RELIABILITY EVALUATION OF COMPOSITE POWER SYSTEMS INCLUDING THE EFFECTS OF HURRICANES

A Dissertation

by

YONG LIU

Submitted to the Office of Graduate Studies of
Texas A&M University
in partial fulfillment of the requirements for the degree of

DOCTOR OF PHILOSOPHY

December of 2010

Major Subject: Electrical Engineering

Reliability Evaluation of Composite Power Systems Including the Effects of Hurricanes

RELIABILITY EVALUATION OF COMPOSITE POWER SYSTEMS

INCLUDING THE EFFECTS OF HURRICANES

A Dissertation

by

YONG LIU

Submitted to the Office of Graduate Studies of
Texas A&M University
in partial fulfillment of the requirements for the degree of

DOCTOR OF PHILOSOPHY

Approved by:

Chair of Committee,	Chanan Singh
Committee Members,	Garng M. Huang
	Alex Sprintson
	Natarajan Gautam
Head of Department,	Costas N. Georghiades

December 2010

Major Subject: Electrical Engineering

ABSTRACT

Reliability Evaluation of Composite Power Systems Including the Effects of Hurricanes.

(December 2010)

Yong Liu, B.E., Huazhong University of Science and Technology;

M.E., Nanyang Technological University

Chair of Advisory Committee: Dr. Chanan Singh

Adverse weather such as hurricanes can significantly affect the reliability of composite power systems. Predicting the impact of hurricanes can help utilities for better preparedness and make appropriate restoration arrangements. In this dissertation, the impact of hurricanes on the reliability of composite power systems is investigated.

Firstly, the impact of adverse weather on the long-term reliability of composite power systems is investigated by using Markov cut-set method. The Algorithms for the implementation is developed. Here, two-state weather model is used. An algorithm for sequential simulation is also developed to achieve the same goal. The results obtained by using the two methods are compared. The comparison shows that the analytical method can obtain comparable results and meantime it can be faster than the simulation method.

Secondly, the impact of hurricanes on the short-term reliability of composite power systems is investigated. A fuzzy inference system is used to assess the failure rate increment of system components. Here, different methods are used to build two types of fuzzy inference systems. Considering the fact that hurricanes usually last only a few days, short-term minimal cut-set method is proposed to compute the time-specific system and nodal reliability indices of composite power systems. The implementation demonstrates that the proposed methodology is effective and efficient and is flexible in its applications.

Thirdly, the impact of hurricanes on the short-term reliability of composite power systems including common-cause failures is investigated. Here, two methods are proposed to archive this goal. One of them uses a Bayesian network to alleviate the dimensionality problem of conditional probability method. Another method extends minimal cut-set method to accommodate common-cause failures. The implementation results obtained by using the two methods are compared and their discrepancy is analyzed.

Finally, the proposed methods in this dissertation are also applicable to other applications in power systems.

To my parents, my wife and my daughter

ACKNOWLEDGEMENTS

Many people deserve my deep gratitude for their support throughout my studies in College Station.

Firstly, I would like to thank my advisor, Dr. Chanan Singh, for his financial support, profound expertise, and excellent guidance throughout these years. Without his help, I could not have made my dream come true.

I would also like to thank other committee members, Dr. Garng Huang, Dr. Alex Sprintson, and Dr. Natarajan Gautam, for their excellent guidance and valuable support throughout the course of this research. I have learned much from them.

Thanks also go to my friends and colleagues and departmental faculties and staff for making my time at Texas A&M University a great experience.

Finally, last but not least, I owe a huge debt of gratitude to my family for their unconditional support and endless love throughout these years. Without them, it would otherwise be impossible for me to be where I am now. My heart is with them forever.

The work reported in this dissertation was supported in part by NSF Grant ECCS 0725823.

NOMENCLATURE

ANFIS	Adaptive Neuro-Fuzzy Inference System
CCF	Common-Cause Failure
FES	Fuzzy Expert System
IMFR	Increment Multiplier of Failure Rate
M-FIS	Mamdani-Type Fuzzy Inference System
NERC	North American Electric Reliability Corporation
RTS	Reliability Test System
S-FIS	Sugeno-Type Fuzzy Inference System

TABLE OF CONTENTS

CHAPTER Page

Page

LIST OF FIGURES

Page

LIST OF TABLES

CHAPTER I

INTRODUCTION

In this chapter, the background of the research reported in this dissertation is introduced first, then the objectives of this dissertation are listed, and the organization of this dissertation is given in the end.

1.1 Introduction

Adverse weather such as hurricanes can significantly affect the operation of power systems, and it can jeopardize system reliability. In recent years, the hurricanes in the United States have caused hundreds of thousands of customers losing power supply. Moreover, due to the interdependency of various infrastructural systems, even brief power interruption may affect communication, water distribution, traffic signaling, and other lifeline systems [1]-[3]. Predicting the impact of hurricanes on power system reliability can help utilities for better preparedness and make appropriate restoration arrangements [2]-[3].

The impact of adverse weather on the reliability of power systems has been investigated in the past decades, i.e. the average effect of adverse weather over a long period of time. Some weather models have been proposed to evaluate power system reliability considering the effect of weather, e.g. two-state weather model [4] and three-state weather model [5].

This dissertation follows the style of *IEEE Transactions on Power Systems*.

When the effect of weather is considered, the states of the components of power systems can become dependent. For instance, when two-state weather model is used, usually a set of linear equations need to be solved by using Markov process [6]. However, this becomes impractical when applied to large power systems considering the fact that the number of system components is large. To solve this problem, usually Monte Carlo simulation can be used. But, due to its inherent nature of random experiments, simulation process can take long time to converge. In [6], Markov cut-set method was proposed to simplify the analytical approach. Its basic idea is that Markov process can be only applied to system minimal cut-sets as well as their unions, and its application to all system components is unnecessary. Although this method was described for some simple transmission configurations, it has not been developed for application to composite power systems, especially the nodal indices.

In this dissertation, algorithms are developed to implement Markov cut-set method and simulation method to evaluate the impact of adverse weather on the long-term reliability of composite power systems including system and nodal indices [7]. The obtained results by using different methods are compared and analyzed [7].

Usually, hurricanes last only a few days but their effect is drastic. Thus, the short-term impact of hurricanes, i.e. their dynamic impact during their durations, need to be investigated as their impact may not be reflected properly in the long term indices. Since a composite power system covers a large area, the weather models in [4]-[5] are not applicable and instead the regional weather model proposed in [8]-[9] can be used.

In this dissertation, the impact of hurricanes on the short-term reliability of composite power systems is investigated and the common-cause failures of system components are also considered. A common-cause failure refers to the simultaneous failures of multiple components due to a common cause [10], e.g. those of transmission lines installed on a same tower.

It has been known for a long time that the failure rate of a transmission or a distribution line is a function of the weather that it is exposed to and the failure rate of the transmission (distribution) line can be much higher in adverse weather than that in normal weather [11]. In this dissertation, a fuzzy inference system combined with regional weather model is used to assess the failure rate increment of system components cause by hurricanes. Additionally, different methods are proposed to build different types of fuzzy inference systems [12]-[14]. After the incremental failure rates of system components are obtained, short-term minimal cut-set method is proposed to compute the time-specific system and nodal reliability indices of composite power systems [13]. Here, only the independent failures of system components are considered.

Next, two methods are proposed to investigate the impact of hurricanes on the short-term reliability of composite power systems including common-cause failures [15]-[16]. One of them uses a Bayesian network to alleviate the dimensionality problem of conditional probability method when numerous common-cause failures are modeled [15]; the other method extends minimal cut-set method to accommodate common-cause failures [16]. The obtained results by using the two methods are compared and the difference is analyzed [16].

Finally, it is shown that the evaluation methods proposed in this dissertation are also applicable to distribution systems [17] and other applications [18], e.g. operational reliability, and intermittent renewable energy.

1.2 Objectives

The objectives of this dissertation are as follows:

1) Investigate the impact of adverse weather on the long-term reliability of composite power systems. The evaluation results can be used in power system planning.
2) Investigate the impact of hurricanes on the short-term reliability of composite power systems. The evaluation results can be used in power system operation.
3) Investigate the impact of hurricanes on the short-term reliability of composite power systems including the common-cause failures of components. Thus, the impact of hurricanes on the operational reliability of composite power systems can be predicted more accurately.

1.3 Organization of the Dissertation

This dissertation is organized as follows: in Chapter II, basic concepts of power system reliability and some weather models are introduced; in Chapter III, the investigation of the impact of adverse weather on the long-term reliability of composite power systems is presented; in Chapter IV, the investigation of the impact of hurricanes on the short-term reliability of composite power systems is presented; in Chapter V, the investigation of the impact of hurricanes on the short-term reliability including the common-cause failures of components is presented; finally, in Chapter VI, the evaluation methods proposed in this dissertation is summarized and their possible extensions are discussed.

CHAPTER II

POWER SYSTEM RELIABILITY EVALUATION

In this chapter, some basic concepts of power system reliability are introduced first; then, two models to consider the effect of weather are introduced.

2.1 Basics of Power System Reliability

Generally, reliability is defined as the ability of a system or component to perform its required functions under stated conditions for a specified period of time [19]. A key element of this definition is that the concerned system or component should operate under stated conditions. Operational environment such as weather is such a condition which should be addressed in reliability evaluation. The effect of adverse weather on power systems and other infrastructural systems are introduced in the next section.

For power systems, North American Electric Reliability Corporation (NERC) defines reliability as "the degree to which the performance of the elements of [the electrical] system results in power being delivered to customers within accepted standards and in the amount desired." Actually, NERC's definition of reliability includes two concepts: *adequacy* and *security*. *Adequacy* is defined as "the ability of the system to supply the aggregate electric power and energy requirements of the consumers at all times." NERC defines *security* as "the ability of the system to withstand sudden disturbance." In other words, *adequacy* refers to that sufficient system resources are available to meet predicted load with reserve for contingencies; *security* refers to that the system remains reliable even in contingent cases.

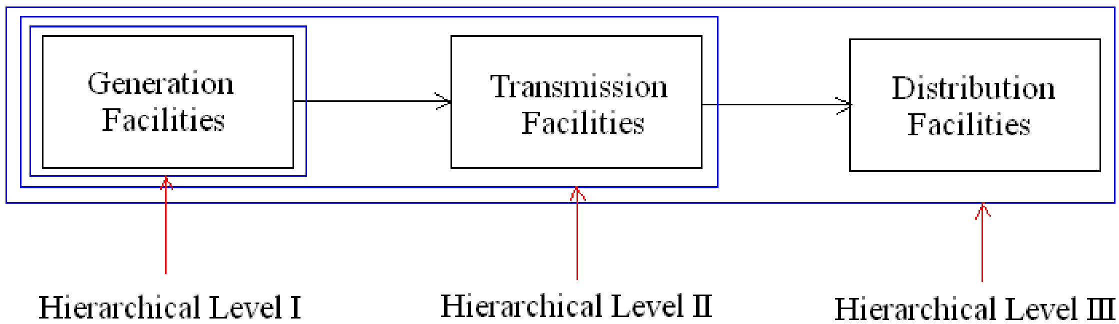

Fig. 1: Functional Zones and Hierarchical Levels

It is noted that most present probabilistic techniques for power system reliability evaluation are used for adequacy assessment. The available probabilistic techniques to assess the security of power systems are limited. Accordingly, most reliability indices used at the present time are adequacy indices.

The reliability evaluation can be implemented in different segments of a power system, i.e. functional zones, as well as the combinations of them which shapes the hierarchical levels shown in Fig. 1 [20].

The evaluation methods at different hierarchical levels of a power system can be different. For instance, at hierarchical level II the configuration of a transmission system is usually in a meshed fashion, and the effects of load flow, overload alleviation, generation rescheduling need to be considered. In contrast, at hierarchical level III the configuration of a distribution system is usually radial. Thus, power flow is usually not considered in distribution systems. Generally, the evaluation methods for power system reliability fall into two categories: analytical and simulation. The details of the algorithms to implement them are described in the following chapters. Clearly, the reliability assessment at hierarchical level III becomes very complex as it involves all three functional zones. Thus, the distribution system is usually analyzed as a separate part.

In this dissertation, the impact of hurricanes is investigated at hierarchical level II which is usually called a composite power system or a bulk power system. But, the proposed evaluation methods in this dissertation are also applicable to distribution systems. This is discussed in detail in Chapter VI.

