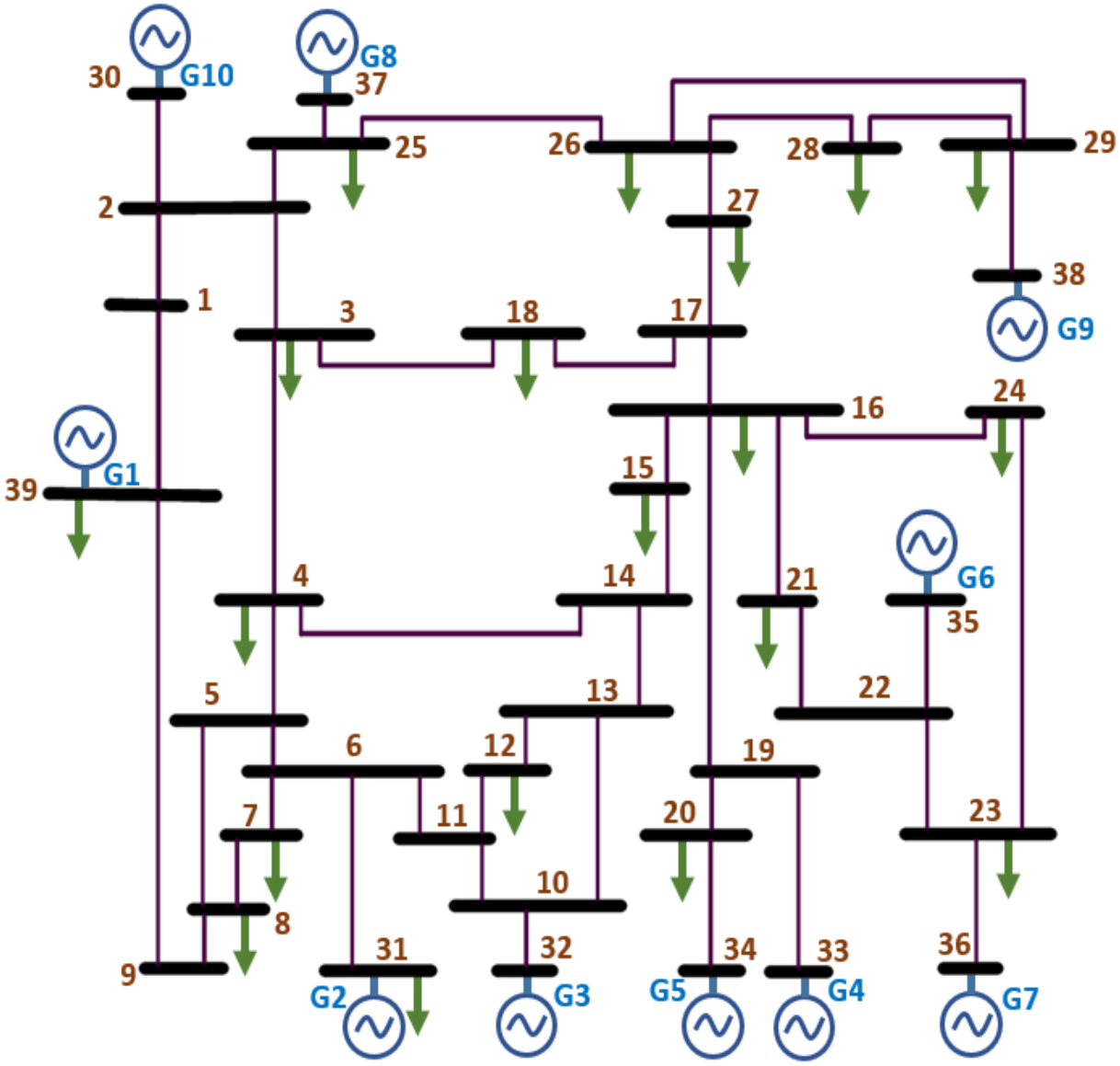

Figure C.1. Single line diagram (SLD) of the IEEE 39-bus system.

Table C.1. Transmission lines' locations.

Line Number	From bus	To bus	Line Number	From bus	To bus
1	1	2	24	16	24
2	1	39	25	17	18
3	2	3	26	17	27
4	2	25	27	21	22
5	3	4	28	22	23
6	3	18	29	23	24
7	4	5	30	25	26
8	4	14	31	26	27
9	5	6	32	26	28
10	5	8	33	26	29
11	6	7	34	28	29
12	6	11	35	12	11
13	7	8	36	12	13
14	8	9	37	6	31
15	9	39	38	10	32
16	10	11	39	19	33
17	10	13	40	20	34
18	13	14	41	22	35
19	14	15	42	23	36
20	15	16	43	25	37
21	16	17	44	2	30
22	16	19	45	29	38
23	16	21	46	19	20

Table C.2. Load bus locations, MW, and MVAR information.

