Identification of Cascading Failure Propagation Under Extreme Weather Conditions

ScuDo

Scuola di Dottorato ~ Doctoral School

WHAT YOU ARE, TAKES YOU FAR

Doctoral Dissertation

Doctoral Program in Energy Engineering (30^{th}cycle)

Identification of Cascading Failure Propagation Under Extreme Weather Conditions

By

Pi Renjian

******

Supervisor(s):

Prof. Ettore Bompard , Supervisor

Dr. Marcello Masera, Co-Supervisor

Doctoral Examination Committee:

Prof. Marco Aiello, Referee, University of Stuttgart

Prof. Elmenreich Wilfried, Referee, Alpen-Adria-Universität Klagenfurt

Prof. Marti Rosas-Casals, Universitat Politecnica de Catalunya

Prof. Zofia Lukszo, Delft University of Technology

Politecnico di Torino

2018

Declaration

I hereby declare that, the contents and organization of this dissertation constitute my own original work and does not compromise in any way the rights of third parties, including those relating to the security of personal data.

Pi Renjian
2018

* This dissertation is presented in partial fulfillment of the requirements for **Ph.D. degree** in the Graduate School of Politecnico di Torino (ScuDo).

I would like to dedicate this thesis to my loving parents

Acknowledgements

I wish I could thank every people who helped me during the three years of perusing my PhD at Politecnico di Torino.

Firstly, I would like to express my deep gratitude to Professor Ettore Bompard and Dr. Marcelo Masera, my research supervisors and co-supervisor, for their patient guidance, enthusiastic encouragement and useful critiques of this research work. Thank them to bring me to the fantastic world of electrical engineering.

I would also like to thank Dr. Tao Huang, for his advice and assistance in keeping my progress on the study of Power System Analysis and the programming of Matlab. My grateful thanks are also extended to Dr. Ye Cai, for her some valuable suggestions on the research topic. I am also grateful to all of those with whom I have had the pleasure to work in our group, including Dr. Abouzar Estebsari, Dr. Enrico Pons, Mr. Yang Zhang, Ms. Shaghayegh Zalzar, Mr. Francessco Arrigo. I think I still need to give my special thanks to Mr. Xun Nan and Mr. Chao Lu who are my roommate in Italy. Thank them to cook Chinese food for me.

Assistance provided by all the administrative staff of the department was greatly appreciated. Thanks for the support of Ms. Mariapia Martino, Ms. Lidia Veglia, Mr. Mauro Gregio, Mr. Quarona Franco.

Finally, nobody has been more important to me in the pursuit of my PhD than the members of my family. I would like to thank my parents, whose love and guidance are with me in whatever I pursue.

Abstract

As a fundamental infrastructure, power systems play a vital role in modern society, but it can be damaged by different adverse events e.g. natural, accidental, and malicious, of which the adverse natural events, especially extreme weathers, with huge destructive force can bring tremendous damages and economic losses. The high exposure and comprehensive geographical coverage of the power system make it highly vulnerable to extreme weathers, resulting in equipment damage which leads to cascading failures and blackouts.

Traditional methods only focus on modeling and analysing the reliability of the power system under extreme weathers, without focusing on the propagation of the cascades. In this thesis, innovative methods of studying the cascading failure were proposed, and further extend to collectively consider the impact of extreme weathers on the transmission networks. The proposed models were further validated by applying them to a study system (IEEE-30 bus system) and a real system (Italian transmission network).

A so called normal failure model based on probabilistic graphs was proposed to describe how a cascading failure propagates under a contingency analysis. This model employed Monte Carlo simulation to consider most of the possible operating conditions to establish directed probabilistic graphs to identify the cascading propagation by tripping all branches one by one under each operating condition. Obviously, the results of the model can clearly and legibly show the main cascading path of a given network without considering the initial operating condition and the triggering contingency. Further, an index based on branch vulnerability was designed to select the triggering event to increase the effectiveness of the failure in the simulation.

Furthermore, by integrating a probabilistic model of extreme weather impact into the normal failure model, the extreme weather model was proposed based on failure networks, which maps a physical electricity network into a graph in the cascading

propagation dimensions. Based on the generated failure networks, a new method based on clustering techniques was proposed to fast track the cascading failure path from any initial contingencies without recalculating the cascading failure in the physical network. The high similarity of the simulation results on the IEEE 30 bus system from the two proposed models indicates the validity of the models.

Further, to demonstrate the extreme weather model, we selected a winter storm, which could happen in Northwest of Italy as an example. The data of snowfall on the Alps was collected and modeled by probability density function and probability mass function. By applying the proposed extreme weather model, the propagation paths can be predicted.

The values of the study provide two powerful tools which can 1) clearly present the inherent characteristic of any one given network, i.e. main propagation paths exist regardless of the initial network and failure condition; 2) fast and reasonably predict the cascading paths in a network under extreme weather conditions.

Contents

List of Figures

List of Tables

Nomenclature

Acronyms / Abbreviations

BF Blackout Factor

CDF Cumulative Distribution Function

CF Cascading Failures

CFP Cascading Failure Propagation

CN Child Nodes

CR Changing Rate

ECF criteria of Ending Cascading Failures

EN End Nodes

FO Failure Occurrence

FP Failure Probability

FSL Frequency of Severity Levels

IF Island Factor

IN Initial Nodes

LBVV Transmission Lines which are relating to the Bus Voltage Violation

MC Monte Carlo

ND Neighbour Distance

OTL Overloading Transmission Lines

PD Probability Distribution

PDF Probability Density Function

PMF Probability Mass Function

PN Parent Nodes

RD Residual Degrees

RFL Random Fault Lines

$RMSE$ Root-Mean-Square Error

SN Spread Nodes

TD Total Degrees

TRL Total Removing Lines

VBF Voltage Collapse Factor

Chapter 1

INTRODUCTION

1.1 Literature review of cascading failures

Cascading failures (CF), as the main reason for almost all blackouts, pose severe threats to the security of power systems. In terms of the process of a cascading failure, it usually begins with the initial failure of one or more components, and then the initial failure leads to a sequence of cascading events and finally ends up with a massive power outage. There are many historical records for the blackout caused by the cascading failures. For example, a planned routine disconnection in Northwest Germany caused a cascading failure in European electrical grids and finally affected about 15 millions of European households [1]. Besides technical problem , some adverse/extreme weather can also lead to a cascading failure, such as the blackout that happened in Italy September 2003. Owing to a storm, a cascade disconnection of the transmission line interconnecting North Europe to Italy eventually led to a blackout, and it affected around 60 million people and the disrupted energy reached 180 GWh [2]. Thus, it can be seen that cascading failures have a huge negative impact on the society.

Currently, there are two main methods which have been developed to study cascading failures: **Graph Analysis of Grid Topology** and **Power Flow Based Analysis** [3].

The essential of **Graph Analysis of Grid Topology** is to map a complex network to the power grid topology, which is to convert substations and generators to nodes while convert transmission lines and transformers to links. There are several graph

analysis techniques to study cascading failures including betweenness centrality, small-world network, scale-free network.

Betweenness centrality is a method to measure the centrality for transport flow in a network based on shortest paths, and the means of betweenness centrality is assumed to be the loads on nodes or links of a network [4–8]. If the betweenness centrality of a link or node exceeds a pre-specified critical value, the link or node will be overloaded and removed from the network, then all betweeness centrality will be redistributed [3, 9, 10]. A cascading failure propagates along with the iteration process carries on.

The small-world network is a network model with relative small average path length and relative large clustering coefficient [11]. A cascading failure based on small-world networks assumed that a node will fail if a given value of its neighbours has failed [3, 12–15]. At the beginning, the failure happens on some isolated nodes, then the initial failure will cause subsequent failures as a result of exceeding of the given value [3, 16–18].

The scale-free network has two significant features: 1) the degree distribution follows the power law distribution; 2) some nodes have a large amount of links while most of the node just has a few links [11]. It was found that the network's electrical structure can share some properties with the scale-free network [19]. Furthermore, cascading failures in the scare-free network was investigated by proposing the nodes with high centrality as defence nodes and then to mitigate cascading failure [20–25].

Based on steady-state modeling, **Power Flow Based Analysis** can be performed to study cascading failures by some typical models such as OPA model, CASCADE model, and Manchester model.

OPA model employed DC load flow and LP dispatch of the generation to present cascading line overloads and outages [26–28]. More specifically, when a line fails, generations and loads are re-dispatched by employing the linear programming method [29–31]. The iteration of generator re-dispatch and power flow redistribution can overload further lines [32]. OPA model represents a simple dynamic process of cascading outages.

CASCADE model estimated the failure propagation by considering the overloads of some components after adding an initial disturbance load [26, 33, 34]. To be specific, adding loads to each component as the initial disturbance to cause some

components to fail. As some components fail because of exceeding their loading limit, further components would fail in the subsequent stages.

According to AC power flow, the Manchester model represents a range of cascading failure reciprocal actions [35–37]. The interactions include generator instability, under frequency load shedding, redistribution of active and reactive resources, and cascading outages of transmission lines [35, 38, 39].

1.2 Tasks encountered

The above reviews the state of the art of cascading failures, thus it can be seen that there is little research to investigate the cascading failures under adverse weather conditions. It is not uncommon to model how the weather will impact on power systems, but most existing studies focus on exploring the potential catastrophic consequence of cascading failures on account of the extreme weather. For example, some researchers previously proposed the reliability model under extreme weather conditions [40–42]. Some difficulties to model cascading failures caused by the extreme weather temporarily restrained researchers to have a further insight into this area. The difficulties can be addressed in three parts:

Firstly, how to model the weather. As power grids can be affected by various types of adverse weather such as earthquakes, ice storms, volcanic eruption, landslide, tsunami, etc. It is rarely possible to model a genera asthmatic model to represent different types of adverse weather.

Secondly, to combine the weather model and cascading failure model. The weather model and cascading failure model belong to different areas, so integrating them seems to be not simple. Besides, because of the uncertainty of adverse weather, sometimes it is hard to predict the failures of electrical components resulting from the adverse weather.

Thirdly, the areas impacted by the adverse weather in power grids are not regular. Traditionally, researchers always assumed that the cascading failure began with the vulnerable component. However, when a power grid is affected by natural disasters, then the initial failure of cascading failures will happen randomly. Therefore, predicting the cascading failure propagation (CFP) with random initial failures at the current is not only hard but also demanding.

1.3 Innovations and contributions

In this thesis, we proposed two models to study the cascading failures: one is the normal failure model and another is the extreme weather model. The normal failure model is subject to a cascading failure that happens in the power transmission network under a normal situation(without considering the factor of weather), while the extreme weather model is a model to study cascading failures under the extreme weather(specifically focusing on the snowfall).

The normal failure model took the engineering characteristics of power systems into account to implement vulnerability assessment, and combined the probabilistic graph to illustrate the CFP. Probabilistic graph is a graph model to present the conditional dependence structure between random nodes (variables) [43]. To put it simply, we adopted the Monte Carlo (MC) method, with consideration of the uncertainty of loads and generations, to produce random operating states to identify vulnerable transmission lines in a power system. Based on the most vulnerable transmission line, the probabilistic graph was applied to simulate the CFP. The proposed methodology employed four large samples size to implement Monte Carlo method to identify and compare the component which has the highest failure probability. Compared with other methods, this result could be more accurate and persuadable as it was concluded from a great deal of simulation result. After implementing Monte Carlo, a table which contained a mass of cascading failure chains can be obtained. Statically analysing this table, we can transfer this table into a probabilistic graph to simulate the CFP. This probabilistic graph can help researcher clearly understand how a cascading failure will propagate.

The extreme weather model consisted of a weather (snowfall) model and a partial method of the normal failure model. The reason why to choose the snowfall is because we wanted to apply this extreme weather model on Italian transmission network and studied how the snowfall on Alps will lead to cascading failures on Italian power grids. At the beginning, we used the historical data of snowfalls on Alps to model the probability density function (PDF) of snowfalls. Based on the PDF of snowfalls and a partial method of the normal failure model, we produced a "Failure Network". "Failure Network" refers to a network which combines all possible CFP in a system. Then we established two parameters to reveal how the cascading failure propagates and when the cascading failure stops in "Failure Network". The extreme

weather model not only integrates the weather factor to analysis cascading failures, but it also can predict the CFP with random initial failures.

As mentioned above, the extreme weather model used a partial method of the normal failure model. Fig. 1.1 shows the detailed relation between those two models. Both of those two models belong to cascading failure model and either of them contains four parts. Furthermore, it can be observed that the extreme weather model shares two parts with the normal failure model.

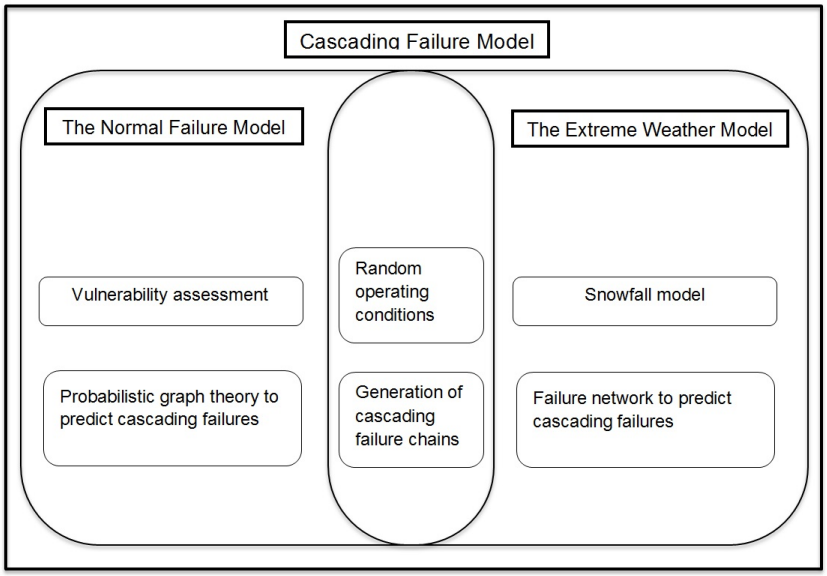


Fig. 1.1 Relation between the normal weather model and the extreme weather model

1.4 Structure of the thesis

The remaining chapters of the dissertation are organised as following:

- Chapter 2 gives a short but comprehensive description about the category of extreme weather and the impact of the extreme weather on power grids. Specifically, this chapter recalls some historical major extreme weather and how they affected the power grids globally.
- Chapter 3 introduces the normal failure model. Starting from the definition of the normal failure model, this chapter then introduces the method of the model, and finally presents how to apply this model to a 30-bus system.

- Chapter 4 describes the extreme weather model based on the normal failure model. This chapter begins with the method of modeling the probability distribution of the snowfall data, and then introduces how to establish the failure networks to display the cascading failure propagation under the extreme weather condition.
- Chapter 5 applies the extreme weather model to the Italian transmission network. It firstly introduces the extreme weather in Italy and the georeferenced model of the Italian transmission network. After that, the process of implementing the extreme weather model on a real transmission network is described in details.
- Chapter 5 applies the extreme weather model to the Italian transmission network. It firstly introduces the extreme weather in Italy and the georeferenced model of the Italian transmission network. After that, the process of implementing the extreme weather model on a real transmission network is described in details.
- Chapter 6 summaries the whole dissertation and briefly introduces how to carry on the research in the future.

Chapter 2

BACKGROUND

2.1 Motivation of this chapter

This chapter introduces the theoretical background for the dissertation, which helps readers understand the remaining chapters. This chapter mainly introduce some basic concepts of two mathematical theories: graph theory and conditional probability which is a part of probability theory. This chapter also mentions the background information of extreme weather. Graph theory is the fundamental theory to establish the normal failure model and the extreme weather model. In terms of conditional probability, it is used to establish the probabilistic graph in the normal failure model. Knowing the basic information of extreme weather can be helpful to understand the extreme weather model.

2.2 Graph theory

2.2.1 Brief introduction

The performance of many physical systems depends not only on the characteristics of the components, but also on the locations of the components. For example, in a structure, if the location of a component is relocated, the structure's properties will become differently. Therefore, the topology of the structure affects the performance of the whole structure. As a result, it is significant to represent a system by using a

graph model so that its topology can be clearly understood [44]. Graph Theory is an important area of Discrete Mathematics, which was used to model pairwise relations between objects [45]. Graphs are very simple to explain discrete structures, but are useful for a basic functional structure. This theory can be used to model numerous discrete things. For instance, a collection of computers and communication links between the machines, and the relationship between each user of Facebook.

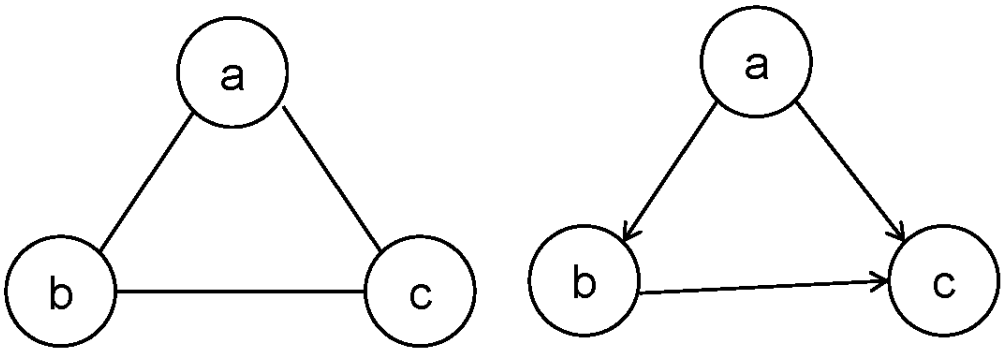
2.2.2 Basic concepts of graph theory

A graph G consists of a finite set V of elements called vertices and a set E of elements called edges [45]. The graph can be represented by Equation (2.1). in Equation (2.2), u and v are two vertices of the graph.

$$G = (V, E) \tag{2.1}$$

$$E = \{\{u, v\} : u, v \in V\} \tag{2.2}$$

There are two main types of graphs: one is undirected graph and another is directed graph. An undirected graph is a graph whose edges have no direction while a directed graph defines as a set of vertices that are connected and all the edges are directed from one node to another [46]. Fig. 2.1 displays the examples of undirected graph and directed graph.



(a) An undirected graph (b) A directed graph

Fig. 2.1 Examples of undirected and directed graphs

Taking Fig. 2.1a for instance, V = { a, b, c} and E = { {a, b}, {a, c}, {b, c}}. If e = { a, b} ∈ E, it means the edge e connects the vertex a and b. Furthermore, a and b are called the endpoints of edge e. In a directed graph like Fig. 2.1b, e = { a, b} and e = { b, a} cannot represent a same connection between two vertices.

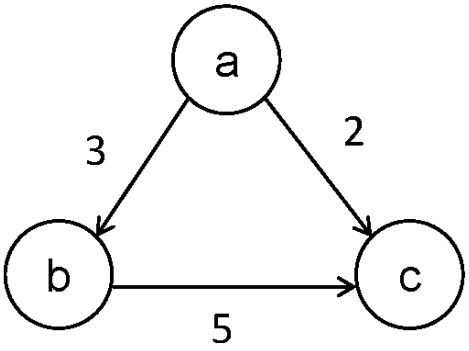


Fig. 2.2 An directed graph with weights

Sometimes, edges of the graph are associated with weights. With each edge of a graph let there be associated a real value w(e), called its weight [47]. Those weights can represent capacity, length etc. of the connection between two vertices [48].The graph with weights is called the network, and it can be defined as Equation (2.3), where w represents the weight of a edge. Fig. 2.2 displays the example of the directed graph with weights.

$$G = (V, E, w) \tag{2.3}$$

Considering Fig. 2.2, where V = { a, b, c} and E = { {a, b}, {a, c}, {b, c}}. The weight of each edge can be understood by Table 2.1.

Table 2.1 Weight of each edge in Fig. 2.2

edge e ∈ E	{a, b}	{a, c}	{b, c}
weight w(e)	3	2	5

There are some basic and important concepts in the graph theory such as degree, path, cycle, acyclic and tree. The degree of a vertex is the total number of of edges connecting to the vertex [47]. The degree of a vertex is denoted as deg(v). If the degree of a vertex equals to zero, the vertex is called an isolated vertex. In case of

the directed graph, the degree can be classified into two types: in-degree (degin(v)) and out-degree (degout(v)) [49]. In terms of in-degree and out-degree, they can be defined as the total number of edges incoming to a vertex and the total number of edges outgoing to a vertex respectively [47]. The relationship between degree, in-degree and out-degree can be described as Equation (2.4).

$$\deg(v) = \deg in(v) + \deg out(v) \tag{2.4}$$

Path in a graph means a sequence of unrepeatable vertices such that two vertices are adjacent [50]. The first vertex of a graph is called start vertex whereas the last vertex is called end vertex [50]. Those two vertices are called terminal vertices of a path. In a path, if the start vertex equals to the end vertex, the sequence is defined as a closed path. A closed path which has the distinct edges and vertices (except that the start vertex is the same as the) end vertex) is called a cycle. For example, in Fig. 2.3, (a, b, c) is a path whereas (a, b, c, d, a) is a cycle

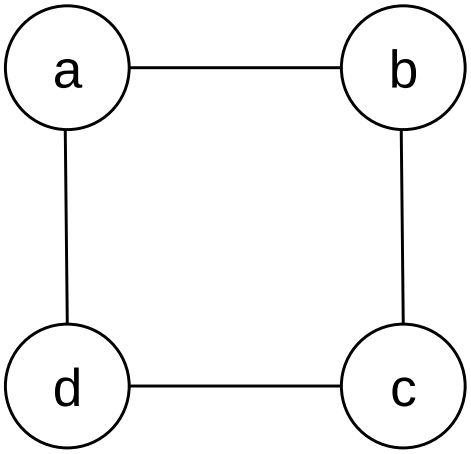


Fig. 2.3 Graph to explain path and cycle

.

There is no cycles in a graph and this graph can be called a acyclic graph. Additionally, a connected acyclic graph can be called a tree [50]. Fig. 2.4 shows an example tree.

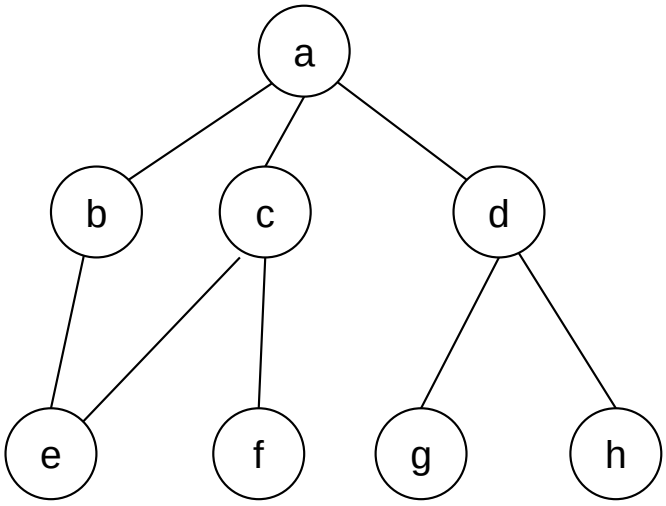


Fig. 2.4 A tree example

.

2.3 Conditional probability

2.3.1 Brief introduction

The probability theory handles patterns that occur in random events [51]. Probability theory is useful in many areas such as the physical, economics, management, computer sciences, etc. It can be used to model complex systems and make decisions when there is uncertainty. Moreover, It also helps in proving theorems in other mathematical fields including but not limited to graph theory, game theory, communications theory, etc. Classical probability applies in situations where there are just a finite number of equally likely possible consequences [52]. However, conditional probability answers the question that how the probability of an event changes if the extra information is previously obtained.

