POLITECNICO DI MILANO

Dipartimento di Energia

Dottorato di Ricerca in Scienze e Tecnologie Energetiche e Nucleari

METHODS FOR THE VULNERABILITY ANALYSIS OF CRITICAL INFRASTRUCTURES

Relatore:
Prof. Enrico ZIO
Co-relatore:
Dr. Giovanni SANSAVINI
Tutor:
Dr. Francesco CADINI

Coordinatore del Corso di Dottorato:
Prof. Carlo Enrico BOTTANI

Tesi di Dottorato di:

Roberta PICCINELLI

XXV CICLO

Ringraziamenti

Ringrazio innanzitutto il Prof. Zio, per essere un eccellente esempio di rigore e passione per la ricerca, per avermi dato fiducia, per la pazienza e per le opportunità che mi ha offerto.

Ringrazio di cuore Giovanni, un team leader fantastico.

Un grazie a tutti i compagni di viaggio: Lucia, Diana, Francesca, Jacopo, Michele, Piero, Nicola, Francesco C. e Francesco D., per i momenti “seri” e per i momenti allegri.

Un grazie speciale a Marco, per esserci, sempre e comunque.

E infine, un grazie enorme a mamma e papà, i miei angeli custodi.

Contents

PART I

PART II (selected papers)

PAPER I

E. Zio, R. Piccinelli, G. Sansavini, "An All-Hazard Approach for the Vulnerability Analysis of Critical Infrastructures", *Proceedings of the Annual Conference ESREL*, pp. 2451 -2458, Troyes, September 2011

PAPER II

E. Zio, R. Piccinelli, G. Sansavini, "A Framework for Ranking the Attack Susceptibility of Components of Critical Infrastructures", *Chemical Engineering Transactions*, vol. 26, pp. 309-314, 2012.

PAPER III

E. Zio and R. Piccinelli, " Randomized flow model and centrality measure for electrical power transmission network analysis", *Reliability Engineering and System Safety* 95, pp. 379 – 385, 2010.

PAPER IV

E. Zio, R. Piccinelli, M. Delfanti, V. Olivieri, M. Pozzi, "A comparison of the load flow and random flow models for the analysis of power transmission networks", *Reliability Engineering and System Safety* 103, pp. 102-109, 2012.

PAPER V

E. Zio, L. Golea, R. Piccinelli, G. Sansavini, "Uncertainty propagation in power transmission networks" (submitted for publication)

1. Introduction

Society has always been dependent on services provided by infrastructures but recently it has become even more dependent on the continuous services that infrastructures offer (the Internet and consumption of electricity as a common good may serve as examples). Today, we cannot allow many such systems to debilitate or collapse, as inconveniences and risks are unacceptable and financial losses are huge.

Moreover, infrastructures are witnessing higher and tighter integration, they that have become more complex and their behavior is hard to understand or predict; research on such complex systems has shown that some elements evolve to become more important and some structures are more susceptible against random failures or targeted attacks than others. Reduction of technological and social vulnerabilities calls for better system understanding and preventive analyses. In this PhD thesis some methods for the analysis of critical infrastructures, with respect to their vulnerabilities to random failures and targeted attacks, are assessed.

1.1 Critical infrastructures (CIs)

Modern society relies on the continuous and secure supply of essential goods (such as energy, data) and services (such as banking, health care). In this scenario, infrastructures are more than just a collection of individual companies engaged in related activities; they are instead a network of independent, mostly privately-owned, manmade systems and processes that function collaboratively and synergistically to produce and distribute a continuous flow of essential goods and services [PCCIP Report, 1997]. Infrastructures are considered critical because their incapacity or destruction would have a debilitating impact on the health, safety, security, economics and social well being, including the effective functioning of governments.

Critical infrastructures are various in nature, e.g. physical-engineered, cybernetic or organizational systems. Engineered, physically networked critical infrastructures (CIs) will be in focus of this thesis; examples are those providing:

- energy (electricity, oil and gas supply as subsectors);
- transportation (by rail, road, air, shipping);
- information and telecommunication (such as the internet);
- drinking water, including wastewater treatment.

Engineered critical infrastructures are characterized as large scale, spatially distributed, complex systems. These systems are more than just an aggregation of their components. As large sets of components of various natures (physical, human, organizational, etc.) are brought together and interact with one another, synergies emerge. Therefore, they can be seen as systems of interacting agents, based on internal processes, of a network that describes which components of a system interact, multiple scales of space and/or time, or symmetry. The components of many complex systems are heterogeneous and form a hierarchy of subsystems" [Guckenheimer and Ottino, 2008]. In this view, critical infrastructures show emergent properties difficult to anticipate from the knowledge of single components, are characterized by a large degree of adaptability to absorb random disruptions and are highly vulnerable to widespread failure under adverse conditions" [Duenas-Osorio and Vemuru, 2009]. Small, unnoticed perturbations can trigger large scale consequences in critical infrastructures, as well as disruptions may also be caused by targeted attacks.

Each infrastructure is vulnerable in varying degrees to natural disasters, component failures, human negligence, and willful human misconduct. Also, the structure of most critical infrastructures is not fixed but can undergo both technological and organizational changes, due to renovation, expansion or maintenance operations. Furthermore, critical infrastructures are not isolated but “are highly interconnected and mutually dependent in complex ways, both physically and through a host of information and communications technologies" [Rinaldi et al., 2001].

Due to their spatially distributed, interconnected nature, complex systems are considered to be subjected to uncertainty, thus, a challenging task of critical infrastructure vulnerability analysis is quantifying this uncertainty and predicting how it propagates throughout the system. Emerging behavior is the essence of a complex system, which distinguishes it from a complicated system. Indeed, the overall behavior emerges from the interactions among single parts of a complex system. Following this line of thought, the system must be analyzed as a whole and decomposing it and analyzing single subsystems does not necessarily give a clue as to the behavior of the whole [Guckenheimer and Ottino, 2008].

Critical infrastructures are placed within an environment such as the geographical, political and economical context, etc. The operating state and condition of each infrastructure influence the environment and the environment in turn exerts pressures on the individual infrastructure (normal system operations, emergency operations, repair and recovery operations). Most of the critical infrastructures show adaptive behavior, in that the system as a whole regulates its internal environment and maintains a stable, constant condition needed to fulfill its specific tasks despite all components are influenced by past experience e.g. degradation from overuse, aging over time and by adjustment to new conditions or disturbances [Rinaldi et al., 2001].

The CIs in the society have undergone, and are undergoing, considerable changes. Zimmermann argues that "technological changes have improved the provision of services of transport, water, electricity, and communications, often transforming the way we live, while, at the same time substantially increasing the fragility and vulnerability of these systems and the services they provide by making them more complex and interdependent" [Zimmermann, 2001].

Several research projects have been initiated in the area of modeling and analyzing critical infrastructure systems: see Pederson et al. (2006) for an overview of approaches. It is argued that methods and models with different perspectives are needed since no single method/model can possibly capture everything of relevance regarding this complex system of systems. The research field is still evolving and Rinaldi (2004), Pederson (2006) and Kroger (2008) suggest that research on method development in this area is highly relevant.

1.2 Synthesis of the contribution of the thesis

Vulnerability is defined as the set of flaws and weaknesses in the design, implementation, operation, and/or management of an infrastructure system or its elements that renders it susceptible to destruction or incapacitation when exposed to a hazard or threat. The likelihood (frequency) of the accident scenarios, and the magnitude of their consequences can be evaluated through specific

elaborations depending on the particular infrastructure considered. As an example, the vulnerability of the electric power system might be assessed in terms of the frequency of major blackouts (number per year), and associated severity (undelivered MW, or MWh).

Critical infrastructures (CIs) exhibit a number of complex system characteristics which call for analyzing them as a whole and make their holistic study highly challenging. A comprehensive vulnerability analysis requires not only the consideration of a large number of spatially distributed, interacting elements with nonlinear behavior and feedback loops, but also a broad spectrum of hazards and threats including failures and threats.

Fig. 1.1 presents a schematic conceptualization of the vulnerability assessment of a CI [Kroger and Zio, 2011]. It implies:

a) hazards and threats identification;
b) physical and logical structure identification and operation modes definition;
c) dependencies and interdependencies identification and modeling;
d) failure cascade dynamics analysis.

The two main outputs of a CI vulnerability assessment are shown to be the quantification of system vulnerability indicators, and the identification of critical elements.

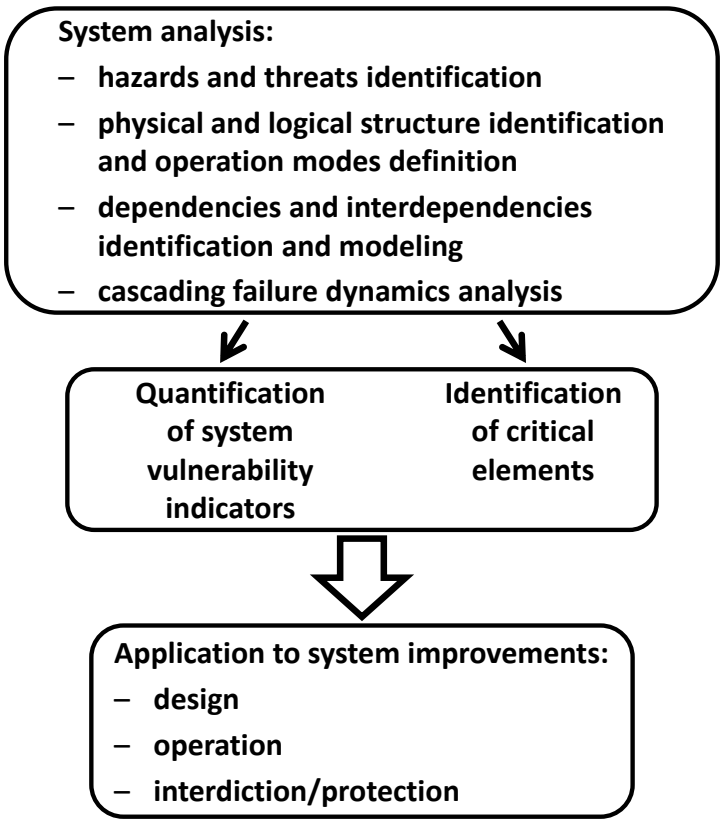


Figure 1.1. Conceptualization of CI vulnerability assessment [Kroger and Zio, 2011].

A number of methods can be undertaken for the vulnerability assessment of CI depending on the type of system, the objective of the analysis, and the available information. It is worth noting that no all-encompassing method exists but rather an interplay of methods is necessary to provide trustworthy information about vulnerabilities within and among critical infrastructures (CIs).

In this view, three methods for the analysis have been devised to perform the vulnerability assessment (Figure 1.2):

- an all-hazard analysis to address issue a);
- topological analysis of electrical power networks to address issue b);
- uncertainty analysis in electrical network systems to address issue a) and d);

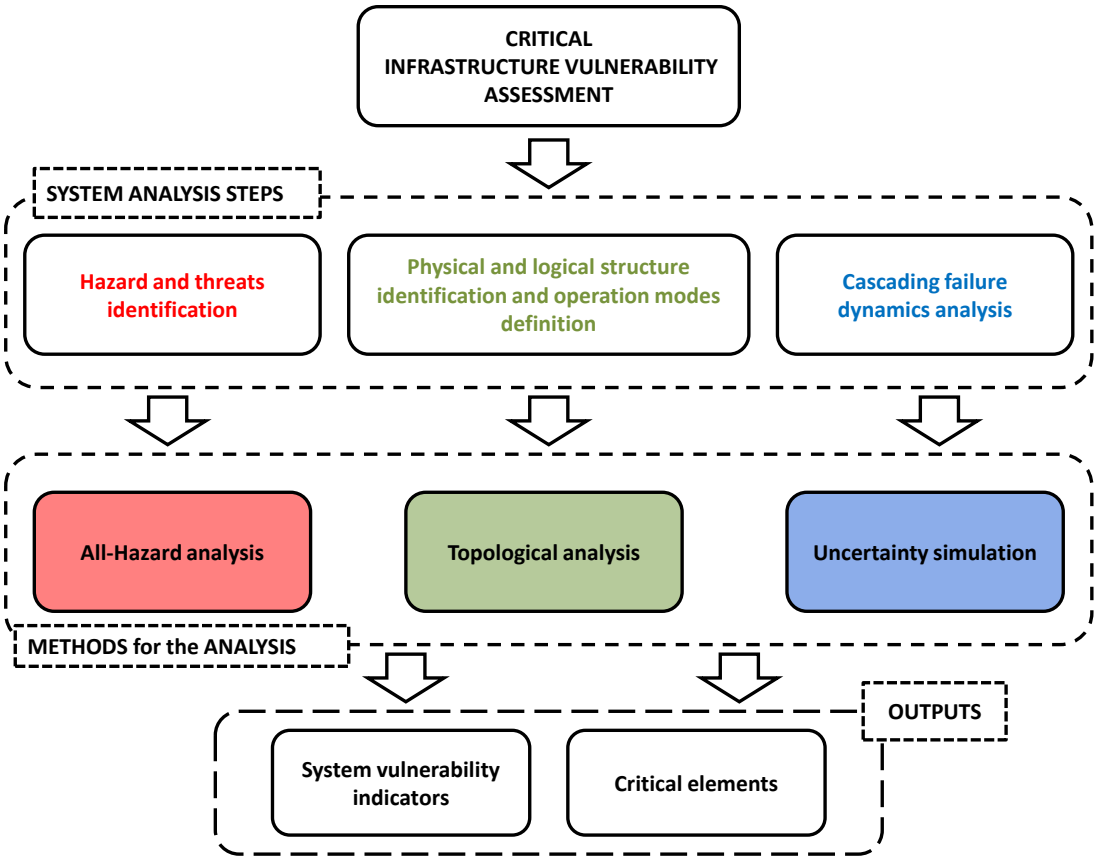


Figure 1.2. Pictorial view of the critical infrastructure vulnerability assessment presented in the current Ph.D. research work carried out at LASAR.

The contributions of this PhD thesis to the above issues are:

- an all-hazard approach intended to provide the basis for addressing unexpected events of any nature such as deterioration and random failures, natural disasters, accidents, and malevolent

acts. CIs are especially attractive targets for malevolent attacks because today's societies operate heavily on their reliance. In risk and vulnerability analysis, random accidents, natural failures and unintentional man-made hazards are typically known and categorized by emergency planners. The likelihood of their occurrence is traditionally addressed within a probabilistic framework. On the other hand, terrorism poses a hazard that eludes a quantification by probability theory due to the intentional and malevolent planning it implies. Therefore, there is the need of an all–hazard approach encompassing a broader view on the hazards, that threaten CIs. In this PhD thesis, an *All-HAZard ANalysis* (A-HAZAN) is developed. It aims at identifying the features, operating conditions and failure modes relevant to CI vulnerability, and capturing the CIs vulnerability sources and issues, given their technical and physical features, and the dependencies and interdependencies on other CIs.

- CIs are engineered complex systems and can be modeled as hierarchies of interacting components. In this view, the actual structure of the network of interconnections among the components is a critical feature of the system. In a topological analysis, a CI is represented by a graph *G(N, K)*, in which its physical constituents (components) are mapped into *N* nodes (or vertices) connected by *K* edges (or arcs), representing the links of physical connections among them. The focus of topological analysis is on the structural properties of the graphs. In order to quantify the structural importance of the network components, several centrality measures have been introduced: commonly used centrality measures identify the most important elements in networks of components, based on the assumption that physical/communication/service among nodes flow follows the shortest paths in the network. In spite of the usefulness and appealing simplicity of the topological analysis of the network underpinning a CI and of the insights it provides, empirical results show that it cannot capture the rich and complex properties observed in a real infrastructure system, so that there is a need for extending the models beyond pure structural topology. While the topological approaches for identifying critical components are capable of highlighting structural vulnerabilities, they are limited from the point of view of the functional vulnerability of the CI. In real network systems, another important dimension to add to the vulnerability characterization refers to modeling the dynamics of flow of the physical quantities in the network where physical law and operational rules drive the physical/communication/service flow. This entails considering the interplay between structural characteristics and the dynamics, in order to provide indications on the elements critical for the propagation process and on the actions that can be performed in order to prevent or mitigate the undesired effects.
- In the final step of this PhD thesis, the characterization of uncertainties related to the physical flow through the network has been undertaken and exemplified with respect to the electric infrastructure. Failing to incorporate uncertainties in system planning may lead to an

overestimation of risk reduction barriers and of system capabilities to maintain acceptable levels of reliability. In order to quantify the impact that the propagation of the identified uncertainties has on the reliability of the electric infrastructure a stochastic model that simulates the operations of an electric transmission network was developed. This event based model, embedded in the Monte Carlo Simulation framework, and has shown the ability to represent daily hourly changes in power requests at customer side of the system, ambient temperature, wind speed and wind power generation. The increasing variability in the operating conditions lead to an increase in the generated power that cannot be supplied to the customers.

1.3 Structure of the thesis

The thesis comprises two parts. Part I, subdivided in five Chapters, introduces the addressed problems in further details and illustrates the methodological approaches developed and employed in this PhD. work. Part II is a collection of five selected papers published (or submitted for publication) as a result of the work and which the reader is referred to for further details.

Chapter 2 presents a state of the art review on vulnerability of critical infrastructures and on the approaches regarded as most important for the vulnerability assessment of critical infrastructures.

Chapter 3 describes the all-hazard approach to model the exposure of a complex systems to hazard and threats and presents to quantitative approaches. Chapter 3 makes reference to Papers I and II of Part II.

Chapter 4 describes the novel approach to the topological analysis of an electric power network when the physical law rule the flow through the system. Chapter 4 makes reference to Papers III and IV of Part II.

Chapter 5 presents the characterization of the uncertainties related to the vulnerability assessment of the system. Chapter 5 makes reference to Paper V of Part II.

2. State of the art review on vulnerability analysis of critical infrastructures

2.1 Vulnerability analysis

The vulnerability analysis of CIs needs to be theoretically assessed and organized. While the concept of risk is fairly mature and consensually agreed, the concept of vulnerability is still evolving and not yet established [Kroger and Zio, 2011]. In general terms, risk refers to a combination of the probability of occurrence (frequency F) of a specific (mostly undesired/adverse) event leading to loss, damage or injury and its extent. These quantities and their associated uncertainties are regarded as being numerically quantifiable. For CIs, the term risk may include the frequency of loss of service with its resulting consequences for the people concerned.

In risk analysis, hazard is defined as "a potentially damaging physical event, phenomenon and/or human activity, which may cause loss of life or injury, property damage, social and economic disruption or environmental degradation. Hazards can be single, sequential or combined in their origin and effects'' [UN/ISDR 2004].