2.2 Weather Models

2.2.1 Two-state weather model

In [4], each transmission line was assumed to operate in a two-state fluctuating environment as shown in Fig. 2, and Markov method for the whole transmission system was used to evaluate its reliability. In Fig. 2, the arrows represent the transition of component states. For simplicity, only the transition of the states of one component is illustrated.

Advantages of two-state weather model are its simplicity and easy implementation. This weather model can be used for long-term applications in power systems, e.g. power system planning. In this dissertation, it is used to evaluate the impact of adverse weather on the long-term reliability of composite power systems. But, for a composite power system the number of components is large, and applying Markov process to all components is impractical. In this dissertation, Markov cut-set method in [6] is used to solve this problem. The details are given in the next chapter.

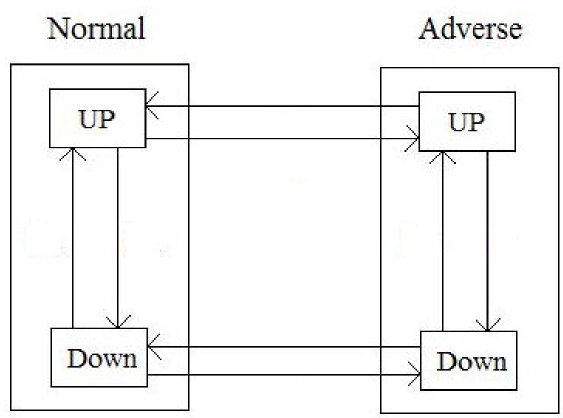

Fig. 2: Two-State Weather Model (Simplified)

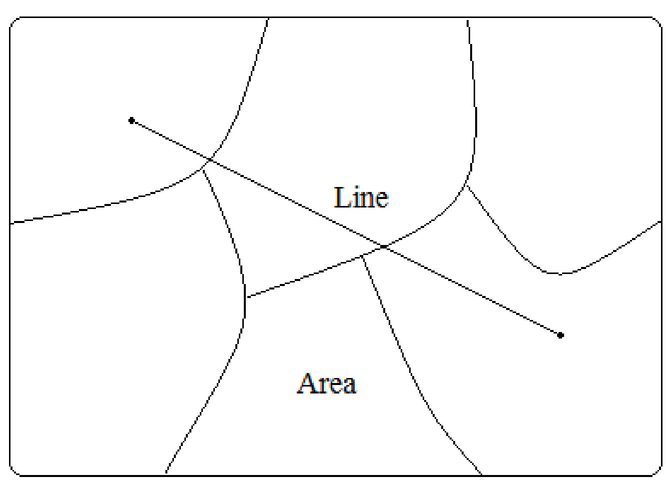

Fig. 3: Regional Weather Model

2.2.2 Regional weather model

An assumption adopted by previous weather model is that all system components are exposed to the same weather at a time. This is not true considering the fact that hurricanes develop and dissipate over time, i.e. the impact of hurricanes on a power system can be different temporally and spatially.

In [8]-[9], the regional weather model as shown in Fig. 3 was used to recognize the regional effects of weather that transmission lines are exposed to and Monte-Carlo simulation was used to evaluate the reliability of composite power systems.

In [2]-[3], similar regional weather model was applied to distribution systems and statistical regression method was used to predict the number of outages caused by hurricanes in each geographic unit.

2.2.3 Disaster impact on infrastructural systems

The impact of adverse weather on other infrastructural systems has been investigated in the literature. For instance, in [21]   a regional weather model similar to that in [2]-[3] was used to evaluate the performance of cellular networks during hurricanes. Moreover, the impact of other natural disasters on power systems has been investigated. For instance, in [22] simulation method was used to investigate the restoration process of power systems after earthquakes.

CHAPTER III

RELIABILITY EVALUATION OF COMPOSITE POWER SYSTEMS USING MARKOV CUT-SET METHOD

The impact of adverse weather on the long-term reliability of power systems have been investigated during the past decades. Usually, two-state weather model [4] is used to evaluate the effects of fluctuating weather on power system reliability. As a result, the states of system components are not independent anymore. To solve this problem, usually simulation method can be used. In this chapter, Markov cut-set method [6] is used to achieve the same goal and algorithms to implement this method are developed. For the purpose of comparison, algorithm to implement sequential simulation is also developed, and the results obtained by using the two methods are compared and analyzed.

This chapter is organized as follows: in Section 3.1 relevant researches using two-state weather model are reviewed; in Section 3.2 the assumptions adopted in this chapter are listed; in Section 3.3 minimal cut-set method is briefly introduced and the developed algorithms for Markov cut-set method are presented; in Section 3.4 Monte Carlo simulation is briefly introduced and the developed algorithm for sequential simulation is presented; in Section 3.5 the analytical and simulation methods proposed are applied to the modified IEEE reliability test system (RTS). The results obtained by using the two methods are presented and compared; finally, in Section 3.6 main conclusions obtained in this chapter are summarized.

3.1 Literature Review

In the reliability evaluation of power systems, usually the states of system components are assumed to be independent, and system reliability indices are calculated by using the methods based on the multiplication rule of probabilities [23]. But, in some cases, for instance, when the effect of fluctuating weather or common-cause failures is considered, the previous assumption is invalid. The main weather models used include two-state weather model [4] and its variant [5]. Generally, two kinds of methods can be adopted, namely analytical [4], [6], [23] and Monte Carlo simulation [24].

Generally, simulation method mimics the operational process of a physical system by using random experiments and obtains system reliability indices using statistical inference. Generally speaking, simulation method is suitable when complex system operational conditions are considered. In [24], basically two kinds of simulation methods are described: random sampling and sequential simulation. Generally, random sampling assumes that component states are independent and system states, i.e. the combinations of component states, are uncorrelated. Sequential simulation is more flexible and is suitable to simulate the effect of fluctuating weather. Relevant details are given in Section 3.4.

However, by the nature of simulation method, its convergence may need acceleration by using other techniques [24]. On the other hand, the Markov process used in [4] is accurate within the distribution assumptions, but it is only applicable to relatively small systems considering that the solution of 2^{n+1} linear equations is required [6]. Here, n is the number of system components.

To alleviate the dimensionality problem, a method was proposed in [23] to reduce the state space by merging system states and systematically deleting low probability states. In [6], Markov cue-set method was proposed to evaluate the reliability of transmission and distribution systems considering the effects of fluctuating weather. In [6], minimal cut-set method was used to compute system reliability indices, and Markov process was applied to the components of a minimal cut-set or a union of minimal cut-sets to alleviate the computational burden. Markov cue-set method is based on the concept that if two-state weather model is used, the reliability indices of a minimal cut-set or a union of minimal cut-sets can be calculated by applying Markov process only to its members, and the application of Markov process to all system components simultaneously is unnecessary. Thus, if the minimal cut-sets up to some order (e.g. third-order, i.e. the maximum number of components) are determined, only a limited number of linear equations need to be solved at a time. For example, considering a system of 500 components, if the entire system is to be modeled by using Markov process, there will be 2^{501} number of states and thus as many equations to be solved. However, if the maximum number of the components in a minimal cut-set or a union of minimal cut-sets is say 6, then by using Markov cut-set method, the highest number of equations to be solved at a time is 2^7 . This can make the difference in the practical applicability of Markov process.

However, in [6] Markov cut-set method was applied to a simple 5-component system, and the minimal cut-sets were determined by using simple enumeration method and the connectivity criterion in transmission and distribution systems. Additionally, nodal reliability indices were not computed, and the comparison with simulation method was not given in [6].

In this chapter, Markov cut-set method is used to investigate the impact of adverse weather on the long-term reliability of composite power systems. From previous discussion, it is clear that a key step of Markov cut-set method is the identification of minimal cut-sets. In this dissertation, this is modeled as a linear constrained optimization problem to shorten computational time. Since enumerating all minimal cut-sets of a power system is impractical and unnecessary, the algorithm of computing the bounds of minimal cut-sets is also developed. An important new feature is the method for computing nodal indices as these indices are important for assessing the impact of adverse weather as well as extensions to distribution systems.

3.2 Assumptions

In this chapter, the following assumptions are adopted:

1) Voltage is assumed as 1pu at each bus and DC power flow is used.

2) The distribution of state residence times is assumed exponential. Thus, Markov process can be used to compute reliability indices.

3) All reliability indices computed are steady state indices. Thus, the probability of a system state can be obtained by using the steady state condition of Markov process.

4) All system components have two possible states: success or failure.

3.3 Markov Cut-Set Method

In this section, the developed algorithms for implementing Markov cut-set method are presented in detail. Firstly, minimal cut-set method is briefly introduced; then, the algorithm for identifying system and nodal minimal cut-sets is presented; finally, the algorithm for computing the bounds of minimal cut-sets is presented.

3.3.1 Minimal cut-set method

A cut set is a set of components whose failures alone could cause system failure. Here, the definition of system failure is rather broad and it can be any kind of anomaly defined. In this dissertation, system failure refers to the load shedding at any node of a composite power system. A minimal cut-set has the further property that it has no proper subset of components whose failures alone could cause system failure. Here, the term "component" is also used in a broad sense. It can be any device in a power system and even can be a condition or a function whose presence or absence could cause system failure.

The basic idea of minimal cut-set method is to first identify the minimal cut-sets of a power system, and then use the following equations of probabilities to compute the reliability indices [6].

$$p_f = \sum_i p\left(\overline{C_i}\right) - \sum_{i<j} p\left(\overline{C_i} \cap \overline{C_j}\right) + \sum_{i<j<k} p\left(\overline{C_i} \cap \overline{C_j} \cap \overline{C_k}\right) - \cdots + (-1)^{m+1} \cdot p\left(\overline{C_1} \cap \overline{C_2} \cap \cdots \cap \overline{C_m}\right) \tag{1}$$

$$f_f = \sum_i p\left(\overline{C_i}\right) \cdot \overline{\mu_i} - \sum_{i<j} p\left(\overline{C_i} \cap \overline{C_j}\right) \cdot \overline{\mu_{i+j}} + \sum_{i<j<k} p\left(\overline{C_i} \cap \overline{C_j} \cap \overline{C_k}\right) \cdot \overline{\mu_{i+j+k}} - \cdots + (-1)^{m+1} \cdot p\left(\overline{C_1} \cap \overline{C_2} \cap \cdots \cap \overline{C_m}\right) \cdot \overline{\mu_{1+2+\cdots+m}} \tag{2}$$

$$d_f = p_f / f_f \tag{3}$$

where

p_f	=	failure probability
C_i	=	minimal cut-set *i*
$\overline{C_i}$	=	event that all members of C_i fail
$\overline{C_i} \cap \overline{C_j}$	=	joint event that all members of both C_i and C_j fail
m	=	number of minimal cut-sets
f_f	=	failure frequency
μ_i	=	repair rate of component *i*
$\overline{\mu_i}$	=	$\sum_{i \in C_i} \mu_i$
$\overline{\mu_{i+j}}$	=	$\sum_{i \in C_i \cup C_j} \mu_i$
d_f	=	mean duration of failure

In practice, enumerating all the minimal cut-sets of a power system and using (1)-(3) to compute the exact values of reliability indices are impractical and unnecessary. Instead, the minimal cut-sets are usually determined up to a desired order and the following equations are used to compute the bounds of the reliability indices to approximate the results of (1)-(3):

$$p_f^u = \sum_i p\left(\overline{C_i}\right) \tag{4}$$

$$p_f^l = \sum_i p\left(\overline{C_i}\right) - \sum_{i<j} p\left(\overline{C_i} \cap \overline{C_j}\right) \tag{5}$$

$$f_f^u = \sum_i p\left(\overline{C_i}\right) \cdot \overline{\mu_i} \tag{6}$$

$$f_f^l = \sum_i p\left(\overline{C_i}\right) \cdot \overline{\mu_i} - \sum_{i<j} p\left(\overline{C_i} \cap \overline{C_j}\right) \cdot \overline{\mu_{i+j}} \tag{7}$$

where

p_f^u = first upper bound of p_f

p_f^l = first lower bound of p_f

f_f^u = first upper bound of f_f

f_f^l = first lower bound of f_f

By using inclusion-exclusion formula, a sequence of increasingly closer bounds of the reliability indices can be obtained [25].

Following the above introduction, Markov cut-set method can be implemented as follows:

1) Determine the minimal cut-sets up to the preset order and the ones of higher order are ignored.