2.3.2 Independent and dependent events

When two events are independent of each other, which means that the probability that one event happens in no way influences the probability of the other event occurring

[53]. Simply, the independent events are not possibly affected by previous events. A typical example of two independent events is the coin tossing. The probability of obtaining any number face on the die in no way affects the probability of getting a tail or a head on the coin.

If A and B are independent evens, then the probability of both events happening can be described as Equation (2.5).

$$P(A \cap B) = P(A) \times P(B) \tag{2.5}$$

When two events are dependent, the probability of one event happening affects the likelihood of the other event. In other words, dependent events are affected by previous events. The typical example is the to pull out marbles from a bag. Supposing there are 2 red marbles and 2 blue marbles in a bag. Firstly, pulling out one marble (might be blue or red). Now only 3 marbles are left in the bag. What is the probability that the second marble will be red? From this example, it can be understood that the outcome of the first affects the outcome of the second if two events are dependent.

If A and B are dependent evens, then the probability of both events occurring can be described as Equation (2.6) [53]. In Equation (2.6), P(B|A) means probability of event B given event A.

$$P(A \cap B) = P(A) \times P(B|A) \tag{2.6}$$

2.3.3 Concepts of condition probability

Conditional probability is the probability of one event occurring with some relationship to another event or more other events [54]. If the interest is A and the event B is assumed to have occurred, the situation can be described as the conditional probability of A given B, and it is usually written as P(A|B). the conditional probability of event A given B can be defined as Equation 2.7.

$$P(A|B) = \frac{P(A \cap B)}{P(B)} \tag{2.7}$$

If event A and event B are independent events, Equation (2.7) can be extended to Equation (2.8).

$$P(A\,|B\,) = \frac{P(A \cap B)}{P(B)} = \frac{P(A) \times P(B)}{P(B)} = P(A) \tag{2.8}$$

If event A and event B are dependent events, Equation (2.7) can be extended to Equation (2.9).

$$P(A\,|B\,) = \frac{P(A \cap B)}{P(B)} = \frac{P(A) \times P(B\,|A\,)}{P(B)} \tag{2.9}$$

2.3.4 Law of total probability

In probability theory, the law of total probability is a fundamental rule to the conditional probability. The law of total probability is a solution to calculate the likelihood of an event whose occurrence is influenced by several other disjoint events [55]. It can be defined as Equation (2.10) or Equation (2.11).

$$P(A) = \sum_{n} P(A \cap B_n) \tag{2.10}$$

$$P(A) = \sum_{n} P(A\,|B_n\,) \times P(B_n) \tag{2.11}$$

2.4 Extreme weather against power systems

2.4.1 Brief introduction

Various definitions of extreme weather have been proposed. It was defined as all atmospheric, hydrologic, geologic and wildfire phenomena that, has the potential to affect the human health, activities or even the constructions [56]. Some researchers suggested that they referred to weather phenomena that are at the extremes of the historical distribution and are rare for a particular time and/or place, especially unseasonal or severe weather [57]. In this thesis, considering the involvement of power systems, we define extreme weather as a potential event or a set of events happened around the world with different scales (local, national, continental) and different short time frames (instantaneously, minutes, days), and it is not directly

involved by human's activities but it would affect the operation of power systems with a large scale disruption of electricity supply.

2.4.2 Classification of extreme weather

Based on the definition of extreme weather, we proposed a classification for studying their impact on electricity infrastructures. Extreme weather can be classified into three categories which are hydrology, meteorology, and climatology. Fig. 2.5 shows the classification of the most typical extreme weather against the secure operation of power systems.

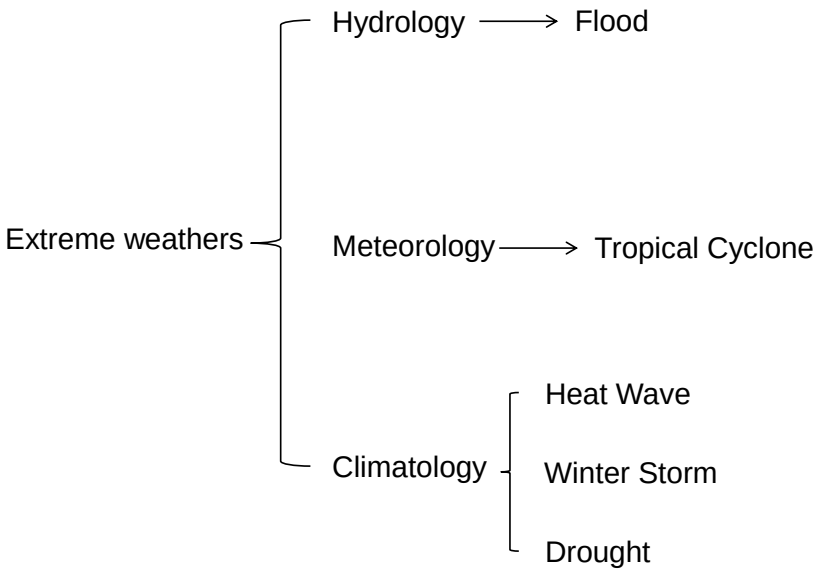


Fig. 2.5 Classification of extreme weather

It should be noted that the difference between "meteorology" and "climatology" mainly lies in the time perspective. Meteorology is intended to indicate the weather conditions over the short-term while climatology employs a long-term perspective [58]. With references to Fig. 2.5, a flood is a land covered by water that is not usually covered by water [59]. A tropical cyclone is a rotating, organized system of clouds and thunderstorms that originates over tropical or subtropical waters and has a closed low-level circulation (hurricane, tornado, and typhoon are the same phenomena in different places) [60]. A heat wave is an extended period of extreme heat and is often accompanied by high humidity [61]. A winter storm is an event in which the main

types of precipitation are snow, sleet or freezing rain [62]. A drought is a lengthy period of time, stretching months or even years in which time land has a decrease in water supply [59].

2.4.3 Impact of extreme weather on power systems

Bulk power systems are easily threatened by extreme weather due to the large exposure to the environment. However, the components of a power system such as generator, transformer, substation, overhead line, cable, control center, etc., can be affected differently by extreme weather [63]. In general, tropical cyclones primarily affect transmission and local distribution systems, while floods could damage the generating equipment as well. Although heat waves and droughts normally cannot straightforwardly destroy elements of the power system, unless the temperature reaches extremely high, they can significantly increase the cooling/air conditioning consumption while decrease the generation capacity of hydrological power plants. However, under extreme situations, flood can damage or put out of service the underground substations or control centers, whereas heat waves may trip transformers or overhead lines due to the temperature protection or short circuits caused by the elongation of the wires. In terms of the winter storm, the ice may result in the damage of transmission towers or short circuits in substations. Table 2.2 shows the potential impact on power systems from almost all common extreme weather, which is based on the assumption that all the events are serious enough.

Table 2.2 Threat and impact of extreme weather

		Flood	Tropical Cyclone	Heat Wave	Winter Storm	Drought
Threats	Generator	H	H	H	M	M
	Transformer	H	H	N	M	N
	Substation	N	H	N	N	N
	Overhead line	M	H	N	H	N
	Cable	N	N	N	N	N
	Control centre	N	H	N	N	N
Possible impact	Equipment damage	✓	✓		✓	
	Short circuit	✓	✓		✓	
	Overload			✓	✓	✓
H:Huge impact M: Moderate impact N:Negligible impact						

Table 2.3 shows some historical instances of extreme weather and how they brought the huge loss to power systems. The tropical cyclones happened in 1982 and 1992 are two of the most severe natural disasters which posed a huge threats to America's power grid. The total economic loss reached to about 2 G$, and the failure of critical equipment of power system caused blackout for a few days [64]-[65]. The winter storm happened on 1st April 1998 in North American caused 7.4 G$ of economic loss, 980 casualties, 5 million people without electricity up to one month [66]. In addition, more than 3.5 thousand poles, 5 thousand transformers, and 1.3 thousand steel pylons were in need of repair [67]. A severe heat wave happened in Europe in June 2003 and continued through July until mid-August, which raised summer temperatures by 20% to 30% higher than the seasonal average temperature [68]. It also affected a large area which was extended from northern Spain to the Czech Republic [68]. With a death toll more than 30 thousands, the heat wave of 2003 become one of the most serious natural disasters in Europe for the last 100 years, and it ended up with a great financial loss at 14.5 G$ [69]. Moreover, four nuclear power plants were forced to shut down because of the dramatic rise temperature of rivers used to cool the reactors, which engendered to loss 4 GW power generation during the summer of 2003 [69]. Although there was no direct

damage to the infrastructures of power systems, electricity demands dramatically soared due to the cooling loads.

Table 2.3 Major extreme weather against power systems

Country	Type	Time	Affected power subscriber (M)	Affected power Transmission line (miles)	Affected power plant	Lossing power (GW)	Duration of power outage (day)
USA	Tropical cyclone	19/11/1982	0.23				30
USA	Tropical cyclone	11/9/1982		280			90
China	Winter storm	10/1/2008	0.054	6500		6209	28
North America	Winter storm	4/1/1998	5	1850			30
Europe	Heat wave	1/6/2003			4	4	
UK	Flood	25/06/2007	0.13		2	10	5
Australia	Tropical cyclone and flood	17/1/2013	0.3		133.8		

A more complicated situation is that sometimes blackouts are attributed to combinations of multiple natural disasters. They are more complex than a single one as they often involve cascading events. Explanatorily, the primary failure triggers a sequence of secondary failures and failures of propagation will finally lead to a blackout in a large area. The event that happened on 17th January 2011 in Australia was a typical instance. A tropical cyclone was first identified in the Gulf of Carpentaria on January 17th, but it became destructive wind and produced over 1000 mm of rainfall in some areas during 48 hours. A major flooding took place within the following weeks caused 133.8 GW power outages, l.7 G$ economic loss and great impact on 0.3 million power customers [70].

2.4.4 Discussion

Although no single failure will have a significant effect on the electricity system considering most utilities maintain sufficient generation and transmission reserves to withstand such failures, if an extreme event happens, it will bring on several damages on substations, the transmission system and even loads would suffer from a great loss. Extreme weather, as the unneglectable risks for power systems, happen rarely but their occurrences will lead to enormous losses for the society.

The large geographical exposure of power systems in the natural environment indicates the vulnerability of power grids when facing destructive adverse natural events. They do not only directly damage power facilities, but may also lead to blackouts through cascading failures. This would further result in huge economic loss and a large number of casualties. The gradually increasing population and electricity demand have a tendency to ascend losses for each country around the world. The extreme weather recorded in the history already alerted humans that the huge damage can pose apparent and considerable threats to the power systems, so it is important to understand the effects of extreme weather and develop proper approaches and tools to efficiently reduce the negative consequences of extreme weather. To improve the security of a power grid, a possible way to achieve this is to consider the impact of extreme weather and model the cascading failure in power grids under the extreme weather. Therefore, it is necessary to have appropriate approaches and models to work on the influences of extreme weather on the power systems.

Chapter 3

CASCADING FAILURES IN THE NORMAL FAILURE MODEL

3.1 Brief introduction

The traditional way to predict the CFP is firstly to implement vulnerability assessment, and then based on the vulnerable part as the starting failure to simulate the CFP. The essential idea of the normal failure model is still unchanged, but the innovation of normal failure model is to adapt the probabilistic graph theory to predict the CFP.

We take the engineering characteristics of power systems into account to implement vulnerability assessment, and combine the probabilistic graph to illustrate the cascading failure propagation. In simple terms, we adopted the Monte Carlo method, with consideration of the uncertainty of loads and generations, to produce random operating states to identify vulnerable transmission lines in a power system. Based on the most vulnerable transmission line, the probabilistic graph was applied to simulate the cascading failure propagation. The proposed methodology employed four large samples size to implement Monte Carlo method to identify and compare the component which has the highest failure probability. Compared with other methods, this result could be more accurate and persuadable as it was concluded from a great deal of simulation result. After implementing Monte Carlo, all cascading failure chains can be obtained. Statically analysing those chains, we can transfer this result into a probabilistic graph to simulate the cascading failure propagation. This

probabilistic graph can help researcher clearly understand how a cascading failure will propagate.

The scheme based on MC is displayed in Fig. 3.1. The initial step is to select a power system condition which can withstand the loss of any one electrical component. The next step is to ensure the sampling size (K) of MC, and use the original power system condition to produce random operating states. The following step is to choose a random operating condition to implement "N-1". After implementing all random conditions, the final step is to assess the vulnerability and simulate the CFP.

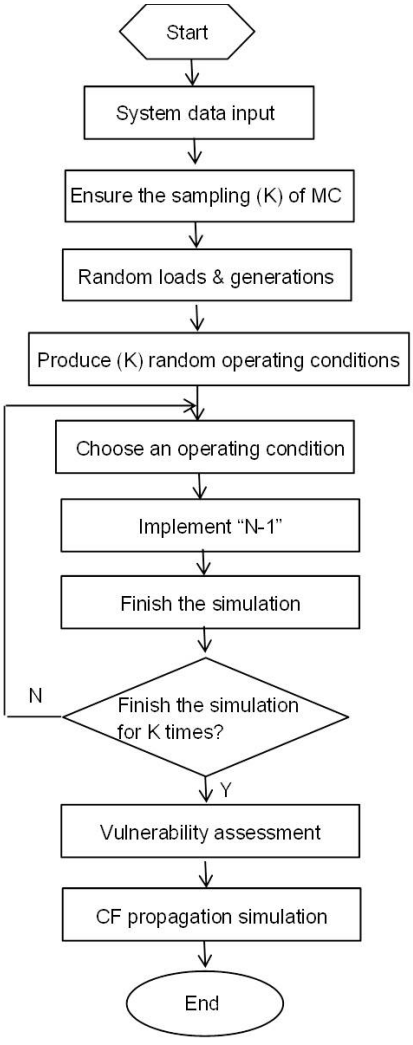


Fig. 3.1 Flow Chart of the normal failure model

3.2 Methodology of the normal failure model

3.2.1 Random operating conditions

With consideration of the uncertainty of loads and generations, the model producing random operating conditions was established by the combination of MC and ZIP model. Specifically, ZIP are the coefficients of a load model which consists of constant impedance, constant current, and constant power loads [71]. The coefficients *Z, I, P* were determined by experiments for modern residential, commercial, and industrial loads [72]. Fig. 3.2 presents details of the model to simulate random operating conditions.

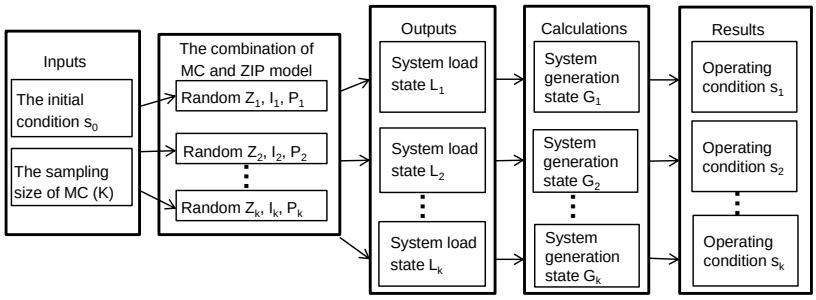


Fig. 3.2 The model to simulate random operating conditions

To simulate random operating states, an initial system condition (s_0) and the sampling size (K) of MC are firstly given as inputs, then randomly selecting ZIP coefficients (Z_k, I_k, Z_k) to produce different system load states (L_k). Based on a system load state, dispatching the active power balance of the system to each generator and then a system generation state (G_k) can be produced. After ensuring a system load state and a system generation state, an new operating condition (s_k) can be acquired. Therefore, the operating conditions are defined, in this thesis, as the combination of system load states and system generation states. A system load state is defined as a vector represented consuming active powers of each load node in a power system, while in a similar way, a system generation state refers to a vector which contains active powers of each generator node. Their relation is shown by Equation (3.1):

$$s_k = [L_k, G_k] \tag{3.1}$$

where
k : the k-th random time of MC
s : an operating state
L : a system load state
G : a system generation state

The set of operating states ($\mathscr{S}$) can be represented by a collection of different random operating states (s_k) as:

$$\mathscr{S} = \{s_1, s_2, ..., s_k\} \tag{3.2}$$

$s.t.$

$$|\mathscr{S}| = K \tag{3.3}$$

where
k : the k-th random time of MC
K: the sample size of MC
s : an operating state
$\mathscr{S}$: the set of operating states

Furthermore, a system load state is expressed mathematically as Equation (3.4). Combining all system load states, the matrix of system load states can be obtained as Equation (3.6). Similarly, a system generation state and the matrix of system generation states can be defined as Equation (3.5) and Equation (3.7):

$$L_k = [l_{k1}, l_{k1}, ...l_{ki}] \tag{3.4}$$

$$G_k = [g_{k1}, g_{k1}, ...g_{kj}] \tag{3.5}$$

$$ML = \begin{vmatrix} L_1 \\ L_2 \\ \vdots \\ L_k \end{vmatrix} = \begin{vmatrix} l_{11} & l_{12} & \dots & l_{1i} \\ l_{21} & l_{22} & \dots & l_{2i} \\ \vdots & \vdots & \ddots & \vdots \\ l_{k1} & l_{k2} & \dots & l_{ki} \end{vmatrix} \tag{3.6}$$

$$MG = \begin{vmatrix} G_1 \\ G_2 \\ \vdots \\ G_k \end{vmatrix} = \begin{vmatrix} g_{11} & g_{12} & \cdots & g_{1j} \\ g_{21} & g_{22} & \cdots & g_{2j} \\ \vdots & \vdots & \ddots & \vdots \\ g_{k1} & g_{k2} & \cdots & g_{kj} \end{vmatrix} \tag{3.7}$$

where
k: the k-th random time of MC
i : the index of load nodes
j : the index of generator nodes
l : the consumed active power of a load node
g : the active power of a generator node
L : a system load state
G : a system load state
ML : the matrix of system load states
MG : the matrix of system generation states

3.2.1.1 System load states

As mentioned above, a system load state is based on the consumed active powers of each load node. In order to generate the random consumed active powers of a lode node, ZIP load model was applied in our research. The coefficients Z, I, P were determined by experiments for modern residential, commercial, and industrial loads [72]. Table 3.1 presents the range of each coefficient (Z, I, P) with different types of loads.

Table 3.1 The range of ZIP coefficients

	Residential Loads		Commercial,Loads		Industrial,Loads	
	Minimum	Maximum	Minimum	Maximum	Minimum	Maximum
Z	0.96	1.57	0.27	0.77	1.21	1.21
I	-2.49	-1.17	-0.84	0.24	-1.61	-1.61
P	1.21	1.93	0.21	1.07	1.41	1.41

According to ZIP coefficients, the loads can be estimated by [72]:

$$l = l_0[Z(\frac{V}{V_0} + I \times \frac{V}{V_0} + P)] \tag{3.8}$$

where
V : operating voltage
V_0 : rated voltage
l : active powers of a load node at operating voltage

l_0 : active powers of a load node at rated voltage
Z, I, P : coefficients for the ZIP model

In this thesis, we assumed that all load nodes contained three types of loads (residential loads, commercial loads and industrial loads). Moreover, the proportions of those three loads at each load node were randomly generated. Based on Equation (3.8), we provided a new method to obtain random active power of each load node as shown in Equation (3.9) and Equation (3.10):

$$\begin{aligned} l_i = & l_{0_i}\{(\alpha_i \times Z_{RL_i} + \beta_i \times Z_{CL_i} + \delta_i \times Z_{IL_i}) \\ & \times [\frac{V_i}{V_{0_i}} + (\alpha_i \times I_{RL_i} + \beta_i \times I_{CL_i} + \delta_i \times I_{IL_i}) \times \frac{V_i}{V_{0_i}} \\ & + (\alpha_i \times P_{RL_i} + \beta_i \times P_{CL_i} + \delta_i \times P_{IL_i})]\} \end{aligned} \tag{3.9}$$

$$s.t.$$

$$\alpha_i + \beta_i + \delta_i = 1 \tag{3.10}$$

where
i : the index of load nodes
l : the active powers of a load node at the current stage
l_0 : the active powers of a load node at the initial stage
V : the operating voltage
V_0 : the rated voltage
Z_{RL}, I_{RL}, P_{RL} : coefficients for the residential loads
Z_{CL}, I_{CL}, P_{CL} : coefficients for the commercial loads
Z_{IL}, I_{IL}, P_{IL} : coefficients for the industrial loads
α, β, δ : proportions of three types of loads

Considering the uncertainty of loads is not enough, but the correlation of each load should also be taken into consideration. After employing ZIP model, the matrix of system load states (ML_{ZIP}) can be obtained, but this step can only solve the "uncertainty issue". In order to indicate the correlation between each load, a random correlation matrix (M_{rc}) was generated, which presented the correlation between each load node. Then Cholesky Decomposition [73] was used to decompose the random correlation matrix to obtain the lower triangular matrix (U). Finally, (U) multiplied (ML_{ZIP}), then the final matrix of system load states (ML) can be finally obtained. Equation (3.11) and Equation (3.12) explain the method to get ML which

can reveal the uncertainty and the correlation of loads:

$$M_{rc} = U^* U \tag{3.11}$$

$$ML = ML_{ZIP} \times U \tag{3.12}$$

where
M_{rc} : the random correlation matrix (this matrix should be a positive definite matrix)
U : the lower triangular matrix
U^* : the conjugate transpose of U
ML_{ZIP} : the matrix of system load states after employing ZIP model
ML : the final matrix of system load states which can reveal the uncertainty and the correlation of loads

3.2.1.2 System generation states

According to Fig. 3.2, it can be understood that a system generation state depends on a system load state. When a new random system load state is produced, the unbalance active power might exist, and it equals to the difference between the total consumption of the new random system load state and the total generation of the initial generation state. Dispatching Δp based on that assuming the droop of generator is around $4\% - 5\%$ [74] and considering both the residual power and the active reserve for the primary control [75]. The method is defined as:

$$\Delta p = \sum_{i=1}^{I} l_i - \sum_{j=1}^{J} g_{0_j} \tag{3.13}$$

$$g_j = g_{0_j} + \Delta p * \frac{\min\{g_{\max_j} - g_{0_j}, 10\% g_{\max_j}\}}{\sum_{j=1}^{J} \min\{g_{\max_j} - g_{0_j}, 10\% g_{\max_j}\}} \tag{3.14}$$

where
j : the index of generator nodes
i : the index of load nodes
Δp : the unbalance active power
g : the generations of a generator node after dispatching Δp
g_0 : the generations of a generator node at the initial stage
l : the consumed active power of a load node

$g_{\max}$: the maximum allowable power output of a generator
J : the total number of generator nodes
I : the total number of load nodes

3.2.2 Generation of cascading failure chains

A CF begins with an initial failure and the initial failure will further lead to the transmission line overload and\or the bus voltage violation in subsequent stages, and eventually to cause the failure of the entire grid. Overloads are determined by the power flow on the transmission line. If the power flow on a transmission line exceeds 10% of its capacity, then we ensure there is overload. In terms of the voltage violation, we assume the acceptable voltage range from 0.9 (p.u.) to 1.1 (p.u.). If the bus voltage is beyond this limit, then it is a voltage violation.