The term vulnerability has been introduced as the hazard-centric perception of disasters that is revealed as being too limited to understand in terms of risks. A hazard of low intensity could have severe consequences, while a hazard of high intensity could have negligible consequences: the level of vulnerability is making the difference [White 1974].

To date, there is no consensus definition of vulnerability. Historically, the reflection on the concept of vulnerability developed in three main steps:

- first definitions of vulnerability focused on the *degree of loss and damages* due to the impact of a hazard, i.e. on the technical dimensions of vulnerability. Proposed measures to reduce vulnerability, i.e. the *sensitivity* of the element at risks to the impact of a given hazard, were therefore limited to engineering and technical measures;
- the second step is linked to the understanding at the beginning of the 1980s that the degree of loss and damages was determined by the *degree of exposure t*o the hazard. Vulnerability was therefore defined as the likelihood of being exposed to hazards and as the susceptibility of an element at risk to suffer losses and damages as a function of its degree of exposure to a given source of hazard. All elements at risk do not show the same level of exposure to a hazard, as a function of their location in space for instance. The assessment of losses and damages as a function of the degree of exposure became the measure of vulnerability" [Dow, 1992];
- finally, as a result of two scientific approaches, a third type of definitions was proposed:
  on the one hand, applied sciences underlined the fact that the degree of loss and damage depends on *internal characteristics of the element at risk.* Vulnerability was thus considered as an internal risk factor, related to the *resistance capacity* of the element at risk: beyond a given level of resistance, the element at risk could suffer damages.
  On the other hand, within social sciences, Susman et al [Susman and al., 1983] introduced the topic of a population's capacity to cope with a disaster, absorb and recover as a measure of their vulnerability. This capacity for adaptation was defined as the *capacity of resilience* of a society [Blaikie and al., 1994].

These three step highlight "two sides of the vulnerability": an external side of shocks and perturbations to which a system is exposed; and an internal side which represents the ability or lack of ability to adequately respond to and recover from external stresses [Chambers and al., 1989]. Exposure and susceptibility refer to the external side and their analysis contributes to a pre-disaster vulnerability analysis, based on the assessment of potential losses and damages, including hazard independent vulnerability factors (state of the system before the disaster) and hazard dependent factors (scenarios as a function of the likelihood of a hazard to happen and its potentially damaging effects).

The internal side of vulnerability refers to the capacity of resilience of a system, once the disaster stroke. The analysis of the internal side of vulnerability constitutes a post-disaster vulnerability analysis, based on the assessment of post-disaster impacts balanced by the capacities of the system to adapt and recover from a disaster [Bouchon, 2006].

These two aspects merges in the operational definition of vulnerability, useful for its systematic assessment as a flaw or weakness (inherent characteristic, including resilience capacity) in the design, implementation, operation, and/or management of an infrastructure system, or its elements, that renders it susceptible to destruction or incapacitation when exposed to a hazard or threat, or

reduces its capacity to resume new stable conditions. The latter can be provided with a likelihood (frequency) while a measure for destruction or incapacitation (loss or damage, respectively) needs specific elaborations depending on the value placed on the asset by its owner/ operator or the customer/government. For example, the vulnerability of the electric power system might be specified in terms of changes in network characteristics following attacks on nodes and the scale (e.g., number of nodes/lines lost) or the duration of the associated loss. More sophistically, it can be expressed in terms of the frequency of major blackouts (number per year) and the associated severity, measured either in power lost or energy unserved (MW or MWh).

2.2 Approaches for the vulnerability assessment of CIs

The two main outputs of a CI vulnerability assessment are the quantification of system vulnerability indicators and the identification of critical elements. The information they provide is complementary: while vulnerability indicators are parameters encompassing the static and/or dynamic characteristics of the whole system, the identification of critical elements comes from their ranking with respect to their individual connectivity efficiency and centralities and/or their contributions to the propagation of failures, with their effects, through the network. A number of approaches can be undertaken for the vulnerability assessment of CIs. The choice of the suitable approach depends on the type of system, the objective of the analysis, and the available information. In this Section, an attempt to defined the characteristics of these approaches is made.

2.2.1 Risk Analysis

This approach can be divided into two lines of analysis: the first entails the qualitative assessment of system vulnerabilities by expert judgment and tabular methods [Moore, 2006; Piwowar et al., 2009], while the second entails the quantitative vulnerability assessment of a CI [Apostolakis and Lemon, 2005; Flammini et al., 2009], with the further aim of ranking systems elements according to their criticality [Koonce et al., 2008]. To a certain extent, the risk analysis approach to the vulnerability of CIs can be considered a general framework of analysis, since it often takes advantage of other approaches and tools, i.e. power flow analysis for electrical transmission networks [Koonce et al., 2008] and network analysis [Apostolakis and Lemon, 2005].

Methods such as fault tree analysis (FTs) and probabilistic risk assessment (PRA) have been applied to the vulnerability analysis of CIs for protecting the systems against malevolent actions [Piwowar et al., 2009]. The approach comprises a step-by-step process typical of PRA: 1 – systemic

analysis, in which the system itself and its surroundings are analyzed to identify all the parameters that could interact with it; 2 – analysis of the interactions between aggressors' profiles and systems, in which the systems are ranked according to the degree of risk of being attacked; 3 – assessment of vulnerabilities and determination of key points of vulnerability, in which a panel of subject matter experts and a PRA process normative expert conduct a study and weigh the importance of the system elements to qualify them according to several criteria concerning the attack; 4 – building of scenarios, taking into account system functionalities, the global environment and the geopolitical context; 5 – updating the security systems, considering the results of the previous steps.

A risk-based quantitative method for the identification and prioritization of vulnerabilities in interdependent CIs has been proposed in [Apostolakis and Lemon, 2005]. The CI is modeled as interconnected digraphs and graph theory is employed to identify the candidate vulnerable scenarios. These scenarios are screened for the susceptibility of their elements to a terrorist attack, and a prioritized list of elements vulnerabilities is produced. The prioritization methodology is based on multi-attribute utility theory [Morgan et al., 2000]. The impact of losing infrastructure services is evaluated using a value tree that reflects the perceptions and values of the decision maker and the relevant stakeholders. These results, which are conditional on a specified threat, are provided to the decision maker for use in risk management. Interestingly, this method embeds the vulnerability quantification into the framework of the stakeholders' perspective, making it suitable for a realistic ranking of vulnerabilities in CIs.

Yet, the approach presents some inconveniences related to the size of the distributed system to analyze, the number of scenarios of attack to be considered, the human subjectivity as a base for all the quantifications to assess vulnerabilities on a given CI and the needs to be updated frequently, because of the evolving geopolitics and the introduction of new factors in the whole system (a new apparel, new ways of productivity, new entrance or exit points, etc).

2.2.2 Complex Network theory

CIs are networked complex system of interacting components, for which the actual structure of interconnection is a relevant feature [Albert et al., 2000].

Complex network theory approach can be applied to the analysis of CIs: in a topological analysis, a CI is represented by a graph $G(N, K)$, in which its physical constituents (components) are mapped into N nodes (or vertices) connected by K unweighted (all equal) edges (or arcs), representing the links of physical connection among them, where the edges between nodes are either present or not.

Topological analysis based on classical graph theory can unveil relevant properties of the structure of a network system [Albert et al., 2000; Strogatz, 2001] by i) highlighting the role played by its components (nodes and connecting arcs) [Crucitti et al., 2006; Zio et al., 2008], ii) making preliminary vulnerability assessments based on the simulation of faults (mainly represented by the removal of nodes and arcs) and the subsequent re-evaluation of the network topological properties [Rosato et al., 2007; Zio et al., 2008].

In spite of the usefulness of the topological analysis of the unweighted network of a CI and of the insights it provides, empirical results show that it cannot capture all rich and complex properties observed in a real infrastructure system, so that there is a need for extending the models beyond pure unweighted, structural topology [Boccaletti et al., 2006; Eusgeld et al., 2009]. Indeed, along with a complex topological structure, many real networks display a marked physical heterogeneity in the capacity and intensity of the connections: for examples, there are different impedance and reliability characteristics of overhead lines in electrical transmission networks [Hines and Blumsack, 2008; Eusgeld et al., 2009]. To describe the inhomogeneities of real physical systems, numerical weights can be assigned to each link of the representative network, measuring the "strength" of the connection. In this way, the functional behavior of the CI is somewhat embedded into a generalized, but still simple, topological analysis framework. Global and local measures can then be introduced for the statistical characterization of "weighted"networks [Latora and Marchiori, 2001]. The resulting generalized setup, in which global and local efficiencies measure the network global and local connectivity features accounting also for the arcs weights, encompasses the topological analysis of unweighted networks in the case that all edges have unit weight. Topological, weighted or unweighted, analyses focus on the static structural properties of network interconnection, looking at the effects on vulnerability indicators caused by the removal of a certain percentage of nodes or links [Latora and Marchiori, 2005; Zio et al., 2008] or identifying the elements whose presence is critical with respect to the network connectedness [Cadini et al., 2009].

However, in real network systems, another important dimension to add to the vulnerability characterization refers to modeling the dynamics of flow of the physical quantities in the network. This entails considering the interplay between structural characteristics and dynamical aspects, which makes the modeling and analysis very complicated since the load and capacity of each component, and the flow through the network are often highly variable quantities both in space and time. Functional models have been developed to capture the basic realistic features of CI networks within a weighted topological analysis framework, i.e. disregarding the representation of the individual dynamics of the CIs elements. These models have shed light on the way complex networks react to faults and attacks, evaluating their consequences when the dynamics of flow of the physical quantities in the network is taken into account.

Finally, complex network theory models allow accounting for dependencies and interdependencies among different CIs, to assess the influences and limitations which interacting infrastructures impose on the individual system operating conditions, for avoiding fault propagation by designing redundancies and alternative modes of operations, for detecting and recognizing threats [Zimmermann, 2001; Duenas-Osorio et al., 2007; Johansson and Jonsson, 2009]. Infrastructure interdependency stems from the functional and logical relations among individual components in different distributed systems. In developing modeling and simulation frameworks that allow the coupling of multiple interdependent infrastructures, it is important to know that simply linking existing infrastructure models together fails to capture the emergent behavior arising in interdependent infrastructures, a key element of interdependency analysis. In order to characterize the extent to which a contingency affecting an infrastructure is going to weaken, and possibly disrupt, the safe operation of an interconnected system, it is necessary to model the relations established through the connections linking the multiple components of the involved infrastructures. The modeling of interdependencies among network systems and of their effects on failure propagation can be carried out within the simulation framework of failure cascade processes; the sensitivity of the coupling parameters defining the interdependency strength is of particular interest for the definition and prescription of cascade-safe operating margins in interdependent CIs [Zio and Sansavini, 2010b].

The methods of complex network theory can provide information useful for the vulnerability assessment of CIs, within a screening analysis that leads off to an adequate system understanding that cannot be superficial for the following detailed analysis. The analysis is supported by structural information provided by system owners, including the general understanding of main functionalities, interfaces, (inter-) dependencies, etc. The evaluation of the statistical indicators derived from the analysis provides indications of obvious vulnerabilities, e.g., structural or reliability bottlenecks, etc.

2.2.3 Probabilistic modeling

Probabilistic risk assessment is another mature methodology that can be applied for analyzing network systems [Patterson and Apostolakis, 2007; Volkanovski et al., 2009]. It integrates deterministic and stochastic tools to carry out a systematic and structured evaluation of the risk associated with every life cycle aspect of a complex engineered technological system, which may lead to undesired consequences triggered by an accident initiating event.

This approach encompasses a variety of methods used for the characterization of CIs, such as Markov chains (MCs), Markov/Petri nets (MPNs), probabilistic dynamics modeling and Bayesian networks (BNs). MCs and MPNs rely on the definition of probabilities of transition of the system components among their reachable states: the behavior of a CI is described by its states and by the

possible transitions between these states. The system states are defined by the states of the components comprising the system. This may pose significant challenges because of the exponential growth in the number of CI configurations to be evaluated [Iyer et al., 2009].

Probabilistic dynamics models can be considered to overcome the computational limitations of the previous methods; yet, their analysis is affected by the drawback that the identification of the system logical structure is not accounted for [Watts, 2002; Dobson et al., 2005]. Also, probabilistic dynamics models allow accounting for interdependencies and interactions among several CIs [Newmann et al., 2005], while MCs and MPNs have been typically used to analyze isolated systems only [Sultana and Chen, 2009].

BN analysis is a probabilistic approach that can be used for modeling and predicting the behavior of a system, based on observed stochastic events. A BN is a model that represents the interactions among the components in a system, from a probabilistic perspective. The representation is illustrated via a directed acyclic graph, where the nodes represent the variables and the links between each pair of nodes represent the causal relationships between the variables.

The disadvantages of this methodology arise from its complexity that leads to significant efforts in logic modeling and quantification, and from the limited capability of providing an exhaustive analysis.

2.2.4 Statistical analysis

Statistical analysis [Casals and Solé, 2011] is suitable when rich data sets about the system operation and performance are available. However, using these data effectively is difficult for a number of reasons: i) data about CIs operation and performance generally come from a variety of past operating conditions that may not fully reflect the situations of interest at present and in the future; ii) the relationships between the measures of the operating conditions (e.g., the loads placed on the different portions of the system) and system performance may be complicated and poorly understood; iii) the data sets may be very large, making it difficult to draw clear insights from them. Moreover, the structure of the network under analysis may be hidden by the fact that the data are often presented in an aggregate form [Dekker, 2005; Debon et al., 2010]. The wealth of statistical models available for the analysis of engineered systems [Lord et al., 2005] can also be a drawback, in that a proper choice must be made of the most suitable model for the specific CI which best fits the physics of the provided service. In this sense, special emphasis must be put on comparing the accuracy and usefulness of the models by means of goodness of fit statistics.

These statistical techniques have been proposed as tools for decision support in the diagnosis and rehabilitation of CIs, e.g. water supply systems [Yamijala et al., 2009], with the additional aim of identifying the system most critical parameters. However, criticality ranking of the systems components is not possible by resorting to statistical techniques only, because the data are typically presented in an aggregate form and no identification of the topological structure is accounted for.

2.2.5 Agent based Modeling simulation

Agent- based modeling (ABM) offers an attractive modeling paradigm for describing the dynamic system operational behavior, with close adherence to the reality of the coupled processes involved [D'Inverno and Luck, 2004]. One of the major advantages of ABM for modeling and simulating critical infrastructures is the possibility to include physical laws into the simulation and to emulate the behavior of the infrastructure as it emerges from the behaviors of the individual agents and their interactions. The conceptual modeling framework consists in the abstraction of the relevant components of the system as individual interacting agents. Agents are used to model both technical components, such as the electric power generators, and non-technical components, such as grid operators. The different agents interact with each other directly , e.g., generator dispatch, or indirectly, e.g. via the physical network [Schlapfer et al. 2008]. This modeling achieves a close representation of the system behavior by integrating the spectrum of different stochastic phenomena which may occur, thus generating a multitude of representative stochastic, time-dependent event chains.

The level of modeling detail offered by the object-oriented approach allows analyzing a multitude of time-dependent availability aspects. The main problems are related to the slow simulation speed and the large number of parameters to be input in the analysis [Eusgeld et al., 2009]. However, by focusing on specific safety aspects, the model can be simplified and the computational burden reduced.

3. All–hazard approach

3.1 Problem statement

The attention on critical infrastructures is evolving from concerns about aging public works (in the 1980s) to redefinition in terms of national security as a result of increased international terrorism (after 11 September 2001) and susceptibility against natural hazards, to unprecedented failure combinations and malicious (cyber) attacks (during mid-2000s). Consequently, the view has broadened from local via national/regional to global while concerns about single failure mechanisms have developed into a full set of potential failures, hazards and threats. As a result, nowadays strategies to reduce and manage vulnerabilities in critical infrastructures and the provision of related analytical instruments have to follow an ‘‘all hazards approach’’ [Kroger and Zio, 2011].

Past years have seen an increased numbers of events affecting vital infrastructures: they were triggered by various sources of hazards but showed the importance of an expanding “spectrum of threats” including terrorism or other manmade disasters, modifying the understanding of the risk’s triggering event. This kind of hazard is highly unpredictable and therefore difficult to assess and to prevent.

An analysis aimed at identifying the causes of damage or disruption of services in CIs is necessary for safe operation and protection: an *all-hazard analysis* encompassing a more general view on the hazards targeting a given system is required. In particular, the approach must handle also malevolent acts, which differ from natural or other man–made hazards and lack of a well-established methodology for uncertainty assessment.

The term “all-hazard” has been adopted from emergency planning where “it means that there are things that commonly occur in many kinds of disasters, such as the need for emergency warning or mass evacuation, that can be addressed in a general plan and that that plan can provide the basis for responding to unexpected events “[Waugh, 2004; Pollet and Cummins, 2009].

However, establishing a common denominator to assess the vulnerability of a system that is exposed to natural and accidental hazards, and threats of malevolent acts it is not straightforward. The need is to capture the CI vulnerability sources and issues, given its technical and physical features, and the dependencies and interdependencies on other systems.

In this PhD thesis, the *All-HAZard Analysis* (A-HAZAN) is presented: a framework of system analysis for identifying the vulnerable elements of CIs, considering natural hazards, random failures and intentional attacks. While the first two types of vulnerability are characterized by stochastic uncertainties and can be analyzed by traditional safety analysis tools, intentional attacks require a new way of analysis. Two methods aiming at quantifying the susceptibility to intentional attacks are proposed.

3.2 Qualitative A-HAZAN

3.1 Hazard and threat

Hazard and threat are fundamental concepts in vulnerability analysis. In literature, several definitions can be found, each one of them highlighting a particular nuance.

In the United Nations’ view, hazard is defined as "a potentially damaging physical event, phenomenon and/or human activity, which may cause loss of life or injury, property damage, social and economic disruption or environmental degradation. Hazards can be single, sequential or combined in their origin and effects" [Turner et al, 2003]. The European Cooperation for Space Standardization (ECSS) defines hazard as “an existing or potential condition of an item that can result in an accident” [White, 1974]; the condition is associated with the design, fabrication, operation or environment of the item, and has the potential for accidents. These two definitions encompass the idea of hazard described as a “condition prerequisite to a mishap” [UNISDR, 2004] and as “a source of potential harm” [ECSS, 2004]. The concept of hazard is strictly tied to the presence of a potential source of difficulty both natural or manmade.

On the other hand, the concept of threat is defined as “a potential intent to cause harm or damage to the system by adversely changing its state” [ISO guide, 2009]. This definition is strictly

linked to intentional and malevolent acts [Apostolakis and Lemon, 2005], and it seems in contrast with the one by the US Homeland Security, which describes threat as a "natural or manmade occurrence, individual, entity, or action that has or indicates the potential to harm life, information, operations, the environment, and/or property" [OHS, 2009]; in the latter, little reference is made to the idea of intention embedded in the former definition of threat.