Load number	Bus	MW	MVAR	Load number	Bus	MW	MVAR
1	3	322.0	2.4	11	23	247.5	84.6
2	4	500.0	184.0	12	24	308.6	-92.2
3	7	233.8	84.0	13	25	224.0	47.2
4	8	522.0	176.6	14	26	139.0	17.0
5	12	8.5	88.0	15	27	281.0	75.5
6	15	320.0	153.0	16	28	206.0	27.6
7	16	329.4	32.3	17	29	283.500	26.9
8	18	158.0	30.0	18	31	9.200	4.6
9	20	680.0	103.0	19	39	1104.000	250.0
10	21	274.0	115.0				

APPENDIX D: CASES' INFORMATION SAMPLE

Each case in the IEEE 39-bus system is constructed using the rules and specifications described in Chapter 3. This table is provided for reproducibility and the verification of used probability distribution where the exact numbers for the first 100 scenarios (out of the total 54,756 scenarios) are presented. The columns are as follows: the scenario index (column 1), the first line with forced outage (0 if none) (column 2), the second line with forced outage (0 if none) (column 3), the generator with forced outage (0 if none) (column 4), the location of the fault (column 5), and the relative loading factor of the 19 loads of the system (columns 6-24).

Table D.1. Cases' information sample (the first 100 scenarios in the IEEE 39-bus system).

#	tl1	Tl2	g	f	load1	load2	load3	load4	load5	load6	load7	load8	load9	load10	load11	load12	load13	load14	load15	load16	load17	load18	load19
1	0	0	0	37	1.014	0.899	1.442	1.421	0.667	0.847	1.500	0.641	0.994	1.477	0.747	0.763	0.824	1.467	1.318	0.620	0.759	1.338	0.635
2	0	0	0	13	1.154	1.262	1.220	1.295	1.247	1.176	1.050	1.489	1.170	1.156	0.698	0.606	1.151	1.167	0.765	1.034	1.014	1.039	0.680
3	0	0	0	23	1.325	0.730	1.107	0.981	0.687	0.663	0.626	0.647	0.783	1.273	0.689	0.859	0.779	0.820	0.918	1.135	1.024	1.247	1.455
4	0	0	0	29	1.360	1.455	0.605	1.040	1.166	0.755	1.159	1.124	1.015	1.210	0.819	1.202	1.376	1.475	1.086	0.905	0.675	1.157	0.603
5	0	0	0	21	0.744	1.159	1.370	0.635	0.837	0.798	1.067	1.123	1.270	1.286	1.400	1.123	1.152	0.646	0.891	0.610	0.874	1.205	1.402
6	0	0	0	31	0.704	0.694	1.351	1.425	0.718	1.255	0.820	1.462	1.062	0.741	0.656	1.251	0.611	0.669	0.708	1.231	0.875	0.993	1.259
7	0	0	0	3	1.121	0.812	0.924	0.967	0.858	1.295	0.628	1.357	0.913	1.005	0.643	1.242	1.278	1.476	0.708	1.172	1.445	0.749	1.116
8	0	0	8	32	0.836	0.694	1.140	1.084	0.799	0.837	1.465	0.624	1.482	1.313	1.491	1.129	1.232	1.408	0.844	1.209	0.902	1.352	0.671
9	34	0	0	26	1.234	0.628	0.881	0.600	0.960	0.917	0.839	1.156	1.034	0.606	1.449	1.241	0.907	1.103	1.175	0.632	1.157	0.858	1.072
10	0	0	0	2	1.287	1.274	1.209	0.956	1.201	0.651	1.230	0.928	0.761	1.120	1.290	0.952	0.883	1.215	0.761	0.715	0.622	0.895	1.114
11	0	0	0	33	0.782	1.075	1.172	1.208	1.347	0.850	1.417	0.817	0.899	0.628	0.958	1.254	0.864	0.990	1.054	1.091	1.319	1.395	1.070
12	0	0	0	34	0.902	0.905	1.237	1.229	1.182	1.192	1.348	1.324	0.919	1.257	1.223	1.290	0.703	0.857	1.332	0.671	1.229	0.909	1.047
13	0	0	10	46	0.831	0.919	1.179	1.082	0.747	1.369	0.671	0.667	1.275	0.935	1.198	0.697	1.488	1.099	0.608	1.374	0.741	0.898	1.048
14	0	0	0	22	1.153	0.674	1.494	1.242	1.023	1.299	1.126	1.364	1.447	0.757	1.424	0.633	0.662	0.617	1.429	1.127	1.366	0.640	0.675
15	0	0	0	8	1.287	0.696	1.130	0.600	1.424	1.423	1.441	1.267	0.954	0.977	1.200	0.852	0.670	1.121	1.161	1.312	0.616	0.666	0.944
16	0	0	0	43	0.724	1.478	0.666	0.681	1.474	0.752	1.239	1.024	0.812	1.495	1.482	1.257	0.763	0.790	0.969	0.785	1.005	0.812	1.087
17	0	0	0	43	1.028	1.379	1.060	1.138	0.891	0.882	1.170	0.940	0.741	1.416	1.199	0.801	0.888	0.969	0.705	1.184	1.104	0.719	0.702