2) Compute the reliability indices of the minimal cut-sets and their unions. Here, the multiplication rule of probabilities is not applicable anymore. Instead, the algorithm developed in Subsection 3.3.4 can be used.

3) Use (3)-(7) to compute the reliability indices.

3.3.2 Identification of minimal cut-sets

In the literature, numerous methods have been proposed to generate minimal cut-sets to evaluate the reliability of large complex systems [26]-[34]. However, these graph-based methods mainly explore the connectivity feature of networks and are not suitable for the reliability evaluation of composite power systems considering the capacity and admittance of transmission lines. In [31], [33], although the link capacity of networks were considered, the proposed algorithms are only suitable for some general networks considering the fact that generation rescheduling and load shedding have to be considered in composite power systems. In [29], although the proposed method was implemented in power systems, the previous issues were not addressed.

Normally, a composite power system can be modeled as a capacitated-flow network subjected to some operational constraints, such as generation-load balance, generator capacity limits and voltage magnitude limits. In reliability evaluation, usually the analysis of failure effects should be implemented after the occurrence of a system event, i.e. determining the resultant system state is success or failure as defined. In a composite power system, after a system event occurs, e.g. the outage of a generator or the tripping of a transmission line, usually the output of generators is rescheduled first. If the violation of system constraints cannot be remedied, usually load shedding is finally executed.

In this dissertation, the identification of minimal cut-sets is modeled as a constrained linear optimization problem to reduce computational time. When the voltage is considered, the proposed algorithm can be easily extended to the AC model.

Mathematically, the objective is to minimize the amount of load shedding M_D if necessary and meantime the following constraints are satisfied:

Balance of active power flow:

$$P_G = P_D - P_{LD} \tag{8}$$

where

P_G = active power of total generator output

P_D = active power of system load

P_{LD} = active power of total load shedding

Transmission line capacity limit:

$$P_{i,j} \leq P_{i,j}^{\max} \tag{9}$$

where

$P_{i,j}$ = active power flow in transmission line from bus *i* to *j*

$P_{i,j}^{\max}$ = upper limit of $P_{i,j}$

Generator capacity limit:

$$P_g \leq P_g^{\max} \tag{10}$$

where

P_g = active power output of generator *g*

$P_g^{\max}$ = upper limit of P_g

Load shedding limit:

$$P_d \leq P_d^{\max} \tag{11}$$

where

P_d = active power shedding of load *d*

$P_d^{\max}$ = upper limit of P_d

The algorithm to determine the minimal cut-sets of a composite power system up to the preset order is as follows:

1) Choose an *n*-order arbitrary combination of system components.
2) Check all the existing lower-order minimal cut-sets to examine if they are the subsets of the combination in Step (1): if yes, go back to Step (1); if not, go to the next step.
3) Run the optimization routine on the condition that these *n* components are out of service simultaneously.
4) Examine if load shedding is needed: if yes, these components make up an *n*-order minimal cut-set; otherwise, not.
5) Check if all the *n*-order combinations of system components have been examined: if not, go back to Step (1); if yes, forward to the next step.
6) Check if the pre-set order of the combinations is reached: if yes, stop; if not, forward to the next step.
7) Set $n = n + 1$ and go back to Step (1).

The proposed algorithm has some advantages as follows:

1) The implementation is simple. Since linear optimization is widely used in various applications in power systems, the proposed algorithm can be implemented by slightly modifying the current software.
2) It is easy to extend this algorithm to incorporate more system operational considerations. For instance, it is simple to extend it to the AC model.
3) It is easy to compare the proposed algorithm with other methods since linear optimization is also used in other analytical and simulation methods to analyze the failure effects. For example, the number of calling the optimization routine can indicate the efficiency of a reliability evaluation method.

4) It can be used to compute system and nodal reliability indices. Although simulation method can achieve the same goal, the computation may be more expensive. This is discussed in detail in the next subsection.

However, a practical composite power system may have a large number of components. Even if the minimal cut-sets are determined to a small order, their number can be still very large. This problem can be further alleviated by using the approaches proposed in [35]. The basic idea is that using learning methods to classify system states as success or failure. Thus, the computational time can be reduced further. It should be pointed out that this problem is shared by analytical and simulation methods. Some intelligent methods can accelerate the algorithms of both of them.

*3.3.3 System and nodal minimal cut-set*s

As mentioned previously, the proposed algorithm can identify nodal minimal cut-sets too. Actually, it can identify system and nodal minimal cut-sets simultaneously. Thus, the computation of nodal reliability indices can be much simplified.

When a minimal cut-set is determined, the information about the nodes which suffer loss of load is saved. Thus, in the end there are two *lists*, first a *list* of all minimal cut-sets and then an additional *list* of nodes that have loss of load corresponding to each minimal cut-set. To compute system reliability indices, all the minimal cut-sets are used, i.e. they are system minimal cut-sets. To compute the reliability indices of a node, only those minimal cut-sets which have it suffering loss of load are used, i.e. they are the nodal minimal cut-sets. It should be pointed out that most of the computation time is spent in identifying the minimal cut-sets. The time taken by the computation of (3)-(7) is relatively small. Since the nodal minimal cut-set are the subsets of system minimal cut-sets, no additional time is needed for nodal reliability indices as far as the identification of minimal cut-sets is concerned. The only additional time needed is for the use of (3)-(7) for the calculation of nodal reliability indices and this is not significant. This point will be further illustrated in Section 3.6.

When simulation method is used to obtain nodal reliability indices, as pointed out in the above discussion, a system failure may have different effects at different nodes, i.e. it may only have some nodes suffering loss of load. Usually, for a power system the number of system states which are success is much greater than that of system states which are failure. Considering the fact that the number of system states that cause a node suffering loss of load is smaller than that of system states which are failure, the converge of simulation method is slower to simulate nodal reliability indices than that simulating system reliability indices.

3.3.4 Calculation of probabilities

Another key step of implementing Markov cue-set method is to compute the probabilities of a minimal cut-set or a union of minimal cut-sets. In this subsection, an improved algorithm is developed to compute the bounds of the reliability indices. This algorithm can automatically generate the transition rate matrix of a minimal cut-set or a union of minimal cut-sets, thus the computation of system and nodal reliability indices can be much easier. This algorithm is an improvement of the method proposed in [36]. The improvements are summarized as follows:

1) Although the algorithm in [36] is applicable to *n*-component system, it is different from that in this dissertation. The index *n* in [36] is 'fixed' whereas in this chapter *n* is 'variable'. In other words, the algorithm in [36] is applicable to fixed-dimension problems whereas the algorithm developed here is applicable to variable-dimension problems. This is needed as the number of the components of a minimal cut-set or a union of minimal cut-sets keeps on changing.
2) The algorithm in [36] is for single-state weather model whereas in this chapter the algorithm is applicable to two-state weather model considering the effects of fluctuating weather. Thus, the transition rate matrix produced here comprises four sub-matrices, and they are generated in sequential steps and finally all the diagonal elements are updated.
3) The core parts of two algorithms are different. The core part of the algorithm in [36] is based on “number” processing whereas in this chapter it is based on “bit” processing. Relevant details are given in the following discussion.

To illustrate the proposed algorithm to compute the bounds of the reliability indices, a simple example is given first. Here, the two-state weather model in Chapter II for a single component is used. It is reproduced in Fig. 4. and necessary parameters are added.

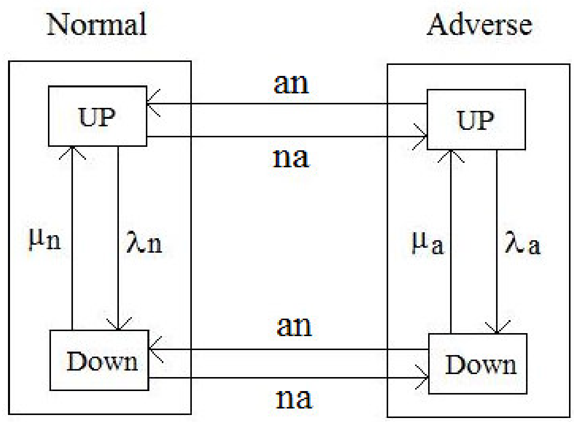

Fig. 4: Two-State Weather Model (Parameterized)

here

λ_n	=	failure rate in normal weather
μ_n	=	repair rate in normal weather
λ_a	=	failure rate in adverse weather
μ_a	=	repair rate in adverse weather
na	=	transition rate from normal weather to adverse weather
an	=	transition rate from adverse weather to normal weather

In this chapter, the impact of adverse weather on the long-term reliability of composite power systems is of interest. Thus, the steady state condition of Markov process can be used to compute the probabilities, i.e. the following equation can be obtained.

$$\begin{pmatrix} -(\lambda_n+na) & \mu_n & an & 0 \\ \lambda_n & -(\mu_n+na) & 0 & an \\ na & 0 & -(\lambda_a+an) & \mu_a \\ 0 & na & \lambda_a & -(\mu_a+an) \end{pmatrix} \cdot \begin{pmatrix} P_{up}^{n} \\ P_{down}^{n} \\ P_{up}^{a} \\ P_{down}^{a} \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \end{pmatrix} \tag{12}$$

where,

P_{up}^{n} = success probability in normal weather

P_{down}^{n} = failure probability in normal weather

P_{up}^{a} = success probability in adverse weather

P_{down}^{a} = failure probability in adverse weather

However, the above equations cannot be directly solved to compute the probabilities because they are linearly correlated, i.e. they are not independent. Now, we have the following equation:

$$P_{up}^{n} + P_{down}^{n} + P_{up}^{a} + P_{down}^{a} = 1 \tag{13}$$

Then, we can replace say the fourth row of (12) as follows:

$$\begin{pmatrix} -(\lambda_n + na) & \mu_n & an & 0 \\ \lambda_n & -(\mu_n + na) & 0 & an \\ na & 0 & -(\lambda_a + an) & \mu_a \\ 1 & 1 & 1 & 1 \end{pmatrix} \cdot \begin{pmatrix} P_{up}^{n} \\ P_{down}^{n} \\ P_{up}^{a} \\ P_{down}^{a} \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 1 \end{pmatrix} \tag{14}$$

Now, the above equations can be solved appropriately to compute the probabilities. The previous discussion shows that a key step to compute the probabilities is to generate the transition rate matrix.

For a minimal cue-set or a union of minimal cue-sets, generally the following equation can be used to compute the steady state probabilities:

$$A'P = B \tag{15}$$

where

A'	=	obtained from $A=\begin{pmatrix} NN & AN \\ NA & AA \end{pmatrix}$ by replacing the elements of an arbitrary row *k* by summing vector 1
NN	=	$2^n \times 2^n$ transition rate matrix in normal weather
AA	=	$2^n \times 2^n$ transition rate matrix in adverse weather
NA	=	$2^n \times 2^n$ transition rate matrix from normal weather to adverse weather
AN	=	$2^n \times 2^n$ transition rate matrix from adverse weather to normal weather
n	=	order of a minimal cut-set or a union of minimal cut-sets
P	=	a column vector whose *i*th element is the steady state probability of system state *i*
B	=	a vector of zeros with the *k*th element set to 1

Actually, only the sum of the state probabilities in two weathers which correspond to the minimal cut-set or the union of minimal cut-sets, needs to be calculated.

Next, the algorithm to generate matrix *A* is presented in detail. The basic idea is as follows:

1) Generate the transition rate sub-matrices in different weathers.

2) Generate the transition rate sub-matrices between two weathers.

3) Update the transition rate sub-matrices in Step(1).

Generating NN

NN is a sub-matrix whose elements are as follows:

$$NN_{i,i} = -\sum_j NN_{j,i}$$

$NN_{i,j} \quad = \quad \lambda^n_{j,i}$

$\lambda^n_{i,j} \quad = \quad$ transition rate from system state *i* to *j* in normal weather

The algorithm used to determine $\lambda_{i,j}$ is as follows: for *NN* the number of system states is 2^n and each system state is represented by an *n*-bit binary vector on the principle - for each bit the binary number is 1 or 0 if the state of the corresponding component is success or failure.

1) Firstly, number $\underbrace{00\cdots0}_{n}$ is assigned to state 1. From state 2 to $2^n - 1$, the binary representation of each system state is as follows:

   - From state 2 to state $\binom{n}{1}+1$, the corresponding binary vectors are in the following form: $\underbrace{10\cdots0}_{n}$, $\underbrace{01\cdots0}_{n}$, $\cdots$, $\underbrace{00\cdots1}_{n}$.