From these definitions, the general concept of hazard emerges as a general condition of potential source of harm that embeds the concept of threats. Complex infrastructures have critical elements that, if disabled, could lead to significant disruptions. Hazards looming over such elements of the system may originate from the outside and inside of the system. Within hazards, threats strongly contain the concept of intentionality. Terrorist acts, for example, are distinguished by a malevolent intelligence directed toward maximum social disruption; conventional or computer-generated attacks could unleash a chain of events in which, for example, a service grid collapses with cascading effects; they also must be characterized by a "degree of likelihood".

In this respect, there is a need for systematic approaches to the identification of the relevant risks from intentional attacks, and to the development of effective measures for managing them. The all-hazard concept must be encompassed within an approach that accounts for both the degree of accessibility of hazards and threats to a potential target, and the system-damaging sequence of events that is initiated after the attack. In an all-hazard view, malevolent acts, accidental and natural hazards are all considered (figure 3.1); yet, they require a different analytical treatment.

The first preliminary step of the HAZAN consists of a qualitative evaluation of all the possible hazards (all-hazard) that can impact the system.

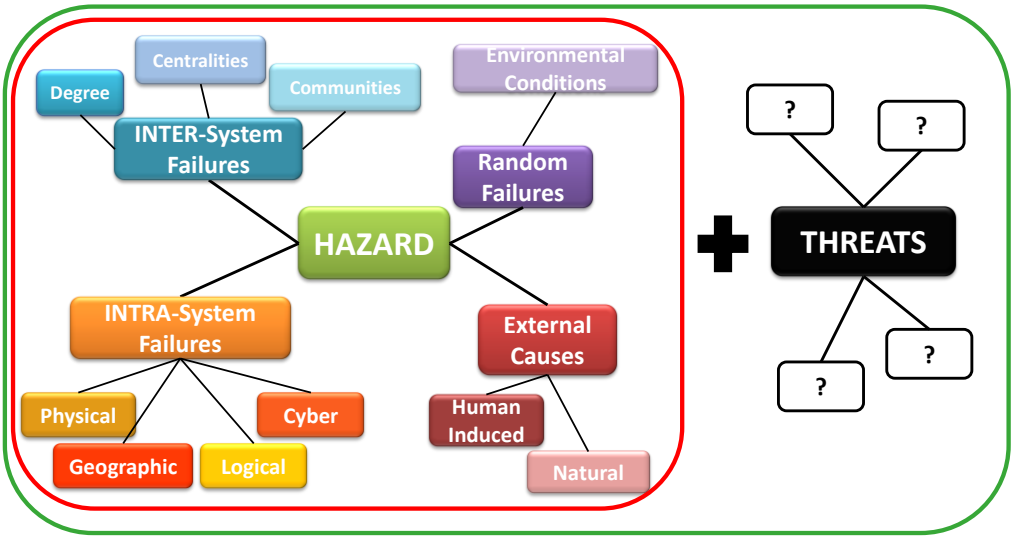


Figure 3.1 Pictorial representation of all-hazard vulnerability analysis.

We assumed the following (figure 3.1) classification:

- *Threats.* These are potential events characterized by the act of a malevolent intelligence directed towards maximum social disruption. Along with the definition by Piwowar et al. [Piwowar et al., 2009], intentional threats can be further declined in three different specifications. Firstly (1), the "malevolence of opportunity", which is brought by one person, possibly an employee of the company; often, the person who decides to enact the attack is motivated by a given fact: it could be an argument with a teammate or a manager, psychological problems, etc.; the difficulties to counter such an act are that the attack could occur everywhere, at any moment and reaction is not necessarily prepared; moreover, the attacker is a person already included in the system, who often knows the safety/security system and is able to act in a short time. Secondly (2), the "claiming malevolence" which is organized and prepared by a group of people in reaction to a concrete event (but not with an over-destruction willingness); often, the clinched event is dismissal plans or a refusal of syndicates' demand for social improvement. It could also be an act to manifest against a political announcement (for example, about sensitive subjects like nuclear, which often leads to demonstrative operations in order to avoid any building of new reactors or site of treatment of radioactive wastes). Thirdly (3), the "massive terrorism" whose goal is to destroy, or create damage to infrastructures and people; a definition by the French philosopher Raymond Aron [Aron, 1966] states that "an act is called terrorism if its psychological effects widely overcome its physical damages" and it is perfectly aligned with the assumption of the "homeland security" [OHS, 2009] , which conceives terrorism as "a premeditated threat or an act of violence against non combatant persons, properties and environmental or economic targets to induce fear, intimidate, coerce, or affect a government, the civilian population, or any segment thereof, in furtherance of political, social, ideological, or religious objectives.
- *Random failures.* These are typically permanent or transient outages due to components' failures, the effects of environmental conditions or random failures due to maintenance operations and may be identified by standard safety analysis techniques (e.g., FMECA, Hazop, and others).
- *INTRA-system failures.* These are failures within the system, typically dependent failures. This term considers how the components are related in the network. In a power transmission network, these are common mode failures or cascading failures.
- *INTER-system failures.* This term accounts for interdependencies, so it considers the connection among different systems. Following failures spreading through (Rinaldi et al., 2001) interdependencies can be classified in:
    - *physical dependencies (*an output from a system is required as an input to another system and vice versa): for example, for a generating unit this could be the supply,

e.g., oil, coal, uranium, etc. An effective supply could be guaranteed if the transportation network works.

  - *cyber dependencies* (the state of a system is dependent on the information transmitted through an information infrastructure): this entails the communication and control system, SCADA (Supervisory Control And Data Acquisition) and EMS (Energy Management System). The SCADA system allows the automation and the control of the network.
  - *geographic dependencies* (two or more systems can be affected by the same local event, i.e., they are spatially proximate): these are due mainly to co-location, that is different components are at the same place and failures happen at the same time to different components. For example, a bus station may host generating units in conjunction with load buses.
  - *logical dependencies* (includes all types of interdependencies other than physical, cyber and geographic dependencies, for example related to human behavior): these encompass the dependencies between financial markets and the power transmission grid. Such interdependencies will not be treated here.

- *External causes.* This term considers natural hazards, such as meteorological or seismic phenomena, and unintentional human-induced hazards, such as the processing or the storage of potentially hazardous materials, or nearby military installations. External causes may be further grouped into *local* external causes, e.g., a lightening striking a building or an aircraft crash, or *diffuse* external causes, such as earthquakes, hurricanes, flooding or the leakage of explosive or toxic materials.

This first classification allows a qualitative evaluation and organization of all the possible sources of hazard a system may be exposed to.

3.2 Susceptibility, attractivity and accessibility modeling

The concept of *susceptibility* has been introduced in [Apostolakis and Lemon, 2005] to measure the degree to which a CI is prone and accessible both to hazards and threats. For example, the susceptibility to threats of a transmission line in a power grid, i.e. the degree of exposure of a pylon, is affected by its location, e.g. if the pylon is located in an urban area or in a remote, isolated, inaccessible region.

Susceptibility refers to the property of the system of being potentially damaged and combines the likelihood of a hazardous event, the differential exposure, the potential sensitivity of a system or element of the system exposed, i.e. the degree to which a system or the element could be potentially

damaged or affected by a given hazard and the existing capacity of this system that could potentially reduce this level of damages (e.g. existing measures of prevention, mitigation, etc.) [Bouchon, 2006].

But, the idea of susceptibility is meant to encompass that of probability for random events, while grasping the intentionality that is behind malevolent attacks. Therefore, it must be a function of those factors that are involved in planning and performing an attack, i.e., the accessibility and the attractivity of a target. The term *attractivity* considers the appeal of the target to intentional attacks. The term *accessibility* considers that components have been designed for efficiency and convenience, yet access must be easy for maintenance staff but difficult for attacks [Apostolakis and Lemon, 2005]:

$$\text{Susceptibility} = f\,(\text{attractivity, accessibility}) = f\,(y_1, y_2)$$

Since vulnerability is the manifestation of the inherent states of the system (e.g., physical, technical, organizational, cultural) that can be exploited by an adversary to harm or damage the system [Haimes and Horowitz, 2004], the identification of the features of the system influencing its susceptibility to attacks, and then attractivity and accessibility, is important (Figure 3.2). We identified three influencing variables, that compose attractivity and accessibility: size, level of protection and social criticality:

$$\text{Attractivity} = f_1\,(\text{size, level of protection, social criticality}) = f_1\,(x_1, x_2, x_3)$$

$$\text{Accessibility} = f_2\,(\text{size, level of protection, social criticality}) = f_2\,(x_1, x_2, x_3)$$

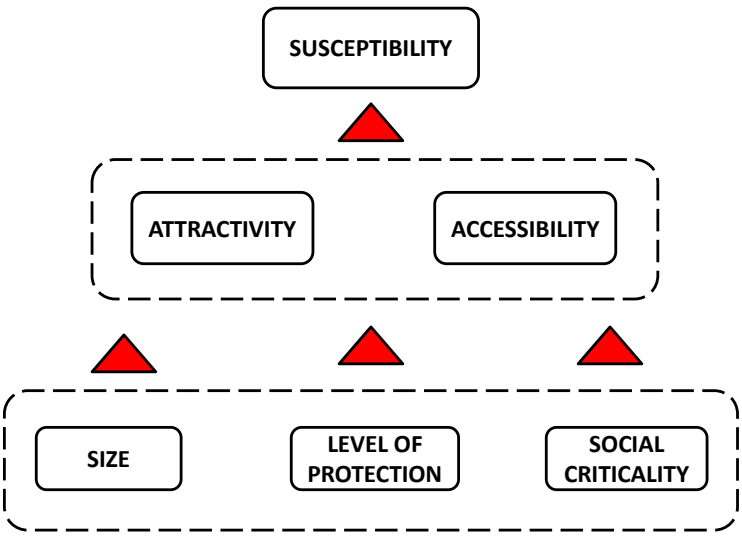


Figure 3.2 Pictorial representation of the susceptibility as a function of the characteristics of the system (size, level of protection and social criticality) and of the factors underpinning an intentional attack (attractivity and accessibility).

The three system's characteristic are here highlighted:

- *size* (x_1) of the component: this variable includes both physical and technical features of the component. On one side, physical, we consider that the importance of the component is given by its physical size; on the other side we consider the importance of the task of the component, i.e., for the elements of a power transmission network, the *level of demand, transmission or supply*. Thus, this influencing variable can be measured both with a characteristic physical dimension and as the consumed, transmitted or supplied power of the component (or as the fraction of the overall consumed, transmitted or produced power with respect to the whole system). As an example, if we consider generating unit in a power transmission system, the size of the component can be measured in MW and can be grouped in three levels:

Level	**Description**
3 – Large (L)	large plants (power production ≥ 300 MW)
2 – Medium (M)	medium plants (100 MW ≤ power production ≤ 300 MW)
1 – Small (S)	small plants (power production ≤ 100 MW)

Table I. Levels of size of infrastructure elements.

- *level of protection* (x_2): these are the logical and physical barriers deployed to prevent or discourage malevolent acts. Following [Apostolakis and Lemon, 2005], six levels of protection can be considered:

Level	**Description**
6 – Extreme(E)	Completely secure
5 – High (H)	Guarded, secured area, alarmed
4 – Moderate (M)	Secure area
3 – Low (L)	Complex barriers, security patrols, video surveillance
2– Very low (VL)	Unlocked, non-complex barriers
1 – Zero (Z)	Completely open, no control, no barriers

Table II. Levels of protections of infrastructure elements.

- *social criticality* (x_3)*:* given that the attack will be successfully accomplished, the impact on public opinion is influenced by the (conditional) effects caused by the achieved intentional act. The most relevant consequences here considered are in terms of human lives and geographic extension of the event:

Level	Description
4 – Severe (S)	Many victims, widely spread extension
3 – High (H)	Many victims, contained extension
2 – Moderate (M)	Few victims, widely spread extension
1 – Low (L)	Few victims or no victims, contained extension

Table III. Levels of social criticality of infrastructure elements.

We assume that the more protected a component, because of the security measures, the less accessible it is, but, the most attractive it is perceived, from a malevolent act point of view. This assumption relies on the idea that a malevolent will think that if a component deserves a good level of protection, then its damage would produce a large disruption, and a double intent is accomplished: a huge damage and a deep psychological effect.

3.3 Quantitative A-HAZAN

Data and information concerning malevolent acts are incomplete, imprecise, ambiguous. Moreover, when dealing with malevolent acts one should explicitly take into account the intelligent and adaptive nature of the threat.

Protecting critical infrastructures against intentional attacks is fundamentally different from protecting against random accidents or acts of nature. Intelligent and adaptable adversaries may try different offensive strategies or adapt their tactics in order to bypass or circumvent protective security measures and exploit any remaining weaknesses. Although engineering risk and reliability analysis are clearly important for identifying the most significant security threats and vulnerabilities to terrorist attacks (particularly in complex engineered systems, whose vulnerabilities may depend on interdependencies that cannot be readily identified without detailed analysis), such analyses do not lead in any straightforward manner to sound recommendations for improvements. A different approach is needed.

3.3.1 Fuzzy inference modeling

When dealing with the evaluation of threats, the information provided is typically rather inhomogeneous and uncertain. In this view, there is the need to model the causal relationship between

the susceptibility to threats and the relevant influencing variables that concur with the occurrence of the intentional attack. Moreover, the complexity and uncertainty of the characteristic parameters render inferential modeling difficult to execute with standard analytical tools.

To partially overcome these difficulties, a fuzzy logic approach [Zadeh, 1965; Cox, 1999] is adopted in this PhD work, which seems suitable to handle the ambiguous and limited information involved in the problem.

Following the fuzzy logic formalism, the input and output variables are described with labels, and corresponding linguistic terms, that quantitatively take into account the inherent uncertainties and ambiguities. The ranges of variability of the values of the input and output variables, called the universe of discourse, are quantitatively divided into subset labeled with the linguist terms as defined in tables I, II, and III in Section 3.2. and in table IV in the following, and each variable is attributed a membership function (MF).

For a given defined value of the parameter, the corresponding MF value (membership grade) quantifies how much that crisp is properly described by the linguistic term associated with that MF.

In this PhD thesis work, triangular, half-overlapped normalized MFs $\mu(\cdot)$ are used, each one having the peak (μ =1) at the midpoint of the supporting subset and the two vertexes of the base in correspondence to the midpoints of the adjacent neighboring fuzzy sets. In Figure 3.3 are shown, as example, the membership function for the linguistic variable x_2 corresponding to level of protection:

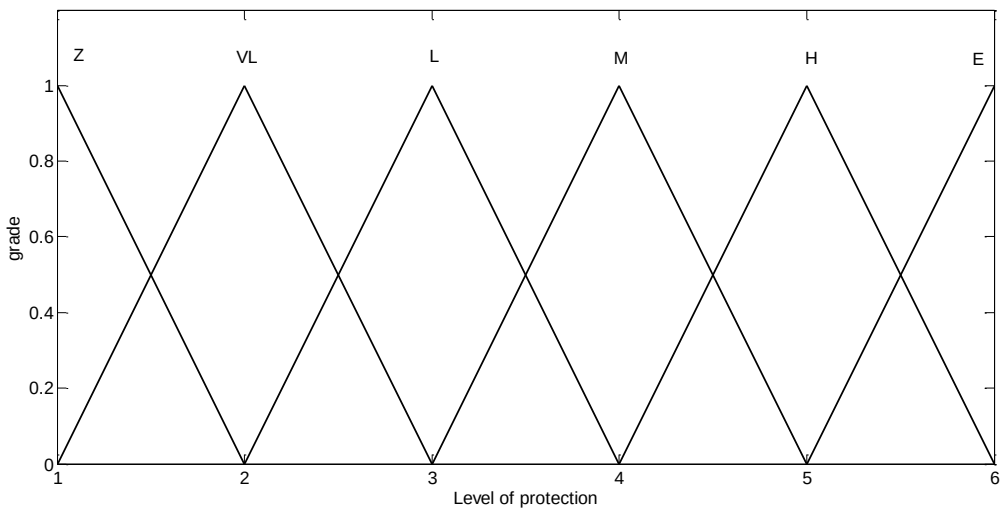


Figure 3.3 Membership functions for the linguistic variable x_2.

The next step of the model definition is the generation of a table of fuzzy rules, i.e., the set of linguistic statements that relate the three input variables x_1, x_2, x_3 (antecedents) to the output y (consequent). The generic l'th fuzzy rule is written as:

IF x_1 is X_1^1 AND IF x_2 is X_2^1 AND IF x_3 is X_3^1 THEN y is Y^1 (3.1)

where X_i^1 is the linguistic term associated with the fuzzy set of the i'th antecedent (i = 1,2,3) and Y^1 is the linguistic term associated with the fuzzy set of the consequent. We propound the following approach to generate the rules:

IF x_1 is X_1^k AND IF x_2 is X_2^m AND IF x_3 is X_3^n THEN y is Y^1 $3 \leq k+m+n \leq 5$ (3.2)

IF x_1 is X_1^k AND IF x_2 is X_2^m AND IF x_3 is X_3^n THEN y is Y^5 $12 \leq k+m+n \leq 13$ (3.3)

$$\text{IF } x_1 \text{ is } X_1^k \text{ AND IF } x_2 \text{ is } X_2^m \text{ AND IF } x_3 \text{ is } X_3^n \text{ THEN } y \text{ is } \begin{cases} Y^2 \text{ if } k+m+n = 6 \text{ or } 7 \\ Y^3 \text{ if } k+m+n=8 \text{ or } 9 \\ Y^4 \text{ if } k+m+n = 10 \text{ or } 1 \end{cases} \quad (3.4)$$

$$6 \leq k+m+n \leq 11$$

where k, m and n are the indexes of the linguistic levels of, namely, variable x_1, size, variable x_2, level of protection, and variable x_3, social criticality. The index k can vary in the range [1,2,3], m in the range [1,2,3,4,5,6] and n in the range [1,2,3,4].In table 3 are presented the linguistic labels for accessibility and attractivity.

Level	Description
Y_5 – Extreme	Extreme degree of attractivity or extreme ease of accessibility
Y_4 – High	Elevated level of attractivity or high ease of accessibility
Y_3 – Moderate	Average degree of attractivity or medium ease of accessibility
Y_2 – Low	Low level of attractivity or low ease of accessibility
Y_1 – Very low	Almost null degree of attractivity or very low ease of accessibility

Table IV. Levels of accessibility (y_1) and attractivity (y_2).

In Tables V and VI, the model rules are represented: they are shown with reference to input x_2 (level of protection) and x_3 (social criticality) and are parameterized with respect to the three different linguistic terms of variable x_1 (size).