Basically, components of each stage could contain two parts: overloads and\or bus voltage violations, but in this thesis, the components of each stage are simplified to one indicator which is the total removing lines (TRL). Therefore, TRL actually consists of two parts: one is the overloading transmission lines (OTL) and another is the transmission lines which are related the voltage violations (LBVV).

$$TRL = [OTL, \text{LBVV}] \tag{3.15}$$

where:
TRL : total removing lines
OTL : overloading transmission lines
$LBVV$: transmission lines which are connected to the bus when it happens voltage violations

Another indicator which is the the criteria of ending cascading failures (ECF) also needs to be explained, as the propagation depends on two indicators: ECF and TRL. ECF is composed of three factors: Island, blackout and voltage collapse, and it is expressed as Equation (3.16). If there is more than one island in a power system, IF equals to 1. Otherwise, it equals to 0. Similarly for BF and VCF, they equal to 1 respectively when there is a blackout or a voltage collapse (the method to determine the voltage collapse was based on the singularity of power flow Jacobian matrix [76]).

Fig. 3.3 presents how the two indicators, ECF and TRL, affect the propagation.

$$ECF = IF + BF + VCF \tag{3.16}$$

$$s.t.$$

$$IF \in \{0,1\} \tag{3.17}$$

$$BF \in \{0,1\} \tag{3.18}$$

$$VCF \in \{0,1\} \tag{3.19}$$

where
ECF : the criterion of ending cascading failures
IF : the island factor
BF : the blackout factor
VCF : the voltage collapse factor

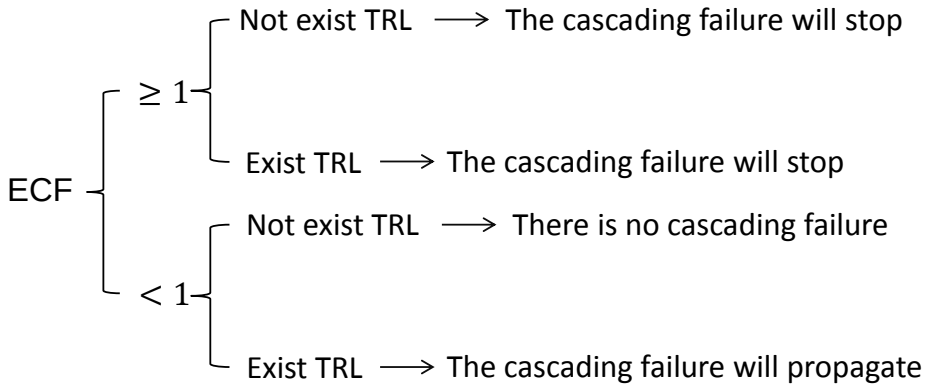


Fig. 3.3 How ECF and TRL affect the cascading failure propagation

Therefore, the chain of a cascading failure will be terminated when ECF is not less than 1. To describe a CF chain, two important issues should be considered: 1) how many stages of the CF there are and 2) what the components of each stage are. Equation (3.21) was used to present the CF chains. d was employed to represent the stages of a CF and the total removing lines (TRL) was proposed to present the components of each stage.

$$F_{kt} = [f_{k1}, f_{k2}, ..., f_{kd}] \tag{3.20}$$

$$s.t.$$

$$f_d = TRL_d \tag{3.21}$$

where
k : the k-th random time of MC
t : the index of transmission lines
d : the stage index of a CF
f : the component of each stage of a CF
TRL : the total removing lines
F : CF chains

Algorithm 1 illustrates the details how to generate CF chains. T is the total number of lines; Δp is the unbalanced power; ECF and TRL are the post-contingency result that will affect the CFP; f_d is the component of each stage of a CF. The first step is to choose a random operating condition, and disconnecting a transmission line. Then it is necessary to re-dispatch Δp and run power flow to obtain the post-contingency result which contains the criteria of ending cascading failures (ECF) and TRL. The propagation depends on those two indicators. When the cascading failure is terminated, recording the CF chain resulted from the disconnecting transmission line. Then implementing "N-1" (it means to disconnecting all transmission lines one by one), we can obtain the CF chains under a random operating condition. Finally, implementing all random operating conditions, all possible CF chains can be obtained.

Algorithm 1 Algorithm for generating cascading failure chains in a random system operating condition

Input: A random operating condition

Output: Cascading failure chains

for $t = 1 \rightarrow T$ **do**
  Disconnect the t-th line
  Re-dispatch Δp
  Run power flow
  Obtain ECF and TRL
  d=1
  if $ECF \geqslant 1$ **then**
    f_d= the disconnecting t-th line
  else if ECF<1 and there is TRL **then**
    f_d= the disconnecting t-th line
  else if ECF<1 and there is no TRL **then**
    f_d= empty
  end if
  while $ECF < 1$ **do**
    if there is TRL **then**
      Disconnect TRL
      Re-dispatch Δp
      Run power flow
      Obtain ECF and TRL
      d=d+1
      if $ECF \geqslant 1$ **then**
        f_d= TRL
      else if ECF<1 and there is TRL **then**
        f_d=TRL
      else if ECF<1 and there is no TRL **then**
        f_d=empty
      end if
    end if
  end while
end for

3.2.3 Vulnerability assessment and cascading failure propagation prediction

The method is based on all CF chains to establish a directed probabilistic graph, which represents the conditional dependence structure between random variables [77]. The probabilistic graph indicating CFP can be defined as Equation (3.22).

$$G = \{N, E\} \tag{3.22}$$

where
G : the probabilistic graph
N : nodes in the probabilistic graph
E : edges in the probabilistic graph

We defined three types of nodes in the probabilistic graph, which are initial node (IN), spread node (SN) and end node (EN). Fig. 3.4 shows an example of probabilistic graph. In Fig. 3.4: node A is IN (it means the initial failure of a CF); node C is SN (it means the CF can be continued from this node); node F is EN (It means the CF is terminated). To introduce the CFP more clearly, other two concepts also need to be notable: parent node (PN) and child node (CN). For example: in Fig. 3.4, node B, C and D are CNs of node A while node A is PN of node B, C and D.

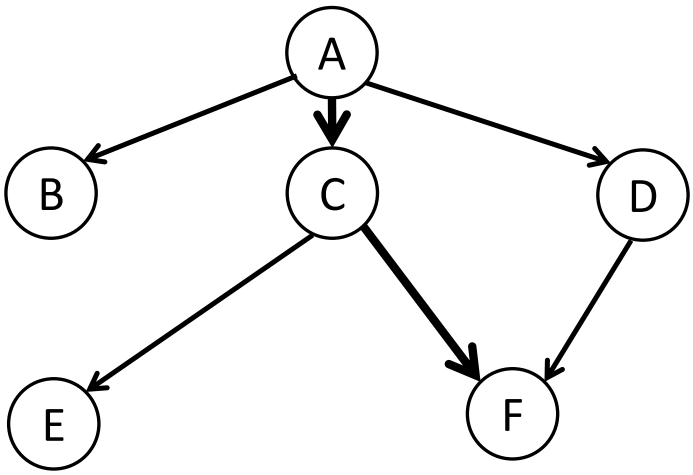


Fig. 3.4 Example of an probabilistic graph

Algorithm 2 Algorithm for predict the cascading failure propagation

Input: All cascading failure chains

Output: Cascading failure propagation

Vulnerability assessment
The most vulnerable transmission line as the IN
CFP=[IN]
d=2
Calculate the weight of all directed links from IN to its CNs
obtain SN_d
CFP=[IN, SN_d]
while SN_d is not EN_d **do**
    d=d+1
    Calculate the weight of all directed links from SN_d to its CNs
    if SN_d is EN_d **then**
        EN_d=SN_d
        CFP = merge (CFP, EN_d)
        CFP stops
    else
        CFP = merge (CFP, SN_d)
    end if
end while
return CFP = [IN, SN_d,..., EN_d]

Algorithm 2 Introduces how to predict CFP. At the beginning, vulnerability assessment is implemented to find the most vulnerable transmission line as IN. After that, calculating the weight of all directed links from IN to its CNs. The CN which has the highest weight can be the next SN. The CF will be not terminated until SN becomes EN. If SN has no CN, then SN will be transferred to EN.

In terms of vulnerability assessment, one of the most effective solution to assess the vulnerability is to employ the risk assessment [78], which considers the comprehensive severity as well as the failure rate. However, in this thesis, the influences of all cascading failure are considered as a same level. The consequence is either to isolate a large number of consumers or to cause the failure of the entire grid. Therefore, we proposed a new indicator which is the vulnerability index (VI) at the second stage (the first stage of a CF implements "N-1", so the failure probability for

each transmission line is the same), and it is defined as the probability of a branch affected by other failures as shown in Equation (3.23). The transmission line which has the highest VI is the most vulnerable point of a power system.

$$VI_t = \frac{FO_t}{K * T} \tag{3.23}$$

where
t : the index of transmission lines
VI : the vulnerability index
FO : the fault occurrence of a branch affected by other failures
K : the sample size of MC
T : the total number of transmission lines

Besides vulnerability assessment, other indicators (weight and probability of CN) and methods (SN and CFP) are necessary to be further explained. The weight is described as Equation (3.24), which shows conditional dependence between two nodes. The probability of each CN is calculated by the law of total probability [79] as shown in Equation (3.25). The method to obtain SN at each stage was mentioned above, and it is defined as Equation (3.26). Collecting all SN and the EN in the probabilistic graph, CFP can be finally obtained, and it is described as Equation (3.27).

$$w(PN, CN_c) = \Pr(CN_c|PN)\Pr(PN) \tag{3.24}$$

$$\Pr(CN_c) = \sum_{m=1}^{M} \Pr(CN_c|PN_m)\Pr(PN_m) \tag{3.25}$$

$$SN_d = \begin{cases} IN & d = 1 \\ \max(w(SN_{d-1}, CN_c)) & d > 1 \end{cases} \tag{3.26}$$

$$CFP = [IN,\ SN_d, ..., EN_d] \tag{3.27}$$

where
d : the stage index
c : the index of CN
m : the index of PN
w : the weight between two nodes
M : the total number of PN of a CN

Notably, for each node, it may contain more than one component. For example, No.2 transmission line is disconnected from a system because of the overload, and it leads to the overloads of No.3 and No.4 transmission line, so the node at the first stage is No.2 line while the node at the second stage consists of No.3 and No.4 line.

3.3 The normal failure model test

The normal failure model was applied to a 30-bus system ("case30.m") in Matpower [80], and the network was modified from IEEE 30-bus system [81]. This network system has 30 buses, 41 transmission lines, 6 generators and 20 loads. The network is shown as Fig. 3.5. To clearly introduce the normal failure model, we numbered the transmission line with red colour.

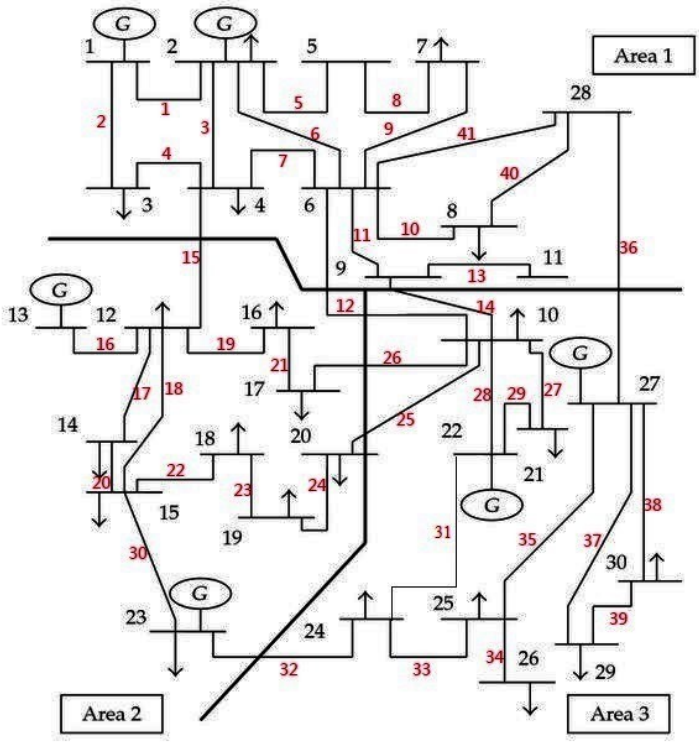
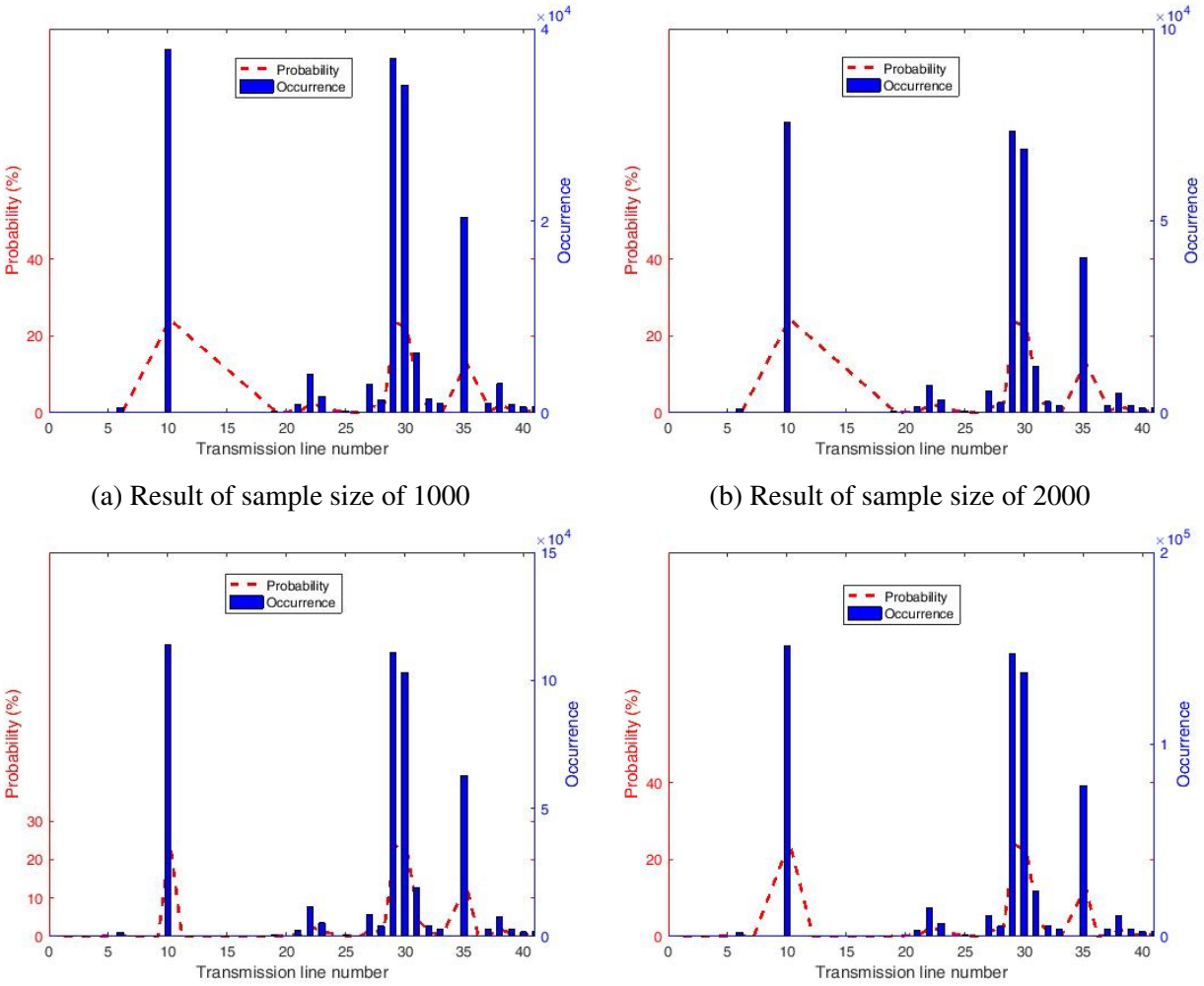


Fig. 3.5 30-bus system in Matpower

To test the normal failure model on the 30-bus system, we firstly employed MC to produce random operating conditions. In order to appropriately execute this methodology and compare the results, four sample sizes (1000, 2000, 3000 and

4000) of MC were chosen. Therefore, the final results will be composed of four results. After that, we implemented "N-1" (focusing on transmission lines) for all random operating conditions, then we can get all CF chains. The result of all CF chains can help us discover the transmission line with the highest VI as the starting point of a CF and establish the probabilistic graphs to predict the cascading failure propagation. To evaluate the vulnerability, VI was calculated by Equation (3.23), whereas the cascading failure propagation can be estimated by Equation (3.24) to (3.27).



(a) Result of sample size of 1000

(b) Result of sample size of 2000

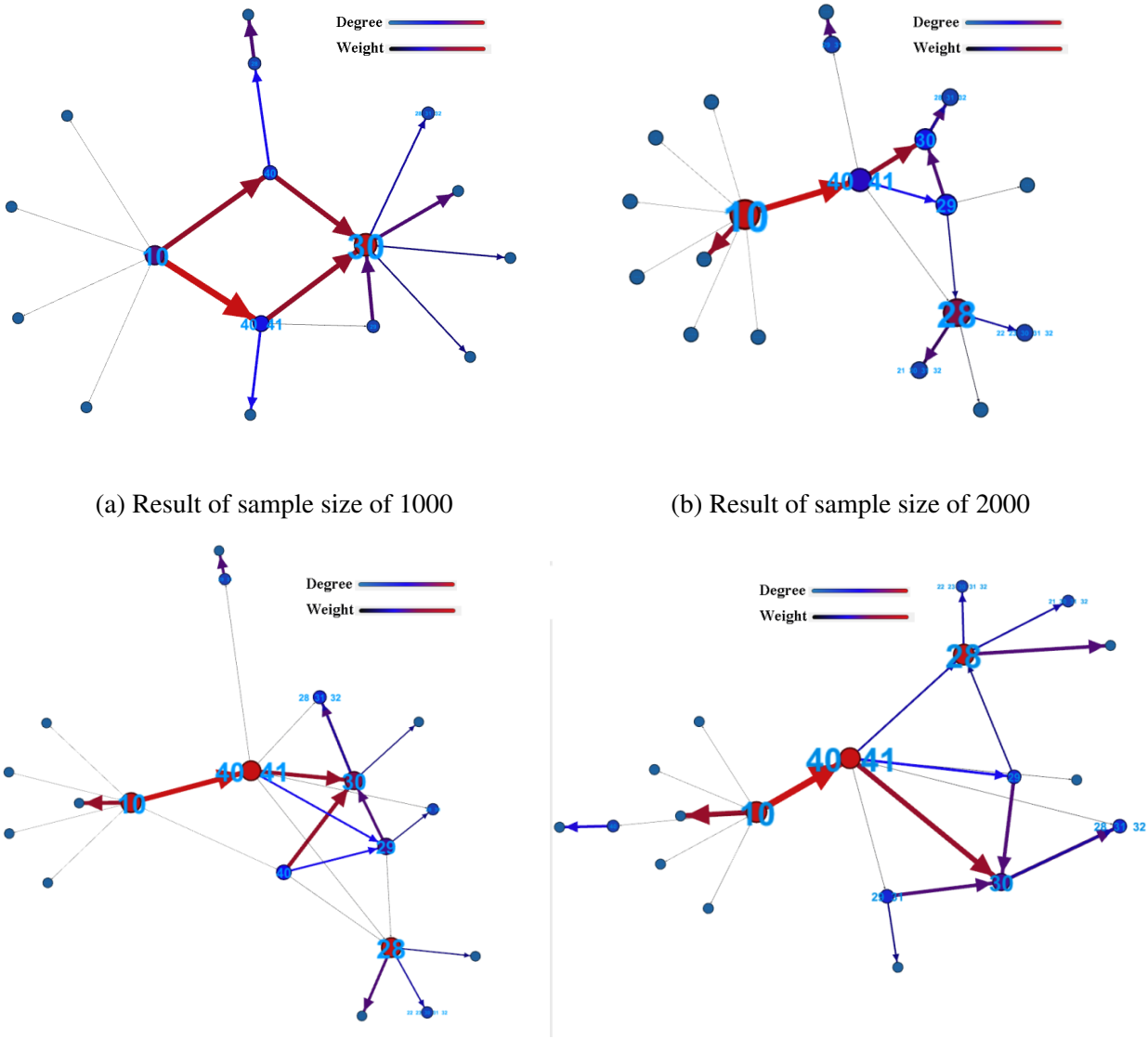
(c) Result of sample size of 3000

(d) Result of sample size of 4000

Fig. 3.6 Vulnerability assessment results of different sample sizes

After evaluating VI for all transmission lines, Fig.3.6 displays vulnerability assessment results of different sample sizes. According to Equation (3.23), there is

relation between VI and FO, so the figure also displays the FO of all transmission lines. From Fig. 3.6, it can be seen that four transmission lines, which are No.10, No.29, No.30 and No.35, have higher VI compared with other transmission lines. Even though the VI of No.10 the VI of No.29 are very close, the VI of No.10 is slightly higher than No.29's VI. Consequently, No.10 is the most vulnerable transmission line in this power system. To simulate the cascading failure propagation, No.10 was selected as the starting point.



(a) Result of sample size of 1000

(b) Result of sample size of 2000

(c) Result of sample size of 3000

(d) Result of sample size of 4000

Fig. 3.7 Probabilistic graphs based on different sample sizes

After choosing the initial failure, probabilistic graphs based on different sample sizes were established. Fig. 3.7 presents the simulation results of different sample sizes. As presented in Fig. 3.7a, the propagation with the sample size of 1000 is different from others' because the sample size is not large enough, but increasing the sample sizes more than 1000, then the CFP are consistent as displayed in Fig 3.7b, Fig. 3.7c and Fig. 3.7d. According to the comparison of four results in Fig. 3.7, the CF due to the failure of No.10 transmission line should propagate as:

$$CFP = [(10),\ (40,41),\ (30),\ (28,31,32)]$$

Another unexpected discovery in Fig 3.7 is that No. 28 has a comparatively higher "degree" (it means connections of a node) than other nodes. We tried to analyse two results (all CF chains and FP) to explain this discovery. 1) as shown in Fig.3.6, No. 28 has a very low FP, which means it is not a vulnerable transmission line in the system, so FP and degree seem to be unrelated. 2) analysing all CF chains, we found that many CF chains with the starting point of No.10 passed through No.28. Moreover, considering the final result of CFP, we could conclude that the node which has a higher degree could be a part of CFP.

3.4 Discussion

After implementing the Monte Carlo method, statistical results were obtained from the power flow analysis, while the study of cascading failure propagation employed the theories of probability and graph. Therefore, the main contribution of this work is to take advantage of random operating states to investigate the vulnerability of a network, and then establishing directed probabilistic graphs to identify the cascading failure propagation if the vulnerable part of a system is attacked, in this way, the large sample size of random operating states can make the results more accurate and reliable, and directed probabilistic graphs are possible to illustrate the cascading failure propagation legibly. In short, the probabilistic graph makes the probability model visualized, so that the relationship between some variables can be easily observed from the figure; at the same time some complex calculations of probability can be understood as the transmitted information between two variables. However, there is an issue with this methodology, which is the computing time. It is necessary to take a large random sampling in this methodology to make the

propagation results consistent. Finding a way to reduce the sample size or using the distributed computing can efficiently decrease the computing time, but the accuracy still needs to be taken into consideration. In the future, the importance of this method is to find an optimized solution which is not only fast but also accurate.

Chapter 4

CASCADING FAILURES IN THE EXTREME WEATHER MODEL

4.1 Brief introduction

Although investigating CFP under the extreme weather is significant, there is little research focusing on this topic. Based on this consideration, we proposed the extreme weather model to have a insight that how the weather will cause outages and lead to a cascading failure propagation in power systems.

Many extreme weather conditions can make a great damage to power grids such as earthquakes, tornadoes, hurricane/tropical storms, lighting, winds/rains, ice storms, etc. The severe weather condition considered in this thesis is the snowstorm, as it is one of typical adverse weather conditions which can cause wide-scale electric power outages. Specifically, we evaluate the snowstorm based on the snowfall amounts.