Table D.1. Continued.

#	tl1	Tl2	g	f	load1	load2	load3	load4	load5	load6	load7	load8	load9	load10	load11	load12	load13	load14	load15	load16	load17	load18	load19
18	0	0	0	43	0.919	0.981	0.773	1.258	0.972	0.915	1.232	1.162	0.761	1.011	1.256	1.416	1.389	0.945	0.734	0.798	1.275	0.952	1.036
19	11	0	0	2	1.109	0.655	0.749	1.304	0.922	0.810	0.805	0.792	0.985	0.999	0.918	1.229	0.716	1.112	1.208	0.707	1.132	1.348	1.185
20	0	0	0	17	1.103	1.423	1.207	0.659	1.246	0.951	1.102	1.225	1.483	1.439	1.479	1.471	0.997	1.188	0.621	0.966	0.839	0.750	0.788
21	0	0	0	15	0.952	0.886	0.966	0.709	1.425	0.647	0.798	0.965	0.979	1.377	1.057	0.910	1.208	1.236	0.959	0.771	0.745	1.365	1.103
22	11	0	0	28	0.993	1.312	0.987	0.717	1.450	1.218	1.301	0.788	0.806	0.762	0.674	1.231	1.492	0.982	0.627	1.031	1.275	1.227	0.853
23	0	0	0	4	0.794	1.290	0.695	0.700	0.614	1.185	1.235	1.243	1.026	1.360	1.174	1.145	1.208	0.739	1.362	1.032	1.246	1.114	1.046
24	0	0	0	38	0.777	1.208	0.606	0.781	1.182	0.650	1.457	1.463	1.239	1.166	0.674	1.374	0.779	1.489	0.612	0.645	1.162	1.458	1.483
25	0	0	0	21	0.806	0.630	0.986	0.667	0.732	1.440	1.236	0.870	1.323	1.175	0.838	1.482	1.131	0.843	1.066	1.415	0.840	1.203	1.205
26	0	0	0	31	1.390	0.783	1.283	0.805	1.308	0.690	1.138	1.165	1.056	0.880	1.457	1.246	0.672	1.008	0.985	0.702	1.010	1.000	1.268
27	0	0	0	25	0.712	0.800	1.006	1.041	0.756	0.976	0.681	0.635	0.910	1.459	1.220	1.365	0.778	1.216	0.972	0.771	0.824	0.674	1.010
28	0	0	0	12	1.079	1.422	1.234	0.612	1.125	1.141	1.155	1.171	1.124	0.871	1.487	1.412	1.094	1.104	0.674	1.474	1.332	0.988	0.741
29	12	0	0	2	1.284	0.717	1.378	0.851	1.431	0.966	1.388	1.374	1.298	1.253	1.132	1.024	0.861	1.259	0.884	1.404	1.256	0.853	0.790
30	0	0	0	7	1.024	1.419	1.394	0.743	1.205	0.840	0.873	0.919	1.073	1.374	0.868	1.361	1.337	0.825	0.685	0.644	1.477	1.208	0.797
31	0	0	0	35	1.033	1.059	1.169	0.766	0.652	0.965	1.279	0.781	1.176	1.438	0.680	0.995	1.111	1.226	0.960	1.463	0.981	0.873	1.087
32	0	0	0	3	1.344	0.952	0.878	1.237	1.471	1.242	1.064	0.842	0.910	1.166	1.175	1.404	0.628	1.058	1.364	1.187	1.435	0.672	0.639
33	0	0	0	42	1.152	0.837	0.612	1.263	0.826	0.766	1.409	0.644	1.316	1.347	1.340	0.846	0.793	1.098	0.894	1.130	0.717	1.353	1.047
34	0	0	0	16	0.943	1.069	0.806	1.284	1.209	1.370	0.697	1.301	0.801	1.257	0.894	0.611	1.303	0.948	0.960	1.080	1.073	1.020	0.727
35	18	0	0	45	1.498	0.688	0.685	0.732	0.685	1.042	1.145	1.331	1.233	1.005	1.278	0.741	1.294	0.611	0.939	0.779	0.662	0.910	1.237
36	0	0	0	7	1.342	0.701	1.262	0.843	1.007	1.047	0.885	1.380	1.257	0.752	0.608	1.249	1.444	1.240	0.754	1.178	0.916	0.641	1.071
37	0	0	0	32	0.792	0.671	1.350	0.678	1.150	0.815	1.230	1.447	1.336	0.605	0.846	1.435	0.721	0.674	0.715	1.228	0.943	0.736	0.820
38	0	0	0	11	0.687	0.657	1.384	0.871	0.801	0.802	1.088	1.024	1.395	0.732	1.249	0.709	1.448	1.221	1.166	1.465	0.805	0.770	1.297
39	0	0	0	17	1.191	0.972	1.160	1.063	1.360	1.378	0.975	1.049	0.835	1.290	1.199	0.611	1.123	1.397	1.462	0.729	1.369	1.184	0.764
40	0	0	0	6	0.980	0.664	0.823	0.808	1.351	1.000	1.226	1.318	1.260	1.128	0.952	0.713	0.986	0.823	1.435	0.853	1.082	0.971	0.836
41	0	0	0	12	0.930	1.248	1.056	0.657	0.750	0.824	0.675	0.893	1.038	1.286	1.377	0.970	1.196	1.090	1.206	1.288	1.123	0.874	1.207
42	0	0	0	42	0.754	0.908	1.295	0.921	1.151	1.268	0.676	1.409	0.829	0.749	1.252	0.714	1.173	0.741	1.489	1.403	1.298	1.311	1.051
43	0	0	0	43	0.621	1.227	1.169	0.646	1.429	1.244	0.832	1.125	1.178	0.770	1.096	1.402	1.250	1.410	1.185	0.863	0.837	0.740	0.730
44	0	0	0	22	1.196	0.838	0.759	1.122	0.604	1.425	1.100	0.687	1.001	1.277	1.074	0.842	1.363	0.807	1.037	0.802	1.441	1.160	1.007
45	0	0	0	31	1.005	0.706	1.257	1.370	0.621	0.646	1.012	1.208	1.106	1.278	1.492	1.006	1.196	0.978	1.299	1.412	1.275	1.294	0.915