	L	M	H	S
Z	Y_3	Y_3	Y_4	Y_4
VL	Y_2	Y_3	Y_3	Y_4
L	Y_2	Y_2	Y_3	Y_3
M	Y_1	Y_2	Y_2	Y_3
H	Y_1	Y_1	Y_2	Y_2
E	Y_1	Y_1	Y_1	Y_2

(a)

	L	M	H	S
Z	Y_3	Y_4	Y_4	Y_5
VL	Y_3	Y_3	Y_4	Y_4
L	Y_2	Y_3	Y_3	Y_4
M	Y_2	Y_2	Y_3	Y_3
H	Y_1	Y_2	Y_2	Y_3
E	Y_1	Y_1	Y_2	Y_2

(b)

	L	M	H	S
Z	Y_4	Y_4	Y_5	Y_5
VL	Y_3	Y_4	Y_4	Y_5
L	Y_3	Y_3	Y_4	Y_4
M	Y_2	Y_3	Y_3	Y_4
H	Y_2	Y_2	Y_3	Y_3
E	Y_1	Y_2	Y_2	Y_3

(c)

Table V. Rules for accessibility variable (y_1) when the size variable x_1 is (a) small, (b) medium and (c) large.

	L	M	H	S
Z	Y_1	Y_1	Y_1	Y_2
VL	Y_1	Y_1	Y_2	Y_2
L	Y_1	Y_2	Y_2	Y_3
M	Y_2	Y_2	Y_3	Y_3
H	Y_2	Y_3	Y_3	Y_4
E	Y_3	Y_3	Y_4	Y_4

(a)

	L	M	H	S
Z	Y_1	Y_1	Y_2	Y_2
VL	Y_1	Y_2	Y_2	Y_3
L	Y_2	Y_2	Y_3	Y_3
M	Y_2	Y_3	Y_3	Y_4
H	Y_3	Y_3	Y_4	Y_4
E	Y_3	Y_4	Y_4	Y_5

(b)

	L	M	H	S
Z	Y_1	Y_2	Y_2	Y_3
VL	Y_2	Y_2	Y_3	Y_3
L	Y_2	Y_3	Y_3	Y_4
M	Y_3	Y_3	Y_4	Y_4
H	Y_3	Y_4	Y_4	Y_5
E	Y_4	Y_4	Y_5	Y_5

(c)

Table VI. Rules for attractivity variable (y_2) when the size variable x_1 is (a) small, (b) medium and (c) large.

The logic approach deployed to combine x_1, x_2 and x_3 into y_1 and y_2 is applied to y_1 and y_2 to yield the susceptibility y. As suggested by the Homeland Security Advisory System for risk (DHS, 2011), five Threat Conditions are identified by means of Roman numerals. From lowest to highest, the levels are:

Low = I. This condition refers to a low susceptibility of terrorist attacks.
Guarded = II. This condition is declared when there is a general susceptibility of terrorist attacks.
Elevated = III. An elevated condition is declared when the susceptibility of attack is significant.
High = IV. A high condition is declared when there is a high susceptibility of malevolent attacks.
Severe = V. A severe condition reflects a severe susceptibility of terrorist attacks.

The higher the threat condition, the greater is the susceptibility of an intentional attack, going from I, low susceptibility to attacks, to V, severe condition of malevolent acts. To evaluate susceptibility to attacks, the set of rules that relate the two input variables, attractivity, y_1 and accessibility, y_2, to the output, susceptibility, y, are assigned in Table VII.

y_2 / y_1	Y_{21}=I	Y_{22}=II	Y_{23}=III	Y_{24}=IV	Y_{25}=V
Y_{11}=I	**I**	**I**	**I**	**II**	**III**
Y_{12}=II	**I**	**I**	**II**	**III**	**IV**
Y_{13}=III	**I**	**II**	**III**	**IV**	**V**
Y_{14}=IV	**II**	**III**	**IV**	**V**	**V**
Y_{15}=V	**III**	**IV**	**V**	**V**	**V**

Table VII. Rules linking the linguistic variables attractivity (y_1) and accessibility (y_2) to the linguistic variable susceptibility y. The Roman numerals refer to the five levels: I, II, III, IV and V.

Table VII shows how the levels of attractivity and accessibility yield the different levels of susceptibility to attacks. The susceptibility to attack increases from the upper left corner where the susceptibility to intentional attacks is low, I level, to the lower right corner where the threat of attacks is severe, V level.

These rules, applied to the different components of the critical infrastructure, identify the different values of susceptibilities that are used in sorting the vulnerabilities of different components.

3.3.2 Game theory modeling

In the scientific community, there is the need to establish an appropriate risk framework for dealing with intentional malevolent acts in vulnerability assessment of the risks related to sabotage and terrorism. Intelligent agents may try different offensive strategies or adapt their tactics in order to bypass or circumvent protective security measures and exploit any open weaknesses.

In particular, risk and reliability analysis generally assumes that threats or hazards are stationary, whereas in the case of security, threats are adaptive and can change in response to the defenses that are implemented. It follows that, the traditional probabilistic risk assessment developed for system safety and reliability engineering is typically not adequate in the security domain and can mislead a vulnerability analysis against malevolent attacks [Brown and Cox, 2011].

Game theory provides useful concepts and computational tools for modeling and quantifying risks and allocating resources to defend targets such as infrastructures against intelligent adversaries [Bier and Azaiez, 2011].

A game is defined as:

$$(X, Y, f: X \times Y \rightarrow R, g: X \times Y \rightarrow R) \qquad (3.5)$$

where X and Y represent the strategies spaces for the first player and for the second player namely, and *f* and *g* are the corresponding payoff, or utility, functions. A play in which two players interact and the payoff of a player can only be the opposite of the payoff of the other player in every outcome of the game is called a *zero-sum game*:

$$(X, Y, f: X \times Y \rightarrow R, g: X \times Y \rightarrow R, f = -g = p) \qquad (3.6)$$

Typically, it is a zero-sum game every game where the final situation for the player is either to win, or to tie or ease to lose. The game is then represented by an $n \times m$ matrix $P = (p_{ij}) \in R$ for all *i, j*, where $i \in X$ and $j \in Y$. The game unfolds as follows (Fig. 3.4): player one chooses to follow the *i* strategy, i.e., row *i*, player two chooses to follow the *j* strategy, i.e., a column *j*, and the entry p_{ij} of the matrix *P* is the amount the second player has to pay to the first player.

	Player 2				
Player 1	p_{11}	p_{12}	…	…	p_{1m}
	…	…	…	…	…
	…	…	p_{ij}	…	…
	…	…	…	…	…
	p_{n1}	p_{n2}	…	…	p_{nm}

Figure 3.4 Matrix of the game.

The fundamental issue is to establish when a pair $(\bar{i}, \bar{j})$, i.e., the choice of a row (strategy) by the first player and a column (strategy) by the second player, can be considered a solution of the game. If the first player selects the *i-th* row, and somehow player 2 is aware of it, then he will react by choosing the column resulting in the value $\min_j p_{ij}$. In response to that, player 1 will react to be able to guarantee himself (at least) the payoff of a value at least equal to $\max_i \min_j p_{ij}$. This is called the conservative value of the first player. In the same way, and taking into account a change of sign, the conservative value of player 2 payoff will be $\min_j \max_i p_{ij}$. When the interests of the players are strictly

opposite, the analysis of the conservatives values identifies an equilibrium for the game, because the choice of the strategies $(\bar{i}, \bar{j})$ guarantees the players to reach the largest payoffs they can get.

In this occurrence, this equilibrium point is a saddle point. The game matrix may not have any saddle point, and the players may chose strategies with some probabilities suggested by optimum rules. For instance, suppose the first player has n possible moves (the rows of the *P* matrix), then he will choose among a vector of strategies $x = (x_1, x_2,\ldots, x_n)$ and similarly, if the second player has m possible moves (the columns of the matrix *P*), he will choose among a vector of strategies $y = (y_1,\ldots,y_m)$. The nth and m^{th} –simplex are the new strategy spaces for the two players, and they are called mixed strategy. The solution of the game entails finding the optimal strategies for the two players. The fundamental theorem by Von Neumann, also known as the "minmax" theorem, guarantees that a two players, finite, zero-sum game has always equilibrium in mixed strategies.

In this PhD work, we represent the interplay between an attacker, player 1, and a defender, player 2, as a zero-sum game. The degree of exposure to an attack, the susceptibility, is the utility or payoff function. The intent of the attacker is searching for the maximum susceptibility of the system to cause the great possible destruction, while the objective of the defender is minimizing the susceptibility function, that is the system degree of exposure to attacks, in order to protect to his best the system from malevolent acts.

Quantitative susceptibility

In order to compute the susceptibility function, i.e., the payoff function, we need to quantify its influencing variables, namely size, level of protection and social criticality. The system under study is a power transmission system in which each component is considered according to its functional role. Three main roles can be identified: generating units, which supply power, load buses, which receive the power supplied by generators and transmission lines which spread the power from the source to the target. For the moment, we neglect the influence of the component physical size on attractivity and on accessibility, and we propose a quantitative expression for susceptibility.

We quantify the social criticality (SC) of a component *i* as proportional to the extent of the damage to the unprotected component, i.e. with a zero level of protection, $LOP_i = 1$:

$$SC_i = m_i \cdot DAMAGE_i + q_i \qquad (3.7)$$

Depending on the role of the component, $DAMAGE_i$ represents the power not supplied, transmitted or received and is expressed in terms of MW. The coefficient m_i in Eq. 3.7 also depends on the functionality of the component *i* and it weights the "importance" of the loss: for example, it is expected that the damage to a single transmission line can cause no loss if the power flow can be

redistributed on other transmission lines (N-1 criterion), so the transmission line coefficient m_i will be lower than m_i for a bus (user) that does not receive power due to its incapacitation. Finally, q_i expresses the perception of the disutility of the component *i* which subsumes the geographic extension and the number of potential victims. As a result, we have:

$$SC_{Transmitter} \leq SC_{User} \leq SC_{\mathrm{Provider}} \qquad (3.8)$$

This inequality considers that an attack launched to a generating unit can have a major impact on the public opinion than an attack to transmission lines or to users.

In Section 3.2, Table II, six levels of protection (LOP) were introduced and identified (Lemon and Apostolakis, 2005) by means of linguistic terms. These levels are here transposed in percentage values as shown in Table VIII:

Level of Protection (LOP)	**Description (Examples)**	**Associated Percentage (%)**
6 - Extreme	Completely secure, inaccessible	100
5 -High	Guarded, secure area, locked, alarmed, complex closure	86
4 – Moderate	Secure area, locked, complex closure	69
3 – Low	Complex barrier, security patrols, video surveillance	35
2 – Very low	Unlocked, noncomplex barriers (door or access panel)	17
1 – Zero	Completely open, no controls, no barriers	0

Table VIII. Level of protection and associated percentage values.

Given the values of social criticality (Eq. 3.8) and of the level of protection (Table VIII), we quantified the susceptibility of the component *i* as:

$$S_i = a \cdot SC_i \cdot e^{\left(-\frac{b}{SC_i} LOP_i\right)} + constant_i \qquad (3.9)$$

In Eq.3.9 it is assumed that a component whose attack has a great social impact on public opinion is more exposed to an attack, that is, has a greater susceptibility. On the other hand, if the component has high level of protection, the exposure to possible attacks should diminish.

For the entire system, the total susceptibility is assumed to be the sum of the susceptibilities of the components:

$$S = \sum_i \left(a \cdot SC_i \cdot e^{\left(-\frac{b}{SC_i} LOP_i \right)} + constant_i \right) \qquad (3.10)$$

The model of the game

Our model of the game considers two players: an attacker and a defender. We assume that the attacker aims at the destruction of the system or of a part of it: his aim is to maximize the losses he can cause to the system. From this point of view, an attacker will aim at incapacitating the component with the highest value of susceptibility which could guarantee a successful attack. On the other side, the defender wants to minimize the damage produced on the system by malevolent acts, and he will try to protect the system by changing the levels of protections of the different components: his aim is to minimize the susceptibility to attacks of the different components. We assume the susceptibility function to represents the utility, or payoff, function for the game: a player, the attacker, wants to maximize it; the other player, the defender, needs to minimize it. Since both players compete on the same quantity, trying to maximize its increase or decrease, the game outlines as a zero-sum game.

As for the strategies, we assume that the attacker takes the offensive and strikes a component. The strategy of the defender will be to dispose all the possible protections to limit the damage caused by the attack. Each component of the system is assigned a value of susceptibility, which will depend on its social criticality (SC_i) and on its current level of protection (LOP_i): these susceptibility values are computed using Eq. 3.9 and they represent the original system configuration.

Then, the defender can act by changing the LOP of the different components. The changes executed on the system identify a strategy. The susceptibility of the system will change with reference to the initial configuration accordingly to the protection strategy *X* carried out:

$$\Delta S_X = S_i\left(SC_i, LOP_i\right)_{STRATEGY(X)} - S_i\left(SC_i, LOP_i\right)_{INITIAL\_CONFIGURATION} \qquad (3.11)$$

We exemplify the power grid with its functional model. In the schematic representation of the system (Figure 3.5), only the operational role of the components is considered: generating units (G), load buses (B) and transmission lines (T).

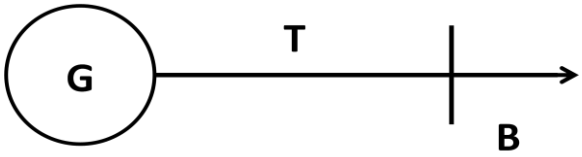


Figure 3.5. Schematic representation of the functional component of a transmission network.

Starting from the initial configuration (Config. 0 in Table X), the level of protection (LOP) of the different components may be increased (+) or decreased (-) by a certain amount or can be left unchanged (0). In this example, we assume that only unitary increases or decreases of LOP are allowed, i.e. $\Delta LOP = \pm 1$. Moreover, we assume that the cost of an increase is equal, but with opposite sign, to the cost for a decrease, and for each configuration we ask the total cost for changing LOP to be zero. This constraint reflects the limitation of constrained budget. The possible configuration variations for the elementary example are represented in Table IX:

Component	Config.0	Config.1	Config.2	Config.3	Config.4	Config.5	Config.6
G	0	+	+	0	0	-	-
T	0	-	0	+	-	+	0
B	0	0	-	-	+	0	+

Table IX. Table of the possible configuration variations for the elementary system.

In order to maintain a zero cost constraint, for an increase in the level of protection on one component we need to have a decrease of the same amount on another element of the system. For example (Table X), in the configuration number 2, the defender chooses to elevate the level of protection for the generating units from level 3 to level 4 and, in order to maintain the constraint on costs, he diminishes the protection on bus from level 3 to level 2. The components with 0, the transmission line in strategy 2, undergo no changes.

Component	Config.0	Config.1	Config.2	Config.3	Config.4	Config.5	Config.6
G	3	4	4	3	3	2	2
T	3	2	3	4	2	4	3
B	3	3	2	2	4	3	4

Table X. Table of the possible variation of the levels of protection in each configuration for the elementary system.

Given the possible choices of levels of protection, for each component of the system, the incremental susceptibilities (Eq. 3.11) are computed and collected in Table XI. Each element of Table

XI represents the incremental susceptibility due to the variation of the degree of exposure to an attack, i.e., the susceptibility function due to a change in the level of protection with respect to the susceptibility computed for the initial configuration. If a component is increased its level of protection with respect to Config. 0, the susceptibility to attack decreases and the incremental value is negative. On the other hand, if a component is decreased its level of protection, with respect to Config. 0, its exposure to attacks increases, the susceptibility to attack raises and its incremental value is positive. Finally, if the level of protection of a component is unchanged, there is no variation in the susceptibility with respect to the initial configuration of the component.

The simulation of the game was performed on the system, and the obtained values are summarized: in table XI:

			DEFENDER						
			0	**1**	**2**	**3**	**4**	**5**	**6**
			0.0000	**0.0000**	**0.4028**	**0.0000**	**0.0000**	**0.1668**	**0.4304**
ATTACKER	**G**	**0.2138**	0	-31	-31	0	0	17	17
	T	**0.4410**	0	11	0	-13	11	-13	0
	B	**0.3452**	0	0	12	12	-17	0	-17

Table XI. Table of the possible strategies for the elementary system

The mixed startegy equilibrium is obtained by mean of the minmax theorem and the strategies that the players could carry out are inserted beside the payoff matrix (Table XI). From the standpoint of the attacker, the best strategy is to attack component T: this move can guarantee the attacker to increase his expected payoff, causing the maximum damage to the system. In this case, the expected payoff for the attacker is -2.2388: this means that the attack does not succed in increasing the susceptibility of the system.

In order to limit the weakening of the system, the defender can put into effect three “best” strategies: strategy 2, strategy 5 and strategy 6. We can see that the more likely strategy for the defender is strategy 6 with a probability p_6 = 0.4304, then strategy 2 with a probability p_2= 0.4028 and strategy 5 with a probability p_5 = 0.1668. With respect to the initial configuration of the system, in strategy 6 the defender has increased the level of protection for the bus unit (corresponding to a decrease in the susceptibility) and has decreased the level of protection of the generating unit. In strategy 2 the defender can do exactly the opposite: he can increase the level of protection of the generating unit and decrease the level of protection of the load bus. Finally, in strategy 5 there is an increase in the level of protection of the transmission lines to the detriment of the protection for the

generating unit. The expected payoff for the defender is +2.2388. The fulfillment of one of these "best" strategies helps the defender improving the level of protection of the system, minimizing the susceptibility to intentional attacks.

4. Topological analysis of electrical power networks

4.1 Problem statement

Topological analysis based on classical graph theory can unveil relevant properties of the structure of a network system by highlighting the role played by its components, nodes and connected arcs, making preliminary vulnerability analysis assessment based on the simulation of faults, mainly represented by the removals of nodes and arcs and the subsequent re-evaluation of the network topological properties.

In the case of critical infrastructures, particularly in the case of electrical transmission networks of interest here, vulnerability analysis largely takes a topological approach to identify the critical components in the network [Albert e al., 2004; Crucitti et al., 2005; Zio et al. 2008]. Even though such analyses are capable of identifying elements of structural vulnerability, they are limited from the point of view of the physical analysis of the electrical transmission system, which the networks represent. These limitations are all related to the fact that the analysis performed focuses only on the topological features of the network, thus neglecting its physical characteristics and the actual physical flowing through it [Cadini et al., 2009; Hines and Blumsack, 2008]; this is not realistic for electrical transmission networks in which:

i) the "electrical" length of a path differs from the topological length, depending on the difficulty (resistance) of transmission; this has been pointed at as a limitation in engineered network infrastructures, where the capacities of the arcs connecting the components limit the flow among them [Hines and Blumsack, 2008];

ii) the electrical power is not necessarily routed through the shortest paths: rather, the transmission of power is determined by physical rules, e.g. Kirchoff's laws, nodal voltages etc.;

In order to capture all the complex and rich properties observed in real CIs, there is the need for extending the models beyond pure structural topology.