To predict the CFP under the snowstorm condition, the basic method is to model the snowstorm. To model this adverse weather, it is necessary to consider two important facts: one is the severity of snowfall and another is the occurrence of snowfall. Solving those two issues, we employing probability distribution (PD) to estimate the probability of different snowfall amounts. Applying the PD of different snowfall amounts into the normal failure model, we can establish a "failure network" to achieve the purpose of estimating the CFP under the snowstorm condition.

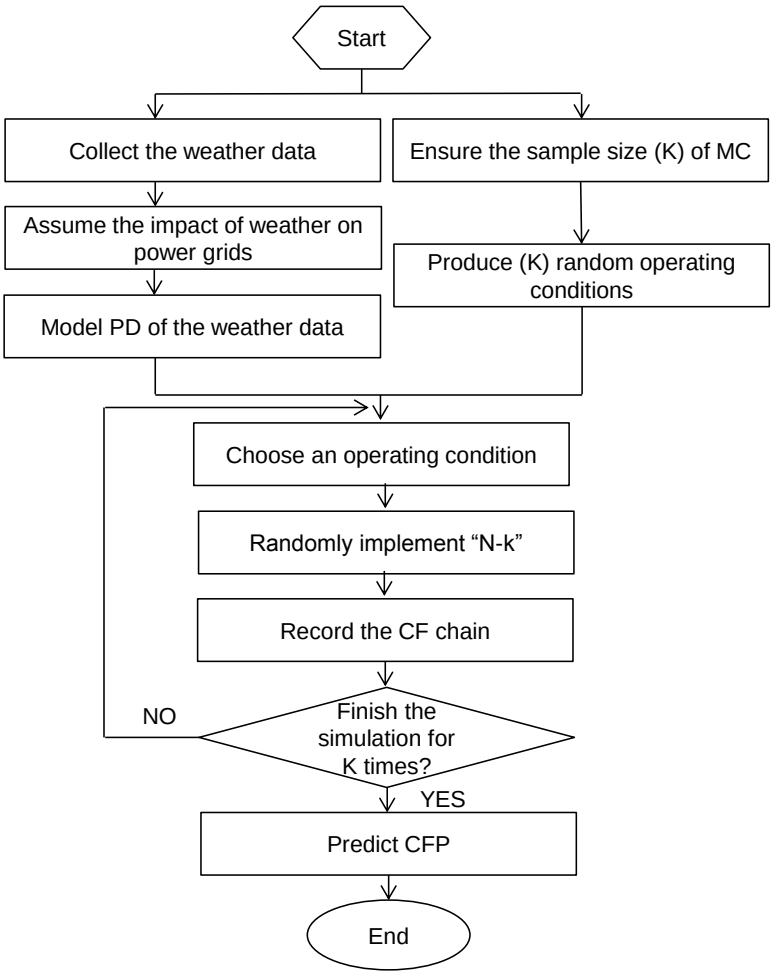


Fig. 4.1 Flow chart of the extreme weather model

The flow chart of extreme weather model is presented as Fig.4.1. Initially, two steps including modeling PD of snowfalls and producing random operating condition are paralleled. The method to produce random operating conditions can be adopted by the same method as mentioned in Section 3.2.1 (so this method will be described below again). To model PD of the snowfall amounts, the first step is to collect a large amount of historical snowfall data in a certain area, and then to make a statistical analysis to establish the PD for the snowfall data. After modeling the PD, it is indispensable to assume that how the different snowfall amounts will affect the power transmission networks. The next process is to generate CF chains and it will run in a loop. This process dose not terminate until all random operating conditions are implemented by "N-k". k is the random disconnecting transmission lines because of the different snowfall amounts. After this process, all CF chains can be obtained

to establish a "failure network". Finally, based on the failure network, it is possible to estimate the CFP under the snowstorm condition.

4.2 Methodology of the extreme weather model

4.2.1 Probability distribution of snowfall amounts

There are two types of probability distribution, which are probability mass function (PMF) and probability density function (PDF). According to [82], they can be defined as:

- if random variables equate finite values, then they are discrete random variables. Probability distribution of the discrete random variables is known as the "probability mass function".
- If random variables equate infinite values, then they are continuous random variables. Probability distribution of continuous random variables are known as the "probability density function".

We summarised some important probability distributions for PMF and PDF in Table 4.1.

Table 4.1 Some important probability distributions

PMF	Bernoulli	It only has two possible outcomes: 0 or 1
	Binomial	It is the sum of identically and independent distributed Bernoulli random variable
	Poisson	It presents the number of occurrences of random events occurring in a fixed interval of space and/or time
PDF	Uniform	It is a distribution that has constant probability.
	Normal	It represents the distribution as a symmetrical bell-shaped graph.
	Exponential	It describes events between the time in a Poisson process

As shown in Table 4.1, Exponential Distribution is suitable to model the PD of snowfall amounts. However, Poisson Distribution can be also adopted to model the PD of snowfall amounts. In order to apply Poisson distribution, the snowfall data needs discretization. According to the National Oceanic and Atmospheric Administration (NOAA), the snowfalls can be categorised into 5 levels [83] as

shown in Table 4.2. The original category is based on the unit of "inch", to be more convenient, we use"mm" as shown in the third column of Table 4.2. Beside the severity category, we made an assumption in the last column on which is the adverse impact of different severity levels on power grids. The "impact" can be explained by examples. For instance, 0% impact means there is no affect on the power grids. 10% impact means 10% of total transmission line in the affected area will fail and finally be removed.

Table 4.2 Severity category of snowfall amounts

Level	Range (inch)	Range (mm)	Impact on power grids (%)
Level 0	[0, 2)	[0, 50)	0
Level 1	[2, 4)	[50, 100)	10
Level 2	[4, 6)	[100, 150)	20
Level 3	[6, 8)	[150, 200)	30
Level 4	[8, 100)	[200, 1000)	40

After ensuing the severity level of snowfall amounts and its impact, Equation (4.1) presents PMF of snowfall amounts:

$$\Pr(X = SL) = \begin{cases} e^{\lambda} \times \frac{\lambda^{SL}}{SL!} & SL = [0,1,2,3] \\ 1 - \sum_{X=0}^{3} \Pr(X) & SL = 4 \end{cases} \tag{4.1}$$

where
e : the Euler's number
λ : the average probability per severity level
SL : the severity levels
$SL!$: the factorial of SL

λ is equal to the expected value of SL [84], as shown in Equation (4.2):

$$\lambda = E(SL) \tag{4.2}$$

4.2.2 Generation of cascading failure chains

To produce CF chains under a snowstorm, two issues need to be figured out:

- how the transmission lines will be affected by the snowfall amounts. For example, if the snowfall amounts are already known, it is necessary to know how many transmission lines will be disconnected because of the severity level of the snowfall amount.
- what the frequency of each severity level is. For example, the sample size of MC is 1000, so it is necessary to know the proportion of each severity level in the samples of 1000. Knowing the frequency of each severity can be useful to implement MC method.

In order to solve the first issue, random fault line (RFL) was proposed and it can be calculated as Equation (4.3):

$$RFL = \begin{cases} 0 & SL = 0 \\ T \times 10\% & SL = 1 \\ T \times 20\% & SL = 2 \\ T \times 30\% & SL = 3 \\ T \times 40\% & SL = 4 \end{cases} \tag{4.3}$$

where
RFL : the randomly disconnecting lines
T : the total transmission lines in a certain area
SL : the severity level

Furthermore, to solve the second issue, the frequency of each severity level (FSL) can be calculated as:

$$FSL = K \times \Pr(X = SL) \tag{4.4}$$

where
FSL : the frequency of each severity level
K : the sample size of MC
SL : the severity levels

To be more understandable, we transformed Equation (4.3)-(4.4) to Table 4.3:

Table 4.3 Details of RFL and FSL at different severity levels

Level	RFL	FSL
Level 0	0	K*Pr(X=0)
Level 1	T*10%	K*Pr(X=1)
Level 2	T*20%	K*Pr(X=2)
Level 3	T*30%	K*Pr(X=3)
Level 4	T*40%	K*Pr(X=4)

Algorithm 3 introduces the procedure to generate CF chains under a severe weather. This is similar to the one introduced in Chapter 3.2.2, but there is still a great difference. Under extreme weather conditions, the method implements "N-RFL" instead of implementing "N-1". Therefore, it is unnecessary to disconnect all transmission line once in this method. At the preparing stage, the process is to estimate PMF of historical snowfall data, and calculate RFL and FSL. The next step is a loop, which is to choose a unrepeatable random operating condition and remove RFL from the system, and then to implement step 1 to step 29 of Algorithm 1. After implementing $K - K \times \Pr(X = 0)$ times, all CF chains can be obtained. In Algorithm 1, the repeated times depend on the sample size of MC, but unlike Algorithm 1, the repeated times of Algorithm 3 are based on FSL, as the first level of snowfall amounts has no impact on power grids.

Algorithm 3 Algorithm for generating cascading failure chains under the extreme weather condition

Input: All random operating conditions, historical snowfall data, and RNSL

Output: Cascading failure chains

1: Estimate the PMF of historical snowfall data
2: Calculate RFL and FSL
3: **for** level=1 $\rightarrow$ 4 **do**
4: **for** repeated times=1 $\rightarrow$ FSL(level) **do**
5: Choose an unrepeatable random operating condition
6: Disconnect RFL(level)
7: Implement step 3 $\rightarrow$ step 29 of Algorithm 1
8: **end for**
9: **end for**

4.2.3 Prediction of the cascading failure propagation

After obtaining all CF chains, implementing two steps can help us to identify the CFP. The first step is to establish a failure network and the second step is to propose a method to ensure how the cascading failure will propagate.

4.2.3.1 Failure network

A failure network was established by transforming all CF chains into an adjacency matrix. An adjacency matrix is used to represent a finite graph and it indicates whether pairs of vertices are not in the graph or adjacent [85]. The failure network, in this thesis, is defined as a undirected graph with weights. We defined the failure network as undirected graph because we found the edges between two nodes in a failure network were always bidirectional if the sample size of MC is large enough.

$$G = (N, E) \tag{4.5}$$

where
G : the failure network
N : nodes in the failure network
E : edges in the failure network

A general adjacency matrix is a (0, 1)-matrix [47]. Zero means there is no connection between two nodes while one means two nodes exist the connection. Moreover, the diagonal of adjacency matrix equals zero. However, in order to transform the adjacency matrix into a failure network, we made two changes for the adjacency matrix: 1)the diagonal of the adjacency matrix is defined as the degree of the node rather than zero; 2)the connection between two nodes is not only euqal to 1 or 0, but it equals to the total number of connections between two nodes (it is defined as weight). The new defining adjacency matrix is shown as Equation (4.6). As the adjacency matrix of a undirected graph is a symmetric matrix [47], Equation (4.7) indicates the character of a symmetric matrix.

$$A_{ij} = \begin{cases} W(N_i,N_j) & if\ i \neq j\ and\ E(N_i,N_j) \in E(G) \\ 0 & if\ i \neq j\ and\ E(N_i,N_j) \notin E(G) \\ Deg(N_i) & if\ i = j \end{cases} \tag{4.6}$$

s.t.

$$A_{ij} = A_{ji} \tag{4.7}$$

$$W(N_i,N_j) = \sum E(N_i,N_j) \tag{4.8}$$

where:
i, j : the index of nodes
G : the failure network
A : the adjacency matrix of failure network
W : weights between two nodes
N : nodes in the failure network
E : edges in the failure network
Deg : degrees of a node

To illustrate the method to transform CF chains into the adjacency matrix, Fig. 4.2 displays an example. The assumptions are that there are 4 nodes (node 1, node 2, node 3 and node 4) and 3 CF chains ($F_1 = [(1,2),\ 3,\ 4]$, $F_2 = [2,\ (3,4),\ 1]$ and $F_3 = [2,\ 3,\ (1,4)]$). The first chain ($F_1 = [(1,2),\ 3,\ 4]$) means that node 1 connects to node 3, node 2 connects to node 3 and node 3 connects to node 4, so (N_1,N_3), (N_2,N_3), and (N_3,N_4) equal to 1. On the other hand, there is no connection between node 1 and node 2, node 1 and node 4, as well as node 2 and node 4, so (N_1,N_2), (N_1,N_4) and (N_2,N_4) equal to 0. The degrees of node 1, node 2, node 3 and node 4 equal to 1, 1, 3 and 1, respectively, then filling in the diagonal of matrix with degrees of each node. The second and the third step are the same to the first step, but the weights between two nodes and the degrees of each node will increase. As the adjacency matrix of undirected graph is a symmetric matrix, we can firstly transform the CF chains into the upper triangular matrix, and then to transpose the upper triangular matrix. After that, the final adjacency matrix can be obtained.

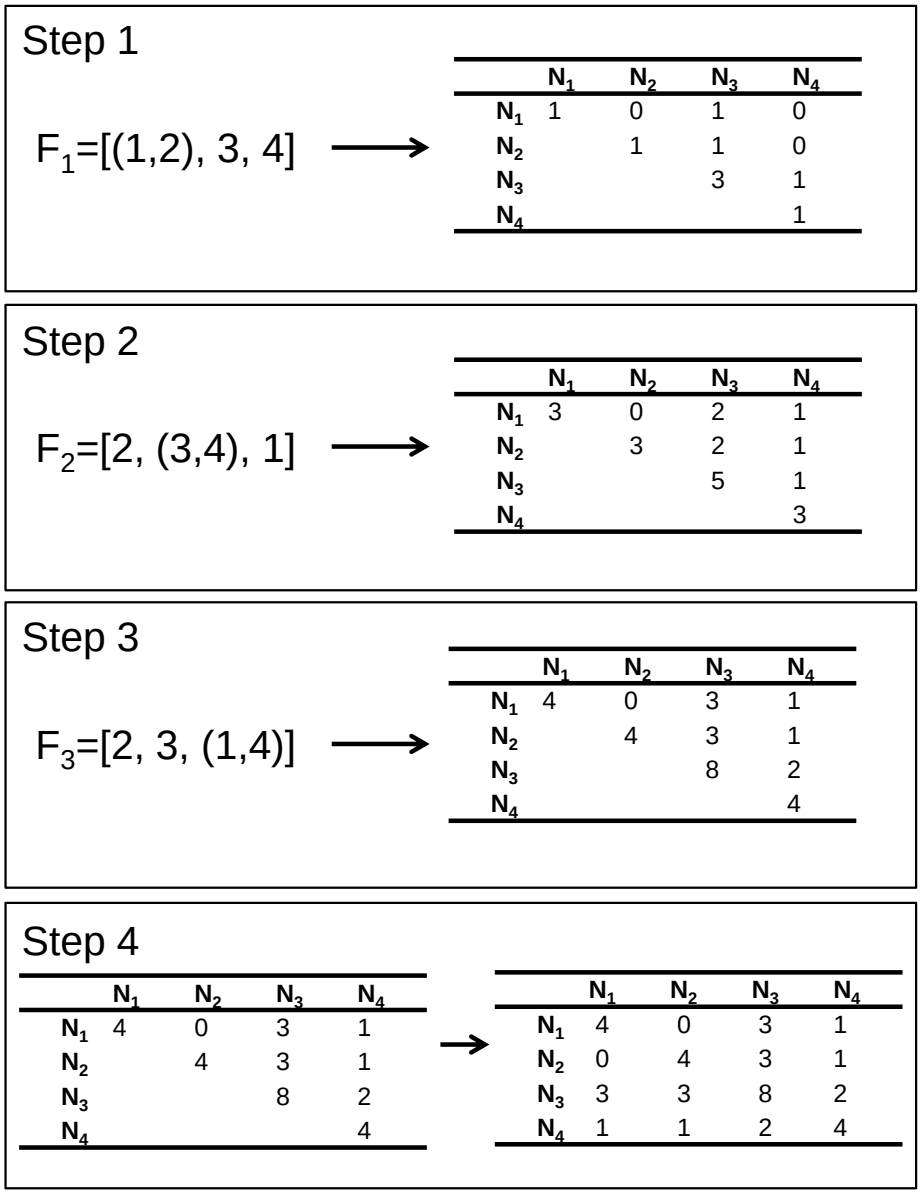


Fig. 4.2 Method of transforming CF chains to the adjacency matrix

After obtaining the adjacency matrix, we can transform it into a failure network. Fig. 4.3 displays the method to transform the result of Fig. 4.2 into a simple failure network. Different from a probabilistic graph, a failure network has undirected edges. Besides, the node in a probabilistic graph may have more than one component, but a node in the failure network only contains one component. Additionally, the size of a node indicates the degree and the size of edge between two nodes indicates the weight.

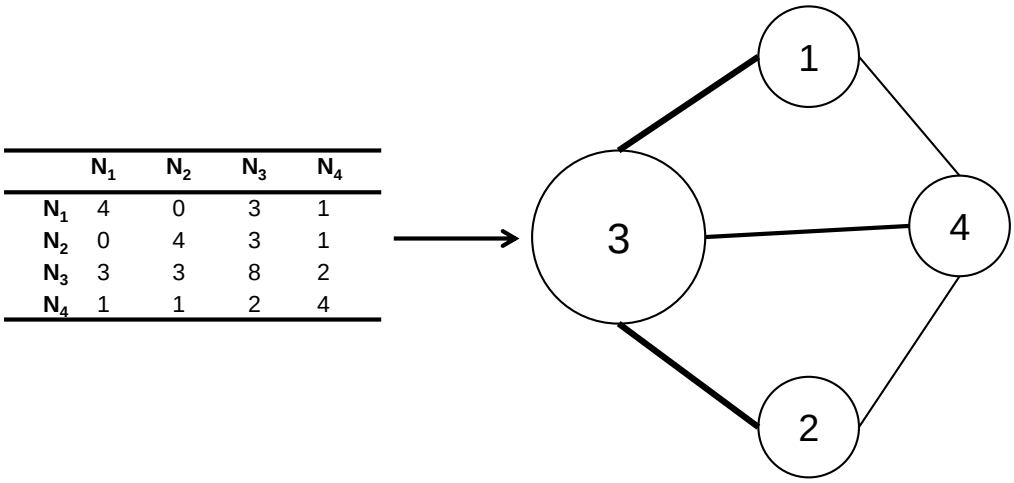


Fig. 4.3 Method of transforming the adjacency matrix into a failure network

4.2.3.2 Method of predicting the cascading failure propagation

Before identifying the CFP based on a failure network, three issues need to be considered:

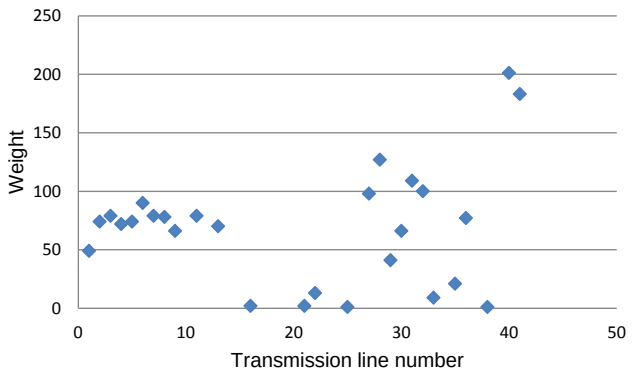
- how to choose the initial failure
- how the CF will propagate
- when the CFP will end

In a normal failure model, the initial failure happens on the most vulnerable part of a power system, but in a extreme weather model, the failure could happen on any transmission line in the affected area. Therefore, the initial failure in the extreme weather model can be randomly chosen in the affected area.

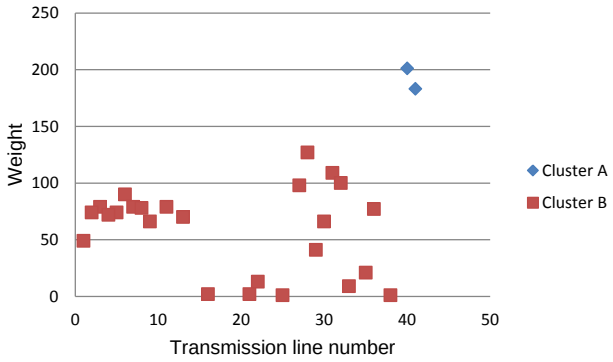
To reveal the propagation of cascading failures, we cluster the CN of SN into two clusters according to their weights. One cluster contains the CN which have higher weights, then the CN in this cluster will be turned into the SN at the next stage of CF. In order to achieve this purpose, we employed Hierarchical Clustering method to cluster the weights of CNs into two clusters. Hierarchical Clustering is a type of clustering methods to build a hierarchy of clusters [86]. To implement Hierarchical Clustering, the first step is to calculate the distance between each weight until the distances between all the pairs have been calculated, and then we can get a matrix which contains the distances between each weight. After that, to search the

closest pair based on the distance matrix to create a hierarchical cluster tree with two hierarchies [87].

However, the traditional methods of Hierarchical Clustering cannot be the comprehensive solutions to cluster the CN in the failure network, as the clustering result might contain the CN with low weight (the situation can be explained by Fig. 4.5). Under some circumstances, the traditioal methods can be very useful to cluster the CN. For example, Fig. 4.4a shows a typical distribution of some CN. After adopting the traditional methods such as singel linkage, compelet linkage and average linkage [88], we can obtain the result as displayed in Fig. 4.4b. The components in "Cluster A" can exactly present the wanted results.



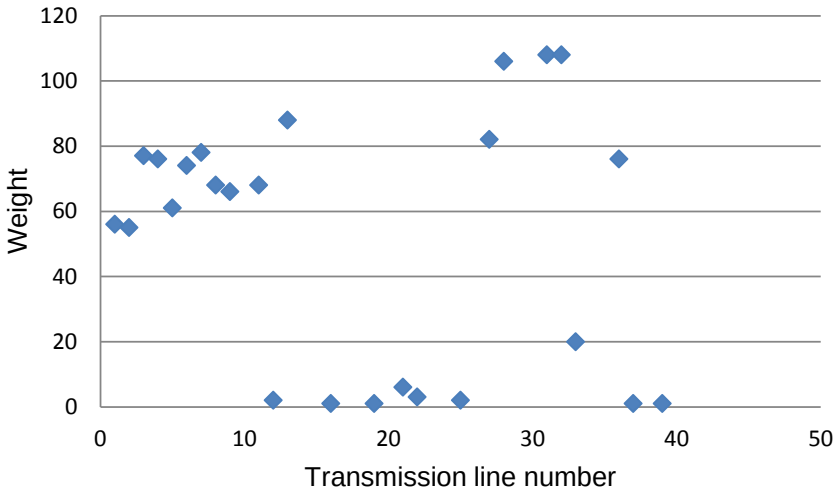
(a) Weight distribution example of some child nodes



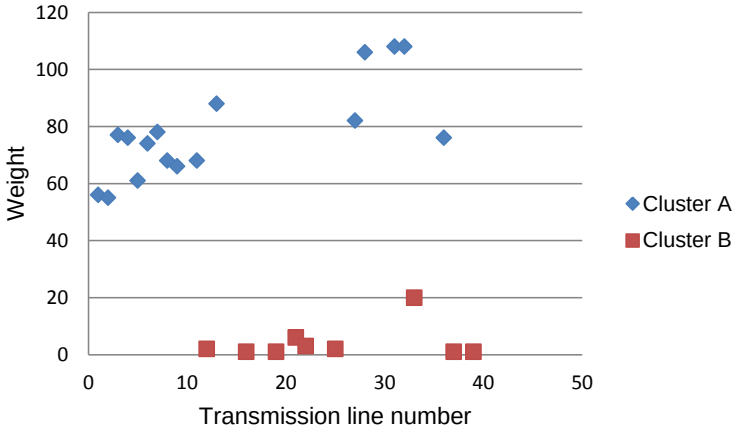
(b) Result of implementing traditional method

Fig. 4.4 Employing the traditional method of Hierarchical Clustering on the typical example

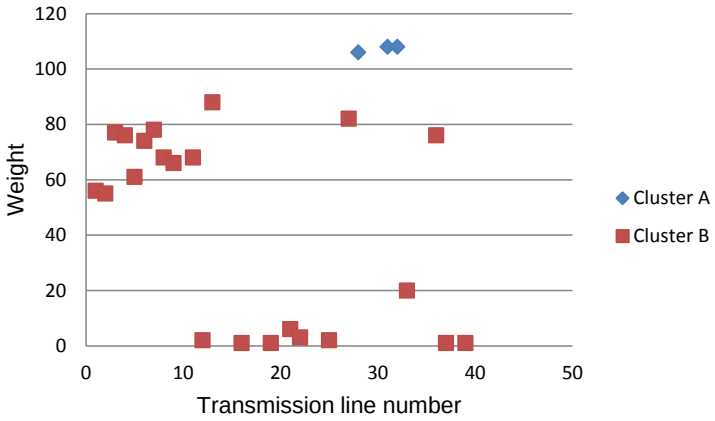
Nevertheless, under some special conditions, the traditional methods of Hierarchical Clustering cannot work properly. For example, if the weight is distributed as shown in Fig. 4.5a, adopting the traditional methods to cluster the CN into two groups, the result will be like in Fig. 4.5a, which is not a satisfied outcome. The result we want should be the same as Fig. 4.5c. Essentially, the traditional method of Hierarchical Clustering is possible to find the "Cluster A" in Fig. 4.5c, but the problem is that the number of clusters will be more than two. To ensure the number of clusters of CN at different stages of the cascading failure propagation is greatly difficult, as the process of cascading failure is dynamic. Therefore, it is necessary to propose a new solution to solve this problem.