Table D.1. Continued.

#	tl1	Tl2	g	f	load1	load2	load3	load4	load5	load6	load7	load8	load9	load10	load11	load12	load13	load14	load15	load16	load17	load18	load19
46	35	0	0	3	0.741	1.255	0.615	1.303	1.023	0.868	0.640	0.657	1.114	1.226	1.234	0.915	1.164	0.628	1.453	0.619	0.637	0.930	0.741
47	0	0	0	15	1.023	0.856	0.611	0.946	1.369	1.451	0.724	0.624	1.156	1.062	1.182	1.378	0.864	0.861	1.006	1.079	1.137	0.775	1.324
48	0	0	0	35	1.163	0.996	0.763	1.330	1.128	1.407	0.753	0.976	0.782	0.982	1.438	1.427	1.420	1.049	0.826	1.405	0.829	1.350	0.960
49	21	0	0	14	1.444	1.349	0.649	0.821	0.935	1.329	0.607	0.841	1.087	0.656	1.070	0.830	1.365	0.915	0.860	1.158	0.704	0.626	0.869
50	0	0	2	33	0.875	0.989	1.209	1.029	1.093	1.289	1.065	1.331	0.968	1.145	0.930	0.909	1.014	1.226	1.485	0.779	0.848	1.072	0.998
51	0	0	0	26	1.193	1.202	0.930	1.465	1.091	1.148	0.703	1.346	0.960	0.668	1.170	1.378	0.700	0.928	0.850	1.487	0.906	0.997	0.900
52	0	0	0	13	1.292	1.003	0.787	1.073	1.135	0.643	1.196	0.941	0.895	1.388	1.163	0.742	1.170	1.056	0.755	1.087	0.632	0.819	1.493
53	0	0	0	44	0.908	1.459	0.781	0.847	1.340	0.882	0.943	0.969	0.812	1.064	1.173	1.104	0.976	1.472	1.275	1.495	1.143	1.093	0.902
54	0	0	0	10	1.430	0.845	0.684	1.477	1.033	1.493	1.186	0.812	1.191	1.207	0.605	0.994	0.978	1.040	1.146	1.112	0.650	0.619	1.084
55	0	0	0	27	1.364	1.131	0.851	1.119	1.452	1.364	1.443	1.071	0.959	0.929	0.642	0.886	1.427	1.389	0.996	1.167	0.931	1.366	1.002
56	0	0	0	41	1.275	1.312	0.920	0.684	1.076	1.288	0.724	0.863	1.092	0.790	0.760	0.737	1.249	0.903	1.110	0.885	0.859	1.219	1.082
57	26	0	0	5	1.168	1.060	1.173	1.043	1.046	0.932	1.279	1.267	0.927	1.275	0.821	1.088	1.258	1.176	0.736	1.405	0.816	1.307	1.037
58	39	0	0	21	0.836	0.635	1.115	0.765	0.696	1.095	0.674	1.275	0.826	1.060	0.674	0.618	0.702	1.093	0.683	1.275	1.360	1.137	0.869