4.2 Network elements importance: centrality measures

Topology-driven analysis of vulnerability is an essential part which allows addressing important questions related to the connectedness of nodes, shortest path lengths, geographical and regional specifics, etc. and to provide the reliable identification of the most critical connections, nodes or areas on which to focus a detailed modeling and simulation analysis. Given the somewhat 'abstract' level of the topological analysis, the results gained with respect to the vulnerable points (or lines) in the system ("first findings") have to be treated with caution and may not necessarily match with the results of an accurate modeling. We therefore focused our attention on how to contribute to reducing the gap between the highly conceptualized (and abstract) analyses based purely on considerations of the system topology and the highly detailed (and computationally demanding) simulations of system behavior, in order to render the overall vulnerability assessment more feasible and robust.

In this view, the formalism of weighed networks was exploited to provide different graph-theoretical representations and analyses of a power transmission system. The "weights" appended to the network elements are intended to capture relevant electrical and reliability properties of the system, so as to overcome the classical simplifying but unrealistic assumption that electrical flow occurs along the shortest and failure-free paths of connections.

When analyzed from a purely topological point of view, the transmission network system can be modeled as a stochastic, weighted, undirected, connected network in which each electric bus is transposed into a node, linked by edges representing the overhead lines connecting consecutive buses. In this respect, this representation focuses on the actual topological structure of the power transmission network.

Mathematically, the topological structure of the network can be represented as an undirected graph $G(V,E)$ where V represents the set of vertexes (or nodes, or components) ($N = \dim(V)$ is the number of nodes) and E represents the set of edges (i, j) ($K = \dim(E)$ is the number of edges). The connections are specified in an $N \times N$ adjacency matrix $\{a_{ij}\}$ whose entries are 1 if there is an edge joining node i to node j and 0 otherwise.

In the topological representation, no specification of the physical length of the edges is given. Each link is considered having a length equal to one and thus the distance between two nodes i and j is represented solely by the number of edges travelled in the path from i to j. On the basis of $\{a_{ij}\}$, it is possible to compute the matrix of the shortest path lengths $\{d_{ij}\}$ whose generic entry d_{ij} is the number

of edges making up the shortest path linking i and j in the network. The fact that G is assumed to be connected implies that d_{ij} is positive and finite $\forall i \neq j$ and that there are $N(N-1)/2$ distinct shortest paths among the N nodes.

From a topological viewpoint, various measures of the importance of a network element (arc or node), i.e., of the relevance of its location in the network with respect to a given network performance, can be introduced. In social networks, for example, so-called 'centrality measures' are introduced as importance measures to qualify the role played by an element in the complex interaction and communication occurring in the network. The term 'importance' is then intended to qualify the role that the presence and location of the element plays with respect to the average global and local connection properties of the whole network. Classical topological centrality measures are the degree centrality [Niemen, 1974; Freeman, 1979], the closeness centrality [Freeman, 1979; Sabidussi, 1966; Wasserman and Faust, 1994], the betweenness centrality [Freeman, 1979] and the information centrality [Latora and Marchiori, 2007]; they specifically rely only on topological information to qualify the importance of a network element.

A relevant outcome of the analysis of the network structure is the identification of the most important groups of elements of different sizes in the network. Indeed, critical groups of components may include components that are not critical when considered individually. Therefore, the criticality of these components may be underestimated if the importance ranking is performed only with respect to individual components.

Depending on the specific definition, a centrality measure describes the way in which a node interacts/communicates with the rest of the network, thus providing a way of prioritization of the importance of the nodes for network communication.

The group degree centrality [Everett and Borgatti, 1999], $C^D(g)$ of a group g in a network of N nodes, can be quantitatively defined as the number of first neighbours of the group nodes, normalized over the number of non-group members:

$$C^D(g) = \frac{\sum_{i \in g} k_i}{N - \dim(g)} \tag{4.1}$$

where k_i is the degree of node i, in group g and dim(g) is the dimension of the group, i.e. the number of member nodes. A node that is connected with multiple group nodes is counted only once.

The group closeness centrality [Everett and Borgatti, 1999], $C^C(g)$, is based on the idea that a node can quickly interact with all other nodes if it is easy accessible (close to) all others. If d_{ij} is the topological shortest path length (i.e., the number of connected arcs) between nodes i and j (also called

geodesics), the group closeness of a group g is the sum of such distances from the group to all vertices outside the group:

$$C^{C}(g) = \frac{N - \dim(g)}{\sum_{i \in g, J \in G} d_{ij}} \tag{4.2}$$

This measure is normalized by dividing the distance score into the number of non-group members, with the result that larger numbers indicate greater centrality. When the group consists of a single node, the group closeness centrality is the same as the individual node closeness centrality [Sabidussi, 1966; Freeman, 1979; Wasserman and Faust, 1994].

When the group consists of a single node, the group closeness centrality is the same as the individual node closeness centrality [Sabidussi, 1966; Wasserman and Faust, 1994].

To capture the failure behavior of the network, the reliability of its connecting edges is included in the framework of analysis by means of the formalism of weighted networks, the weight w_{ij} associated to the edge between the pair of nodes i and j being its reliability:

$$p_{ij} = e^{-\lambda_{ij} T} \tag{4.3}$$

where λ_{ij} is the failure rate of edge ij linking nodes i and j and T is a reference time ($T = 1$ year, in this work).

The distance from other nodes may not be the only important property in a network of components; of relevance is also which nodes lie on the shortest paths among pairs of other nodes, because such nodes have control over the flow of information in the network. To capture this feature, the betweenness centrality is defined [Freeman, 1979], such that a node is central if it lies on several shortest paths among other pairs of nodes. If g is a subset of a graph with vertex set V, let s_{ij} be the number of geodesics connecting i to j and $s_{ij}(g)$ be the number of geodesics connecting i to j passing through g. Then, the group betweenness centrality of g [Everett and Borgatti, 1999], denoted by $C^{B}(g)$, is given by:

$$C^{B}(g) = \frac{\sum_{i,j \in G, i<j} \frac{s_{i,j}(g)}{s_{i,j}}}{(N - \dim(g) - 1)(N - \dim(g))} \tag{4.4}$$

In the above equation, the sum is taken over all pair of nodes.

This measure is normalized by dividing by the theoretical maximum value, which occurs for a group of a given size when the result of identifying all the group vertices (i.e., shrinking them to a single vertex) is a star with the group in the center.

In the case of electrical power systems, the existing literature on vulnerability analysis largely takes a topological approach to identify the critical components in the network [Albert et al., 2004; Crucitti et al., 2005; Zio et al., 2008]. Such analyses are capable of identifying elements of structural vulnerability, i.e. network edges and nodes whose failure can induce a severe structural damage to the network through the physical disconnection of its parts. Such analysis is very fast from a computational point of view and only requires the information of the topology of the network.

On the contrary, this kind of analysis is limited by the fact that it focuses only on the topological features of the network, thus neglecting its physical characteristics. In this respect, it is important to verify the extent of these limitations and possibly overcome them by additional more detailed physical analyses on critical parts of the network [Eusgeld et al., 2009; Bier et al, 2007].

The use of betweenness centrality of groups of nodes as importance measures can be extended to consider the betweenness centrality of groups of edges and the variation in network connection efficiency [Latora and Marchiori, 2001] for identifying the critical groups. For power transmission systems this point of view is more realistic since transmission lines (edges) are more exposed to attacks than nodes (substations).

In most networks however, information or power flow does not flow only along geodesic paths (Sthephenson and Zelen, 1989; freeman et al, 1991), because the flow from one node of a network to another is typically a global phenomenon which does not depend only on the links on the direct and shortest paths, since it is quite possible that information will take a more circuitous route; this is true both in social networks, where information may travel by random communication or be intentionally channeled through intermediaries, and in network infrastructures, where flow is channeled through selected routes, following the specific operative rules and constraints which apply to the system; in particular for electric power network in which the flow obeys the Kirchhoff's laws. A realistic betweenness measure should include non-geodesic paths in addition to geodesic ones.
To address this issue a more sophisticated measure, the flow betweenness that includes contribution from non geodesic path has been proposed (Freeman, 1991).

From the modeling point of view, associated with each link (i, j) is a $K \times K$ matrix $\{\gamma_{ij}\}$, describing the capacities of the links (i, j). Additionally, a node may be a generating source (i.e., power source) or a load (i.e., power demand) or simply a junction node (neither generating nor demanding power).

The global efficiency of the graph representing the network is defined as [Latora and Marchiori, 2001]:

$$E = \frac{1}{N(N-1)} \sum_{i,j \in G, i \neq j} \frac{1}{d_{ij}} \qquad (4.5)$$

where $1/d_{ij}$ is the efficiency of the connection between nodes i and j in terms of the number of edges on the shortest path linking the two nodes. It relates the importance of an edge to the impact on the network transmission performance of losing to failure the edges of a group. The relative variation of the global efficiency due to the removal of a group of edges is computed as the difference between the global efficiency of the network with all the edges of the group removed and the global efficiency of the original network, normalized to the latter value. This value can be interpreted as a measure of importance of the group of edges removed [Crucitti et al., 2005]. As for the betweenness centrality, all the information required to evaluate this measure is contained in the adjacency matrix $\{a_{ij}\}$ of the network.

The outcome of the pure topological analysis that underline structural vulnerabilities from the unweighted perspective, is compared to the results of the vulnerability analysis complemented by functional information on the network. This comparison is exemplified with respect to the high-voltage electric transmission network, and it aims at validating the use of pure topological analysis as a preliminary screening tool that requires minimal information but it is capable of identifying major vulnerabilities and guiding in-depth functional vulnerability analysis.

The graph under study is modeled as a capacitated network, i.e., a network in which each edge (i, j) has a flow capacity γ_{ij} and each node is crossed by the flow carried by all its incoming edges (except when the node is a source, in which case flow is only outgoing, or a sink, in which case it is only incoming). The amount of flow being transmitted on an edge cannot exceed the capacity of the edge.

Given the generating sources and load nodes, the Ford-Fulkerson algorithm [Ford and Fulkerson, 1962] can be used to determine the maximum flow in the network, i.e., the largest possible total flow from sources to target nodes in the network, assuming that the flow at a node can be split among the edges in each node. The amount of flow through a node *i* when the maximum flow is transmitted from a source *(s)* to a to target *(t)*, averaged over all *s* and *t* is expressed by the flow betweenness measure (Freeman, 1991):

$$FC_i^B = \frac{\sum_{j=1}^{N}\sum_{k=j}^{N} m_{jk}(i)}{\sum_{j=1}^{N}\sum_{k=j} m_{jk}} \tag{4.6}$$

where m_{jk} be the maximum flow from a node i to a node k and let $m_{jk}(i)$ be the maximum flow from node j to node k that passes through node i. Each edge in the network can be thought of as a transmission line carrying a flow of current. In general, more than a single unit of current can be carried between source and target by making simultaneous use of several different paths through the network.

In practical terms, the flow betweenness measures the betweenness of nodes in a network in which a maximal amount of flow is continuously pumped between all sources and targets. Necessarily that flow still needs to "know" the ideal route (or one of the ideal routes) from each source to each target in order to realize the maximum flow. This still seems an unrealistic definition, in that it is often the case that flow does not take any sort of ideal path from source to target. To account for this, a more appropriate betweenness centrality measure, the random walk betweeness has been introduced (Newman, 2003):

$$RWC_i^B = \frac{\sum_{i=1,s<t}^{N} I_i^{st}}{\frac{1}{2}N(N-1)} \tag{4.7}$$

This measure is appropriate to a network in which information wanders about essentially at random until it finds its target, and it includes contributions from many paths that are not optimal in any sense. Note that this measure is more physically detailed but requires not only the information on the system network connection pattern, contained in the adjacency matrix, but also the edges capacities, contained in the capacity matrix, the values of the generating sources and loads, and an algorithm to evaluate the random walk of the flow through the network.

5. Simulation of electric power transmission networks: uncertainty analysis

5.1 Problem Statement

Modern society depends on electric power as an essential resource for communication, transportation, heating and cooling systems, lighting, and the powering of computers and electronics. Electric power systems are pervasive in our everyday's life, they are made of a large number of interconnected elements (wires and machines), which link the electricity generators to the customers, for satisfaction of their diverse needs.

Providing electricity in a reliable fashion is a complicated and technically challenging task. It involves real-time assessment, coordination, and control of thousands of generating units, the transfer of electric power over networks of transmission lines and, finally, the delivery of electric power to the consumers.

Originally developed as loosely interconnected networks of local systems, electric power grids have extended on large scales, across regional and national boundaries. The extent of the interconnectedness, the number and variety of power sources and generators, of controls and loads make electric power grids among the complexes engineered systems [Zio and Aven, 2011].

Another relevant complexity attribute relates to the increased integration of electrical power grid with other critical infrastructures, e. g. driven by the use of computer based communication and control systems, which is beneficial in many respect but on the other hand introduces additional vulnerabilities due to the increased interdependences. The re-conceptualization of the electric power grid to allow the integration of large shares of electricity produced by renewable energies at the most suitable sites calls for a system with decentralized generation and new devices for increased controllability, which will convert the existing power grid from static infrastructure to be operated as designed into a flexible, adaptive infrastructure [Zio and Aven, 2011].

Besides the mentioned technological challenges, a number of emerging issues are daunting the electric power grid systems and increasing the stress of the environment in which these are to be operated. These are:

- deregulation in electric power systems, which has resulted in the system being operated closer to their capacity and limits; the electricity market is becoming more competitive and reliability assessment is becoming increasingly important;
- the prospected demand for electricity in the next 20-50 years, which results in the need for technically respond by increasing capacity and efficiency;
- the sensed increase in exposure to malevolent attacks which calls for effective protection to different type of hazard and threats, much more difficult to predict than random failures.

In this scenario of increased stress on the electric power grids, concern naturally arise on the vulnerability and reliability of the transmission power grid and on the associated uncertainties.

5.2 Main sources of uncertainty

One of the main purposes of a power system is to satisfy the demands of customer loads in a reliable and economical manner. Failing to properly address planning problems and constraints will eventually yield operation problems and constraints, and, therefore, will affect power system reliability. For example, failing to incorporate uncertainties in system planning may lead to an overestimation of risk reduction barriers and of system capabilities to maintain acceptable levels of reliability.

The appropriate incorporation and presentation of the implications of uncertainty are widely recognized as fundamental components in the analyses of complex systems [Billington and Huang, 2008]. There are two fundamentally different forms of uncertainty in power system reliability assessment [Hoffman and Hammonds, 1994]. Aleatory uncertainty arises because the study system can potentially behave in many different ways. The component failure and repair processes are random and create variability known as aleatory uncertainty. Another type of uncertainty enters the system reliability assessment, due to the incomplete knowledge and information on the system and related phenomena which leads to imprecision in the model representation of the system and in the evaluation of its parameters. This type of uncertainty is often referred to as subjective, epistemic state-of-knowledge [Apostolakis, 1990]. In the field of power system research, the epistemic uncertainty has already been considered in the fuzzy power flow analysis [Matos and Gouveia, 2008; Gouveia and Matos, 2009] where the power injections of all loads and generations are regarded as fuzzy variables.

Uncertainty in demand, transmission and generation parameters, line ratings, extreme weather, and other environmental factors, introduce uncertainty in operation and planning of electric networks.

In general, the degree of uncertainty increases significantly from a shorter time frame in system planning to a longer time frame in system operation. Therefore, it is of paramount importance to identify and quantify these sources of uncertainty during the design phase of electric networks. Table I summarizes the uncertainties identified in the electrical transmission system. In this study, we represent and propagate the uncertainties related to (I) consumption variability, (II) ambient temperature variability, (III) wind speed variability and (IV) wind power generation.

Element	*Parameter*		*Source of uncertainty*	*Type of available information*	*Uncertainty representation*
Load bus	**Load value**		Consumption variability	Historical data	**Probabilistic** (Normal pdf)
Wind generating unit	Output power	**Wind speed**	Wind speed variability	Historical data	**Probabilistic** (Weibull pdf)
		Operation parameters	Wind power variability	Experimental data	**Probabilistic** (Normal pdf)
Weather	**Wind speed**		Wind speed variability	Historical data	**Probabilistic** (Weibull pdf)
	Ambient temperature		Temperature variability	Historical data	**Probabilistic** (Normal pdf)
Transmission Line	Line temperature		Material properties incomplete knowledge	Experts' judgment	Possibilistic
			Line diameter	Historical data	Probabilistic
			Line resistance	Historical data	Probabilistic

Table I. Uncertainties sources and their representation in the electrical transmission system.

When the uncertainty in the variables is mainly due to their inherent randomness (aleatory uncertainty) and there is sufficient information to assign probability distributions and estimate their parameters, probabilistic modeling is embraced. The model output is represented by a function of n random variables, $Y = f(X_1,...,X_i,...,X_n)$, where X_i denotes the i-th probabilistic input variable with PDF $p_{X_i}(x)$. The probabilistic model defines the probability distribution of the output random variable Y as a function of the probability distribution of the inputs. Such distribution can be found analytically in simple cases, or by MCS for more realistic setting.

In power system studies, the MCS is typically embraced, given the large number of variables involved and their complex relationships, which make analytical models difficult or even impossible

to derive [Billinton and Gao, 2008; Karki et al. 2010]. The operative procedure of MCS calls for a large number *m* of iterations: at each *e*-th iteration, an input vector of values $\left(x_1^s, x_2^s, ..., x_n^s\right)$ is sampled from the PDFs of the input variables and a realization of the output value y^e is computed solving the system model. After *m* repetitions, an empirical estimate of the distribution of the system output is obtained.

5.3 Uncertainty propagation

As seen in Chapter 4, the advantages of using abstract topological approaches for modeling the dynamics of CIs are that these models require small amount of data and have a relatively low complexity. However, given the somewhat abstract level of the modeling supporting complex network analysis, the insights gained with respect to the vulnerable areas in the system (first findings) may not be clear-cut, and additional vulnerabilities may still be expected, and more detailed information about the system and its operating environment needed. Then, for practical uses the analysis can serve for guiding more detailed approaches that include operational aspects related to the specific system. For example, for a power system it would be necessary to account for: the line impedances and Kirchhoff's laws in assessing the power flows from the generators to the loads, the fact that the line capacities are engineered to accommodate the usual patterns of line flows, the effects of automatic protection, power flow redistribution after transients and after re-dispatch of generation and shedding of load. In order to draw firm conclusions about the significance of topological methods, e.g. in electricity infrastructure vulnerability, current research focuses on systematically comparing the results from power-flow based vulnerability models with those from graph theoretic models of vulnerability [Hines et al., 2010; Rosato et al., 2009]. The general impression is that there is only a partial superposition of functional and structural criticalities. On the other hand, it is very difficult to model the full complexity of the network's dynamics as it requires a large number of long simulations. One way to approach this difficult problem is to include simple, but representative, models for each component of this large and complex system as is required in order to understand the global dynamics of power systems.