   - From state $\binom{n}{1}+2$ to state $\binom{n}{1}+\binom{n}{2}+1$, the corresponding binary vectors are in the following form: $\underbrace{11\cdots0}_{n}$, $1\underbrace{01\cdots0}_{n}$, $\cdots$, $\underbrace{00\cdots11}_{n}$.

   $\vdots$

   - From state $\binom{n}{1}+\binom{n}{2}+\cdots+\binom{n}{n-2}+2$ to state $\binom{n}{1}+\binom{n}{2}+\cdots+\binom{n}{n-1}+1$, the corresponding binary vectors are in the following form: $\underbrace{11\cdots10}_{n}$, $\underbrace{11\cdots01}_{n}$, $\cdots$, $\underbrace{01\cdots11}_{n}$.

- Finally, vector $\underbrace{11\cdots1}_{n}$ is assigned to state 2^n.

2) From system state *i* to *j*, if there is one and only one bit of their binary vectors being different, forward to the next step; if not, $\lambda_{i,j}=0$. Here only the change of the state of one component at a time is considered, i.e. common-mode failures are not considered.

3) Suppose that the change of the state takes place at the *l*th bit of two binary vectors: if it is $0\rightarrow1$, $\lambda_{i,j}=\mu_l$; otherwise, $\lambda_{i,j}=\lambda_l$.

4) If all the pairs of system states are examined, stop; if not, go back to Step (2).

Generating AA

AA is a sub-matrix whose elements are as follows:

$AA_{i,i}$	=	$-\sum_j AA_{j,i}$
$AA_{i,j}$	=	$\lambda^a_{j,i}$
$\lambda^a_{i,j}$	=	transition rate from system state *i* to *j* in adverse weather

The algorithm to generate *AA* is the same as that to obtain *NN* except that the transition rates in adverse weather are used instead.

Generating NA and AN

Both *NA* and *AN* are diagonal sub-matrices and they are easy to produce. *NA* is a sub-matrix whose elements are as follows:

$NA_{i,i}$	=	$\lambda_{N\rightarrow A}$
$\lambda_{N\rightarrow A}$	=	transition rate from normal weather to adverse weather

AN is a sub-matrix whose elements are as follows:

$AN_{i,i}$ = $\lambda_{A \to N}$

$\lambda_{A \to N}$ = transition rate from adverse weather to normal weather

Update NN and AA

Finally, *NN* and *AA* are updated as: $NN = NN + NA$ and $AA = AA + AN$.

In previous discussion, the relevant reliability parameters, i.e. the transition rates, are assumed to be known. Here, these parameters can be obtained as follows.

Parameters in different weathers

Usually, average reliability parameters are ready to use or can be easily obtained by using simple conversion. For example, usually the mean time to failure or mean time to repair of a component is known. Then, the failure rate or repair rate is just the reciprocal of mean time to failure or mean time to repair. But, the average reliability parameters are undistinguished in different weathers. In the next chapter, a simple approach is proposed to differentiate the reliability parameters in different weathers.

Parameters between two weathers

These parameters can be obtained by using a method similar to that obtaining the average parameters, i.e. computing the reciprocal of mean time in normal weather or mean time in adverse weather to obtain the corresponding transition rate.

3.4 Simulation Method

In this section, the algorithm to simulate the impact of fluctuating weather on composite power system reliability is presented. Firstly, the basic concepts of sequential simulation are introduced; then, the proposed simulation algorithm is presented; finally, some possible improvements to simulation method are discussed.

3.4.1 Sequential simulation

Generally, the evaluation techniques of power system reliability fall into two categories: analytical and simulation. Analytical method is usually based on some mathematical models and calculates reliability results using mathematical derivation. Basically, analytical method enumerates some dominant system states in state space which have non-trivial probabilities. In contrast, simulation method usually does not depend on specific mathematical models. Instead, it simulates the operational process of a physical system, and repeat the simulation till termination criterion is satisfied. Finally, the reliability indices are obtained by using statistical inference [37].

As mentioned in the beginning of this chapter, basically there are two main simulation techniques: random sampling and sequential simulation [24]. The major difference of these two methods is as follows: random sampling assumes that component states are independent and consecutive simulations are also independent with each other; in contrast, sequential simulation simulates system operation literally over time. Actually, this method simulates a Markov chain chronologically. Thus, it is more flexible than random sampling and it is used to simulate the effects of adverse weather on the reliability of composite power systems in this chapter.

Generally, there are two methods to control the advance of sequential simulation: fixed time interval method and next event method [24]. Fixed time interval method advances the simulation in step of a constant time interval Δt . Next event method advances the simulation in a temporal sequence which is determined by the occurrence order of system events. Here, a system event refers to the change of the state of a component or weather. In this chapter, the next event method is used.

A main step of sequential simulation is to generate the random residence time of a component at a state and then determine the next most imminent event. The former issue is discussed next and the latter issue is explained in Subsection 3.4.4.

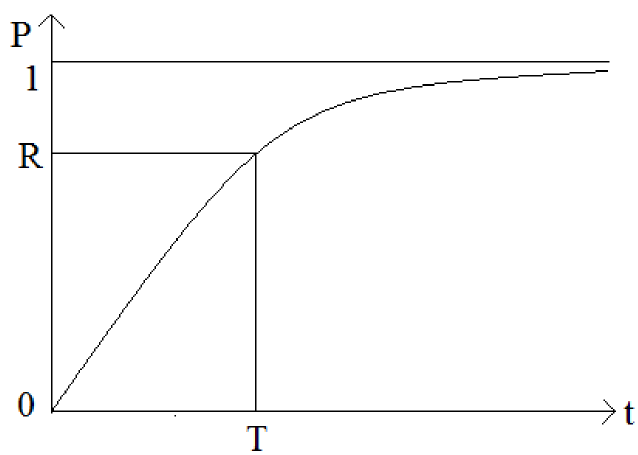

Fig. 5: Function Inversion

To generate the random time that a component resides at a state, usually the function inversion approach as shown in Fig. 5 can be used [24]. In this chapter, the residence time of component state is exponentially distributed. Firstly, a random number within $[0,1]$ is generated; then, the value of the corresponding residence time on the horizontal axis can be obtained by inversing the exponential function. Mathematically, the cumulative distribution function of an exponential distribution is as follows:

$$P = 1 - e^{-\alpha t} \tag{16}$$

here,

α = rate parameter

Then, the residence time can be computed as follows:

$$t = -\frac{\ln(1-P)}{\alpha} \tag{17}$$

When α is replaced by failure rate λ or repair rate μ, the residence time of a component at success state or failure state can be obtained accordingly.

3.4.2 Estimation and convergence

As in the analytical method, here frequency and duration indices are simulated. An advantage of sequential simulation is that the estimation of reliability indices is simpler than that in random sampling. The estimates of p_f, f_f are as follows:

$$\overline{p_f} = \frac{\sum_{i=1}^{N} T_i}{N} \tag{18}$$

$$\overline{f_f} = \frac{\sum_{i=1}^{N} f_i}{N} \tag{19}$$

where,

$\overline{p_f}$ = estimate of p_f

$\overline{f_f}$ = estimate of f_f

N = a sufficiently large number (e.g. the number of years)

T_i	=	system failure time during the *i*th cycle
f_i	=	system failure frequency during the *i*th cycle (e.g. the frequency from success to failure)

Then, the mean duration of system failure d_f can be computed as follows:

$$d_f = \frac{\overline{P_f}}{\overline{f_f}} \tag{20}$$

Apparently, $p_f = E(T_i)$ and $f_f = E(f_i)$.

In this dissertation, the coefficient of variation of an estimate is used to terminate the simulation, i.e. when it is less than a preset value. The coefficients of variation of $\overline{p_f}$, $\overline{f_f}$ are as follows:

$$COV_p = \frac{\sqrt{Var\left(\overline{p_f}\right)}}{\overline{p_f}} = \frac{\sqrt{\frac{1}{N}\overline{Var\left(p_f\right)}}}{\overline{p_f}} \tag{21}$$

$$COV_f = \frac{\sqrt{Var\left(\overline{f_f}\right)}}{\overline{f_f}} = \frac{\sqrt{\frac{1}{N}\overline{Var\left(f_f\right)}}}{\overline{f_f}} \tag{22}$$

where,

COV_p	=	coefficient of variation of $\overline{p_f}$
COV_f	=	coefficient of variation of $\overline{f_f}$
$Var\left(\overline{p_f}\right)$	=	variance of $\overline{p_f}$
$Var\left(\overline{f_f}\right)$	=	variance of $\overline{f_f}$
$Var\left(p_f\right)$	=	variance of p_f

$Var(f_f)$ = variance of f_f

$\overline{Var(p_f)}$ = estimate of $Var(p_f)$ and is equal to $\frac{1}{N}\sum_{i=1}^{N}\left(T_i-\overline{p_f}\right)^2$

$\overline{Var(f_f)}$ = estimate of $Var(f_f)$ and is equal to $\frac{1}{N}\sum_{i=1}^{N}\left(f_i-\overline{f_f}\right)^2$

3.4.3 Confidence interval

As discussed previously, T_i, f_i are random variables and their expected values are p_f, f_f respectively. Now, suppose that their variances are σ_T^2, σ_f^2 respectively. Then, their sample variances are as follows:

$$S_T^2 = \frac{1}{N-1}\sum_{i=1}^{N}\left(T_i-\overline{p_f}\right)^2 \tag{23}$$

$$S_f^2 = \frac{1}{N-1}\sum_{i=1}^{N}\left(f_i-\overline{f_f}\right)^2 \tag{24}$$

where,

S_T^2 = sample variances of T_i

S_f^2 = sample variances of f_i

Then, $Z_T = \frac{\overline{p_f}-p_f}{S_T/\sqrt{N}}$, $Z_f = \frac{\overline{f_f}-f_f}{S_f/\sqrt{N}}$ have *t*-distribution [23], and the $100\%(1-\alpha)$ confidence intervals of p_f, f_f are as follows:

$$P_r\left(\overline{p_f} - A_{\alpha/2}\frac{S_T}{\sqrt{N}} < p_f < \overline{p_f} + A_{\alpha/2}\frac{S_T}{\sqrt{N}}\right) = (1-\alpha) \tag{25}$$

$$P_r\left(\overline{f_f} - A_{\alpha/2}\frac{S_f}{\sqrt{N}} < f_f < \overline{f_f} + A_{\alpha/2}\frac{S_f}{\sqrt{N}}\right) = (1-\alpha) \tag{26}$$

where,

$A_{\alpha/2}$ = $100(1-\alpha/2)th$ percentile of *t*-distribution

Actually, confident interval is an interval estimation in contrast to the point estimation introduced in the last subsection, and it can provide another perspective on the estimates.

3.4.4 Simulation algorithm

The algorithm of sequential simulation to assess the effects of fluctuating weather is as follows:

1) Each system state is represented by an (n+1)-bit binary vector. From bit 1 to n, each binary number is 1 or 0 if the state of the corresponding component is success or failure. The last bit indicates the state of weather and it is 1 or 0 if weather is normal or adverse.
2) For an arbitrary combination of n+1 binary numbers, examine the last bit: if it is 1, the transition rates of all components in normal weather are used; otherwise, the transition rates of all components in adverse weather are used.
3) Compare the residence times of all components and weather at their current states, and the smallest one determines the next most imminent event. Here, the change of weather state is also treated as an event.
4) Update all residence times on the principle: each one minus the smallest one, and the one being 0 will get a new time.
5) Check the type of the event: if it is the change of the state of weather, go back to Step (2); if it is the change of the state of a component, go to the next step.

6) After the event has happened, check if the obtained system state is failure (here the same criterion as that in the analytical method is used): if true, the corresponding event time is saved; otherwise, go to the next step directly.
7) Check if the system state before this event is failure: if false, the transition of system state is counted; otherwise, go to the next step directly.
8) Update all values: estimates, coefficients of variation, confidence intervals.
9) Check if termination criterion is matched: if true, stop; if not, go back to Step (1).

3.4.5 Possible improvements

As in the analytical method, most of the computational time is spent on analyzing failure effects, i.e. determining a system state is failure or not. This can be improved by using the methods mentioned in the last section. Due to the characteristics of simulation method, there are two other improvements which can be implemented. One is that some intelligent methods can be used to improve the selection of system states [24]; the other is that variance reduction method can be used to accelerate the convergence of simulation [23].

3.5 Implementation

In this section, the analytical and simulation methods proposed are applied to the modified IEEE reliability test system [38], and the results obtained by using the two methods are compared and analyzed.