59	0	0	0	29	1.178	0.842	1.489	1.230	0.741	0.852	1.415	1.227	1.026	1.310	1.330	1.192	1.055	1.200	1.290	1.179	1.088	0.859	0.619
60	0	0	0	24	1.327	1.087	0.750	0.869	0.910	0.746	0.631	1.320	0.796	0.863	0.630	0.738	1.185	1.382	0.932	1.392	1.194	1.330	1.138
61	0	0	0	43	0.643	1.043	0.840	0.755	1.387	1.293	0.982	1.009	1.325	1.185	1.418	1.410	0.758	1.342	0.895	1.482	0.998	0.771	1.132
62	0	0	0	22	0.737	1.346	0.684	1.267	0.792	1.309	1.265	1.288	1.368	1.272	0.778	1.469	0.974	1.410	0.842	1.182	1.367	1.346	0.632
63	0	0	0	23	1.081	0.855	0.940	0.748	0.994	1.060	1.063	1.355	1.111	0.747	1.397	0.762	0.606	0.896	0.798	0.895	1.215	0.866	0.652
64	0	0	0	10	1.330	0.745	1.315	1.242	0.946	0.862	1.206	1.219	1.161	0.921	1.157	0.694	1.241	0.992	0.736	1.286	1.254	0.878	0.807
65	0	0	0	7	1.174	0.807	1.235	0.998	0.779	1.345	1.296	1.450	0.881	1.357	1.196	0.729	0.915	0.725	0.666	0.905	0.724	0.698	0.894
66	0	0	0	2	0.807	1.057	1.129	1.077	1.153	1.070	1.315	0.923	1.330	0.731	1.008	0.906	1.229	0.713	1.189	0.973	0.748	0.994	1.198
67	0	0	0	27	1.390	1.365	0.697	0.702	0.924	1.242	0.942	0.877	1.048	0.935	0.872	0.763	1.338	1.362	0.815	0.619	1.249	1.049	0.919
68	0	0	0	14	0.770	0.620	1.366	0.803	0.811	0.700	1.391	0.676	0.819	0.993	0.716	0.946	0.934	0.725	0.731	1.366	1.001	1.004	0.900
69	0	0	0	2	1.105	1.031	0.977	0.672	0.779	1.296	0.983	1.291	0.794	1.109	0.935	0.612	0.637	0.729	0.605	0.745	0.829	1.065	1.324
70	0	0	0	46	1.421	1.134	0.787	1.220	1.450	1.092	0.774	1.484	0.757	1.439	0.906	1.063	1.422	0.615	1.429	0.952	1.288	0.835	1.036
71	0	0	0	31	1.238	0.839	1.026	0.773	0.995	0.962	0.942	0.688	1.099	0.819	1.112	1.426	1.339	0.642	1.092	0.884	0.600	1.103	0.629
72	7	0	0	31	0.832	0.890	1.123	1.443	1.403	1.202	1.272	0.918	1.239	1.105	1.347	1.045	0.686	1.205	1.467	1.232	0.908	0.807	0.604
73	0	0	0	30	1.247	1.161	1.053	1.344	0.762	0.702	1.046	1.133	0.889	1.281	0.675	1.112	1.246	1.122	1.226	0.680	1.089	0.675	1.197

Table D.1. Continued.