In order to study the effects of the uncertainties related to load forecast on one side, and uncertainties in the weather parameters on the other side, that introduce disturbances in the grid and induce line outages due to overloads, we developed a stochastic model that simulates the operations of the electric network. Our approach, inspired by the model introduced in [Dobson et al., 2001; Anghel et al., 2007 Giorsetto, 1983], combines: 1) a DC load flow algorithm that computes the distribution of power flows using a linear load flow approximation, 2) the contribution of wind generation power in a transmission power grid, 3) a strategy for generation dispatch in order to balance the power production and consumption throughout the network, 4) the dynamics of line temperatures as function of the

power flow and environmental conditions (wind speed and ambient temperature), 5) the event of automatic line disconnection when the rated line temperature is reached, and 6) the event of line reconnection.

In order to describe the evolution of cascading events in their slow initiating stages, transmission lines failures in our model are caused by line overheating due to excessive power flows. To describe this effect, we monitor the evolution of the line temperature, and its slow dependence on electric flow redistributions, using the model of heat conduction in rods of small cross section in which an electric current of constant intensity flows. Further contributing to the evolution of cascades is a line restoration model which prevents a damaged line to be put back in service before a fixed restoration time has passed.

The model of transmission line failure due to loading over their transmission capacity and following restoration, is part of the developed event-based stochastic framework which has also the ability to represent daily hourly changes in power requests at customer side of the system, room temperature and wind speed variations.

The stochastic framework is based on sequential Monte Carlo Simulation (MCS) in which the combination of load requests, room temperature, wind power generation, wind speed and network topology is a system realization. Due to the yearly periodicity of the load request and the room temperature average values, each year is considered to be statistically equivalent to one another and the results are provided on the basis of yearly averages. The simulation begins by establishing the load demand, the room temperature and wind speed values. If no line disconnection due to excessive heating occurs, the next event corresponds to the occurrence of the next hourly time step ("next hour" event) with updated load demand, room temperature and wind speed conditions. If the temperature of a line exceeds the critical temperature set for that line, the "line disconnection" event may occur before the scheduled "next hour" event. A DC load flow is performed following the occurrence of each event. The "line reconnection" event occurs after a time chosen a priori for each line that is disconnected.

After each event, we solve the DC load flow equations in order to determine the line temperatures and the type of the next event. The change time to the next event is computed as the minimum between the time to the next hour change, the minimum failure time among all lines, and the minimum time to reconnection of all lines.

We assume that the electrical transmission system operates in steady-state conditions and that this assumption holds even during the evolution of major disturbances in the system. This is obviously an approximation which is violated during the late stages of major disturbance events. It can be relaxed if voltage dependent phenomena are modeled during these events.

In order to determine the steady-state operating conditions of the power grid, we should solve the full nonlinear power flow equations that provide information about the voltage magnitudes and phases and the active and reactive power flows along each transmission line. Unfortunately, since our simulations involve numerous power flow solutions for a power grid system that evolves in time, solving repeatedly the full non-linear power flow equations becomes computationally prohibitive. Moreover, the full nonlinear equations pose very difficult nonlinear optimization problems. We have therefore chosen to linearize the power-flow equations and to solve instead the so-called DC power flow equations that connect the flow of real power to the voltage phases of the system's buses, which results in a completely linear, non iterative, power flow algorithm (Wood and Wollenberg, 1996).

Uncertainty representation of the power demand at load buses

The average hourly peak power demand follows the hourly load curve based on data from [Grigg et al. 1999]. The curve accounts for customer power need variations from day to night and from season to season. An example of a daily load peak curve, in different days and seasons, is given in Figure 5.1.

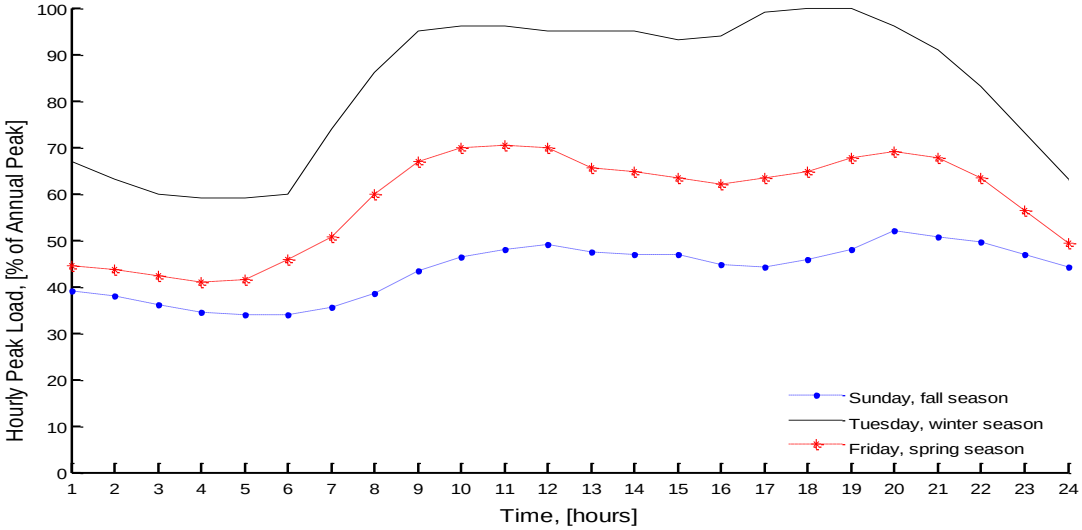


Figure 5.1. 24-hours load curve (Grigg et al. 1999). First hour corresponds to 12 a.m. – 1 a.m. interval of each day.

Uncertainties in the hourly peak power demand arise from because power consumption by users is not exactly uniform and simultaneous, i.e. we assume that the power needed at the customers side may experience stochastic fluctuation from the average hourly peak power demand. Following [Billinton and Li, 1994], it is assumed that load uncertainty is well described by a normal distribution. Therefore, the load hourly values are sampled from normal distribution $N(\mu,\sigma^2)$ with mean μ equal to the

hourly peak load considered in the deterministic case (Fig.3) and standard deviation σ assigned according to the perceived load forecast uncertainty, such as 10% of the mean value, $\sigma = 0.1\mu$ [Billinton and Li, 1994].

Uncertainty representation of the ambient temperature

In order to compute the annual ambient temperature curve (Fig. 5.2), the daily minimum and maximum values during one year in a specific location of the United States were collected and analyzed. A linear variation of the temperature values between the daily minimum and maximum values is assumed, with the minimum and maximum peak registered at 5 a.m. and 4 p.m., respectively.

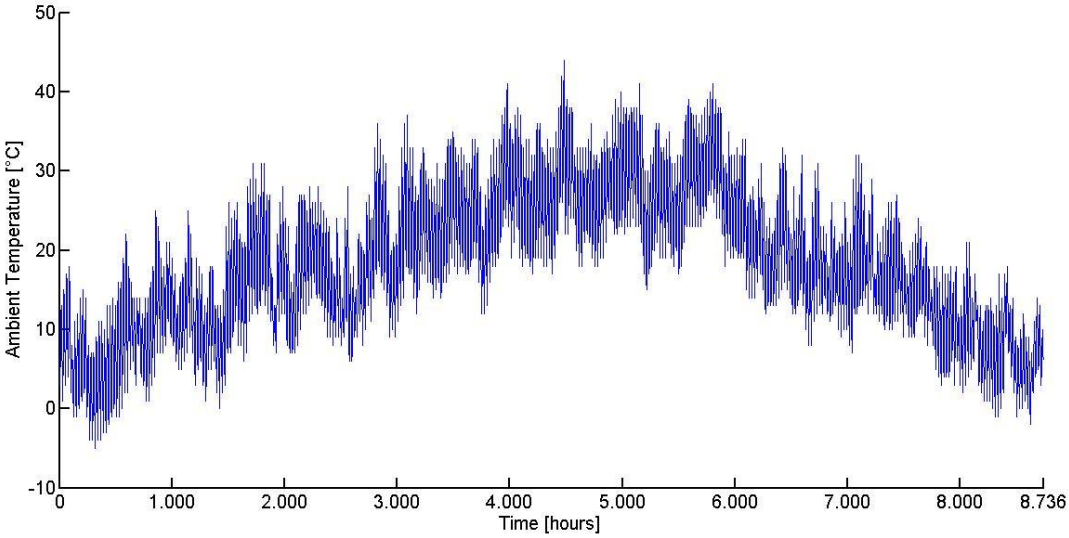


Figure 5.2. Ambient temperature curve used in the deterministic case. The maximum and minimum data values have been collected at the location of Beckersfield, CA, USA.

We assume that the uncertainty associated with the ambient temperature is well described by a normal distribution. Therefore, the temperature hourly values are sampled from normal distribution $N(\mu, \sigma^2)$ with mean μ equal to the hourly value from the annual ambient temperature curve (Fig. 5.2) considered in the deterministic case and standard deviation σ equal to 5% of the mean value, $\sigma = 0.05\mu$. This value has been identified by computing the standard deviations of the minimum of the median and of the maximum temperature values that are recorded within each month of the year by choosing the maximum among them.

Uncertainty representation of the wind speed

In order to compute the annual wind speed curve, the hourly values during a year in a specific location were collected and analyzed. In the deterministic case it is assumed, for each day of the year, that the wind speed is constant throughout the day and it is equal to the daily average value.

Following [Johnson, 1985], the Weibull distribution has been used to represent the wind speed randomness within a yearly time frame. It is shown that data collected at many locations around the world can be reasonably well described by the Weibull probability density function if the collection time frame is not too short, i.e. longer than several weeks. Figure 5.3 shows the distribution of the hourly wind speed collected at Bakersfield, CA, USA, and the Weibull distribution whose parameters are calculated by maximum likely estimation based on the hourly values. From the collected data, we notice that either the used anemometer has a lower bound of measuring at about roughly 6 ms^{-1}, or a wind speed below 6 ms^{-1} is an unlikely event at the considered location. Moreover, we notice that the Weibull approximations holds beyond the maximum of the distribution while it is a rough approximation for lower speed values. Therefore, we expected that the wind speed values sampled from the Weibull distribution will be biased towards lower values if compared to the collected data. Nonetheless, this bias does not affect the modeling framework that is the main objective of the work.

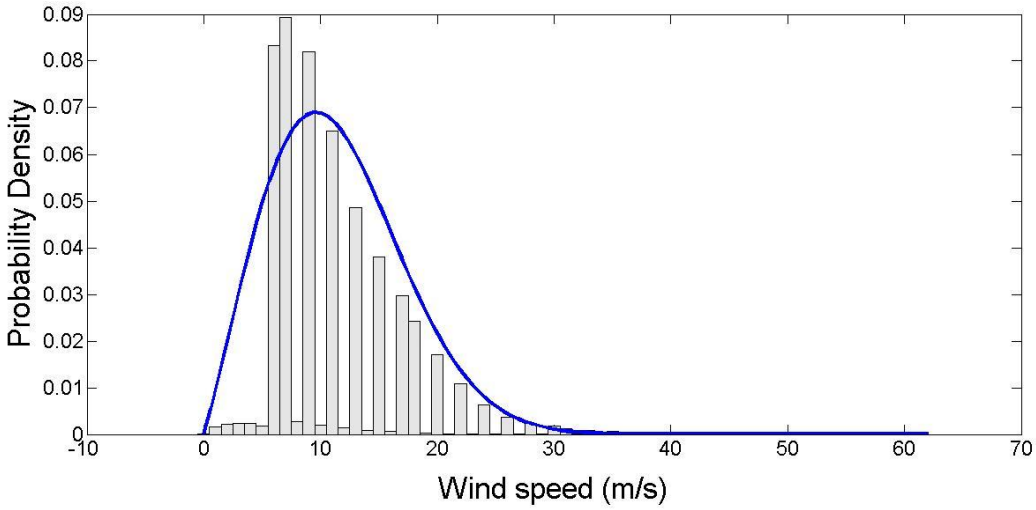


Figure 5.3. Distribution of the wind speed values collected at Bakersfield, CA, USA, and the corresponding Weibull distribution evaluated through maximum likelihood estimation of the distribution parameters.

Uncertainty representation of the wind power generation

Finally, the variability of wind speed propagates to the power output of wind generators. The power output of a WTG depends strongly on the wind regime as well as on the performance characteristics and the efficiency of the generator. A fundamental assumption is made when considering the deterministic power curve (figure 4): the relationship between the wind speed and the output power is fixed given the same type of WTG systems. In other words, the output power of the WTG is always the same at a specific wind speed. In reality, the output power for a fleet WTG of the same type always exhibits considerable variations even when they are operating at the same wind

speed [6]. Moreover, Thiringer and Linders [12] analyzed the relationship between the wind speed and the output power based on a group of wind turbines. They found that the powers generated from individual wind turbines of the same type actually vary even at the same wind speed. These research findings suggest that a probabilistic model incorporating the power variations may be more appropriate to characterize the relationship between the wind speed and the actual output powers. Following [Jin and Tian, 2010], the actual output power P_d is proposed as a random variable which is characterized by the mean power output and its standard deviation:

$$P_d(x) = P(x) + \varepsilon \qquad (5.1)$$

where $P_d(x)$ represents the actual WTG power output, $P(x)$ represents the deterministic output governed by the equation (5.1) and ε represents the variation of the power output with $\varepsilon \sim N(0, \sigma_\varepsilon^2)$. Following [Jin and Tian, 2010], we assume $\sigma_\varepsilon = 0.1 \cdot P_r$, i.e. 10% of the rated power output. Since we also considered the uncertainty in wind speed, the function for power curve actually contains two random parameters: the wind speed $x \sim Weibull$ and the variation of the power output $\varepsilon \sim N(0, \sigma_\varepsilon^2)$.

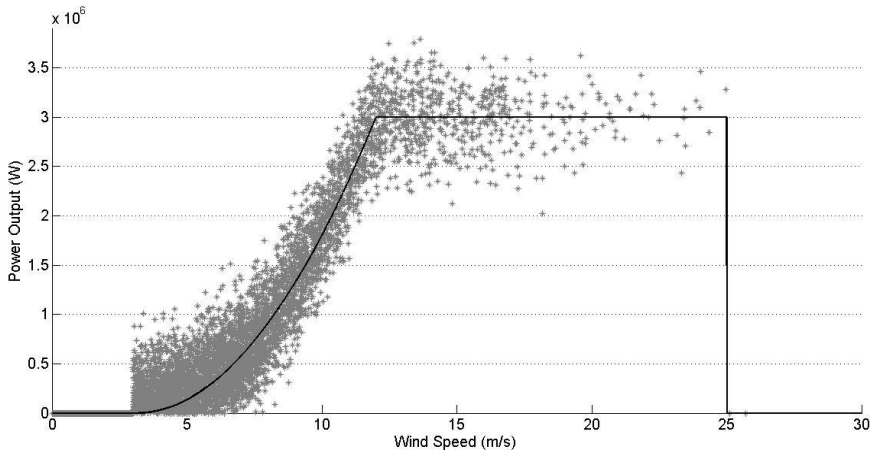


Figure 5.4. Power output realizations for the wind regime described by the Weibull distribution of figure 3 and the performance characteristic and efficiency of the generator.

Figure 5.4 shows the power output realizations (grey points) of the WTG of figure 2 for the wind regime described by the Weibull distribution of figure 5 and the performance characteristics and the efficiency of the generator described by $N \sim (0, \sigma_\varepsilon^2)$ [Jin and Tian, 2010]. Wind turbine starts producing power when wind speed equals the cut-in speed of 3 ms^{-1}: the majority of wind power generation concentrates during the nonlinear part of the output curve. This effect is consistent with the uncertainty in wind speed distribution.

The developed stochastic framework allow the identification and quantitative propagation of uncertainties related to consumption, wind power generation, ambient temperature and wind cooling of lines, aiming at assess their impact on a system adequacy assessment.

6. Conclusions

The present Ph.D. research work has dealt with the development of methods for critical infrastructure vulnerability analysis. The proposed analysis framework can be divided into three main steps:

- all-hazard analysis;
- topological network analysis;
- uncertainty analysis in electrical power transmission systems.

For the above developments, three approaches have been employed at different levels of detail: (I) all-hazard approach modeling, that provides a framework for the identification and quantification of hazard and threats, (II) abstract graph-theoretic modeling that allows analyzing structural properties and the extent to which the failure propagation process affects a network system when its physical and functional characteristics are abstracted, and (III) power flow modeling based on the physical laws of electric networks, for the analysis of power grids.

The all-hazard analysis is conceived as a general framework of analysis divided in two lines: the first line entails the qualitative assessment of system vulnerabilities by expert judgment and tabular methods while the second line entails the quantitative vulnerability assessment of a CI according to the system's components susceptibility to all hazards. In the qualitative assessment all the characteristic of the system's elements that could result as possible sources of vulnerability are identified and organized. From this standpoint, the concept of susceptibility is used as a paradigm to encompass the concepts of hazard, and its stochastic nature, and the concept of threat, which entail the idea of intentional harm. The quantitative assessment considers that threats have no established framework and define the relevant quantities and the approaches to account for intentional attacks. In the present PhD work, a fuzzy logic approach and a game theory approach are propound. These techniques allows to account for imprecise information and intentionality issues when dealing with malevolent attacks.

Topological structural analysis based on classical graph theory has proven able to unveil relevant properties of the structure of the network underlying a CI to i) highlight the role played by its components (nodes and connecting arcs), ii) make preliminary vulnerability assessments based on the simulation of faults (mainly represented by the removal of nodes and arcs) and the subsequent re-evaluation of the network topological properties. Yet, in spite of its usefulness and of the insights it provides, empirical results show that the topological analysis of the unweighted network of a CI can capture only partially the rich and complex properties observed in a real infrastructure system, so that there is a need for extending the models beyond pure unweighted, structural topology. In this view, the

formalism of weighed networks has been exploited to provide different graph-theoretical representations and analyses of a power transmission system. The aim is to contribute towards reducing the gap between the highly conceptualized (and abstract) analyses merely based on considerations of the system topology and the highly detailed (and computationally demanding) simulations of system behavior, in order to render the overall vulnerability assessment more feasible and robust. A further important dimension to add to the vulnerability characterization of CI refers to modeling the dynamics of flow of the physical quantities in real network systems. This has entailed considering the interplay between structural characteristics and dynamical aspects, which makes the modeling and analysis very complicated because the load and capacity of each component, and the flow through the network are often highly variable quantities both in space and time. From the abstract modeling of a cascading failure propagation process, it has been possible to identify the safe and critical failure-prone working conditions for a single CI. This indication can be used as a system vulnerability indicator for CIs.