(a) Weight distribution example of some child nodes



(b) Result of employing traditional methods



(c) Desirable result

Fig. 4.5 Comparing the result of employing traditional methods and the desirable result

The new proposed clustering method is still on account of distance. The idea is to calculate the distances between two neighbour weights and find the first distance which had hugest increase. To be preaise, we defined the distances between two neighbour weights as neighbour distance (ND), and we used changing rate (CR) to indicate the change between two neighbour distances. The normal method to calculate the distance is the Euclidean distance [89]. The basic idea to calculate the distance still relied on Euclidean distance, but we ignored the x-axis of CN as it indicates the transmission line number and it is meaningless to calculate the distance. Therefore, we assumed the value of x-axis is equal to zero for all CN. The formula to calculate the neighbour distances of weights is shown in Equation (4.9)). Based

on the neighbour distances, the changing rates can be calculated as Equation (4.10):

$$ND(W_i, W_{i+1}) = \begin{cases} \sqrt{(W_i - W_{i+1})^2} & i < TN \\ 0 & i = TN \end{cases} \tag{4.9}$$

$$CR_i = \begin{cases} \frac{W_i - W_{i+1}}{ND_i - ND_{i+1}} & ND_i \neq ND_{i+1} \\ 0 & ND_i = ND_{i+1} \end{cases} \tag{4.10}$$

where

i : the index which can represents the locations of CN, weights and neighbour distances

ND : the neighbour distances between two weights

TN : the total number of weights

W : weights between CN and their SN

CR : the changing rates between two nearest neighbour distances

Algorithm 4 Algorithm for new clustering method

Input: Weights of all CN

Output: Two clusters of weights

1: Sort the weights from the largest to the smallest
2: Calculate neighbour distances
3: **for** i=1 $\rightarrow$ TN **do**
4: calculate CR_i
5: **if** $CR_i > 1$ **then**
6: The first i is the required location
7: **end if**
8: **end for**
9: Cluster weights from W_1 to W_i into "Cluster A"
10: The remaining weights will be grouped into "Cluster B"

Algorithm 4 introduces the steps of the new clustering method. As we want to find the cluster for the the components that have higher weights, the first step was to sort the weights in descending order. Based on the new location of each weight, we calculated the distance between two neighbour weights and calculated the change rate between two neighbour distances. Then we looked for the weight which its changing rate is firstly more than 1. After that, we can cluster the first weight to the found weight in the last step into "Cluster A". Consequently, the rest weights will be

grouped into another cluster. Finally, the CN which its weight in "Cluster A" will become the SN at the next stage.

In order to point out the termination of cascading failure, we proposed an indicator named residual degree (RD). The inspiration to propose this indicator is the process of transmitting energy. At the beginning, there is a total energy and it will transmit to other nodes, but when it reaches a node, the node will consume energy, and the energy will become increasingly less. When the remaining energy is less to a value, the transmission process will stop. As described in this example, the degrees of all nodes in a failure network are the total energy; the residual degree is the remaining energy; the degree of a node is the consumption of the node. When the residual degree is less than the average degree of the failure network, cascading failure will be terminated. The methods to calculate the total degree and the residual degree and are respectively presented as Equation (4.11) and Equation (4.12). It is important to note that, in Equation (4.12), PN refers specifically to the PN of SN and PN should be SN at the previous stage.

$$TD = \sum Deg(N_i) \tag{4.11}$$

$$RD_d(SN_k) = \begin{cases} TD - Deg(IN_j) & d = 1 \\ \frac{\sum_{m=1}^{M} RD_{d-1}(PN_m)}{M} - Deg(SN_k) & d > 1 \end{cases} \tag{4.12}$$

where:
k, i, j, m : the indexes of CN, nodes, IN, and PN respectively
d : the stage of CF
TD : the total degrees
RD : the residual degrees
Deg : degrees of a node
M : the total number of PN
N : the node

Algorithm 5 introduces the method to identify the CFP from a failure network. At the initial stage, a failure network should be established by transforming all CF chains. When the failure network is formed, the next step is to randomly choose IN, and then IN will become SN. The next step is to calculate RD for SN. If RD is greater than half to TD, the CFP will continue. After that, it is necessary to find the next SN. Employing Hierarchical Clustering method, the next SN can be obtained.

to continue the CFP, RD needs to be calculated again for the new SN. CFP will not terminate until RD is less than half of the total degrees.

Algorithm 5 Algorithm for identifying cascading failure propagation under the extreme weather condition

Input: All cascading failure chains

Output: A failure network and the cascading failure propagation

```
Transform all CF chains to a adjacency matrix
Transform the adjacency matrix to a failure network
Randomly choose IN in the failure network
CFP=[IN]
SN=IN
Calculate RD for SN and calculate AD
while RD > TD\ do
    Find the CN of SN
    Cluster the CN into two clusters
    Choose the components in the cluster with higher degrees as SN
    CFP=merge(CFP, SN)
    Calculate RD for SN
end while
```

4.3 The extreme weather model test

The process of the extreme weather model is composed of three steps. The first step is the preparatory stage which is to produce random operating conditions and model the PD of snowfall amounts. The next step is to generate all possible CF chains. At last, the failure network can be established, based on the CF chains, to identify the CFP.

To produce random operating conditions, we also employed the 30-bus system ("case30.m") in Matpower [80], and the example system was introduced in Section 3.3. We additionally assumed that Area 1 in the 30-bus system suffered from the snowstorm.

To model PD of the snowfall amounts, we firstly collected the data from National Snow and Ice Data Center (NSIDC), and it recorded monthly snowfall amounts for

18 stations in the western Italian Alps [90]. The recording period in the collecting data varies with each station, with the longest station record of 119 years from 1877 to 1996, and the average station record duration is around 61 years. Table 4.4 presents the 18 stations and their record periods. Besides, Fig. 4.6 illustrates the accurate location of all stations (the locations were marked with the yellow pins).

Table 4.4 Details of snowfall data

Station	Record Period	Station	Record Period
Lago Toggia	1932-1996	Gressoney D'Ejola	1928-1996
Lago Alpe Cavalli	1931-1996	Rimasco	1925-1996
Lago Goillet	1947-1996	Ceresole Reale	1926-1996
Lago Gabiet	1928-1996	Lago Valsoera	1959-1996
Lago Cignana	1927-1996	Balme	1929-1996
Lago Serru'	1955-1996	Lago della Rossa	1938-1996
Lago Moncenisio	1931-1996	Lago Castello	1943-1996
Lago Rochemolles	1925-1996	Cuneo	1877-1966
Bardonecchia	1925-1996	Lago Chiotas	1979-1996

Fig. 4.6 Locations of the 18 stations

4.3.1 PMF of the snowfall data

To estimate the PMF of snowfall amounts in the western Italian Alps, two steps were implemented. The first step was to clean and integrate data. From the original data, we found that there was barely snow from June to October for all stations, so the data during those 5 months were excluded. Moreover, we deleted extraordinarily huge values such as "9999". To simplify the method, we ignored two factors, record period and elevation, to integrate the data of all stations into a data set. After the preparation, we made Fig. 4.7 to present the distribution of snowfall amounts. Approximately, the larger snowfall amounts, the less frequencies it seems.

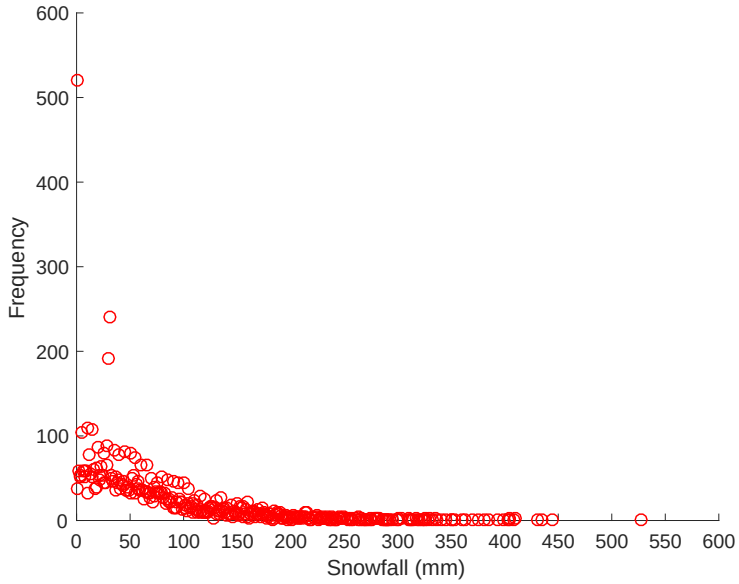


Fig. 4.7 Scatter graph of the snowfall amouts

The next step was to estimate λ. Based on Equation (4.2), λ can be calculated as the expect value of severity levels and it is about equal to 0.8. After obtaining λ, the PMF of snowfall amounts can be present as Equation (4.13).

$$\Pr(X = SL) = \begin{cases} e^{0.8} \times \frac{0.8^{SL}}{SL!} \times 100\% & SL = [0,1,2,3] \\ 0.908\% & SL = 4 \end{cases} \tag{4.13}$$

Transforming the formula to a graph, Fig. 4.8 visually displays the probability function of snowfall amounts. Comparing Fig. 4.7 and Fig. 4.8, we can reach

the same conclusion, which is that the probabilities of snowfall amounts gradually decrease with the increase of the snowfall severity levels.

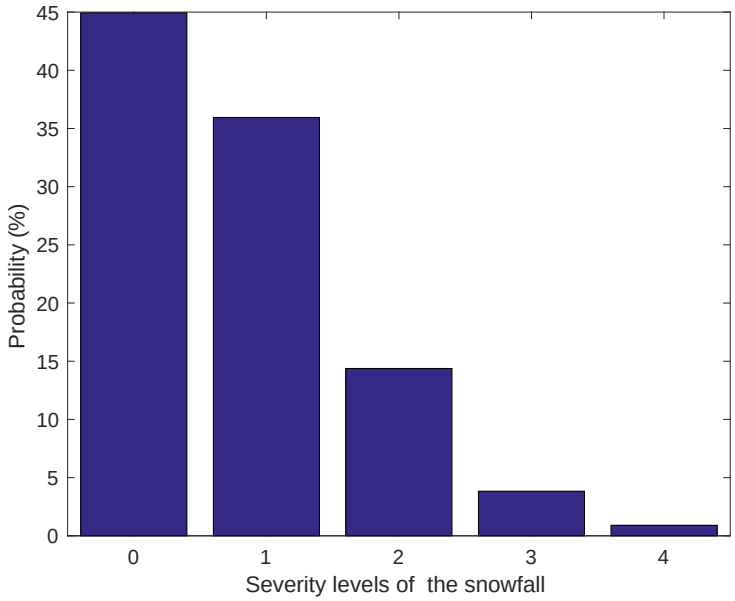
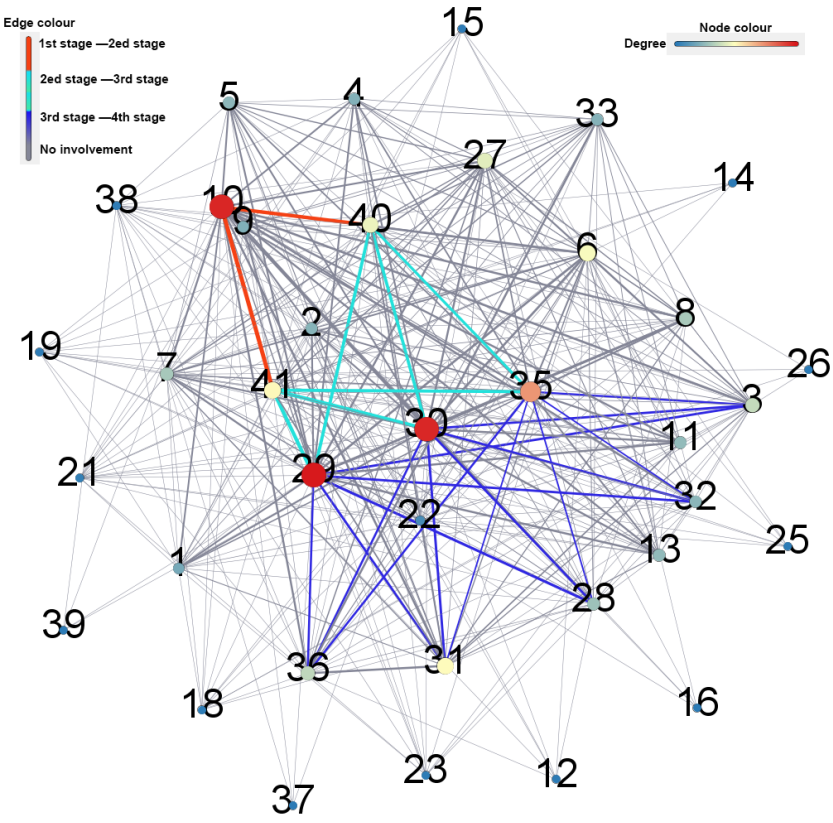


Fig. 4.8 PMF of the snowfall amounts

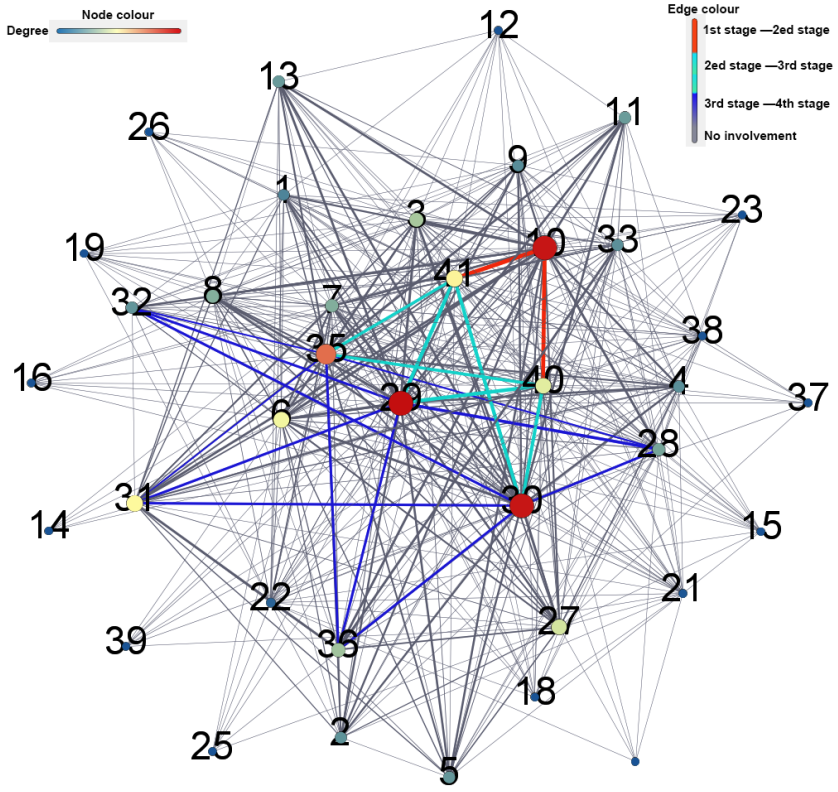
4.3.2 Failure network and cascading failure propagation

Based on the probability distribution of snowfall amounts and the their different impact on power grids, we can run the simulation to get all CF chains. According to the method mentioned in Fig. 4.2, all CF chains can be transformed to the failure network. Adopting the method of predicting the CFP in Chapter 3.2.3.2, we can finally estimate the CFP. As we cannot ensure how many random samples are sufficient, we used four sample sizes (1000, 2000, 3000, 4000) to compare the results. If the results become stable, it indicates that the employed sample sizes are sufficient.

For the purpose of comparing the normal failure model and the extreme weather model, we still chose No. 10 line as the initial failure in the extreme weather. Fig. 4.9 displays the results of choosing No. 10 as the initial failure, and the results are composed of four failure networks after employing four different sample sizes. In those four failure networks, it also indicates the propagation of cascading failures with different edge colours.

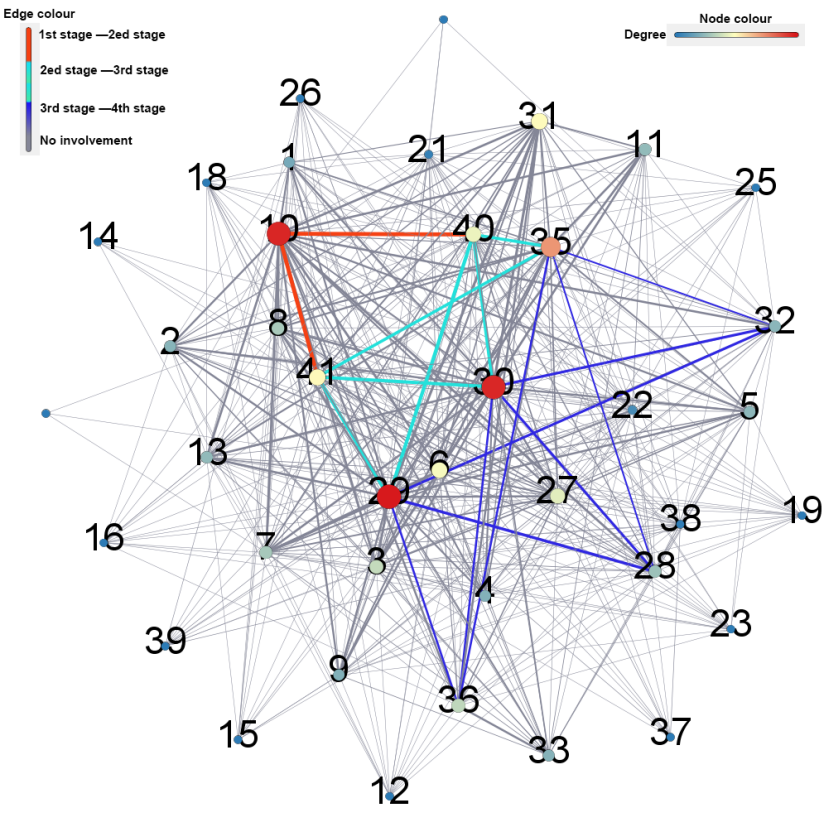


(a) Failure network when samples are 1000

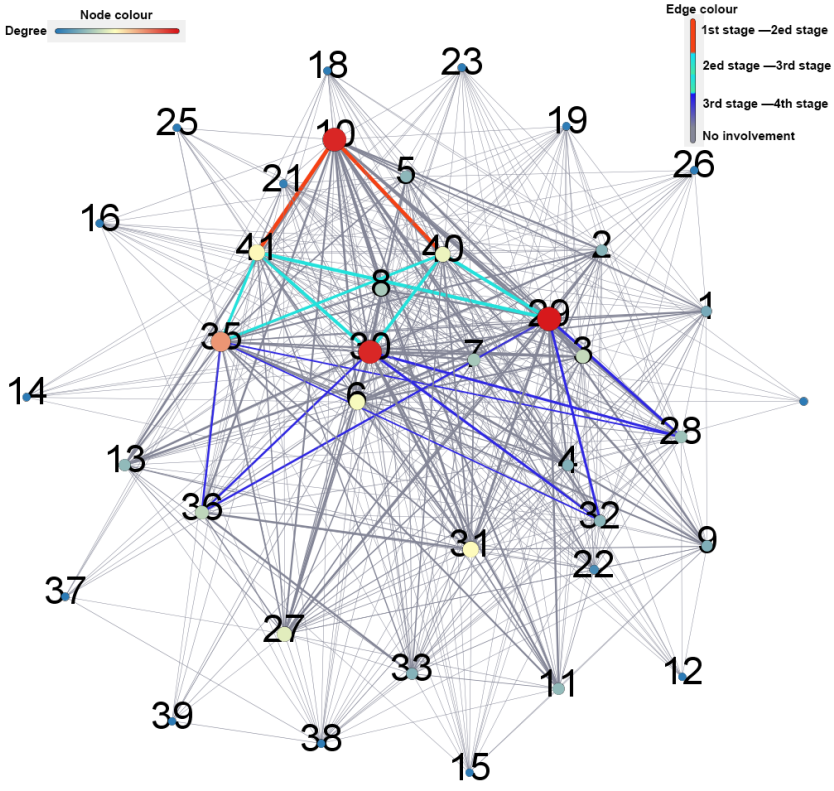


(b) Failure network when samples are 2000

Fig. 4.9 First part of failure network with the initial failure of Node 10



(c) Failure network when samples are 3000



(d) Failure network when samples are 4000

Fig. 4.9 Second part of failure networks with the initial failure of Node 10

The different edge colours in Fig. 4.9 mean:

- the red colour indicates that the cascading failure propagates from the first stage to the second stage,
- the light blue colour indicates that the cascading failure propagates from the second stage to the third stage,
- the dark blue colour indicates that the cascading failure propagates from the third stage to the fourth stage, and
- the grey colour indicates that one or both of connecting nodes will not participate the cascading failure.

To be more clear, we used Table 4.5 to present the CFP in four failure networks. From Table 4.5, it can be seen that the previous three stages of different sample sizes are exactly the same, and only at the last stage, there are some differences. However, when increasing the random samples from 3000 to 4000, then the results keep consistent.