#	tl1	Tl2	g	f	load1	load2	load3	load4	load5	load6	load7	load8	load9	load10	load11	load12	load13	load14	load15	load16	load17	load18	load19
74	0	0	0	35	0.764	1.124	1.419	0.970	0.729	0.844	0.606	1.299	1.214	0.623	1.296	1.404	0.819	1.115	1.095	1.126	0.888	1.044	0.940
75	0	0	0	16	1.482	0.982	1.258	0.952	1.384	1.491	0.806	1.058	0.967	1.446	0.951	0.921	1.087	1.186	0.622	0.631	1.208	1.113	0.696
76	36	7	0	9	1.447	0.952	0.673	0.734	0.644	0.715	1.454	0.779	0.863	0.765	0.632	0.860	1.333	1.488	1.391	1.310	1.374	0.655	0.870
77	0	0	0	12	0.810	1.496	0.784	1.046	0.685	0.722	0.736	0.674	1.364	0.950	0.869	1.004	1.208	0.805	1.232	1.122	1.397	0.992	1.009
78	0	0	0	29	1.021	1.397	1.358	1.241	1.305	0.742	1.306	1.456	0.965	0.789	1.426	0.880	0.690	0.956	0.672	0.739	1.091	0.984	1.172
79	0	0	0	3	1.185	1.000	0.702	1.351	1.238	1.038	0.799	0.893	1.469	0.963	1.464	1.209	1.259	0.751	0.797	0.707	0.819	0.970	0.890
80	0	0	1	36	0.644	0.836	1.436	1.395	0.780	1.300	0.772	1.182	1.448	1.486	0.995	1.372	1.289	1.401	1.048	0.733	0.979	0.717	0.667
81	0	0	0	23	0.697	1.227	0.658	0.751	1.307	0.623	0.610	0.902	1.294	1.211	0.836	1.024	1.396	0.770	0.889	0.720	1.426	0.638	1.147
82	0	0	0	14	0.928	1.451	1.051	0.786	1.263	0.955	0.741	0.821	1.100	0.688	1.187	1.060	1.408	0.732	1.423	0.864	0.905	1.174	0.643
83	0	0	0	43	1.042	0.707	1.327	1.064	1.092	1.083	1.106	0.826	1.019	0.872	0.698	0.967	0.736	0.748	0.742	1.222	1.436	0.605	1.332
84	0	0	0	20	1.074	0.689	1.470	1.186	0.692	1.152	1.251	1.382	1.014	0.755	0.690	1.010	1.280	1.209	0.846	1.367	0.729	0.919	0.787
85	0	0	0	46	0.739	1.247	1.211	1.079	0.622	1.296	0.863	0.959	0.877	0.725	1.224	1.165	0.663	0.940	1.056	1.341	1.409	1.292	0.615
86	0	0	0	39	1.131	0.847	1.343	0.868	0.794	0.764	1.173	0.718	1.423	0.977	1.144	1.054	1.239	0.962	1.374	1.458	0.657	0.673	0.946
87	0	0	0	17	1.031	1.085	1.226	0.888	1.039	1.443	0.852	0.606	0.922	1.001	0.683	0.636	1.222	1.312	1.415	1.106	1.005	1.227	0.808
88	0	0	0	44	0.790	1.461	1.073	0.744	0.790	1.316	0.802	0.872	0.769	0.888	1.110	0.844	1.424	0.694	1.268	1.353	0.756	0.819	0.881
89	0	0	0	29	1.002	0.903	0.655	0.672	1.164	1.050	1.286	1.148	0.705	1.169	1.354	1.176	1.465	0.788	0.997	0.829	1.372	1.254	0.643
90	0	0	0	14	0.805	1.168	1.333	1.425	1.484	1.394	1.217	1.224	1.140	1.074	0.689	1.045	0.756	1.309	1.383	1.129	0.940	1.215	0.740
91	0	0	0	45	1.334	1.310	0.879	0.931	1.184	1.006	1.430	0.930	0.829	0.642	0.782	0.644	1.409	0.680	1.217	0.738	0.661	1.312	1.013
92	0	0	0	44	1.229	1.188	1.349	1.100	1.310	1.075	0.844	0.671	1.013	0.935	0.878	1.301	1.095	0.602	1.083	0.879	1.160	0.759	0.945
93	0	0	0	11	1.418	1.120	0.902	0.966	0.926	1.460	1.344	1.199	0.740	0.732	0.856	1.445	0.697	0.987	0.654	1.289	1.360	0.914	1.271
94	0	0	0	41	0.680	0.727	1.137	0.672	1.163	1.458	1.150	0.836	1.441	0.959	0.770	1.323	1.462	0.863	1.033	0.755	1.064	1.222	0.647
95	0	0	0	8	1.292	1.169	0.754	1.280	0.945	1.462	1.056	1.372	0.934	0.809	0.907	0.984	0.753	0.886	1.045	1.054	1.141	0.945	0.864
96	0	0	0	6	1.282	0.976	0.859	1.099	1.061	0.754	0.670	1.218	1.430	0.753	1.466	0.882	1.262	0.678	1.095	1.283	0.717	1.050	1.214
97	0	0	0	6	1.272	0.794	0.754	1.275	1.211	0.909	1.281	0.737	1.019	0.665	0.851	0.736	0.928	1.211	1.091	0.783	1.078	1.124	0.605
98	0	0	10	36	0.866	1.317	0.777	1.108	1.251	1.090	0.861	0.804	0.752	1.210	0.613	1.170	0.806	1.244	1.185	0.619	0.760	1.072	1.389
99	0	0	0	26	0.622	1.090	0.864	0.645	1.419	1.474	0.604	1.358	1.417	0.878	1.498	1.131	0.678	1.229	1.454	0.963	0.727	1.316	0.961
100	0	0	0	38	0.760	0.853	1.284	1.117	0.618	1.102	1.485	1.383	0.987	0.755	0.627	1.112	1.393	0.633	0.661	1.337	0.879	1.455	0.751