In the final step of the vulnerability analysis of CIs, the characterization of uncertainties related to electric transmission networks has been undertaken, and the impact that the propagation of the identified uncertainties has on the reliability of the electric infrastructure has been quantified. Parameters whose uncertainties affect power system operation to large extent have to be paid a special attention during the design and the management of power systems. Failing to incorporate uncertainties in system planning may lead to an overestimation of risk reduction barriers and of system capabilities to maintain acceptable levels of reliability.

The results produced in this thesis work strengthen the belief that the adopted methods provide information useful for the vulnerability assessment of CIs, within a screening analysis of the CI behavior. Following this line of thought, a methodical framework for the vulnerability analysis of CIs can be devised which encompasses a screening analysis, carried out with the methods of the all-hazard analysis, complex network theory, uncertainty modeling, and a subsequent detailed analysis. The screening analysis can be supported by structural information provided by system owners and operators, including the general understanding of main functionalities, interfaces and interdependencies. The evaluation of the statistical indicators derived from the three approaches can highlight preliminary vulnerabilities, e.g. protection of components, and structural or reliability bottlenecks, which must be the focus of the following detailed system analysis.

The all-hazard analysis has proven essential for the screening of the vulnerability assessment of CIs, able to provide the identification of components of the network that may be more exposed to intentional attack. Moreover, the identification of the most vulnerable parts of a CI should always be complemented by the structural and dynamical analysis of the failure propagation process. Yet, due to the ‘abstract’ level of the modeling supporting the methods of complex network, the results gained

with respect to the vulnerable components in the system may not be ‘clear-cut’ and major hidden vulnerabilities may still be expected. If this would be the case, to better predict the failure behavior and clearly identify the most critical parts of the CI, system understanding has to be further developed and more detailed information about the system and its operating conditions may be needed, paying additional attention to interdependencies among several systems. The assessment of simplifications made in the early stage may call for more sophisticated methods of the successive detailed analysis. For practical uses the screening analysis can serve for guiding more detailed approaches that include operational aspects related to the specific system. For example, for a power system it would be necessary to account for: the line impedances and Kirchhoff's laws in assessing the power flows from the generators to the loads, the fact that the line capacities are engineered to accommodate the usual patterns of line flows, the effects of automatic protection, power flow redistribution after transients and after re-dispatch of generation and shedding of load.

In conclusion, the work developed in the Ph.D. thesis has substantiated that the specific added value of the application of complex network theory methods to the analysis of CIs is to fulfill two main objectives: helping a) to identify preliminary vulnerabilities of critical infrastructures by topology-driven and dynamical analysis and b) to guide and focus further detailed analyses of critical areas of the CIs.

References

[Albert et al., 2000] Albert, R., Jeong, H. and Barabási, A.-L., Error and attack tolerance of complex networks, *Nature*, Vol. 406, pp. 378-382, 2000.

[Amaral et al., 2000] Amaral, L. A. N., Scala, A., Barthélémy, M. and Stanley, H. E., Classes of small-world networks, *Proceedings of the National Academy of Sciences*, 97, 11149-11152, 2000.

[Anghel et al., 2007] Angel, M., Werley, A.-K. and Motter A. E., Stochastic Model for Power Grid Dynamics, *Proceedings of the 40th Hawaii International Conference on System Sciences*, 2007.

[Apostolakis, 1990] Apostolakis G. E., The concept of probability in safety assessments of technological systems. *Science*, 250(4986), 1359–1364, 1990.

[Apostolakis and Lemon, 2005] Apostolakis, E.-G. and Lemon, M.-D., A Screening Methodology for the Identification and Ranking of Infrastructure Vulnerabilities Due to Terrorism, *Risk Analysis*, Vol. 25, No. 2, 2005.

[Aron, 1966] Aron, R., *Press and War*, London, Weinfeld and Nicholson, p.170, 1966.

[Aven, 2007] Aven, T., A unified framework for risk and vulnerability analysis covering both safety and security, *Reliability Engineering and System Safety*, 92, pp. 745-754, 2007.

[Bier and Azaiez, 2011] Bier, V., and Azaiez, M.N., *Game Theoretic Risk Analysis of Security Threats*, ISBN-13: 978-0-387-87766-2, Springer NY, 2009.

[Bier et al, 2007] Bier, V.M., Gratz, E.R:, Haphuriwat, N.J., Magua, W., and Wierzbicki, K.R., Methodology for identifying near-optimal interdiction strategies for a power transmission system, *Reliab. Eng. Syst. Saf.*, vol. 92, no. 9 , pp. 1155–1161, Sep. 2007.

[Billington and Li, 1994] Billinton, R. and Li, W., *Reliability Assessment of Electric Power Systems Using Monte Carlo Methods*, New York: Plenum Press, pp.229-308, 1994.

[Billington and Huang, 2008] Billington, R., and Huang D. Aleatory and Epistemic Uncertainty Considerations in Power System Reliability Evaluation, Probabilistic Methods Applied to Power Systems. PMAPS '08. *Proceedings of the 10th International Conference on*, 25-29 May 2008.

[Billinton and Gao, 2008] Billinton, R., and Gao, Y., Adequacy assessment of composite power generation and transmission systems with wind energy, *Interantional Journal of Reliability and Safety*, vol. 2, no. ½, pp. 79-98, 2008.

[Blaikie and al., 1994] Blaikie, P., Cannon, T., Davis, I., and Wisner, B., *At risk: natural hazards, People's vulnerability and disasters*, Routledge, London, 1994.

[Boccaletti et al., 2006] Boccaletti, S., Latora, V., Moreno, Y., Chavez, M. and Hwang, D.-U., Complex Networks: structure and Dynamics, *Physics Reports*, 424, pp. 175-308, 2006.

[Bouchon, 2006] Bouchon, S., The vulnerability of interdependent critical infrastructure systems: epistemological and conceptual state-of-the art, EUR 22205 EN, 2006.

[Brown and Cox, 2011] Brown, G.C., and Cox, L.A. Jr, How probabilistic risk assessment can mislead terrorism risk analysis, *Risk Analysis*, 31, pp. 196-204, 2011.

[Cadini et al., 2009] Cadini, F., Zio, E. and Petrescu, C.-A., Using Centrality Measures to Rank the Importance of the Components of a complex Network Infrastructure, in *Critical Information Infrastructure Security*, Proceedings of the 3rd International Workshop on Critical Information Infrastructures Security, CRITIS 2008, Rome, Italy, October 13-15, 2008, pp. 155-167, 2009.

[Casals and Solé, 2011] Casals, M. R. and Solé, R. V., Analysis of major failures in Europe's power grid, *International Journal of Electrical Power & Energy Systems*, 2011.

[Chambers and al., 1989] Chambers, R., Pacey, A., Thrupp, L., *Farmer first*, Intermediate Tecnologt Publications, London, 1989.

[(COM (2004) 702 final)] *Critical Infrastructure Protection in the fight against terrorism*, 702 final Communication from the Commission of the Council and the European Parliament, 2004.

[Cox, 1999] Cox, E., *The Fuzzy System Handbook*, Academic Press, 1999.

[Crucitti et al., 2006] Crucitti, P., Latora, V. and Porta, S., Centrality in networks of urban streets, *Chaos*, 16, 015113, 2006.

[Debon et al., 2010] Debon, A., Carrion, A., Cabrera, E., and Solano, H., Comparing risk of failure models in water supply networks using ROC curves, *Reliability Engineering and System Safety*, vol. 95, pp. 43–48, 2010.

[Dekker, 2005] Dekker A. H., Simulating Network Robustness for Critical Infrastructure Networks, *Conferences in Research and Practice in Information Technology*, Proceedings of the 28th Australasian Computer Science Conference, The University of Newcastle, Newcastle, Australia, Vol. 38, V. Estivill-Castro, Ed., 2005.

[D'Inverno and Luck, 2004] D'Inverno, M. and Luck, M., *Understanding Agent Systems*, Springer, Berlin, 2004.

[Dobson et al., 2005] Dobson, I., Carreras, B.A. and Newman, D.E., A loading-dependent model of probabilistic cascading failure, *Probability in the Engineering and Informational Science*, **19**, pp. 15-32, 2005.

[Dobson et al., 2001] Carreras Lynch Newmann An initial model for complex dynamics in electric power systems blackouts, *Hawaii international Conference on System Sciences*, January, 3-6, Maui, Hawaii, 2001.

[Dow, 1992] Dow, K., Exploring differences in our common future(s): the meaning of vulnerability to global environmental change, *Geoforum*, 23, (3), pp. 417-436, 1992.

[Duenas-Osorio and Vemuru, 2009] Dueňas-Osorio, L. and Vemuru ,S.-M.,Cascading failures in complex infrastructure systems, *Structural Safety*, Vol.31, pp. 157-167, 2009.

[Duenas-Osorio et al., 2007] Dueňas-Osorio, L., Craig, I.J., Goodno, J.B. and Bostrom, A., Interdependent Response of Networked Systems, *J. Infrastruct. Syst.*, vol. 13, Issue 3, pp. 185-194, 2007.

[ECSS, 2004] ECSS European Cooperation for space standardization ECSS P-001B, 14 july 2004.

[Eusgeld] Eusgeld I., Kröger W., Sansavini G., Schläpfer M. and Zio E., The role of network theory and object-oriented modeling within a framework for the vulnerability analysis of critical infrastructures, *Reliability Engineering & Systems Safety*, Vol. 94, No 5, pp. 954-963, 2009.

[Everett and Borgatti, 1999] Everett, M.G., Borgatti, S.P., The centrality of groups and classes, *Journal of Mathematical Sociology*, Vol. 23, N. 3, pp. 181-201, 1999.

[Flammini et al., 2009] Flammini, F., Gaglione, A., Mazzocca, N. and Pragliola, C., Quantitative Security Risk Assessment and Management for Railway Transportation Infrastructures, *Critical Information Infrastructure Security*, vol. 5508/2009, 1611-3349, 2009.

[Ford and Fulkerson, 1962] Ford, L. and Fulkerson, D., *Flows in Networks*, Princeton University Press, Princeton, NJ, 1962.

[Freeman, 1979] Freeman, L. C., Centrality in social networks conceptual clarification, *Social Networks*, 1(3), 215-239, 1979.

[Freeman et al, 1991] Freeman, L.C., Borgatti, S.P., and White, D.R., Centrality in valued graphs: A measure of betweenness based on network flow, *Social Networks* 13(2): 141-154,1991.

[Gouveia and Matos, 2009] Gouveia E. M., Matos M. A., Symmetric ac fuzzy power flow model. *European Journal of Operational Research*, 197(3), 1012–1018, 2009.

[Guckenheimer and Ottino, 2008] Guckenheimer, J. and Ottino, J. M., *Foundations for Complex Systems Research in the Physical Sciences and Engineering*, report from NSF Workshop, Cornell University, September 2008.

[Haimes and Horowitz, 2004] Haimes, Y.Y., and Horowitz, B.M., Modeling interdependent infrastructures for sustainable counterterrorism, *Journal of Infrastructure Systems*, 8, 67 -75, 2004.

[Hines et al., 2010] Hines P., Cotilla-Sanchez E., Blumsack S. Do topological models provide good information about electricity infrastructure vulnerability? *Chaos* 20 (3), 033122 (2010).

[Hines and Blumsack, 2008] Hines, P. and Blumsack, S., A centrality measure for electrical networks, *Proceedings of the 41st Hawaii International Conference on system Science*, 2008.

[Hoffman and Hammonds, 1994] Hoffman, F.O., and Hammonds, J.S., Propagation of uncertainty in risk assessments: the need to distinguish between uncertainty due to lack of knowledge and uncertainty due to variability, *Risk Analysis*, 14, pp. 707-712, 1994.

[ISO guide, 2009] ISO guide 73:2009. Risk management – Vocabulary. ISO Concept database. https://cdb.iso.org/.

[Iyer et al., 1999] Ivey M., Akhil A., Robinson D., Stamber K., Stamp J., Consortium for Electric Reliability Technology Solutions Grid of the Future White Paper on Accommodating Uncertainty in Planning and Operations, Prepared for the Transmission Reliability Program Office of Power Technologies Assistant Secretary for Energy Efficiency and Renewable Energy U.S. Department of Energy, 1999.

[Ivey et al., 2009] Iyer, M. S., Nakayama, K. M. and Gerbessiotis, V. A., A Markovian Dependability Model with Cascading Failures, *IEEE Transactions on Computers*, vol. 58, no. 9, 2009.

[Jeong et al., 2001] Jeong, H., Mason S. P., Barabasi A.-L. and Oltvai Z. N., Lethality and centrality in protein networks, *Nature*, 411, 41-42, 2001.

[Johansson and Jonsson, 2009] Johansson, J., and Jonsson, H., A model for vulnerability analysis of interdependent infrastructure networks, *Safety, Reliability and Risk Analysis: Theory, Methods and Applications* – Martorell et al. (eds), 2009.

[Jönsson et al. 2008] Jönsson, J., Hassel, H., and Tehler, H., Identifying critical components in technical infrastructure networks, *Proceeding of the Institution of Mechanical Engineers, part O*, vol. 222, *Journal of Risk and reliability*, 2008.

[Koonce et al., 2008] Koonce, A. M., Apostolakis, G. E. and Cook, B. K., Bulk power risk analysis: Ranking infrastructure elements according to their risk significance, *Int J Electr Power Energ Syst*, vol. 30, pp. 169–183, 2008.

[Kröger (2008)] Kröger, W., Critical Infrastructures at risk: A need for a new conceptual approach and extended analytical tools, *Reliability Engineering and System Safety*, 93, 1781-1787, 2008.

[Kröger and Zio, 2011] Kröger, W., Zio, E., *Vulnerable systems*, Springer-Verlag London Limited 2011.

[Karki et al. 2010] Karki, R., Billinton R., Incorporating wind power in generating system reliability evaluation, *Int Journal of system assurance Engi. Manag*, 2, pp. 120-128, 2010.

[Latora and Marchiori, 2005] Latora, V. and Marchoiri, M., Vulnerability and protection of infrastructure networks, *Physical Review E* 71, 015103 (1-4), 2005.

[Latora and Marchiori, 2001]] Latora, V. and Marchiori, M., Efficient Behavior of Small-World Networks, *Physical Review Letters*, vol. 87, 19, pp. 198701(1-4), 2001.

[Latora and Marchiori, 2007] Latora, V., and Marchiori, M., A Measure of Centrality Based on the Network Efficiency, *New Journal of Physics*, 9, 188, 2007.

[Little, 2002] Little, R.G., Controlling Cascading Failure: Understanding the Vulnerabilities of Interconnected Infrastructures. *Journal of Urban Technology* 9(1): 109 – 123, 2002.

[Lord et al., 2005] Lord, D., Washington, P.S. and Ivan N.J., Poisson, Poisson-gamma and zero-inflated regression models of motor vehicle crashes: balancing statistical fit and theory, *Accident Analysis and Prevention*, vol. 37, pp. 35–46, 2005.

[Matos and Gouveia, 2008] Matos, M.A., Gouveia, E.M., The Fuzzy Power Flow Revisited. *IEEE Transactions on Power Systems*, 23(1), 213 – 218, 2008.

[Moore, 2006] Moore, A.D., Application of the API/NPRA SVA methodology to transportation security issues, *Journal of Hazardous Materials*, vol. 130, pp. 107–121, 2006.

[Morgan et al., 2000] Morgan, M., G., Florig, H., K., DeKay, M., L., Fischbeck, P., Categorizing Risks for Risk Ranking, *Risk Analysis*, Vol. 20, No.1, 2000.

[Motter and Lai, 2002] Motter, A.-E. and Lai, Y.-C., Cascade-based attacks on complex networks, *Physical Review E* 66, pp.065102(1-4), 2002.

[Motter, 2004] Motter, A.-E., cascade control and defense in complex Networks, *Physical Review Letters*, vol. 93, nr 9, pp. 098701(1-4), 2004.

[NERC, 2009] NERC North American Electric Reliability Corporation, *Special Report: Accommodating High Levels of Variable Generation*, April 2009.

[Newmann et al., 2005] Newman, D.-E., Nkei, B., Carreras, B.A., Dobson, I., Lynch, V.E. and Gradney, P., Risk Assessment in Complex Interacting Infrastructure Systems, *Proceedings of the 38th Hawaii International Conference on System Sciences*, 2005.

[Niemen, 1974] Nieminen, J., On the centrality in a graph, *Scandinavian Journal of Psychology*, 15(1), pp. 332-336, 1974.

[OHS, 2009] Office of Homeland Security, National Infrastructures Protection Plan, 2009.

[Patterson and Apostolakis, 2007] Patterson, S. Apostolakis, G., Identification of critical locations across multiple infrastructures for terrorist actions, *Reliability Engineering & System Safety*, 92(9), pp. 1183-1203, 2007.

[PCCIP Report, 1997] *Critical Foundations - Protecting America's Infrastructures*, The Report of the President's Commission on Critical Infrastructure Protection, October 1997.

[Pederson et al. (2006)] Pederson, P., Dudenhoeffer, D., Hartley, S., Permann, M., Critical Infrastructure Interdependency modeling: A Survey of U.S. and International Research, INL/EXT – 06-11464, 2006.

[Piwowar et al., 2009] Piwowar, J., Chatelet, E. and Laclemence, P., An efficient process to reduce infrastructure vulnerabilities facing malevolence, *Reliability Engineering and System Safety*, vol. 94, pp. 1869–1877, 2009.

[Pollet and Cummins, 2009] Pollet J., and Cummins, J., "All-Hazard approach for Assessing Readiness of Critical Infrastructure", *IEEE Conference on Technologies for Homeland Security*, 2009.

[Rinaldi (2004)] Rinaldi, S.M., Modeling and simulating critical infrastructures and their interdependencies, *Proceeding of the 37th Hawaii International Conference on System Science*, 2004

[Rinaldi et al., 2001] Rinaldi, S.M., Peerenboom, J.P. and Kelly, T.K, Identifying, understanding and analyzing critical infrastructures interdependencies, *IEEE Control System Magazine*, 21(6), 11-25, 2001.

[Rosato et al., 2007] Rosato, V., Bologna, S. and Tiriticco, F., Topological properties of high-voltage electrical transmission networks, *Electric Power System Research*, vol. 77, pp. 99-105, 2007.

[Rosato et al., 2009] Rosato V., Issacharoff L., Gianese G., Bologna S., Influence of the topology on the power flux of the Italian high-voltage electrical network, arXiv.org, physics, arXiv:0909.1664 2009.

[Sabidussi, 1966] Sabidussi, G., The centrality index of graphs, *Psychometrika*, 31(4), pp. 581-603

[Schlapfer et al. 2008] Schläpfer, M., Kessler, T. and Kröger, W., Reliability Analysis of Electric Power Systems Using an Object-oriented Hybrid Modeling Approach, *Proceedings of the 16th Power Systems Computation Conference*, Glasgow, 2008.