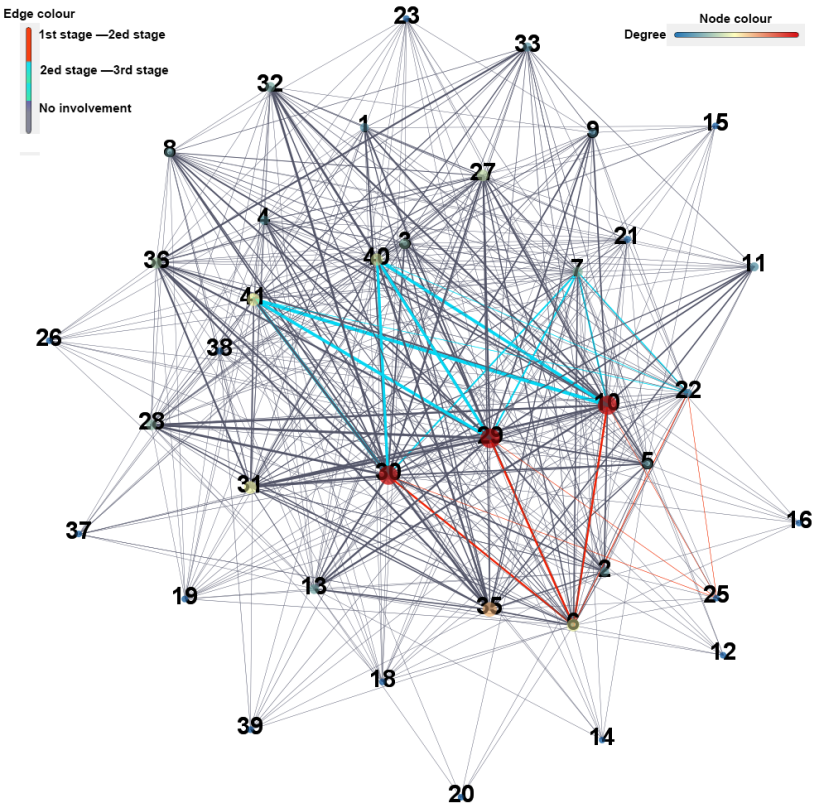
Table 4.5 Cascading failure propagation in four failure networks

Sample size	First stage	Second stage	Third stage	Fourth stage
1000	(10)	(40, 41)	(29, 30, 35)	(3, 28, 31, 32, 36)
2000	(10)	(40, 41)	(29, 30, 35)	(28, 31, 32, 36)
3000	(10)	(40, 41)	(29, 30, 35)	(28, 32, 36)
4000	(10)	(40, 41)	(29, 30, 35)	(28, 32, 36)

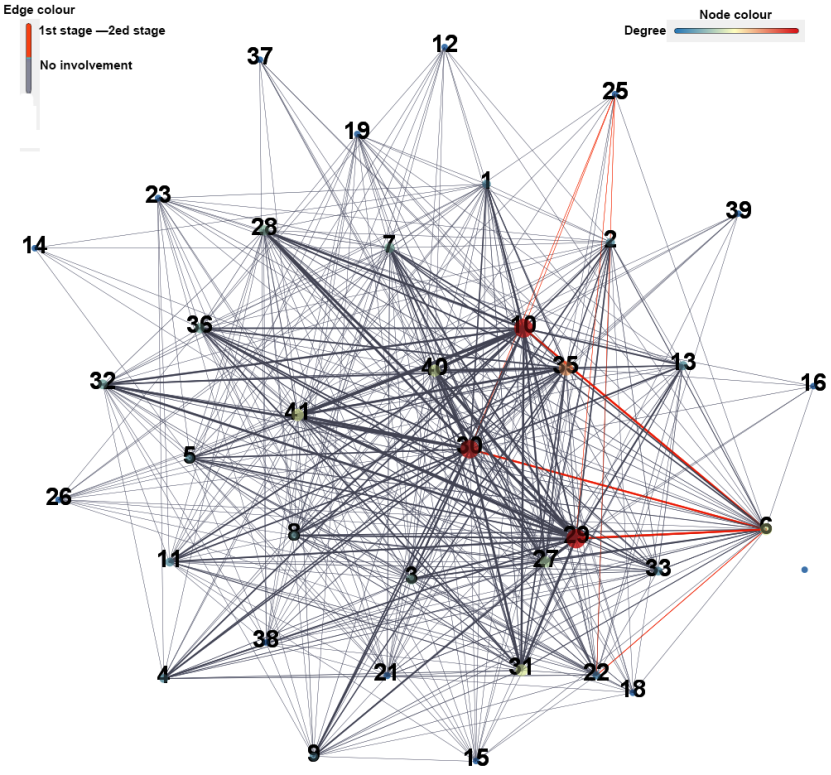
Therefore, the final CFP can be concluded as:

$$CFP = [(10),\ (40,41),\ (29,30,35),\ (28,32,36)]$$

The above example is intent to compare the normal failure model and the extreme weather model, so only No. 10 transmission line was chosen as the initial failure in both models but under the extreme weather condition, the initial failures can randomly happen in the affected area, which means the initial failures might contain more components in the extreme weather model. Therefore, we implemented another simulation to present the results of cascading failure propagation if the initial failures are more than one component. We randomly chose two invulnerable transmission lines, No. 6 and No. 25, as the initial failures.

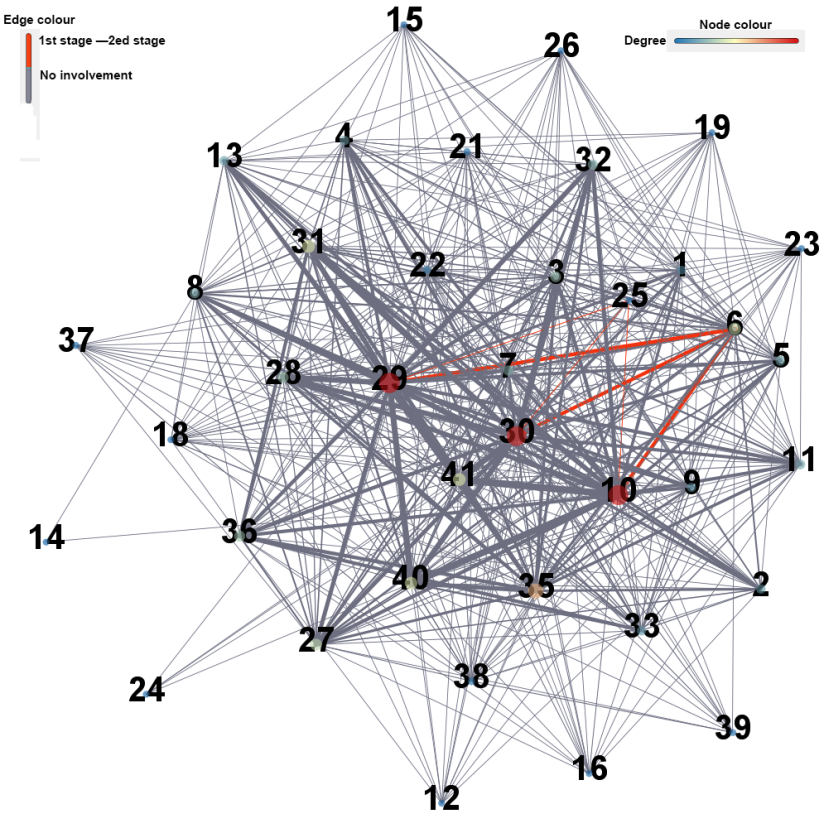


(a) Failure network when samples are 1000

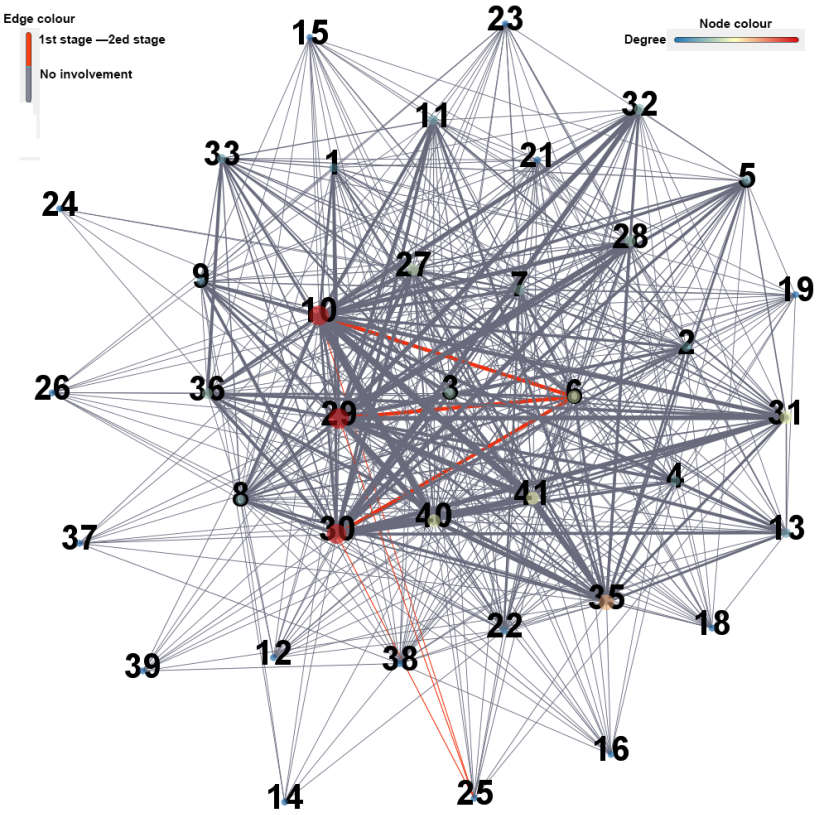


(b) Failure network when samples are 2000

Fig. 4.10 First part of failure network with the initial failures of Node 6 and 25



(c) Failure network when samples are 3000



(d) Failure network when samples are 4000

Fig. 4.10 Second part of failure networks with the initial failure of Node 6 and 25

Fig. 4.10 displays the failure networks with initial failures of node 6 and 25. Analysing the degrees of each node in Fig. 4.10, it can be concluded that the nodes which have higher degrees are mode vulnerable. In terms node 6 and node 25, node 6 has much higher degrees than node 25, which means node 25 are much more secure than node 6. As a result, the connections of node 25 are barely to be observed.

In order to see the cascading failure propagation more clearly, we transformed the result of Fig. 4.10 into Table 4.6. When the random samples are 1000, the cascading failure has three stages, but the cascading failures only have two stages after increasing the random samples more than 1000. Especially for the random samples of 3000 and 4000, the results of cascading failure propagation become consistent with two cascading stages.

Table 4.6 Cascading failure propagation with the initial failures of two nodes

Sample size	First stage	Second stage	Third stage
1000	(6, 25)	(22, 10, 29, 30)	(7, 40, 41)
2000	(6, 25)	(22, 10, 29, 30)	
3000	(6, 25)	(10, 29, 30)	
4000	(6, 25)	(10, 29, 30)	

Consequently, the final cascading failure propagation only has two stages. the components at the initial stages are obviously included node 6 and node 25, whereas the components at the second stages consist of node 10, node 29 and node 30. The details are shown below:

$$CFP = [(6,\ 25),\ (10,\ 29,\ 30)\ ]$$

4.4 Discussion

Comparing with the normal failure model, the first advantage of the extreme weather model is to solve the issue of the impact of extreme weather on power grids, and this model can further identify the cascading failure propagation under the extreme weather condition. Employing the extreme weather model, it can truly help us discover how the cascading failure will propagate in power grids under the extreme weather condition. Even though we used snowfall amounts as an example in the

extreme weather model, this model can be extended to other types of weather such as rains, winds, thunders, etc.

The second contribution of the model is that the initial failures to simulate the cascading failure can be randomly chosen. In the normal failure model, the initial failure of cascading failure can only choose one transmission line. However, in the extreme weather model, the initial failure can choose more than one transmission line (the limitation is that the choosing transmission lines must be within the area affected by the extreme weather).

The third contribution is that we proposed a new propagation mechanism in the complex network. Traditionally, the route problems can be solved by Hamiltonian path [91], Minimum spanning tree [92], Shortest path problem [93] and so on. Moreover, many researchers previously proposed some efficient algorithms to solve the issue of propagation such as Bellman–Ford algorithm [94], Dijkstra's algorithm [95], Prim's algorithm [96] and so on. We even tried the epidemiological models such as SIR model [97] and SIS model [98] to reveals the propagation in a network. However, those algorithms seems to be inappropriate for the complex network which is transformed by all cascading failure chains. Considering the features of the cascading failures in power grids, we proposed new indicators to reveal to cascading failure propagation in the failure network.

Table 4.7 Comparing the results of two models with the same initial failure

	The normal failure model	The extreme weather model
First stage	(10)	(10)
Second stage	(40, 41)	(40, 41)
Third stage	(30)	(29, 30, 35)
Fourth stage	(28, 31, 32)	(28, 32, 36)

Comparing the results of the normal failure model and the extreme weather model (both results are due to choose No. 10 line as the initial failure), we find that those two results are very similar. The comparison results are displayed in Table 4.7. At the first stages, the components in both models are exactly the same. The difference happens from the third stage, but the components of both models at the third and fourth stage are not completely different. For example, No. 30 appears at the third stage of both models, and No. 28, No.32 appears at the fourth stage of

both models. Therefore, we believe the results in the extreme weather model are acceptable and reliable.

However, the extreme weather model still has much to improve. Without adequate information, we assumed that the impact of different severity levels of snowfall amounts on power grids. If the data which presents the impact of the extreme weather on power grids can be collected, it is possible to predict the number of fault transmission lines because of the different snowfall amounts. Finally, the accuracy and reliability of the extreme weather model can be further improved.

Chapter 5

APPLYING THE EXTREME WEATHER MODEL TO ITALIAN TRANSMISSION NETWORK

5.1 Brief introduction

The work presented in the chapter aims to predict cascading failure propagation in Italian power grids by considering the snowfall amounts on western Italian Alps. The original datasets for building the transmission systems were from a few of EU projects and open resources, such as ENTSO-E System Study Model (STUM), the FP-project Pan European Grid Advance Simulation and state Estimation (PEGASE), Platts, etc. In terms of the extreme weather, we still used the snowfall data mentioned in Section 4.3.

5.2 Extreme weather in Italy

Italy is one of southern European countries, which is adjacent to France, Switzerland, Austria, and Slovenia. Italy has a large amout mountainous; Alps in the north and Apennines along the peninsula [99]. Landslides, earthquakes, and volcanic eruptions happen frequently in Italy because of the special geography, additionally the frequencies of some extreme weather conditions, such as floods and storms, are

also highly frequent. Collecting the data from The Emergency Events Database (EM-DAT) [100], Table 5.1 presents extreme weather happened in Italy from 1990 to 2017. It can be seen from Table 5.1 that floods and storms are the most frequent extreme weather conditions in Italy. In the past hundred years, the severe floods happened more than 40 times, and they caused around 1000 deaths. Even though the extreme temperatures happened less frequently, it has brought the greatest damage to this country comparing with other extreme weather conditions. Extreme weather in Italy not only occurs frequently but also leads to a great damage. Therefore, extreme weather, in Italy, is a issue that cannot be ignored.

Table 5.1 Extreme weather in Italy

Type	Subtype	Events count	Total deaths
Drought	Drought	4	0
Extreme temperature	Cold wave	3	45
	Heatwave	4	20115
	Severe winter storm	1	9
Flood	Flash flood	28	893
	Riverine flood	18	208
Storm	Convective storm	19	242
	Extra-tropical storm	1	3
	Tropical cyclone	1	35

5.3 Georeferenced model of Italian transmission network

In order to study how the snowfall on Alps will affect Italian power systems, especially for the northwestern part of Italy, we developed fully georeferenced models of the Italian transmission system with four typical power/demand snapshots (winter peak, winter off-peak, summer peak, summer off-peak) in 2014, then the worst case (in terms of post-contingency severity) “winter peak” was chosen.

Georeferenced models in power systems are used for planning, reinforcing, monitoring, and managing the transmission networks. The sophisticated spatial analysis is greatly useful for formulating scenarios, determining optimum generation potential, studying environmental impact, and managing facility assets. Owing to the

geographically-oriented view which combines the electric generation with transmission structures, devices, and network, a georeferenced model not only can be applied to the stability, protection and coordination, contingency analysis, and economic modeling, but also helps utilities to discover new issues about the investments and risks of building a transmission network, and allows the simultaneous assessment of technical, financial, and environmental factors [101].

The georeferenced model improves visualization of power systems by associating spatial data with transmission assets to display geographically referenced real-time power system data such as the voltage and line flow monitoring. Geographical information is stored in geographical map layers making it easy to integrate relevant information such as weather, vegetation growth, and road networks with related transmission network. Data of real-time weather integrated into a geographical map of power system increases the operator's situational awareness. For example, with the help of such model, the identification of a natural threat front moving towards a given area enables operators to rapidly determine transmission facilities with increased risks of outage

The georeferenced model of "winter peak" snapshot is shown in Fig. 5.1. This snapshot contains a complete set of buses and branches (lines, transformers) of the 220-380 kV Italian transmission network. Besides, network structures of important neighboring countries are simplified (the total numbers of buses, generators, lines, and transformers are around 1.2, 0.24, 1.4, and 0.2 thousands respectively). The maximum error of the power flow results, compared with the real network situation, is less than 2%. In order to combine the power system with the geographic information [102], the longitude and latitude of each bus, generator, and transmission tower were found from Google maps and mapped into the Italian power transmission network. Therefore, this model gives the precise location of each element in the system. Overall, the georeferenced model is highly similar to the network that is being used by the Italian transmission system operator (Terna S.p.A) for operation from the perspective of the static power flow. [103].

Fig. 5.1 Georeferenced model of Italian transmission network

5.4 Extreme weather model tests on the Italian transmission network

5.4.1 PDF of the snowfall data

In Section 4.2.1, we estimated the probability distribution of snowfall amounts based on the Poisson distribution. In order to employ Poisson distribution, we transform the snowfall data into discrete data, which was to used severity levels to represent

all snowfall data. In this section, we tend to use another solution to estimate the probability function of snowfall data. This solution is to treat snowfall data as continuous data and estimate its PDF. By evaluating the integral of its PDF, we can obtain the probability of each severity level of snowfall amounts.

We want to construct a curve (mathematical function) that has the best fit to the snowfall data, and the process is called curve fitting [104]. Many functions can be used for fitting curves to data such as linear function, exponential function, power function, logarithmic function, polynomial function, etc. [105]. Based on the distribution of snowfall data (Fig. 4.7 already shown the distribution of snowfall data on the western Italian Alps), two functions, polynomial function and exponential function, can be adopted.

Given a polynomial of degree n [106] (Euqation (5.1)), where a's are the coefficients. Curve fitting is the estimation of such coefficients. When fitting the data using a polynomial function, a polynomial of degree 3 in a normal situation is enough as increasing the degree cannot efficiently improve the accuracy after degree 3 [107].

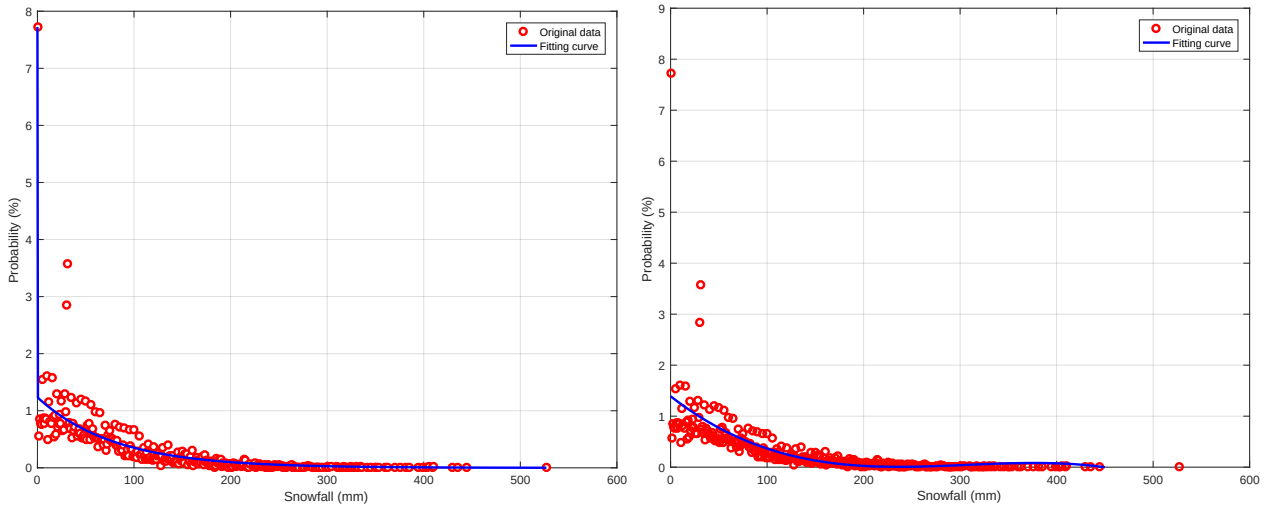
$$f(x) = a_n x^n + a_{n-1} x^{n-1} + ... + a_2 x^2 + a_1 x + a_0 \tag{5.1}$$

A exponential function [108] can be described as Equation (5.2), and a and b are coefficients of the exponential function. Sometimes in order to improve the accuracy of curve fitting, the exponential function can be extended as shown in Equation (5.3).

$$f(x) = a \times e^{bx} \tag{5.2}$$

$$f(x) = a \times e^{bx} + c \times e^{dx} \tag{5.3}$$

By using Matlab, we can estimate the two fitting functions of the snowfall data. One fitting function is based on a polynomial function of degree 3, and another is based on Equation (5.3). The two fitting functions can present the PDFs of the snowfall data, and the estimated fitting functions are presented as Equation (5.4) and Equation (5.5). The fitting of the curves to the snowfall data is displayed in Fig. 5.2.



(a) Exponential function to fit the snowfall data (b) Polynomial function to fit the snowfall data

Fig. 5.2 Fitting curves to the snowfall data

$$f_{polynomial}(x) = -5.5 \times 10^{-8}x + 5.08 \times 10^{-5} - 0.015x + 1.4 \tag{5.4}$$

$$f_{exponential}(x) = 6.5 \times e^{-23.6x} + 1.23 \times e^{-0.013x} \tag{5.5}$$

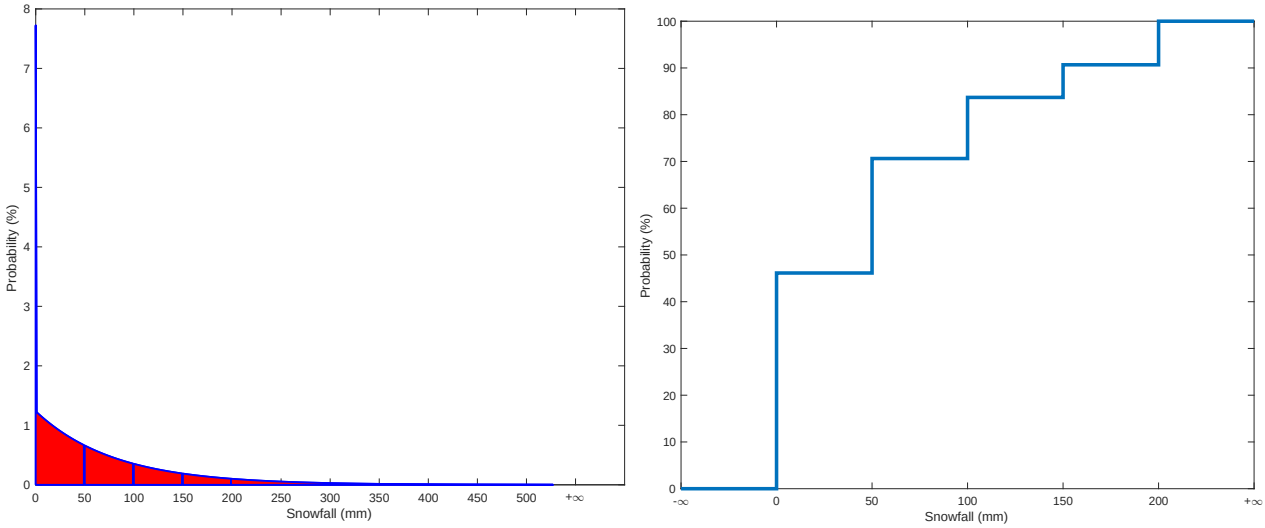
From Fig. 5.2, it is difficult to distinguish which fitting curve is more accurate. The root-mean-square error (RMSE) is a frequent solution to measure how close the observed data points are to the predicted values [109]. A better fit can be indicated by lower values of RMSE. RMSE, as one of the most important criteria for the fit, is a good measure of how precisely the model predicts the response [109]. RMSE can be describe as Equation (5.6), where n is the samples; $\widehat{y}_t$ and y_t are the predictive values and the original values respectively.

$$RMSE = \sqrt{\frac{1}{n}\sum_{t=1}^{n}(\widehat{y}_t - y_t)^2} \tag{5.6}$$

Based on Equation (5.6), we calculated RMSE for both fitting functions. RMSE of the polynomial fitting function equals to 0.43 while RMSE of the exponential fitting function equals to 0.24. Therefore, we have chosen the exponential fitting function as PDF of the snowfall data.