APPENDIX E: ALGORITHMS

Algorithm 1 (Cascading Failure Simulation)

INPUT: desired loading level, fault location(s), line repair state, stopping time, and power system information.

OUTPUT: cascading failure scenario data.

1. **START.**
2. **LOAD** data of the specified power system.
3. **SET** loading level to the desired loading level.
4. **SET** line state to line repair state.
5. **RUN** the power system simulation until it reaches steady state.
6. **IF** system did not converge **OR** became unstable **THEN**:
    7. Raise error flag: system did not converge.
    8. **GO TO END**
9. Apply the fault(s).
10. **WHILE** $t < stopping\ time$:
    11. **RUN** the simulation until stopping time is reached.
    12. **IF** system did not converge **OR** became unstable **THEN**:
        13. Raise state flag: the system had a total blackout.
        14. **END WHILE LOOP.**
15. **OUTPUT** and **STORE** time domain data of the system.
16. **END.**

Algorithm 2 (Generate Random Cascading Failure Scenarios)

INPUT: probability distribution function for loading (PDF_l), probability distribution function for failures (PDF_f), number of cases (s), number of elements (e), number of faults (f), number of load buses (b), and desired probability (p).

OUTPUT: loading level, and fault location(s).

1. **START.**
2. **INITIALIZE** loading level as an $s \times b$ matrix as all zeros.
3. **INITIALIZE** fault location as an $s \times e$ matrix as all **FALSE BOOLEAN**.
4. **FOR** i **FROM** 1 **TO** s:
    5. **FOR** j **FROM** 1 **TO** b:
        6. **SET** loading level(i, j) = random sample from the PDF_l.
    7. **FOR** k **FROM** 1 **TO** e:
        8. **IF** random sample from the PDF > threshold **THEN:**
            9. **SET** fault location(i, k) = **TRUE.**
10. **OUTPUT** and **STORE** cascading failure scenarios system.
11. **END.**

Algorithm 3 (Validate Cascading Failure Scenarios)

INPUT: desired loading level, line repair state, and power system information.

OUTPUT: valid simulation flag.

1. **START.**
2. **LOAD** data of the specified power system.
3. **SET** loading level to the desired loading level.
4. **SET** line state to line repair state.
5. **RUN** the power system simulation until it reaches steady state.
6. **IF** system did not converge **OR** became unstable **THEN**:
    7. **SET** valid simulation flag = **FALSE.**
8. **ELSE:**
    9. **SET** valid simulation flag = **TRUE.**
10. **OUTPUT** and **STORE** valid simulation flag.
11. **END.**

Algorithm 4 (Generate dataset for cascading failure)

INPUT: probability distribution function for loading (PDF_l), probability distribution function for failures (PDF_f), number of cases (s), number of elements (e), number of faults (f), number of load buses (b), desired probability (p), stopping time (t_f), and power system information (ps).