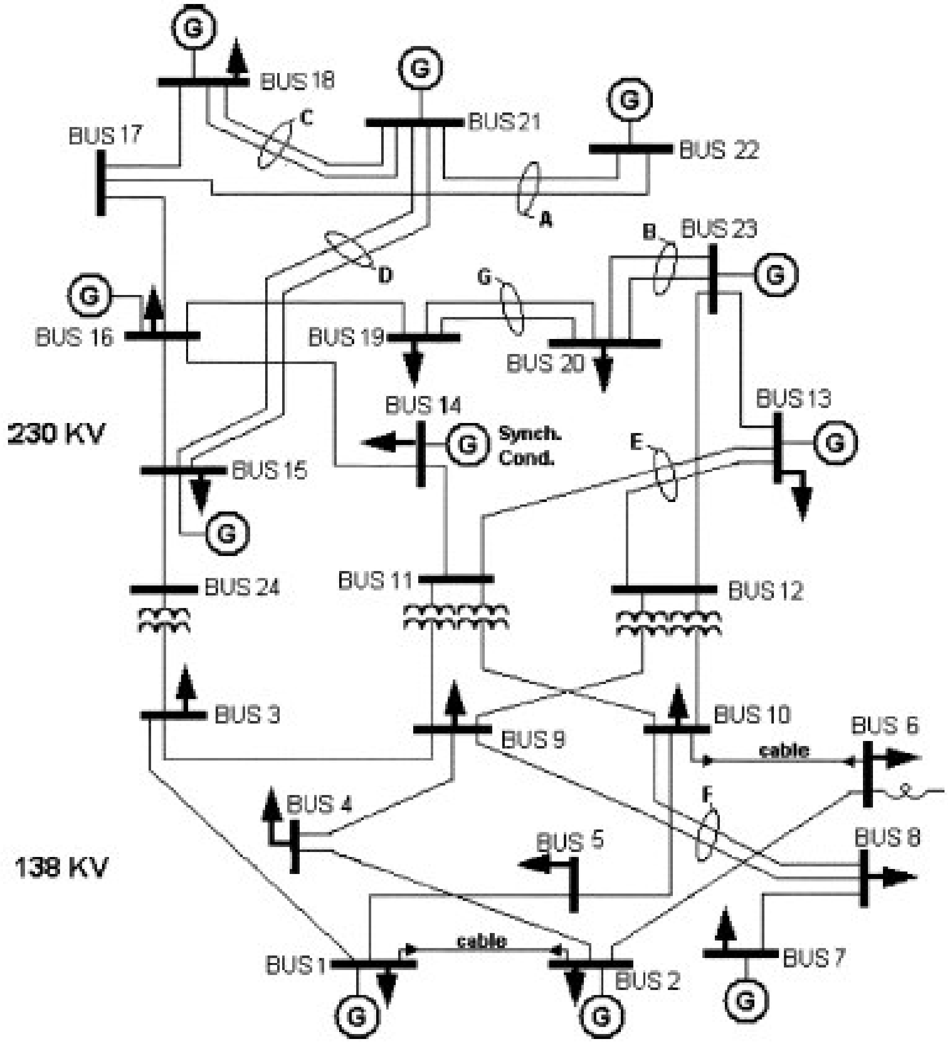

Fig. 6: IEEE Reliability Test System [39]

3.5.1. Test system

The single-line diagram of IEEE reliability test system is shown in Fig. 6. There are 32 generators ranging from 12 MW to 400MW, 24 buses, and 38 transmission lines and transformers. The transmission part of the test system generally consists of two voltage levels: 138 KV and 230 KV.

Considering that the transmission part of the test system is relatively over-reliable [40]-[41], the test system is modified as: the installed capacities of all generators and the load at each bus are increased 1.5 times. Accordingly, the annual peak load 4275 MW is used as the system load, i.e. the system load is constant. But, the proposed methods are also applicable when varied system load is used. Additionally, for the purpose of illustration, all the generators in Table 7 and all the transmission lines in Table 11 in [38] are assigned integer numbers starting from 1 in an ascending order respectively. In this dissertation, the relevant data of IEEE reliability test system is listed in Appendix.

3.5.2 System reliability indices

In this subsection, the system minimal cut-sets identified by using the proposed analytical method and the evaluation results obtained are presented.

3.5.2.1 System minimal cut-sets

Here, the minimal cut-sets are determined up to second-order. The minimal cut-sets determined are as follows: first-order minimal cut-sets of generation and transmission parts, second-order minimal cut-sets of generation and transmission parts, and second-order mixed type which consists of a generator and a transmission line. A mixed minimal cut-set is represented in the form of {generator, transmission line}. The system minimal cut-sets determined are listed in Tables 1-3. It is pointed out that distinguishing the minimal cut-sets of different orders and different types in three tables is just for the purpose of illustration. In programming, actually they are processed indistinguishably as one *table* by using the algorithm developed in Section 3.3.

Table 1: System Minimal Cut-Sets (Generation)

Type	1-Order	2-Order
Generator	None	{12,13},{12,14},{13,14},{12,22},{13,22}
		{14,22},{20,22},{21,22},{12,23},{13,23}
		{14,23},{20,23},{21,23},{22,23},{22,30}
		{23,30}{22,31},{23,31},{12,32},{13,32}
		{14,32},{20,32}{21,32}{22,32},{23,32}
		{30,32},{31,32}

Table 2: System Minimal Cut-Sets (Transmission)

Type	1-Order	2-Order
Transmission Lines	{5},{10} {11}	{1,7},{2,7},{6,7},{4,8},{3,9},{7,9}
		{12,13},{7,14},{7,15},{3,16},{7,16}
		{12,16},{15,16},{3,17},{7,17},{12,17}
		{15,17},{16,17},{7,18},{15,18},{17,18}
		{18,20},{7,21},{18,21},{20,21},{21,22}
		{7,23},{15,23},{17,23},{18,23},{19,23}
		{21,23},{1,27}{2,27},{6,27},{7,27}
		{8,27},{9,27},{14,27},{15,27},{16,27}
		{17,27}{18,27},{21,27},{23,27},{31,38}

Table 3: System Minimal Cut-Sets (Mixed)

Type	1-Order	2-Order
Mixed	N/A	{7,1},{8,1},{1,7},{2,7},{3,7},{4,7}
		{5,7},{6,7},{7,7},{8,7},{12,7},{13,7}
		{14,7},{32,7},{32,25},{32,26},{1,27}
		{2,27},{3,27},{4,27},{5,27},{6,27}
		{7,27},{8,27},{12,27},{13,27},{14,27}
		{32,27},{32,29}

3.5.2.2 System reliability indices

The system reliability indices and the computational time are listed in Table 4. The average value is the average of the upper and lower bounds. For simplicity, only the system reliability indices in normal weather are calculated. If the relevant data is available, the effects of adverse weather can be easily incorporated.

3.5.3 Nodal reliability indices

As discussed in Section 3.3, the proposed analytical method can also compute nodal reliability indices. The algorithm is the same as that computing system indices except that the nodal minimal cut-sets are used instead. For illustration, in Tables 5-7 the minimal cut-sets identified for bus 19 of the test system are listed. As mentioned in Section 3.3, the minimal cut-sets of bus 19 are the subsets of system minimal cut-sets. In Table 8, the reliability indices obtained at bus 19 are listed. The indices for all the nodes are listed in Table 9. For clarity, only the average values of p_f, f_f are computed. The computational time for system and all 20 bus indices is approximately 138 seconds as compared with the only system indices (Table 4) of 134 seconds. So, the additional computational time for the nodal indices is only about 4 seconds, about 3% of the time for the system indices. The reason, as explained earlier, is that the determination of minimal cut-sets where most CPU time is spent, is the same for the algorithms for computing system and nodal indices.

Table 4: Long-Term System Reliability Indices

Index	Upper Bound	Lower Bound	Average Value
p_f	0.1443	0.0853	0.1148
f_f (/yr)	35.04	12.264	23.652
d_f (yr)	N/A		0.0049
Computation time (s)	133.592		

Table 5: Minimal Cut-Sets of Bus19 (Generation)

Type	1-Order	2-Order
Generator	None	{12,13},{13,22}, {14,22},{20,22},{21,22}
		{12,23},{13,23},{14,23},{20,23},{21,23}
		{22,23},{22,30},{23,30}{22,31},{23,31}
		{12,32},{13,32},{14,32},{20,32}{21,32}
		{22,32},{23,32},{30,32},{31,32}

Table 6: Minimal Cut-Sets of Bus 19 (Transmission)

Type	1-Order	2-Order
Transmission Lines	{11}	{4,8},{3,9},{19,23},{7,27},{31,38}

Table 7: Minimal Cut-Sets of Bus 19 (Mixed)

Type	1-Order	2-Order
Mixed	N/A	{32,25},{32,26},{32,29}

Table 8: Reliability Indices at Bus 19 (Long-Term)

Index	Upper Bound	Lower Bound	Mean Value
p_f	0.1333	0.0815	0.1074
f_f (/yr)	30.66	12.264	21. 4506
d_f (yr)	N/A		0.005

Table 9: Long-Term Nodal Reliability Indices

Bus	p_f	f_f (/yr)	d_f (yr)
1	0.1123	22.8456	0.0049
2	0.1121	22.7399	0.0049
3	0.0923	18.0322	0.0051
4	0.1122	22.8178	0.0049
5	0.1122	22.8213	0.0049
6	0.0339	5.1403	0.0066
7	0.0002	0.1966	0.001
8	0.0998	19.6637	0.0051
9	0.0321	4.3344	0.0074
10	0.1121	22.7205	0.0049
13	0.1122	22.8087	0.0049
14	0.0398	6.9801	0.0057
15	0.091	17.6535	0.0052
16	0.1073	21.4319	0.005
18	0.0908	17.4674	0.0052
19	0.1074	21.4506	0.005
20	0.1072	21.3369	0.005
Computation time (s): 138.408			

3.5.4 Simulation results

Table 10: Simulation Results: Part 1 (Long-Term)

Iteration	*up*	*lp*	p_f	*uf*(/yr)	*lf*(/yr)	f_f (/yr)
50	0.2013	0.0787	0.14	53.9589	18.4706	36.2147
250	0.1649	0.1121	0.14	57.5648	33.4846	45.5247
500	0.164	0.124	0.144	40.0002	26.4266	33.2134
2500	0.1258	0.1094	0.1176	27.6618	22.2996	24.9807
5000	0.1245	0.1127	0.1186	27.2848	23.4432	25.364

Table 11: Simulation Results: Part 2 (Long-Term)

Iteration	p_f	f_f (/yr)	d_f (yr)	Computation time(s)
50	0.14	36.2147	0.0039	3.01
250	0.14	45.5247	0.0031	14.835
500	0.144	33.2134	0.0043	29.796
2500	0.1176	24.9807	0.0047	148.34
5000	0.1186	25.364	0.0047	294.54

The reliability indices obtained by using the simulation method are listed in Tables 10-11. Here, for clarity only the results after some number of iterations are listed. Here, the 90*th* percentile of *t*-distribution is used to compute the confidence intervals. Corresponding to the analytical results, only the reliability indices in normal weather are simulated. For simplicity, only the system reliability indices are simulated. But the proposed algorithm is also applicable to simulating the nodal indices. The abbreviations used in Table 10 are as follows:

up = upper bound of the confidence interval of p_f

lp = lower bound of the confidence interval of p_f

uf = upper bound of the confidence interval of f_f

lf = lower bound of the confidence interval of f_f

3.5.5 Comparison of results from two methods

The results of the two methods are compared in Figs. 7-8. Here, the straight lines represent the bounds and the average value of the analytical results, and the curves represent the confidence intervals and the estimates of the simulation results. Here, the legends used are as follows:

UBAM: upper bound of the analytical method

LBAM: lower bound of the analytical method

MVAM: mean value of the analytical method

UBCI: upper bound of the confidence interval

LBCI: lower bound of the confidence interval

EFP: estimate of system failure probability

EFF: estimate of system failure frequency

From the comparison, the following conclusions can be made:

1) The simulation results fall into the bounds of the analytical results, and the bounds of the analytical results is wider than the confidence intervals of the simulation results except in the beginning of the simulation.