The PDF of the snowfall data is displayed in Fig. 5.3a. The probabilities of different severity levels can be obtained by calculating the red areas in Fig. 5.3a. In other words, calculating the red areas can be transformed to evaluate the integral of the exponential fitting function. The solution is shown in Equation (5.7), where SL presents the severity levels and $f_{exponential}(x)$ is the exponential function to fit the snowfall data. Accumulated those probabilities, we draw the cumulative distribution function (CDF) of the snowfall data as shown in Fig. 5.3b.

$$\Pr(X = SL) = \begin{cases} \int_{0}^{50} f_{\exponential}(x)d(x) & SL = 0 \\ \int_{50}^{100} f_{\exponential}(x)d(x) & SL = 1 \\ \int_{100}^{150} f_{\exponential}(x)d(x) & SL = 2 \\ \int_{150}^{200} f_{\exponential}(x)d(x) & SL = 3 \\ 1 - \sum_{SL=0}^{3} \Pr(X = SL) & SL = 4 \end{cases} \tag{5.7}$$



(a) PDF of the snowfall data (b) CDF of the snowfall data

Fig. 5.3 Estimate the probabilities of different severity levels

5.4.2 The prediction of the cascading failure propagation in the Italian transmission network under extreme weather

Considering the probabilities of different severity levels, as shown in Fig. 5.3, we randomly disconnected the transmission lines in the affected area (as the snowfall data was recorded on the western Italian Alps, the northwestern part of Italy was chosen as the affected area). Fig. 5.4 displays the affected area of the Italian transmission network, and the blue area is the affected area. To observe the transmission lines in the affected area, we zoomed in Fig. 5.4 and showed the affected transmission lines in Fig. 5.5

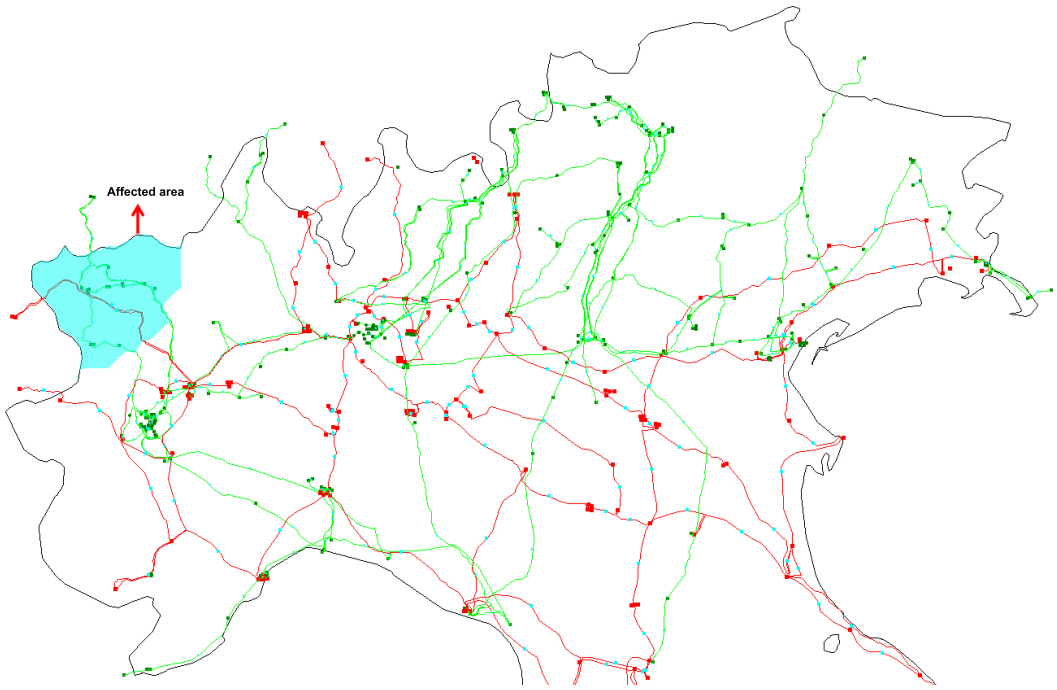


Fig. 5.4 Affected area in the Italian transmission network

To distinguish the affected transmission lines, we numbered those transmission lines based on their locations in the original data. The numbers of transmission lines in the affected area are 739, 740, 466, 747, 908, 514, 283, 285, 284, 518, 345, 84, 83, 85 and 896. If a line has two numbers, it means that this is a transmission line with two circuits. Overall, there are total 15 transmission lines in the affected area.

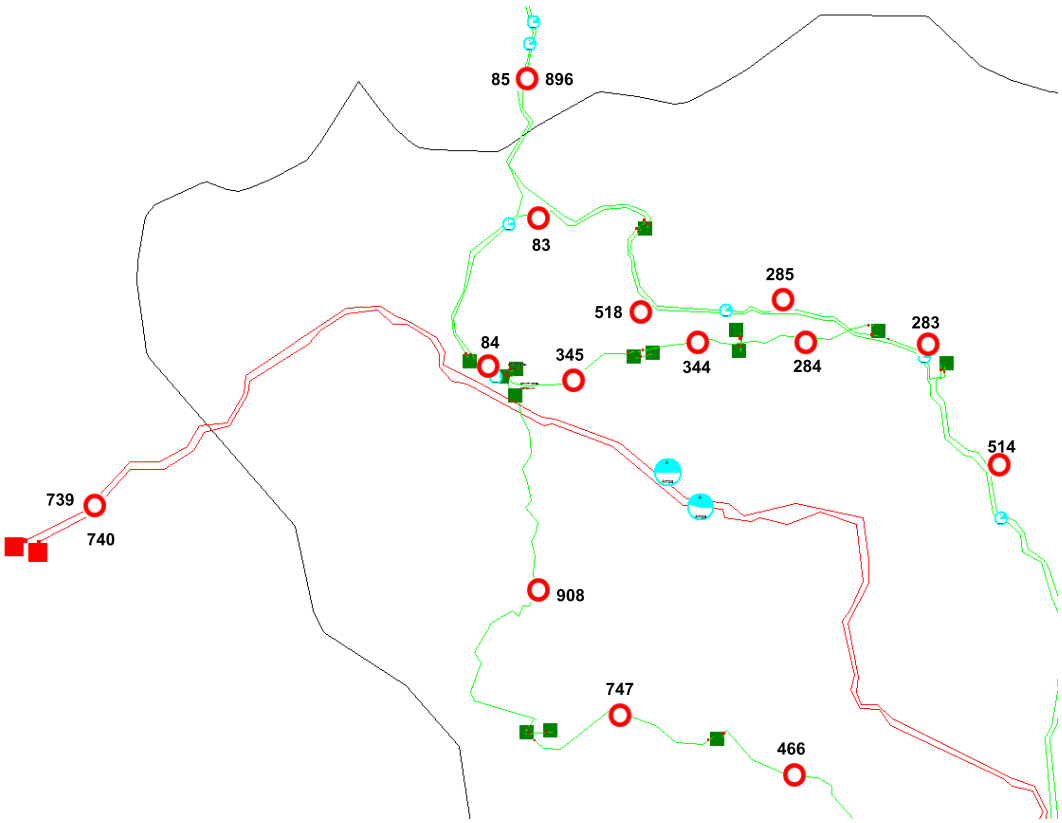


Fig. 5.5 Transmission lines in the affected area

The method to establish failure networks to predict the cascading failure propagation in Italian transmission network is similar with Section 4.2.3. The difference is that the new established failure networks will be the directed graphs. In Chapter 4, the extreme weather model was implemented on a small system, and we found that, increasing the sample size, the edges between two nodes were always bidirectional. Therefore, we assumed the failure networks in Chapter 4 were undirected graphs. However, the situation is different from establishing a failure network based on real power grids. A real system contains a large number of components, so the edges between two nodes are rarely bidirectional in the new established failure network. Consequently, failure networks based on the Italian transmission network are directed graphs.

As the failure networks are directed graphs, we made another difference which was that we used "in-degree" instead of degree to present the size of a node in the failure networks. In case of directed graphs, the degree of a node can be classified into two types: "in-degree" and "out-degree". In-degree of the corresponding node is

number of edges going into a node and number of edges coming out from a node is known as out-degree of the corresponding node [110]. To illustrate the importance of nodes in failure networks, in-degree is much more persuadable. If we presented the size of a node by using degree, the result would be that the initial nodes had the highest degrees, which was meaningless for the failure network. However, using in-degree to present the size of a node can indicate that a node with a higher in-degree would have a higher possibility to involve the cascading failure propagation.

Fig. 5.6 displays the failure networks based on the Italian transmission network. As the Italian transmission network is much larger than the 30-bus system, we increased the minimum sample sizes from 1000 to 5000. As a result, the samples sizes to implement the extreme weather model were 5000, 6000, 7000 and 8000. In terms of the initial failures, we assumed that two transmission lines happened failures because of the heavy snowfall, and the two transmission lines were 908 and 285.

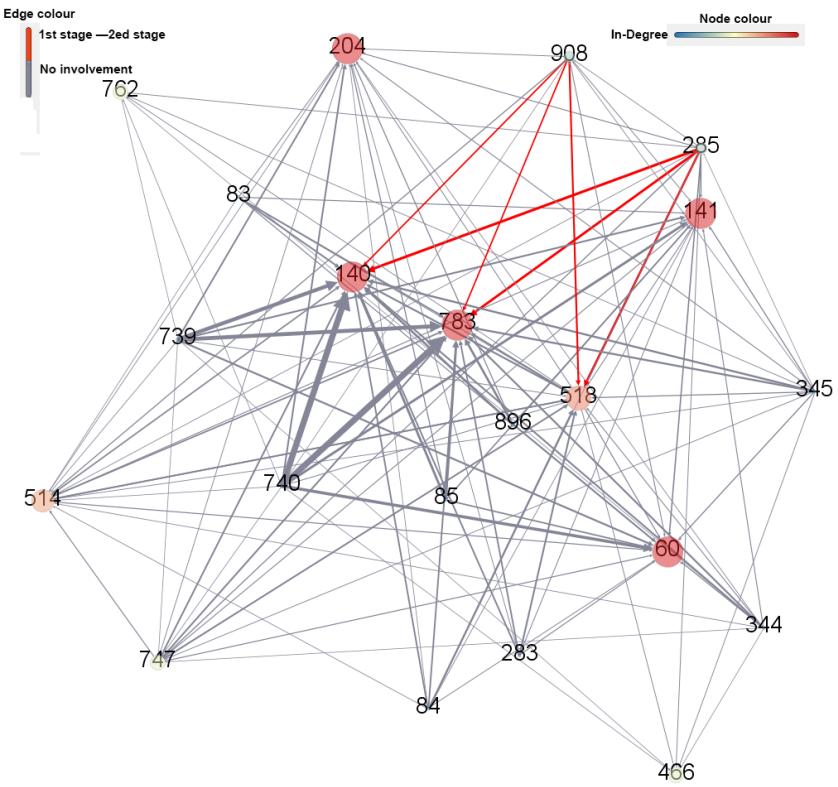
As shown in Fig. 5.6, we can conclude that the cascading failure propagates through three stage, and then the system happens blackout. In Fig. 5.6a, there are just two stages when the sample size is 5000 and the components at the second stage contain 140, 518 and 783. Increasing the samples to 6000 as shown in Fig. 5.6b, the propagation is extended to four stages. 514 and 60 are the components at the third stage and the fourth stage respectively. After increasing the sample size more than 6000, the results become consistent. To be more clearly, we establish Table 5.2 to compare the results of different sample sizes.

Table 5.2 Cascading failure propagation in Italian transmission network

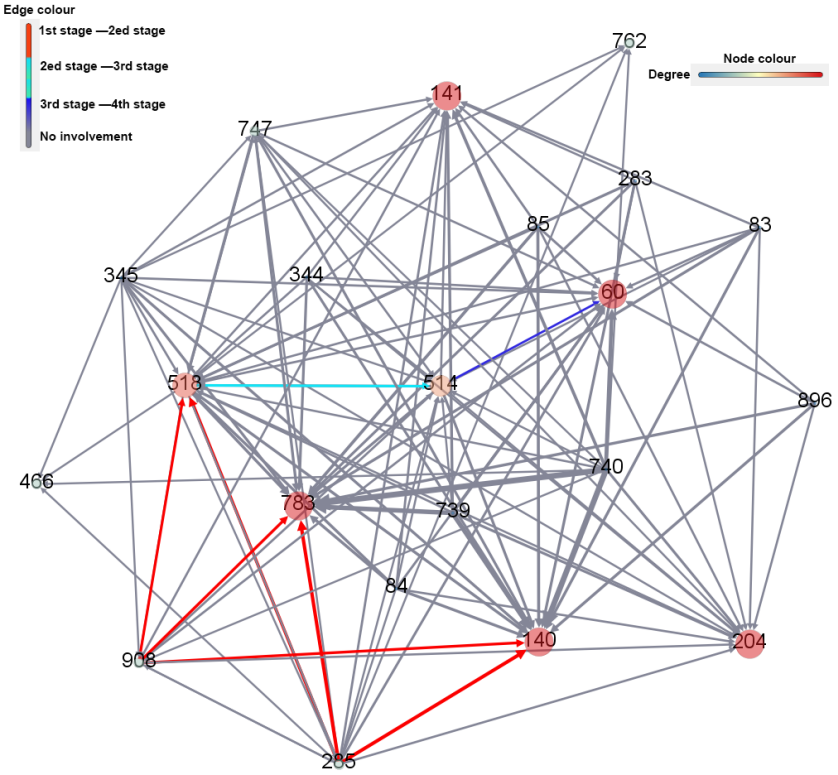
Sample size	First stage	Second stage	Third stage	Fourth stage
5000	908, 285	140, 518, 783		
6000	908, 285	140, 518, 783	514	60
7000	908, 285	140, 518, 783	514	60, 141, 204
8000	908, 285	140, 518, 783	514	60, 141, 204

The final result is displayed as below. Based on the result, we illustrate the propagation of cascading failure in the georeferenced model as shown in Fig. 5.7

$$CFP = ([908, 285],\ [140, 518, 783], [514], [60, 141, 204])$$

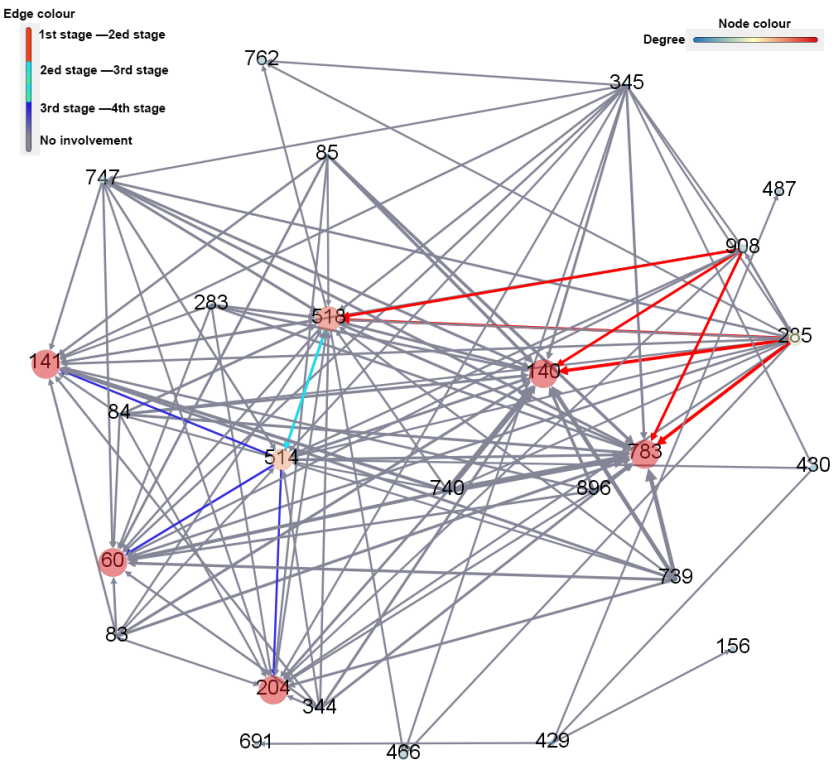


(a) Failure network when samples are 5000

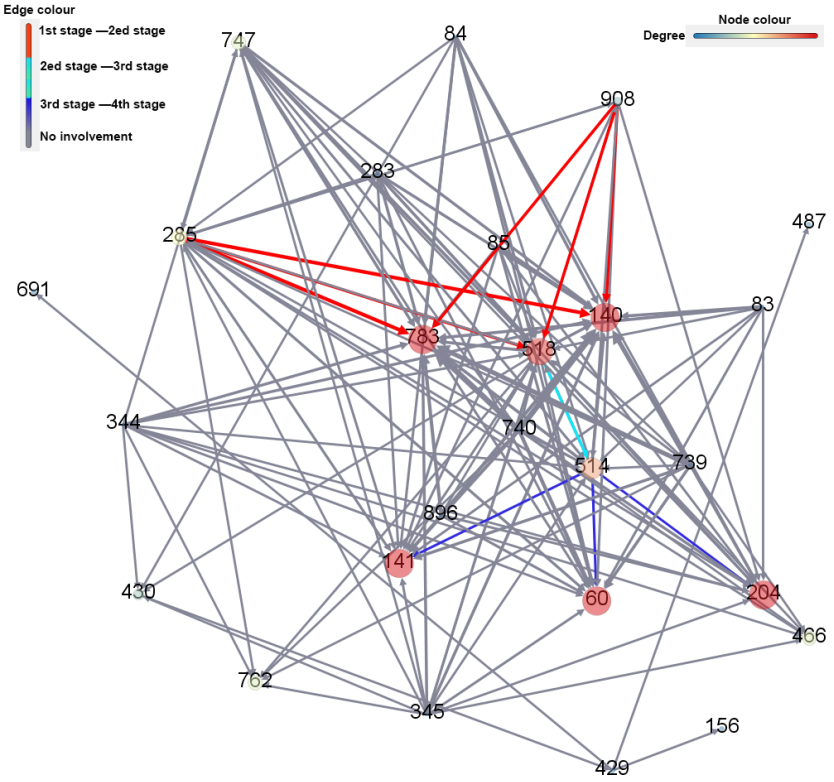


(b) Failure network when samples are 6000

Fig. 5.6 First part of failure network with the initial failures of 908 and 285



(c) Failure network when samples are 7000



(d) Failure network when samples are 4000

Fig. 5.6 Second part of failure networks with the initial failures of 908 and 285

(a) First stage of cascading failure

(b) Second stage of cascading failure

Fig. 5.7 First part of cascading failure propagation

(c) Third stage of cascading failure

(d) Fourth stage of cascading failure

Fig. 5.7 Second part of cascading failure propagation

As evidenced by Fig. 5.7, the cascading failure did not propagate through the near areas. At the first stage, the failures happened in western part of Italian transmission network, but at the next stage, not only did the west happened a failure but the middle and the east also happen failures. At the third stage, the cascading failure only propagated through the west, but at the last stage, the failures happened around the middle and the east again.

5.5 Discussion

Comparing the application of the extreme weather model to the 30-bus system, we mainly made two differences to employ the extreme weather model to estimate the cascading failure propagation in the Italian transmission network.

The first difference is the method to model the snowfall data. In Chapter 4, we transformed the snowfall data into discrete data, and then using Poisson distribution to model the PMF of the snowfall data. However, in this chapter, we modeled the snowfall data without any data transformation and the snowfall data still was continuous data. Next, we employed the techniques of curve fitting to estimate the PDF of snowfall data. Overall, the results of adapting those two different methodologies to estimate the probability distribution of the snowfall data are very similar.

Another difference is the method to establish the failure network. In Chapter 4, the established failure networks based on a small system were undirected graphs, since we found that combining all cascading failure chains and they would propagate through all transmission lines if the sample size was large enough. Nevertheless, in this chapter, the failure networks were established as directed graphs. We assumed that the affected area of the extreme weather only a small part of the power grid, so all the cascading failure chains cannot propagate all transmission lines. If the assumption that the entire power grid would be affected by the extreme weather was made, then the failure networks can be established as undirected graphs.

It is noteworthy that, even though the Italian transmission network is much larger than the 30-bus system, the failure networks based on the 30-bus system are more complicated, in terms of the number of links and nodes, than the failure networks established by Italian transmission network. This means that the complexity of the

extreme weather model is not related to the size of a electrical system. Another evidence is the computing time. The time to establish the failure networks based on the two electrical systems is similar. As a consequence, to identify the cascading failure propagation, the extreme weather model is also suitable for the large power transmission system.

Chapter 6

CONCLUSION AND FUTURE WORK

This thesis mainly focused on investigating the cascading failure propagation under the normal situation and the extreme weather condition. At the beginning of this thesis, a normal failure model was proposed to study the cascading failure propagation without considering the extreme weather. Based on the normal failure model, the following paragraphs introduced the extreme weather model to have an insight into the cascading failure under the extreme weather.

Taking no account of the extreme weather, the normal failure model is an accurate solution to predict the propagation of cascading failure in a power grid. At the same time, it can evaluate the vulnerable parts of a system. As the original intention of proposing this model, it can only handle the situation without considering the weather. Another significant drawback was the necessary time to finish the whole process. The sacrifice of the accuracy was due to the computing time, so at the current stage, the normal failure model is much more suitable for academic research. Now to estimate the cascading failure in a 30-bus system, it still needs several hours.

In the future, to apply the normal failure model to a real power grid, the facing problem is to shorten the computing time. To solve this problem, new advanced computer techniques, such as cloud computing, distributed computing, paralleled computing, etc., will be indispensable.

Considering the factor of extreme weather, the extreme weather model is an effective solution to predict the cascading failure propagation under the extreme

weather condition. As mentioned before, there are no fewer references to investigate the propagation of cascading failure under a extreme weather. The extreme weather model proposed in this thesis can fill this gap. Moreover, The extreme weather is less time consuming than the normal failure model, as the computing time will not greatly increase when the size of a system becomes larger

To implement the extreme weather model, two points need to be noticed. One is the method to model the weather data. There were two methods to model the weather data. One was to transform the weather data into the discrete data and estimated its PMF by using Poisson distribution. Another was to use the techniques of curve fitting to directly estimate the PDF of the weather data. There is no obvious advantage or disadvantage for those two methods, but employing the two different methods can validate the results of modeling the probability distributions.