OUTPUT: loading level, and fault location(s).

1. **START.**
2. **RUN** Algorithm 2 (Generate Random Cascading Failure Scenarios).
3. **FOR** i **FROM** 1 **TO** s:
    4. **RUN** Algorithm 3 (Validate Cascading Failure Scenarios) for case i.
    5. **IF** valid simulation flag = **FALSE THEN:**
        6. Discard datapoint.
    7. **ELSE**:
        8. **RUN** Algorithm 1 (Cascading Failure Simulation).
9. **OUTPUT** and **STORE** cascading failure scenarios and time domain simulations.
10. **END.**

Algorithm 5 (Cascading failure prediction framework)

INPUT: all inputs of Algorithm 4 (Generate dataset for cascading failure).

OUTPUT: prediction model, and prediction model accuracy.

1. **START**
2. **RUN** Algorithm 4 (Generate dataset for cascading failure).
3. **SET** RNN input = concatenation of voltage, current, line state, and fault location.
4. **SET** RNN output = loss of load at the end of the simulation ($= 0$ if there is no loss).
5. Apply weighting to RNN input.
6. **RUN** training of the prediction model.
7. **OUTPUT** and **STORE** prediction model and prediction model accuracy.
8. **END.**

Algorithm 6 (Validation loop)

INPUT: accuracy threshold (acc_{Th}), and all inputs of Algorithm 4 (Generate dataset for cascading failure).

OUTPUT: loading level, and fault location(s).

1. **START.**
2. **INITIALIZE** k=0.
3. **RUN** Algorithm 4 (Generate dataset for cascading failure).
4. Train prediction model M[k] using the dataset from 3.
5. **SET** M = M[k].
6. **SET** acc(M[k]) = prediction accuracy of model M[k].
7. **WHILE TRUE**:
    8. **Predict CF using model M.**
    9. Generate new cascading failure cases.
    10. **SET** k=k+1.
    11. Train prediction model M[k] using the dataset from 3 and 8.
    12. **SET** acc(M[k]) = prediction accuracy of model M[k].
    13. **IF** acc(M[k]) >acc_{Th}**:**
        14. **SET** M = M[k].
    15. **ELSE**:
        16. **SET** M = M[k-1].
    17. **OUTPUT** and **STORE** M, M[k], acc(M[k]).
18. **END.**

PERMISSION TO REPRODUCE

A Prediction Interval Based Cascading Failure Prediction Model for Power Systems

Conference Proceedings:
2020 IEEE Electric Power and Energy Conference (EPEC)
Author:
Mohamed O. Mahgoub
Publisher:
IEEE
Date:
9 Nov. 2020

Thesis / Dissertation Reuse

The IEEE does not require individuals working on a thesis to obtain a formal reuse license, however, you may print out this statement to be used as a permission grant:

Requirements to be followed when using any portion (e.g., figure, graph, table, or textual material) of an IEEE copyrighted paper in a thesis:
1) In the case of textual material (e.g., using short quotes or referring to the work within these papers) users must give full credit to the original source (author, paper, publication) followed by the IEEE copyright line © 2011 IEEE.
2) In the case of illustrations or tabular material, we require that the copyright line © [Year of original publication] IEEE appear prominently with each reprinted figure and/or table.
3) If a substantial portion of the original paper is to be used, and if you are not the senior author, also obtain the senior author's approval.

Requirements to be followed when using an entire IEEE copyrighted paper in a thesis:
1) The following IEEE copyright/ credit notice should be placed prominently in the references: © [year of original publication] IEEE. Reprinted, with permission, from [author names, paper title, IEEE publication title, and month/year of publication]
2) Only the accepted version of an IEEE copyrighted paper can be used when posting the paper or your thesis on-line.
3) In placing the thesis on the author's university website, please display the following message in a prominent place on the website: In reference to IEEE copyrighted material which is used with permission in this thesis, the IEEE does not endorse any of [university/educational entity's name goes here]'s products or services. Internal or personal use of this material is permitted. If interested in reprinting/republishing IEEE copyrighted material for advertising or promotional purposes or for creating new collective works for resale or redistribution, please go to http://www.ieee.org/publications_standards/publications/rights/rights_link.html to learn how to obtain a License from RightsLink.

If applicable, University Microfilms and/or ProQuest Library, or the Archives of Canada may supply single copies of the dissertation.