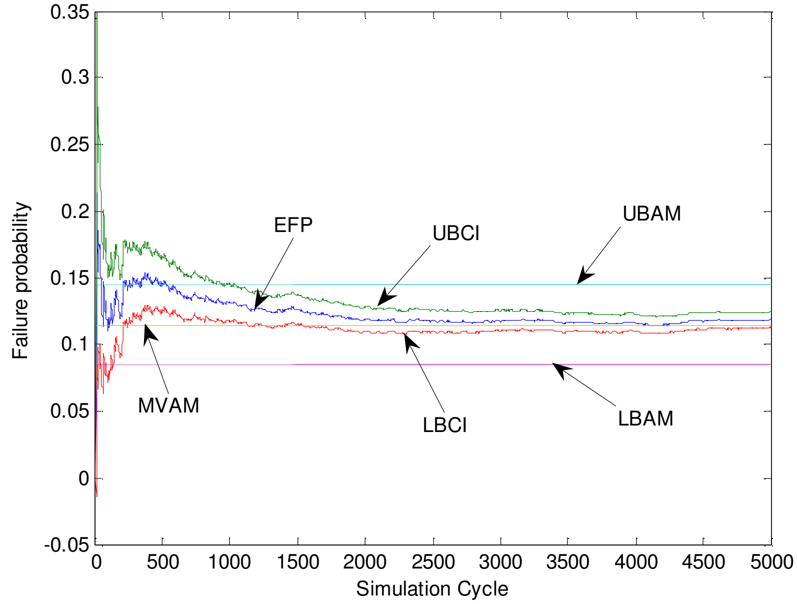

Fig. 7: Long-Term System Failure Probability

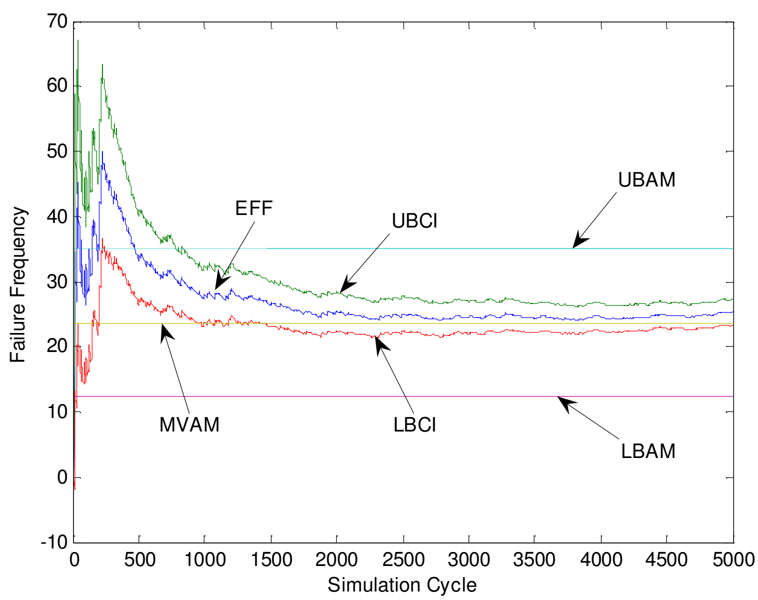

Fig. 8: Long-Term System Failure Frequency

2) The average values of the analytical results fall into the confidence intervals of the simulation results, and they are close to the estimates of the simulation, i.e. the average values of the analytical results can approximate the system reliability indices and they are comparable to the simulation results.
3) When the simulation is proceeding, the confidence intervals become narrower and the bounds of a confidence interval become parallel. Therefore, the variation tendency of the confidence intervals can be used as the termination criterion of the simulation, e.g. setting the difference of the bounds of a confidence interval being less than a small value.
4) In this dissertation, no special technique is used to accelerate the convergence of the simulation. The comparison shows that in the current case the proposed analytical method is faster and comparable results can be obtained.
5) In the current case, the computational time of the proposed analytical method is acceptable, and the additional computational burden in computing nodal reliability indices and storing corresponding data is not significant. For real-world applications, further investigation and improvements of implementation could be done. For instance, usually both analytical and simulation methods use linear optimization to analyze failure effects and this is time-consuming. Some heuristics combined with this approach can improve the performance of reliability evaluation methods.

3.6 Summary

In this chapter, an improved analytical method is proposed to evaluate the impact of adverse weather on the reliability of composite power systems. This method proposes an algorithm to identify system and nodal minimal cut-sets, and proposes an improved algorithm to compute the reliability indices of the bounds of minimal cut-sets. An algorithm for using sequential simulation to assess the effects of fluctuating weather is also developed. These two methods are applied to the modified IEEE reliability test system. The evaluation results obtained by using different methods are compared and analyzed.

From the implementation, the following conclusions are made:

1) The proposed analytical method is effective and efficient. In the current case, it is fast and the evaluation results can be comparable to those of the simulation method.
2) The proposed analytical method has the advantages of easy implementation, convenience of incorporating more system operational considerations, and easy interpretation of the obtained results.
3) The variation tendency of the confidence intervals can be used to terminate the simulation.
4) The additional computational burden in computing nodal reliability indices and storing corresponding data is not significant. The reason is that the proposed analytical method can identify system and nodal minimal cut-sets simultaneously, and this process is time-consuming compared to the calculation of reliability indices.
5) For real-world applications, further investigation and improvements of implementation could perhaps be done. For instance, some intelligent methods can be used to improve the performance of reliability evaluation techniques.

CHAPTER IV

EVALUATION OF HURRICANE IMPACT ON THE SHORT-TERM RELIABILITY OF COMPOSITE POWER SYSTEMS

In the last chapter, the impact of adverse weather on the long-term reliability of composite power systems is investigated. Typically, hurricanes only last a few days but their impact on life-line systems is drastic. Therefore, the impact of hurricanes on the short-term reliability of composite power systems needs to be investigated as their impact on the long term reliability is likely to be diluted. In this chapter, a methodology is proposed to investigate the impact of hurricanes on the short-term reliability of composite power systems. Firstly, a fuzzy inference system is combined with regional weather model [8]-[9] to assess the effect of hurricanes on the failure rates of system components. Here, different methods are used to build two types of fuzzy inference systems [12]-[14]. Then short-term minimal cut-set method is developed to compute time-specific system and nodal reliability indices [13], [17]. The proposed methodology is also applied to the modified IEEE reliability test system. The evaluation results obtained by using different methods are compared and analyzed. The implementation demonstrates that the proposed methodology is effective and efficient and is flexible in its applications [13].

This chapter is organized as follows: in Section 4.1 relevant researches about the effect of weather on the short-term reliability of power systems are reviewed; in Section 4.2 the overall evaluation scheme of the short-term reliability of composite power systems affected by hurricanes is presented; in Section 4.3 the basic concepts of fuzzy sets and fuzzy inference systems are introduced; in Section 4.4 the different methods to build different fuzzy inference systems are presented; in Section 4.5 the main steps of short-term minimal cut-set method are presented; in Section 4.6 the proposed methodology is applied to the modified IEEE reliability test system; finally, in Section 4.7 the main conclusions obtained in this chapter are summarized.

4.1 Literature Review

Many power system components, such as transmission and distribution lines, are exposed to external environment, and it can have a significant impact on the reliability parameters of system components. For instance, it is known for a long time that the failure rate of a transmission or a distribution line is a function of the weather that it is exposed to, and the failure rate of the transmission or distribution line can be much higher in adverse weather than that in normal weather [11]. Thus, one of the challenges of assessing the impact of hurricanes on the short-term reliability of composite power systems is to evaluate how hurricanes affect the reliability parameters of system components, i.e. failure and repair rates.

In [1], it was pointed out that there is rough correspondence between the severity level of hurricanes and the number of power outages. Since the failure rate of a component is close to its failure frequency, i.e. the number of outages during a period of time, the preceding observation can be interpreted that there is some functional relationship between the severity level of hurricanes and the failure rates of system components. Thus, a regression method can be used to assess the relationship between them.

In [42], the impact of vegetation on the failure rates of overhead distribution feeders was assessed by using some parametric methods and an artificial neural network. In [43], multiple linear regression was used to evaluate the impact of weather on the failure rates of transmission lines. In [44], a Bayesian network was used to assess the impact of weather on the failure rates of overhead distribution lines.

In this dissertation, a fuzzy inference system is combined with regional weather model [8]-[9] to assess the functional relationship between the severity level of hurricanes and the failure rate increment of system components. Here, different methods are used to build two types of fuzzy inference systems [12]-[14]. These methods include artificial method and data-driven methods. An advantage of the proposed approach is that these methods can be used in different situations.

After the incremental failure rates of system components are determined, the short-term reliability of composite power systems can be evaluated as follows. Firstly, the reliability indices of system components are calculated. The steady state results of Markov process are not suitable here and so the short-term indices need to be calculated.

In this chapter, a method is proposed to use the minimal cut-set approach described in the previous chapter to evaluate the short-term reliability of composite power systems affected by hurricanes. Here, both system and nodal reliability indices can be computed.

4.2 Overall Evaluation Scheme

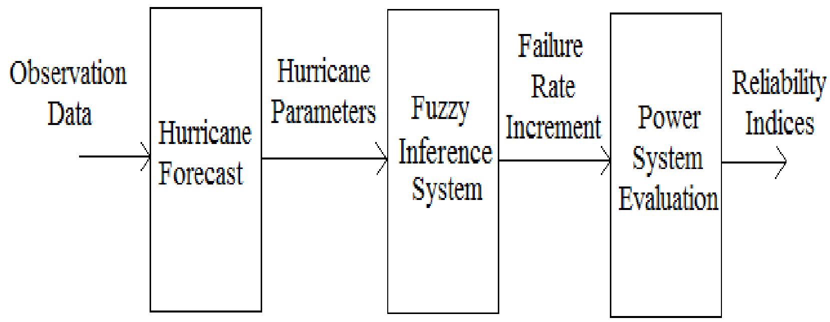

Fig. 9: Short-Term Reliability Evaluation Scheme (Independent Failures)

The overall scheme for investigating the impact of hurricanes on the short-term reliability of composite power systems is shown in Fig. 9. Generally, it consists of three steps: hurricane forecast, assessing the incremental component parameters, and system reliability evaluation. The relevant details are presented in the following subsections. Finally, the collection of required data for data-driven methods to build the fuzzy inference system is discussed.

4.2.1 Hurricane impact

A hurricane is a stormy weather which develops over large bodies of the oceans and then loses its strength after moving over land. Its main effects are strong wind and heavy rainfall when it moves over land. The movement track and strength of a hurricane can be forecast by using observation data and prediction model.

Since hurricanes develop and dissipate with time, their impact on a composite power system can be described in two aspects: *temporal* and *spatial*. *Temporal* refers to the fact that the impact of a hurricane in a given region is different at different times; *spatial* refers to the fact that that the impact of a hurricane is different in different regions at a given time. In this dissertation, the duration of a hurricane is partitioned into some small time intervals to investigate the temporal effects of hurricanes; the affected composite power system is partitioned into several regions to investigate the spatial effects of hurricanes. The temporal partition can be determined by the dissipation rate of the hurricane, and the spatial partition can be determined by the geographical conditions of the composite power system.

The severity level of hurricanes can be represented by some defined parameters. In this dissertation, wind speed and rainfall are used as two parameters of hurricanes. It is noted that in a given region, the parameters of hurricanes are assumed to be identical.

4.2.2 Fuzzy inference

In this dissertation, a fuzzy inference system is used to map the functional relationship between hurricane parameters and the increment multipliers of the failure rates (IMFR) of transmission lines, i.e. the ratios of the failure rates during hurricanes and those in normal weather. Similarly, the IMFR of transmission lines in a given region is assumed to be identical. Since a long transmission line may traverse different regions, its overall IMFR can be determined by using weighted average method which is described in detail in Section 4.5. For the purpose of comparison, in this step different methods are used to build different types of fuzzy inference systems.

4.2.3 System reliability evaluation

After the incremental failure rates of system components are obtained, short-term reliability indices can be computed by using analytical or simulation method. Here, short-term minimal cut-set method is proposed to compute system and nodal indices. Since the states of components are assumed to be independent here, firstly the reliability indices of components are computed by using the transient results of Markov process, then system and nodal indices are calculated by using the multiplication rule of probabilities.