Another interesting aspect is the established failure networks. In this thesis, the method to establish failure networks was proposed and those failure networks can be understandable how a cascading failure will propagate in a network. Additionally, the failure network could be the undirected graph or the directed graph. It depends on the size of the system. If the failure network is established based on a small system, then the failure network would be a undirected graph. Otherwise, the failure network would be a directed graph. The special case is that the failure network based on a large network also could be a directed graph if the extreme weather would affect the entire grid.

Predicting the cascading failure propagation under the extreme weather can actually help the design of power systems. Installing intelligent sensors and measuring equipment for the transmission lines which have high possibility to cause cascading failures, it can efficiently improve the protection of those transmission lines and finally prevent the occurrence of cascading failures. Strengthen the monitoring and early warning of dangerous line running conditions in extreme weather, and make prediction plans. The thesis can be helpful for the engineer to forecast cascading failures, diagnose sudden incidents, and then make corresponding health maintenance and decision scheme, which will provide theoretical support for ensuring the realisation of intelligent functions in Smart Grid, such as anti-accident ability.

Although the extreme weather model proposed in this thesis revealed the propagation of cascading failure from a new perspective, some aspects also can be improved to make this model as good as possible.

In the extreme weather model, an assumption was made to assume about the impact of the extreme weather on power grids. To improve the accuracy of the model, in the future, a new method can be provided to solve this problem. The idea of this method can be implemented as following steps: 1) the most important step is to collect the reliable data which indicates the impact of the extreme weather on power grids; 2)the next step is to use the techniques of machine learning to train the collecting data; 3) the final step is to establish a learning model to predict how the extreme weather will affect the power grids

As only one extreme weather was considered in this thesis, more extreme weather could be addressed to reveal the mechanism from other different angles. For example, references [111–113] mentioned that wind and lighting are the two major weather conditions to cause the failures in distribution systems. Therefore, it is significantly important to extend the extreme weather model to consider other major weather conditions in the future.

References

[1] J. W. Bialek. Why has it happened again? comparison between the ucte blackout in 2006 and the blackouts of 2003. In *2007 IEEE Lausanne Power Tech*, pages 51–56, July 2007.

[2] S. Corsi and C. Sabelli. General blackout in italy sunday september 28, 2003, h. 03:28:00. In *IEEE Power Engineering Society General Meeting*, pages 1691–1702, 2004.

[3] R. Baldick, B. Chowdhury, I. Dobson, Dong Zhaoyang, Gou Bei, D. Hawkins, Huang Zhenyu, Joung Manho, Kim Janghoon, D. Kirschen, Lee Stephen, Li Fangxing, Li Juan, Li Zuyi, Liu Chen-Ching, Luo Xiaochuan, L. Mili, S. Miller, M. Nakayama, M. Papic, R. Podmore, J. Rossmaier, K. Schneider, Sun Hongbin, Sun Kai, D. Wang, Wu Zhigang, Yao Liangzhong, Zhang Pei, Zhang Wenjie, and Zhang Xiaoping. Vulnerability assessment for cascading failures in electric power systems. In *2009 IEEE/PES Power Systems Conference and Exposition*, pages 1–9, 2009.

[4] Petter Holme. Edge overload breakdown in evolving networks. *Phys. Rev. E*, 66:036119, Sep 2002.

[5] L. Liu, Y. Yin, and Z. Zhang. Cascading failure of interdependent networks with traffic: Using a redundancy design to protect influential nodes. In *2016 11th International Conference on Reliability, Maintainability and Safety (ICRMS)*, pages 1–6, Oct 2016.

[6] P. F. Petersen, H. Jóhannsson, and A. H. Nielsen. Unweighted betweenness centrality for critical fault detection for cascading outage assessment. In *2016 IEEE International Energy Conference (ENERGYCON)*, pages 1–6, April 2016.

[7] S. Tauch, W. Liu, and R. Pears. Measuring cascading failures for smart grids vulnerability assessment. In *2015 IEEE International Conference on Data Science and Data Intensive Systems*, pages 384–389, Dec 2015.

[8] Baharan Mirzasoleiman, Mahmoudreza Babaei, Mahdi Jalili, and MohammadAli Safari. Cascaded failures in weighted networks. *Physical Review E*, 84(4):046114, 2011.

[9] Shuangshuang Jin, Zhenyu Huang, Yousu Chen, Daniel Chavarría-Miranda, John Feo, and Pak Chung Wong. A novel application of parallel betweenness centrality to power grid contingency analysis. In *Parallel & Distributed Processing (IPDPS), 2010 IEEE International Symposium on*, pages 1–7. IEEE, 2010.

[10] Paul Hines and Seth Blumsack. A centrality measure for electrical networks. In *Hawaii International Conference on System Sciences, Proceedings of the 41st Annual*, pages 185–185. IEEE, 2008.

[11] Xiao Fan Wang and Guanrong Chen. Complex networks: small-world, scale-free and beyond. *IEEE Circuits and Systems Magazine*, 3(1):6–20, 2003.

[12] J Ash and D Newth. Optimizing complex networks for resilience against cascading failure. *Physica A: Statistical Mechanics and its Applications*, 380:673–683, 2007.

[13] M. Babaei, H. Ghassemieh, and M. Jalili. Cascading failure tolerance of modular small-world networks. *IEEE Transactions on Circuits and Systems II: Express Briefs*, 58(8):527–531, Aug 2011.

[14] Y. Zheng, W. y. Liu, Z. w. Wen, and D. m. Ping. A real-time searching system for cascading failures based on small-world network. In *2010 Asia-Pacific Power and Energy Engineering Conference*, pages 1–5, March 2010.

[15] Lu Zongxiang, Meng Zhongwei, and Zhou Shuangxi. Cascading failure analysis of bulk power system using small-world network model. In *Probabilistic Methods Applied to Power Systems, 2004 International Conference on*, pages 635–640. IEEE, 2004.

[16] Meng Zhongwei, Lu Zongxiang, and Song Jingyan. Comparison analysis of the small-world topological model of chinese and american power grids [j]. *Automation of Electric Power Systems*, 15:004, 2004.

[17] Lu Zongxiang, Meng Zhongwei, and Zhou Shuangxi. Cascading failure analysis of bulk power system using small-world network model. In *2004 International Conference on Probabilistic Methods Applied to Power Systems*, pages 635–640, Sept 2004.

[18] X. Zhang and C. K. Tse. Assessment of robustness of power systems from the perspective of complex networks. In *2015 IEEE International Symposium on Circuits and Systems (ISCAS)*, pages 2684–2687, May 2015.

[19] P. Hines and S. Blumsack. A centrality measure for electrical networks. In *Proceedings of the 41st Annual Hawaii International Conference on System Sciences (HICSS 2008)*, pages 185–185, Jan 2008.

[20] K. Y. Wong, C. I. Wong, and H. H. Lao. Mitigation and recovery of cascading failures in scale-free networks. In *IET International Conference on Information and Communications Technologies (IETICT 2013)*, pages 253–257, April 2013.

[21] ZJ Bao, YJ Cao, LJ Ding, ZX Han, and GZ Wang. Dynamics of load entropy during cascading failure propagation in scale-free networks. *Physics Letters A*, 372(36):5778–5782, 2008.

[22] Y. Zhang, Z. Bao, and Y. Cao. Long-term effect of relay protection operation on cascading failures in growing scale-free small-world power grid. In *2012 IEEE Power and Energy Society General Meeting*, pages 1–6, July 2012.

[23] Zhang Jianli, Sun Jianhua, Yao Feng, Yang Nan, and Liu Wenying. Research on cascading failure spreading mechanism in the henan power grid with scale-free network characteristic. In *2011 2nd International Conference on Artificial Intelligence, Management Science and Electronic Commerce (AIMSEC)*, pages 5735–5738, Aug 2011.

[24] Yang Nan, Liu Wenying, and Guo Wei. Study on scale-free characteristic on propagation of cascading failures in power grid. In *IEEE 2011 EnergyTech*, pages 1–5, May 2011.

[25] Liang Zhao, Kwangho Park, and Ying-Cheng Lai. Attack vulnerability of scale-free networks due to cascading breakdown. *Physical review E*, 70(3):035101, 2004.

[26] I. Dobson, B. A. Carreras, V. E. Lynch, B. Nkei, and D. E. Newman. Estimating failure propagation in models of cascading blackouts. In *2004 International Conference on Probabilistic Methods Applied to Power Systems*, pages 641–646, Sept 2004.

[27] 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 Transactions on Power Systems*, 31(5):4159–4160, Sept 2016.

[28] Ian Dobson, Benjamin A Carreras, Vickie E Lynch, and David E Newman. Complex systems analysis of series of blackouts: Cascading failure, critical points, and self-organization. *Chaos: An Interdisciplinary Journal of Nonlinear Science*, 17(2):026103, 2007.

[29] Y. Gong, S. Mei, D. Peng, W. Long, and M. Guo. An improved opa model in power system considering planning. In *Proceedings of the 33rd Chinese Control Conference*, pages 2817–2822, July 2014.

[30] S. Mei, F. He, X. Zhang, S. Wu, and G. Wang. An improved opa model and blackout risk assessment. *IEEE Transactions on Power Systems*, 24(2):814–823, May 2009.

[31] B. A. Carreras, D. E. Newman, M. Zeidenberg, and I. Dobson. Dynamics of an economics model for generation coupled to the opa power transmission model. In *2010 43rd Hawaii International Conference on System Sciences*, pages 1–9, Jan 2010.

[32] D. E. Newman, B. A. Carreras, M. Kirchner, and I. Dobson. The impact of distributed generation on power transmission grid dynamics. In *2011 44th Hawaii International Conference on System Sciences*, pages 1–8, Jan 2011.

[33] H. Wu and I. Dobson. Analysis of induction motor cascading stall in a simple system based on the cascade model. *IEEE Transactions on Power Systems*, 28(3):3184–3193, Aug 2013.

[34] J. Bialek, E. Ciapessoni, D. Cirio, E. Cotilla-Sanchez, C. Dent, I. Dobson, P. Henneaux, P. Hines, J. Jardim, S. Miller, M. Panteli, M. Papic, A. Pitto, J. Quiros-Tortos, and D. Wu. Benchmarking and validation of cascading failure analysis tools. *IEEE Transactions on Power Systems*, 31(6):4887–4900, Nov 2016.

[35] Dusko P. Nedic, Ian Dobson, D.s Kirschen, Benjamin Carreras, and V Lynch. Criticality in a cascading failure blackout model. *Electrical Power and Energy Systems*, 28:627–633, 03 2006.

[36] Dusko P Nedic, Ian Dobson, Daniel S Kirschen, Benjamin A Carreras, and Vickie E Lynch. Criticality in a cascading failure blackout model. *International Journal of Electrical Power & Energy Systems*, 28(9):627–633, 2006.

[37] Shengwei Mei, Yixin Ni, Gang Wang, and Shengyu Wu. A study of self-organized criticality of power system under cascading failures based on ac-opf with voltage stability margin. *IEEE Transactions on Power Systems*, 23(4):1719–1726, 2008.

[38] Shengwei Mei, Fei He, Xuemin Zhang, Shengyu Wu, and Gang Wang. An improved opa model and blackout risk assessment. *IEEE Transactions on Power Systems*, 24(2):814–823, 2009.

[39] Sheng-wei Mei, Xiao-feng Weng, An-cheng Xue, et al. Blackout model based on opf and its self-organized criticality. In *Control Conference, 2006. CCC 2006. Chinese*, pages 1673–1678. IEEE, 2006.

[40] Francesco Cadini, Gian Luca Agliardi, and Enrico Zio. A modeling and simulation framework for the reliability/availability assessment of a power transmission grid subject to cascading failures under extreme weather conditions. *Applied Energy*, 185(Part 1):267 – 279, 2017.

[41] R. Billinton and G. Singh. Application of adverse and extreme adverse weather: modelling in transmission and distribution system reliability evaluation. *IEE Proceedings - Generation, Transmission and Distribution*, 153(1):115–120, Jan 2006.

[42] K. Alvehag and L. Soder. A reliability model for distribution systems incorporating seasonal variations in severe weather. *IEEE Transactions on Power Delivery*, 26(2):910–919, April 2011.

[43] D. Koller and N. Friedman. *Probabilistic Graphical Models: Principles and Techniques*. Adaptive computation and machine learning. MIT Press, 2009.

[44] Ali Kaveh. *Structural mechanics: graph and matrix methods*. Number 1. Research Studies PressLtd, 2004.

[45] Richard Trudeau. *Introduction to graph theory*. Dover Pub, New York, 1993.

[46] Claude Berge. *Graphs*, volume 6. North-Holland, 1985.

[47] John Adrian Bondy and Uppaluri Siva Ramachandra Murty. *Graph theory with applications*, volume 290. Macmillan London, 1976.

[48] Delbert Ray Fulkerson. Packing rooted directed cuts in a weighted directed graph. *Mathematical Programming*, 6(1):1–13, 1974.

[49] Andrey A Dobrynin and Amide A Kochetova. Degree distance of a graph: A degree analog of the wiener index. *Journal of Chemical Information and Computer Sciences*, 34(5):1082–1086, 1994.

[50] Michel Gondran and Michel Minoux. *Graphs and algorithms*. Wiley, 1984.

[51] Robert Ash. *Basic probability theory*. Dover Publications, Mineola, N.Y, 2008.

[52] Ruma Falk. *Understanding probability and statistics : a book of problems*. Ak Peters, Wellesley, 1997.

[53] Dimitri Bertsekas. *Introduction to probability*. Athena Scientific, Belmont, Mass, 2008.

[54] Allan Gut. *Probability: A Graduate Course : a Graduate Course*. Springer New York Imprint Springer, New York, NY New York, NY, 2013.

[55] Daniel Zwillinger. *CRC standard probability and statistics tables and formulae*. Chapman & Hall/CRC, Boca Raton, 2000.

[56] T Kanakubo and S Tanioka. Natural hazard mapping. *GeoJournal*, 4(4):333–340, 1980.

[57] Institute of Medicine. *Global Climate Change and Extreme Weather Events: Understanding the Contributions to Infectious Disease Emergence: Workshop Summary*. The National Academies Press, Washington, DC, 2008.

[58] Robert E Gabler, James F Petersen, L Trapasso, and Dorothy Sack. *Physical geography*. Nelson Education, 2008.

[59] Ready.gov. Natural disasters. http://www.ready.gov/natural-disasters. Accessed May 07, 2015.

[60] Peter J Webster, Greg J Holland, Judith A Curry, and H-R Chang. Changes in tropical cyclone number, duration, and intensity in a warming environment. *Science*, 309(5742):1844–1846, 2005.

[61] Martin Beniston. The 2003 heat wave in europe: A shape of things to come? an analysis based on swiss climatological data and model simulations. *Geophysical Research Letters*, 31(2), 2004.

[62] Matthew E Hirsch, Arthur T Degaetano, and Stephen J Colucci. An east coast winter storm climatology. *Journal of Climate*, 14(5):882–899, 2001.

[63] Ettore Bompard, Tao Huang, Yingjun Wu, and Mihai Cremenescu. Classification and trend analysis of threats origins to the security of power systems. *International Journal of Electrical Power & Energy Systems*, 50:50–64, 2013.

[64] Arthur NL Chiu. *Hurricane Iwa, Hawaii, November 23, 1982*. National Academies, 1983.

[65] National Oceanic and Atmospheric Administration. National weather service, "hurricane iniki september 6 - 13,1992,". Technical report, National Oceanic and Atmospheric Administration, 1992.

[66] M. Abley and J. Robinson. The ice storm. toronto. *M&S*, 1998.

[67] RMS. The 1998 ice storm: 10-year retrospective. Technical report, RMS, 2008.

[68] World Meteorological Organization. The global climate 2001-2010, a decade of climate extremes summary report. Technical report, Geneva: WMO, 2013.

[69] Andréa De Bono, Pascal Peduzzi, Stéphane Kluser, and Gregory Giuliani. Impacts of summer 2003 heat wave in europe. 2004.

[70] Energex. Energex power restoration fast facts at 1pm, tuesday, 29 january - queensland energy. https://www.energex.com.au/media-centre/media-releases/releases/2013/energex-power-restoration-fast-facts-@-1pm,-tuesday,-29-january. Accessed May 07, 2015.

[71] K. Hatipoglu and I. Fidan. Matlab based gui to investigate effect of voltage changes on static zip load model in a microgrid. In *IEEE SOUTHEASTCON 2014*, pages 1–5, March 2014.

[72] A. Bokhari, A. Alkan, R. Dogan, M. Diaz-Aguiló, F. de León, D. Czarkowski, Z. Zabar, L. Birenbaum, A. Noel, and R. E. Uosef. Experimental determination of the zip coefficients for modern residential, commercial, and industrial loads. *IEEE Transactions on Power Delivery*, 29(3):1372–1381, June 2014.

[73] Hephaestus Books. *Articles on Matrix Decompositions, Including: Cholesky Decomposition, Singular Value Decomposition, Matrix Decomposition, Qr Decomposition, Jordan Nor*. Hephaestus Books, 2011.

[74] Woodward. Governing fundamentals and power management. Technical report, Woodward, 2004.

[75] Shaghayegh Zalzar, Mohammad-Agha Shafiyi, Arzhang Yousefi-Talouki, and Mohammad-Sadegh Ghazizadeh. A smart charging algorithm for integration of evs in providing primary reserve as manageable demand-side resources. *International Transactions on Electrical Energy Systems*, pages n/a–n/a, 2016.

[76] B. Gao, G. K. Morison, and P. Kundur. Voltage stability evaluation using modal analysis. *Ieee Transactions on Power Systems*, 7(4):1529–1542, 1992.

[77] Daphne Koller. *Probabilistic graphical models*. Cambridge, Mass. [u.a.]: MIT Press, 1st edition, 2012.

[78] H. Qi, L. Shi, Sun Qiming, and Yao Liangzhong. Risk assessment of cascading failures based on entropy weight method. In *IEEE Power and Energy Society General Meeting (PESGM)*, pages 1–5, 2016.

[79] Robert J Beaver Mendenhall William. *Introduction To Probability And Statistics*. MA: CL-Wadsworth, 1st edition, 2013.

[80] Carlos E. Murillo-Sanchez Ray D. Zimmerman. Matpower 6.0 user's manual. Technical report, Power Systems Engineering Research Center (PSerc), 2016.

[81] Shahidehpour Mohammad and Wang Yaoyu. *Appendix C: IEEE30 Bus System Data*, pages 493–495. Wiley-IEEE Press, 2003.

[82] L.M. Surhone, M.T. Timpledon, and S.F. Marseken. *Student's T-Distribution: Probability, Statistics, Probability Distribution, Normal Distribution, Student's T-Test, Generalised Hyperbolic Distribution, William Sealy Gosset, Guinness Brewery*. Betascript Publishing, 2010.

[83] NOAA. Mav mos 24h snow best category. Technical report, NOAA, 2017.

[84] Frank Avery Haight. Handbook of the poisson distribution. 1967.

[85] Dragoš M Cvetković, Michael Doob, and Horst Sachs. *Spectra of graphs: theory and application*, volume 87. Academic Pr, 1980.

[86] H.M. Halff. *Graphical Evaluation of Hierarchical Clustering Schemes*. Eric reports. University of Illinois at Urbana-Champaign, Laboratory for Cognitive Studies in Education, 1975.

[87] Matlab. Hierarchical clustering. Technical report, Matlab, 2017.

[88] Don Wunsch and Rui Xu. *Clustering ((IEEE Press Series on Computational Intelligence))*. Wiley, Oxford, 2009.

[89] Per-Erik Danielsson. Euclidean distance mapping. *Computer Graphics and image processing*, 14(3):227–248, 1980.

[90] National Snow and Ice Data Center. Western Italian Alps monthly snowfall and snow cover duration. http://nsidc.org/data/docs/noaa/g01186_west_ital_alps_snow_fall/index.html. Accessed September 4, 2017.

[91] David S Johnson. The np-completeness column: an ongoing guide. *Journal of Algorithms*, 6(3):434–451, 1985.

[92] Ronald L Graham and Pavol Hell. On the history of the minimum spanning tree problem. *Annals of the History of Computing*, 7(1):43–57, 1985.

[93] Refael Hassin. Approximation schemes for the restricted shortest path problem. *Mathematics of Operations research*, 17(1):36–42, 1992.

[94] Baruch Awerbuch, Amotz Bar-Noy, and Madan Gopal. Approximate distributed bellman-ford algorithms. *IEEE Transactions on Communications*, 42(8):2515–2517, 1994.

[95] S Skiena. Dijkstra's algorithm. *Implementing Discrete Mathematics: Combinatorics and Graph Theory with Mathematica, Reading, MA: Addison-Wesley*, pages 225–227, 1990.

[96] Santiago Manen, Matthieu Guillaumin, and Luc Van Gool. Prime object proposals with randomized prim's algorithm. In *Proceedings of the IEEE international conference on computer vision*, pages 2536–2543, 2013.

[97] C Connell McCluskey. Complete global stability for an sir epidemic model with delay—distributed or discrete. *Nonlinear Analysis: Real World Applications*, 11(1):55–59, 2010.

[98] Alison Gray, David Greenhalgh, L Hu, Xuerong Mao, and Jiafeng Pan. A stochastic differential equation sis epidemic model. *SIAM Journal on Applied Mathematics*, 71(3):876–902, 2011.

[99] Marco Armiero. *Nature and history in modern Italy*. Ohio University Press, Athens, 2010.

[100] EM-DAT. The disaters list. http://emdat.be/emdat_db/. Accessed October 05, 2017.

[101] T.J. Overbye and J.D. Weber. Visualization of power system data. In *Proceedings of the 33rd Annual Hawaii International Conference*, page 7 pp., 01 2000.

[102] Terna s.p.a. Dati statistici: Rete elettrica. http://www.terna.it/LinkClick.aspx?fileticket=zrY9HJFeFng%3d&tabid=418&mid=2501. Accessed September 25, 2017.

[103] Baqer M AL-Ramadan. Applications of gis in electrical power system. 2013.

[104] Sandra Arlinghaus. *Practical handbook of curve fitting*. CRC Press, Boca Raton, Fla, 1994.

[105] Matlab. List of library models for curve fitting. https://uk.mathworks.com/help/curvefit/list-of-library-models-for-curve-and-surface-fitting.html. Accessed October 07, 2017.

[106] Math Centre. Polynonmial functions. http://www.mathcentre.ac.uk/resources/uploaded/mc-ty-polynomial-2009-1.pdf. Accessed October 07, 2017.

[107] W.M. Kolb. *Curve Fitting for Programmable Calculators*. Syntec, Incorporated, 1984.

[108] Math Centre. Exponential and logarithm functions. http://www.mathcentre.ac.uk/resources/uploaded/mc-ty-explogfns-2009-1.pdf. Accessed October 07, 2017.

[109] Tianfeng Chai and Roland R Draxler. Root mean square error (rmse) or mean absolute error (mae)?–arguments against avoiding rmse in the literature. *Geoscientific Model Development*, 7(3):1247–1250, 2014.

[110] Douglas Brent West et al. *Introduction to graph theory*, volume 2. Prentice hall Upper Saddle River, 2001.

[111] M Darveniza, C Arnold, B Holcombe, and P Rainbird. The relationships between weather variables and reliability indices for a distribution system in south-east queensland. In *CIRED (International Conference on Electricity Distribution), Vienna, Austria, 2007*, 2007.

[112] Anil Pahwa, Mark Hopkins, and TC Gaunt. Evaluation of outages in overhead distribution systems of south africa and of manhattan, kansas, usa. In *Proceedings of International Conference on Power Systems Operation and Planning, Cape Town, South Africa*, 2007.

[113] Thomas Allen Short. *Electric power distribution handbook*. CRC press, 2014.