4.2.4 Data collection and preprocessing

In this dissertation, data-driven method is also used to build the fuzzy inference system. Thus, the required data need to be collected and preprocessed. These data include hurricane parameters and the failure rates of system components affected. Usually, hurricane parameters can be collected by referring to historical meteorology records, and the average failure rates of system components can be obtained by referring to historical records of utilities, i.e. the failure rates in normal weather and during hurricanes are not distinguished. The failure rates of system components in different weathers can be obtained by using the transformation techniques like those in [43]-[44]. Basically, the relationship between failure frequency (f) and failure rate (λ) is used: $f \approx \lambda$. Here, f is the number of the failures of a system component during a period of time and it can be obtained from historical records.

Since in this dissertation the output of the fuzzy inference system is the regional IMFR of system components, the obtained failure rates need to be preprocessed. Here, the aggregated failure rates of system components during hurricanes and those in normal weather are compared to get the regional IMFR of system components.

Due to the unavailability of relevant data for confidentiality reasons, the data used in this dissertation is generated by using the fuzzy expert system in [12]. The details are given in Section 4.6.

4.3 Introduction of Fuzzy Inference Systems

4.3.1 Basic concepts of fuzzy sets

4.3.1.1 Crisp sets and fuzzy sets

The concept of fuzzy sets is the generalization of that of crisp sets, i.e. classic sets. Usually, whether an element *x* is a member of a crisp set *A* or not is classified by using the characteristic function as follows:

$$CF_A(x) = \begin{cases} 1, x \in A \\ 0, x \notin A \end{cases}, \text{ i.e. } CF_A : X \to \{0,1\} \tag{27}$$

where,

CF	=	characteristic function
X	=	universe of discourse of *x*, i.e. all the possible values that *x* can take: discrete or continuous

However, for a fuzzy set *B* the degree of whether or not an element *x* is its member can be between 0 and 1, and this is described by using the membership function as follows:

$$MF_B : X \to [0,1] \tag{28}$$

Apparently, the expression of membership functions is more flexible than that of characteristic functions, and this makes membership functions more descriptive of how the real world is perceived. Actually, we always encounter many objects that partially belong to a category and maybe belong to other categories at the same time. In practice, a membership function can be any function that satisfies the relationship defined in (28). It can be in triangular, trapezoidal, Gaussian, and many other forms.

4.3.1.2 Operations on fuzzy sets

Correspondingly, the operations of fuzzy sets are the generalization of those on crisp sets. For instance, the characteristic function of the intersection of two crisp sets *A* and *B* can be expressed as follows:

$$CF_{A \cap B}(x) = \min(CF_A, CF_B) \qquad \text{for } x \in X \tag{29}$$

where,

min = minimum operation

Using membership functions instead of the characteristic functions, the membership function of the intersection of two fuzzy sets *C* and *D* can be expressed as follows:

$$MF_{C \cap D}(x) = \min(MF_C, MF_D) \qquad \text{for } x \in X \tag{30}$$

Similarly, other operations on crisp sets can be extended to those on fuzzy sets. Here, it is noted that the laws of *noncontradiction* and *excluded middle* are applicable to crisp sets but not to fuzzy sets. This is described in Table 12 where usually the universe of discourse of interest is the set of real numbers. More generally, the intersection operation on fuzzy sets can be realized by using triangular norms (*t*-norms) [45]. It presents a group of operations, e.g. minimum and product operations. In the same way, the union operation on fuzzy sets can be realized by using *t*-conorms (*s*-norms), e.g. maximum and probabilistic sum operations.

Table 12: Comparison of Crisp Sets and Fuzzy Sets

Operation	Crisp Sets	Fuzzy Sets
Noncontradiction	$A \cap \overline{A} = \varnothing$	$A \cap \overline{A} \neq \varnothing$
Excluded Middle	$A \cup \overline{A} = X$	$A \cup \overline{A} \neq X$

4.3.1.3 Fuzzy relations

A relation captures the association between objects. Generally, a relation *R* defined over the Cartesian product of *X* and *Y* is a collection of selected pairs (x, y), $x \in X$, $y \in Y$. Here, the two-dimensional case is illustrated and the definition is also applicable to multi-dimensional case. Mathematically, *R* is a mapping as follows:

$$R(x, y) = \begin{cases} 1, x, y & related \\ 0, x, y & unrelated \end{cases}, \text{ i.e. } R : X \times Y \rightarrow \{0,1\} \tag{31}$$

A fuzzy relation generalizes the above concept by recognizing the partial degree of association between objects, i.e. a fuzzy relation R_F is a mapping such that:

$$R_F : X \times Y \rightarrow [0,1] \tag{32}$$

Actually, a fuzzy relation is a multi-dimensional fuzzy set or a fuzzy rule, and the aggregation of them forms a key part of a fuzzy inference system.

4.3.1.4 Cylindrical extension and projection

Cylindrical extension and projection are two important notions in fuzzy theory. Generally, they can be regarded as two operations on fuzzy sets. Cylindrical extension on a fuzzy set *A* is defined as follows:

$$MF_A : X \to [0,1] \Rightarrow MF_{Ce(A)} : X \times Y \to [0,1] \quad (33)$$

where,

$Ce(A)$ = cylindrical extension on fuzzy set *A*

Basically, cylindrical extension is an operation which extends a low-dimensional fuzzy set to a high-dimensional one. Oppositely, projection is an operation that reduces a high-dimensional fuzzy set to a low-dimensional one. For instance, the projection of a fuzzy set *B* from space $X \times Y$ to space *X* is as follows:

$$MF_B : X \times Y \to [0,1] \Rightarrow MF_{\Pr oj_Y(B)} : X \to [0,1] \quad (34)$$

here,

$\Pr oj_Y(A)$ = projection of fuzzy set *B* on Y

4.3.1.5 Fuzzy inference

The inference refers to the derivation of the fuzzy set *B* that if a fuzzy set *A* and a fuzzy relation *R* between them are known. Mathematically, it is as follows:

$$MF_B = \Pr oj_Y(Ce(A) \cap R) \quad (35)$$

Alternately, equality (35) can be expressed as follows:

$$MF_B = MF_A \circ R \quad (36)$$

where,

“∘” = composition operation.

The above equation is the composition rule of fuzzy inference.

4.3.2 Fuzzy inference systems

In this subsection, the basic concepts of fuzzy inference systems are introduced. Firstly, the reasoning mechanism is explained; then, the inference procedure of fuzzy inference systems is described.

4.3.2.1 Reasoning mechanism

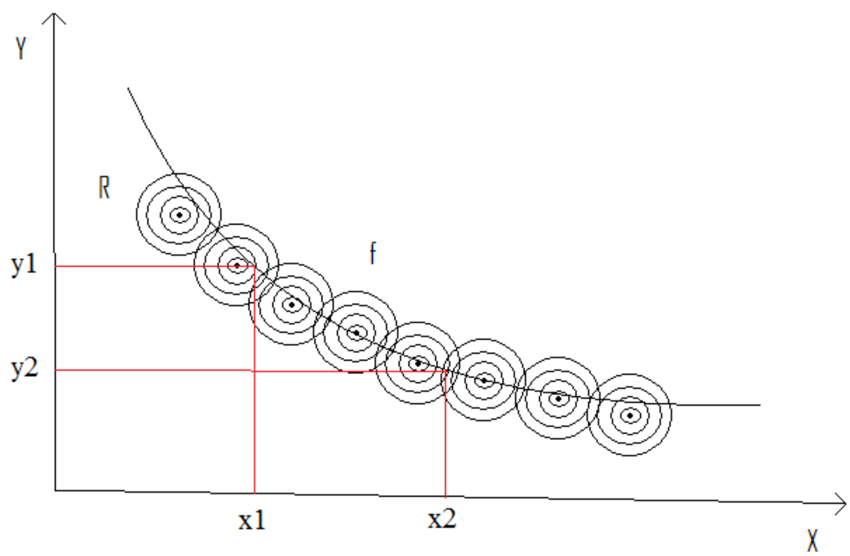

Fig. 10: Reasoning of Fuzzy Inference Systems

Actually, a fuzzy inference system is a rule-based system. Here, a fuzzy rule is actually a fuzzy relation. Thus, the reasoning process of a fuzzy inference system is just the extension of that in the last subsection, and it is as follows:

$$MF_B = \Pr oj_Y \left(Ce(A) \cap \oplus R_i \right) \qquad i \in I \tag{37}$$

here,

"$\oplus$"	=	aggregation operation
R_i	=	ith fuzzy rule
I	=	index of the set of fuzzy rules

The inference process of fuzzy inference systems is shown in Fig. 10. Here, f represents the functional relationship described by a set of fuzzy rules.

4.3.2 .2 Fuzzy inference systems

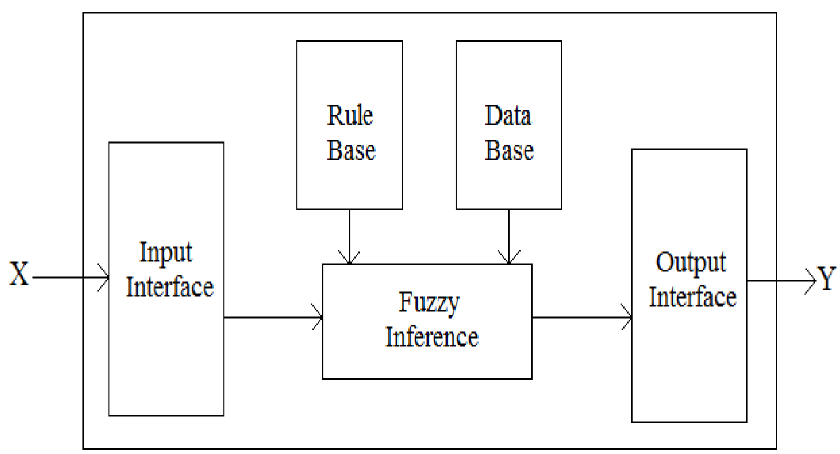

Fig. 11: Architecture of Fuzzy Inference Systems

Usually, a fuzzy inference system consists of five parts as shown in Fig. 11 [45]: input interface, rule base, data base, fuzzy inference, and output interface. Here, X and Y represent the input and output of fuzzy inference systems respectively. Generally, there are two types of fuzzy inference systems: Sugeno-type [46] and Mamdani-type [47]. Their differences are described as follows.

A. Input interface

In this dissertation, the input *X* is the time-specific regional hurricane parameters and it is a vector. Thus, it needs to be converted into the form that the fuzzy inference system can deal with by *fuzzification*, i.e. finding the corresponding membership value of the input. This is shown in Fig. 12. Here, MF represents membership function.

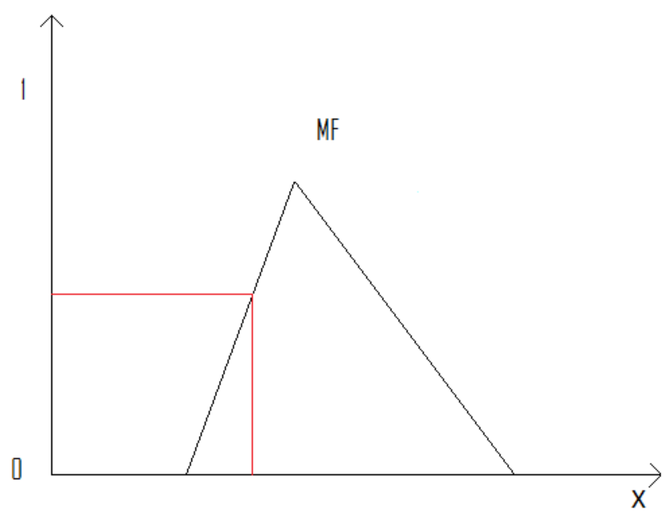

Fig. 12: Fuzzification

B. Rule base

Rule base is a set of fuzzy rules that describe the relationships between the input and output variables of fuzzy inference systems. Generally, for a Mamdani-type fuzzy inference system a fuzzy rule can be in the following form:

$$\text{If } X_1 \text{ is } A_1 \text{ and } X_2 \text{ is } A_2 \text{ and } \cdots \text{ and } X_n \text{ is } A_n \text{ then } Y \text{ is } B$$

here,

X_i	=	ith input variable, $i \le n$
A_i	=	value of X_i, $i \le n$
Y	=	output variable
B	=	value of Y

For instance, in this dissertation a Mamdani-type fuzzy rule can be as follows:

If H_1 is High and H_2 is Medium and $\cdots$ and H_n is Low then IMFR is High

where

H_i	=	ith hurricane parameter, $1 \le i \le n$

For a Sugeno-type fuzzy inference system, a fuzzy rule can be in the following form:

$$\text{If } X_1 \text{ is } A_1 \text{ and } X_2 \text{ is } A_2 \text{ and } \cdots \text{ and } X_n \text{ is } A_n \text{ then } Y \text{ is } f(X_1, X_2, \cdots X_n)$$

here,

$f(X_1, X_2, \cdots X_n)$	=	linear function of X_i, $i \le n$

C. Data base

The type and parameters of the membership functions of input and output variables as well as other parameters of a fuzzy inference system are stored in data base. Generally, there are two kinds of methodologies to construct the rule base and data base of a fuzzy inference system: knowledge-based (expert systems) and data-driven [45]. In this dissertation, both methodologies are used. The relevant details are given in the next section.

D. Fuzzy inference

Inference process is the most important part of a fuzzy inference system. This is the procedure that the fuzzy inference system processes input data and implement the function of reasoning using the information in rule base and data base. Its main steps are as follows:

1) Input matching: for each rule, determine the membership value of each element of input vector.
2) Input aggregation: for each rule, compute rule activation degree, i.e. the intersection of all the membership values of the input obtained in the last step. It is noted that different *t*-norm operations can be applied to different types of fuzzy inference systems. Here, product operation is used for the Sugeno-type fuzzy inference system and minimum operation is used for the Mamdani-type fuzzy inference system.
3) Output derivation: for each rule, a *t*-norm operation is used to compute the intersection of the rule activation degree and the output. Here, product operation is used for the Sugeno-type fuzzy inference system and minimum operation is used for the Mamdani-type fuzzy inference system.
4) Output aggregation: finally, normalized weighted method is used to compute the overall output of the Sugeno-type fuzzy inference system and the details can be found in the next section; for the Mamdani-type fuzzy inference system, a *s*-norm operation can be used to compute the union of all the obtained output in the last step. Here, maximum operation is used.

E. Output interface

In this dissertation, the output of the fuzzy inference system is the time-specific regional IMFR of system components. Whereas, the output of the Mamdani-type fuzzy inference system is a fuzzy set. Thus, it needs to be converted into a numeric value by *defuzzification*. There are many defuzzification techniques available [45]. Here, centroid method is used [45]. Basically, it determines the gravity center of the aggregated membership function of the overall output and is as follows:

$$y = \frac{\int_Y y \cdot MF(y)d(y)}{\int_Y MF(y)d(y)} \tag{38}$$

where,

MF = membership function

For the Sugeno-type fuzzy inference system, the normalized weighted method can be regarded as a method of defuzzification.

4.3.2.3 Comparison of different fuzzy inference systems

From previous discussion, there are differences between the Sugeno-type and Mamdani-type fuzzy inference systems in terms of rule base, inference procedure, and difuzzification. Accordingly, different methods can be used to build them and this is described in the next section.

4.4 Building Fuzzy Inference Systems

Generally, there are two kinds of methodologies to build a fuzzy inference system: knowledge-based (expert system) and data-driven [45]. When sufficient data is available, a data-driven method can be used to build fuzzy inference system; if the data is not available or is insufficient, knowledge-based method can be used. Moreover, other intelligent methods can be used to improve the performance of fuzzy inference systems. In this section, different methods used to build the fuzzy inference system are introduced: expert system [12], fuzzy clustering methods [13], and a hybrid method which combines a neural network and a fuzzy inference system [14].

4.4.1 Fuzzy expert systems

A fuzzy inference system can be built by using artificial method, i.e. the domain expertise of experts are collected and processed to form a rule system. Actually, fuzzy expert systems are the generalization of deterministic expert systems, and they can handle the uncertainty and vagueness that traditional expert systems cannot deal with [45]. For the fuzzy inference system used in this dissertation, the domain knowledge of experts can be helpful to determine fuzzy rules and the parameters of membership functions.

4.4.2 Fuzzy clustering methods

A fuzzy inference systems can be built by using some data-driven methods too, e.g. clustering methods. In this chapter, two fuzzy clustering methods are used to build two types of fuzzy inference systems [13]: subtractive clustering [48] is used to build the Sugeno-type fuzzy inference system and fuzzy *c*-mean clustering [49] is used to build the Mamdani-type fuzzy inference system. In the end, different fuzzy clustering methods are compared.

4.4.2.1 Fuzzy clustering methods

Usually, some relationships exist among a set of variables and the corresponding data can be collected via observation. Fuzzy clustering methods group these data into different clusters. Each fuzzy cluster is represented by a cluster center and the membership degrees of the data belonging to this cluster, and it represents a fuzzy rule. At the same time the membership functions can be obtained by projecting the fuzzy clusters on the spaces of input and output variables. Thus, fuzzy clustering methods can construct the membership functions and fuzzy rules of a fuzzy inference system automatically and simultaneously, and the number of fuzzy rules obtained can be reduced compared to that of a fuzzy expert system [45].

4.4.2.2 Subtractive clustering

The basic idea of subtractive clustering is as follows [48]:

1) The potential value of each data point as a cluster center is computed based on its distances to other data points.
2) The data point with the highest potential value is chosen as the first cluster center, and the potentials of all data points (including the cluster center) are reduced according to their distances to this cluster center.
3) For other data points, the one with the highest remaining potential value is chosen as the next cluster center.
4) The above procedure goes on till the potential values of all data points fall below some threshold.

Mathematically, the main steps of subtractive clustering are as follows:

1) For a collection of data points $\{x_1, x_2, \cdots, x_n\}$, the following equation is used to compute the potential value of each data point as a cluster center:

$$P_i = \sum_{j=1}^{n} e^{-\alpha \left\| x_i - x_j \right\|^2} \tag{39}$$

where,

P_i	=	potential value of data point x_i
n	=	number of data points
α	=	$\frac{4}{r_\alpha^2}$, and r_α is a positive constant
$\left\|\cdot\right\|$	=	Euclidean distance

The above measure shows that a data point with many neighboring data points nearby will have a high potential value. The neighborhood radius is defined by r_α and the data points outside r_α have little influences on the potential of the data point. Generally, a large r_α results in fewer clusters and a small r_α results in more clusters.

2) Suppose now the *k*th cluster center has been determined, then the potential values of all data points (including the cluster centers) are in the following form:

$$P_i^{(k)} = P_i^{(k-1)} - P_k^{*} \cdot e^{-\beta \left\| x_i - x_k^{*} \right\|^2} \tag{40}$$

where,

$P_i^{(k)}$	=	potential of data point x_i, and $P_i^{(0)} = P_i$
x_k^{*}	=	*k*th cluster center
P_k^{*}	=	potential value of x_k^{*}

$\beta \quad = \quad \frac{4}{r_\beta{}^2}$, and r_β is a positive constant

The above equation shows that the data point near the cluster center has greatly reduced potential and therefore is unlikely to be chosen as the next cluster center. The radius of reduction neighborhood is defined by r_β and it is selected as being greater than r_α in order not to produce two close adjacent cluster centers.

3) The following equation is used to compute the membership degree of each data point belonging to the fuzzy cluster with cluster center $x_k{}^*$:

$$m_i = e^{-\alpha\left\|x_i - x_k{}^*\right\|^2} \tag{41}$$

where,

$m_i \quad = \quad$ membership degree of x_i.

Actually, the membership function is in the form of Gaussian function.

4) Since subtractive clustering is used to build the Sugeno-type fuzzy inference system, its output functions are determined by using linear regression [48].

4.4.2.3 Fuzzy c-mean clustering

Fuzzy *c*-mean clustering is one of the most used fuzzy clustering methods in pattern recognition [49]. Its basic idea is to assign each data point to several pre-determined cluster centers and a constrained objective function is solved to determine the optimal partition.

Mathematically, the objective function is as follows:

$$\text{Min} \quad \sum_{i=1}^{c}\sum_{k=1}^{n}(u_{ik})^m \cdot \|x_k - v_i\|^2 \tag{42}$$

$$\text{St.} \quad 0 \le u_{ik} \le 1 \tag{43}$$

$$\sum_{i=1}^{c} u_{ik} = 1 \tag{44}$$

$$0 < \sum_{k=1}^{n} u_{ik} < n \tag{45}$$

where,

c	=	number of cluster centers
n	=	number of data points
x_k	=	kth data point, $1 \le k \le n$
v_i	=	ith cluster center, $1 \le i \le c$
u_{ik}	=	membership value of x_k belonging to the ith fuzzy cluster
m	=	a constant representing fuzzification degree, and $m \ge 1$

Normally, small or big value of m leads to small or big number of cluster centers. The above formulation is a non-linear constrained optimization problem and its optimality condition is as follows:

$$v_i^* = \frac{\sum_{k=1}^{n}(u_{ik})^m \cdot x_k}{\sum_{k=1}^{n}(u_{ik})^m} \tag{46}$$

$$u_{ik}^* = \frac{1}{\sum_{j=1}^{c}\left(\frac{\|x_k - v_i\|}{\|x_k - v_j\|}\right)^{\frac{2}{m-1}}} \tag{47}$$

where,

v_i^* = ith optimal cluster center

u_{ik}^* = optimal membership value of data point x_k belonging to the ith fuzzy cluster

Equations (46)-(47) show that the update of u_{ik}/v_i can only be done if the value of v_i/u_{ik} has been known. Thus, an alternate optimization algorithm can be used to solve the problem (42)-(45), and it is as follows. Here, v_i is assumed being known. Similarly, we can assume u_{ik} being known first.

1) Using (47) to compute u_{ik};
2) Using (46) to update v_i and suppose now it is $\overline{v_i}$;
3) Test if $\left\| v_i - \overline{v_i} \right\|$ is not greater than a pre-set threshold: if true, stop; otherwise, go back to step (1).

4.4.2.4 Comparison of two methods

From previous discussion, subtractive clustering and fuzzy *c*-mean clustering are compared as follows:

A. Initialization problem

For subtractive clustering, the initialization is simple and it considers each data point as a potential cluster center; for fuzzy *c*-mean clustering, this problem is more complex. There are several methods to get the initial cluster centers, e.g. randomly choosing some points, choosing cluster centers according to some modeling or choosing the results of another clustering method as cluster centers, and the chosen cluster centers can be data points or not. Here, the results of subtractive clustering are chosen as the initial cluster centers of fuzzy *c*-mean clustering.

B. Implementation procedure

The comparison of their algorithms shows that the implementation of subtractive clustering is simpler than that of fuzzy *c*-mean clustering. Since fuzzy *c*-mean clustering solves a constrained optimization problem, it has the inherent weaknesses in initialization and convergence, i.e. the sensitivity to initialization and the possibility of trapping at a saddle point, i.e. a local minimum.

C. Complementation of two methods

By contrast, subtractive clustering is a simpler method. But, these two methods can be complementary rather than competitive. For example, as mentioned earlier the results of subtractive clustering can be used as the initial cluster centers of fuzzy *c*-mean clustering. Thus, the performance of fuzzy *c*-mean clustering can be improved.

4.4.2.5 Mamdani-type membership functions

The input membership functions of Sugeno-type fuzzy inference systems are in the form of Gaussian function, whereas the membership functions of Mamdani-type fuzzy inference systems are not in any specific form. Here, the membership functions of the Mamdani-type fuzzy inference system are obtained by using the projection method and are approximated by using two-side Gaussian function [50]. It is actually a combination of two Gaussian functions, and each one represents a side of the membership function.

4.4.3 A hybrid method

After a fuzzy inference system is built, its parameters can be fine tuned by using some intelligent methods. In this chapter, a neural network is used to improve the performance of the Segeno-type fuzzy inference system. Here, adaptive neuro-fuzzy inference system (ANFIS) is used [51]. In this subsection, firstly the underlying motivation of combing fuzzy inference systems and neural networks is given. Then, the algorithm of ANFIS is introduced.

4.4.3.1 Neuro-fuzzy systems

Table 13: Comparison of Neural Networks and Fuzzy Inference Systems

	Neural Networks	Fuzzy Inference Systems
Advantages	No mathematical model	No mathematical model
	No rules required	Prior knowledge used
	Learning ability	Inference and interpretability
Disadvantages	Black box	Rules required
	No prior knowledge	No learning
	Iteration needed to determine parameters	Difficulty in tuning parameters

Neural networks and fuzzy inference systems have some common merits, e.g. no need for establishing mathematical model in advance and universal approximation feature. However, they have their own advantages. As compared in Table 13 [52], neural networks have a learning ability but it is difficult to incorporate prior knowledge into them and to interpret their processing procedure; fuzzy systems can utilize prior