[Sthephenson and Zelen, 1989] Sthephenson, K.A., and Zelen, M., Rethinking centrality: Methods and examples, *Social Networks*, 11, 1–37, 1989.

[Strogatz, 2001] Strogatz, S. H., Exploring complex networks, *Nature*, Vol. 410, pp. 268-276, 2001.

[Sultana and Chen, 2009] Sultana, S. and Chen Z., Modeling flood induced interdependencies among hydroelectricity generating infrastructures, *Journal of Environmental Management*, vol. 90, pp. 3272–3282, 2009.

[Susman and al., 1983] Susman, P., O'Keefe, P., Wisner, B., Global disaster, a radical interpretation, *Interpretations of calamity*, Allen &Unwin, London, pp.263-283, 1983.

[Turner et al, 2003] B.L. Turner et al., A framework for vulnerability analysis in sustainability science, *PNAS*, vol. 100, no.14, 2003.

[UN/ISDR 2004] United Nations International Strategy for Disaster Reduction, *Living with Risk. A Global Review of Disaster Reduction Initiatives* – 2004 version, ISDR, Geneva, 2004.

[Volkanovski et al., 2009] Volkanovski, A., Cepin, M. and Mavko, B., Application of the fault tree analysis for assessment of power system reliability, *Reliability Engineering & System Safety*, Vol. 94 (6), pp. 1116-1127, 2009.

[Wasserman and Faust, 1994] Wasserman, S., and Faust, K., *Social Networks Analysis*, Cambridge U.P., Cambridge, 1994.

[Watts, 2002] Watts J.D., A simple model of global cascades on random networks, *Proceedings of the National Academy of Sciences*, vol. 99, no. 9., pp. 5766-5771, 2002.

[Watts and Strogatz, 1998] Watts, D.-J. and Strogatz, S.H., Collective dynamics of 'small-world' networks, *Nature*, Vol. 393, pp. 440-442, 1998.

[Waugh, 2004] Waugh, W., Terrorism and the All-Hazards Model, *IDS Emergency Management On-Line Conference*, June 28-July 16, 2004.

[White 1974] White G.F., *Natural hazards: local, national and global*. Oxford University Press, New York, p. 288, 1974.

[Wood and Wollenberg, 1996] A. J. Wood and B. F. Wollenberg, *Power Generation, Operation, and Control*, 2nd ed. John Wiley & Sons, New York, 1996.

[Yamijala et al., 2009] Yamijala, S., Guikema, D.S. and Brumbelow, K., Statistical models for the analysis of water distribution system pipe break data, *Reliability Engineering and System Safety*, vol. 94, pp. 282–293, 2009.

[Zadeh, 1965] Zadeh, L., Fuzzy Sets, *Information and Control*, vol. 8, 3, pp. 338-353, 1965.

[Zimmermann, 2001] Zimmerman, R., Social Implications of Infrastructure Network Interactions, *Journal of Urban Technology*, Volume 8, Number 3, pages 97-119, 2001.

[Zio et al., 2008] Zio, E., Sansavini, G., Maja, R. and Marchionni, G., An analytical approach to the safety of road networks, *International Journal of Reliability, Quality and Safety Engineering*, Vol. 15 Issue: 1, Page: 67 - 76 February 2008

[Zio and Sansavini, 2009]] Zio, E. and Sansavini, G., Modeling failure cascades in networks systems due to distributed random disturbances and targeted intentional attacks, in *Safety, Reliability and Risk Analysis: Theory, Methods and Applications* – Martorell et al. (eds), Proceedings of ESREL 2008 and 17th SRA Europe Annual Conference, 22-25 September 2008, Valencia, Spain, Taylor & Francis Group, London, 2009.

[Zio and Sansavini, 2010a]] Zio, E. and Sansavini, G., Component criticality in failure cascade processes of network systems, *Risk Analysis*, 31(8), 1196-1210, 2011.

[Zio and Sansavini, 2010b] Zio, E. and Sansavini, G., Modeling interdependent network systems for identifying cascade-safe operating margins, *IEEE Transactions on Reliability*, 60(1), 94-101, 2011.

[Zio and Aven, 2011] Zio, E., and Aven, T., Uncertainties in smart grids behavior and modeling: what are the risks and vulnerabilities? How to analyze them?, *Energy Policy*, 39, pp. 6308-6320, 2011.

[Zio and Golea, 2012] Zio, E., Golea, L.R., Analyzing the topological, electrical and reliability characteristics of a power transmission system for identifying its critical elements, *Reliability Engineering and System Safety*, Volume 101, Pages 67-74, May 2012.

Part II

Paper I

An All-Hazard Approach for the Vulnerability Analysis of Critical Infrastructures

Enrico Zio, R. Piccinelli, Giovanni Sansavini

Advances in Safety, Reliability and Risk Management, Bérenguer, Grall & Guedes Soares (eds)

An All-Hazard Approach for the Vulnerability Analysis of Critical Infrastructures

Enrico Zio*[a], Roberta Piccinelli[b], Giovanni Sansavini[b]

[a] Chair on Systems Science and the Energetic challenge, European Foundation for New Energy-Electricite' de France Ecole Centrale Paris and Supelec, Grande Voie des Vignes, F92-295, Chatenay Malabry Cedex

[b]Politecnico di Milano, Energy Department, Nuclear Section, c/o Cesnef, via Ponzio 33/A , 20133, Milan, Italy

enrico.zio@ecp.fr

Abstract

In this paper, a framework is proposed for the *All-HAZard ANalysis* (A-HAZAN) of Critical Infrastructures (CIs). Starting from the identification of the task of each component in the infrastructure, we use tabular procedures to organize the information on the susceptibility to attacks, to single and cascading failures. All variables and states are identified that may impact on the component's role as a possible source of vulnerability within the CI and towards interdependent CIs. This is a starting point for a quantitative evaluation of the degree of exposure to intentional acts. A case study of literature is taken as an exemplary demonstration of the procedures of the analysis.

1. Introduction

Critical infrastructures (CIs) like the electricity, oil & gas supply, rail, road, air, sea transport, internet networks, are highly interconnected and mutually dependent in complex ways, both physically and through a multitude of information and communication technologies (so called cyber-based systems) used for data acquisition and control. The coupling of CIs leads to the concept of "systems-of-systems", which implies that single CIs cannot be studied in isolation from other CIs; rather, it is necessary to assess the limitations that interacting CIs pose on the operating conditions of the individual infrastructures so as to implement adequate protections for preventing failures in one CI from cascading to other dependent CIs.

The 2001 prolonged power crisis in California demonstrates the importance of coupling in interdependent CIs (Rinaldi et al., 2001). The crisis took place when electric power disruptions at various times curtailed natural gas production (first order effects); the latter generated a shutdown of steam injection in heavy oil production (second order effects).

A common denominator is needed to assess the vulnerability of a system that is exposed to natural and accidental hazards, and threats of malevolent acts. The need is to capture the CI vulnerability sources

and issues, given its technical and physical features, and the dependencies and interdependencies on other systems. This requires an evaluation of the exposure to different hazards, including threats of malevolent acts.

An analysis aimed at identifying the causes of damage or disruption of services in CIs needs to embrace an *all-hazard approach* (Waugh, 2004), (Pollet and Cummins, 2009), encompassing a general view on the hazards targeting a given system.

We propose a framework for an *All-HAZard ANalysis* (A-HAZAN) which relies on tailored tabular procedures to organize the qualitative and quantitative features of the system relevant for revealing and highlighting its vulnerabilities. The A-HAZAN framework is intended as a tool for managers, analysts and stakeholders of CIs to carry out the identification of all the sources of vulnerability in an all-hazard perspective.

The paper is organized as follows. In Section 2, some methodologies to assess the vulnerabilities of CIs are reviewed. In Section 3, the concepts of vulnerability and susceptibility are outlined from the perspective of CIs. In Section 4, the A-HAZAN tabular procedure and a methodology are presented. In Section 5, the A-HAZAN methodology is applied to a literature case study. Conclusions are drawn in Section 6.

2. Methodologies of vulnerability assessment

Screening methodologies for prioritizing scenarios of terrorist threats and identifying vulnerabilities in single and interdependent CIs have been proposed. Apostolakis and Lemon (Apostolakis and Lemon, 2005) and Patterson and Apostolakis (Patterson and Apostolakis, 2007) focus on the identification of critical locations in infrastructures; these are seen as geographical points that are exposed to intentional attacks. Critical locations are not limited to individual infrastructures but may affect multiple infrastructures: for example, water and electrical distribution systems may occupy the same service tunnels. In the proposed scheme, the conditional probabilities that the terrorists will successfully exploit a vulnerability need to be evaluated. The procedure for this relies on extensive use of expert judgment and may be challenging in practice. The impacts of attacks are treated without including the probabilities of various levels of damage, and without consideration of any intervention by first responders: for this reason, it can be defined conservative. The vulnerabilities and their ranking according to potential impact are eventually obtained by Multi-Attribute Utility Theory (MAUT) (Morgan et al. 2000).

Konce et al. (Konce et al., 2008) have proposed a methodology for ranking components of a bulk power system with respect to its risk significance to the involved stakeholders; the likelihood and the extent of power outages when components fail to perform their designed functions are analyzed; the consequences associated with the failures are determined by considering the type and number of customers affected.

Johansson and Hassel (Johansson and H. Hassel, 2010) have proposed a framework for considering structural and functional properties of interdependent systems and developing a predictive model in a vulnerability analysis context. Piwowar et al. (Piwowar et al., 2009) have proposed a systemic analysis which accounts for malevolence, i.e., the willingness to cause damage.

The aim of the present work is to develop a systematic framework of system analysis for identifying the vulnerable elements of CIs, considering natural hazards, random failures and intentional attacks. While the first two types of vulnerability are characterized by stochastic uncertainties and can be analyzed by traditional safety analysis tools, intentional attacks require a new way of analysis.

3. Hazards, threats and vulnerability

In the United Nations' view, hazard is defined as "a potentially damaging physical event, phenomenon and/or human activity, which may cause loss of life or injury, property damage, social and economic disruption or environmental degradation. Hazards can be single, sequential or combined in their origin and effects" (Turner et al., 2003). The European Cooperation for Space Standardization (ECSS) defines hazard as "an existing or potential condition of an item that can result in an accident" (White, 1974); the condition is associated with the design, fabrication, operation or environment of the item, and has the potential for accidents. These two definitions encompass the idea of hazard described as a "condition prerequisite to a mishap" (UNISDR, 2004) and as "a source of potential harm" (ECSS, 2004). The concept of hazard is strictly tied to the presence of a potential source of difficulty both natural or manmade.

On the other hand, the concept of threat is defined as "a potential intent to cause harm or damage to the system by adversely changing its state" (ISO guide, 2009). This definition is strictly linked to intentional and malevolent acts (Apostolakis and Lemon, 2005), and it seems in contrast with the one by the US Homeland Security, which describes threat as a "natural or manmade occurrence, individual, entity, or action that has or indicates the potential to harm life, information, operations, the environment, and/or property" (OHS, 2009). In the latter, little reference is made to the idea of intention embedded in the former definition of threat. From these definitions, the broader concept of hazard emerges as a general condition of potential source of harm. Therefore, it encompasses the intentionality of threats, e.g., terrorist acts that are distinguished by a malevolent intelligence directed toward maximum social disruption.

In the all-hazard approach, malevolent acts, accidental and natural occurrences are all considered. Yet, they require a different analytical treatment. Random accidents, natural failures and unintentional man-made hazards are typically known and categorized by emergency planners. Their occurrence can be typically addressed within a probabilistic framework (Figure 1). Conversely, terrorism is a hazard that eludes a quantification by probability theory due to the intentional and malevolent planning it implies (Figure 1).

Figure 1. All–Hazard Approach overview.

The concept of vulnerability follows the degree of impact that an hazard has on the CI. In (Konce et al., 2008), vulnerability is defined as "the degree to which a system, a subsystem or a system component is likely to experience harm due to exposure to a hazard, either a perturbation or stress" or, equivalently, in (Apostolakis and Lemon, 2005) as the "manifestation of inherent states of the system (e.g., physical, technical, organizational, cultural) that can be exploited by an adversary to harm or damage the system". Along the same line of thought, vulnerability is also defined as a physical feature or operational attribute that renders an entity open to exploitation or susceptible to a given hazard (OHS, 2009). The United Nations define vulnerability as the conditions determined by physical, social, economic, and environmental factors or processes, which increase the susceptibility of a community to the impact of hazards (UNISDR, 2004).

Quantitatively, vulnerability focuses on three aspects (Konce et al., 2008):

- degree of losses and damages due to the impact of a hazard;
- degree of exposure to the hazards, i.e., likelihood of being exposed to hazards of a certain degree and susceptibility of an element at the risk of suffering losses and damages;
- degree of resilience, i.e., ability of a system to anticipate, cope with/absorb, resist and recover from the impact of a hazard or disaster.

Practically, vulnerability comes from flaws or weaknesses in the design, implementation, operation and management of an infrastructure that makes it susceptible to destruction or incapacitation when exposed to a hazard or threat, or reduces its capacity to resume stable conditions. For example, the vulnerability of the electric power system might be quantified in terms of changes in network

characteristics following attacks on nodes, and the scale (e.g., number of nodes/lines lost) or duration of the associated losses. In a somewhat more intuitive interpretation, vulnerability characterizes a system component or an aspect of a system (Jönsson et al., 2008). A component is said to be a vulnerability of a system if its failure causes large negative consequences. In this sense, the component is said to be critical and the term vulnerability describes a property which can be employed for ranking the system's components with respect to their criticality.

Given the above, the goals of vulnerability analysis becomes then (Kroger and Zio, 2011):

1. identifying the set and sequences of events that cause damages and losses;
2. identifying the relevant set of "initiating events" and evaluate their cascading impact on a subset of elements, or the system as a whole;
3. determining and elaborating on (inter)-dependencies (within the system and among systems) and on coupling of different orders, given the set of initiating events and observed outcomes.

The ultimate goal is to identify obvious and, most important, hidden vulnerabilities in infrastructure systems to act for managing and reducing them. The achievement of these goals relies on the analysis of the system, of its parts and of their interactions. The analysis must account for the environment where the system operates, and for the objectives the system is designed to achieve. During the development of such basic system understanding, first vulnerabilities may already emerge.

In this paper, an *All-HAZard ANalysis* (A-HAZAN) framework is proposed to grasp the complementary aspects of random failures, or unintentional or natural occurrences, and malevolent attacks (Figure 2).

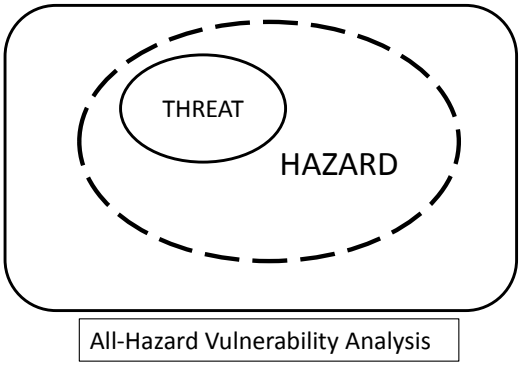


Figure 2. All-hazard vulnerability analysis including the assessment of random, natural and unintentional occurrences, as well as threats of intentional, malevolent acts.

4. The A–HAZAN analysis

We propose a tabular methodology for the all-hazard vulnerability analysis of CIs. It aims at identifying the features, operating conditions and failure modes relevant to CI vulnerability. A tabular

procedure is developed to collect and organize the relevant information to the vulnerability characterization of the system's components (Table 1).

Task (Funtion)	Component	Features	Geographic Location	Hazards	
USER	Load Bus	Size	Area	THREATS	
		Level of demand	Bus Name	RANDOM FAILURES	
		Environmental conditions		INTRA-System	Degree
					Centralities
					Communities
		Maintenance		INTER-System	Physical
		Level of protection			Logical
		Social criticality			Cyber
					Geographic
				EXTERNAL CAUSES	Diffuse
				Natural hazards Human activity	Local

Table 1. The A-HAZAN Table.

Starting from their functional role, the components of the system are broadly divided into three main categories: namely, *user*, *transmitter* and *provider*. A *user* is the target or the recipient of a particular task or service, e.g., a load bus in the power transmission network (or the consumer in a water supply system). A *transmitter* functions as a spreader of the task or service, e.g., the transmission lines in the power grid (or the pipelines in the water supply system). A *provider* is the component which originates that particular task or service, e.g., the generating units in the power transmission network (or the waterworks in the water supply system).

For each of the components, the relevant features that impact on its role as source of vulnerability are listed. The *size* is the physical dimension (if the component is large, it may need large protections). The *level of supply, of transmission or of demand* are the quantities of power supplied, transmitted or consumed by the component, and the fraction of the service produced, transmitted or consumed by the

component with respect to the overall power produced, transmitted or demanded by the whole system. The *level of protection* refers to the logical and physical barriers deployed to prevent or discourage malevolent acts. The *social criticality* anticipates the impact on public opinion of the effects of the intentional attack, given that it is successfully accomplished. The most relevant consequences here considered are measured in terms of human losses and geographic extension. Other features are i.e., permanent outages and transient outages, the effect of environmental conditions and temperature, or maintenance operations.

Other critical aspects of a system's component are its logical position and geographic location (Patterson and Apostolakis, 2007). It is important to know the position of the component with respect to the system and to the environment, and the connections and interconnections between the component and other systems.

Other specific features related to the physics of the provided service are accounted for in this qualitative analysis step. Examples are given in Section 5 with reference to a power transmission grid.

Then, vulnerability characterization in the all-hazard approach considers the following (Table 1):

- *Threats.* These are potential events characterized by the act of a malevolent intelligence directed towards maximum social disruption (Section 3).
- *Random failures.* These are typically permanent or transient outages due to components' failures and may be identified by standard risk analysis techniques (e.g., FMECA, Hazop, and others).
- *INTRA-system failures.* These are failures within the system, typically dependent failures.
- *External causes.* This term considers natural hazards, such as meteorological or seismic phenomena, and unintentional human-induced hazards, such as the processing or the storage of potentially hazardous materials, or nearby military installations. External causes may be further grouped into *local* external causes, e.g., a lightening striking a building or an aircraft crash, or *diffuse* external causes, such as earthquakes, hurricanes, flooding or the leakage of explosive or toxic materials.

5. A-HAZAN of the IEEE RTS-96

The IEEE 1996 reliability test system (RTS-96) (IEEE RTS-96, 1999) is considered (Figure 3). It contains 24 buses and 38 transmission lines. The buses consist of nine load-only buses, eight load/generation buses, three generation-only buses and four transmission buses (no load or generation on the bus). The test system does not refer to any particular geographic location; however, in the following, we suppose and characterize the locations of its components in order to contextualize the all-hazard framework.

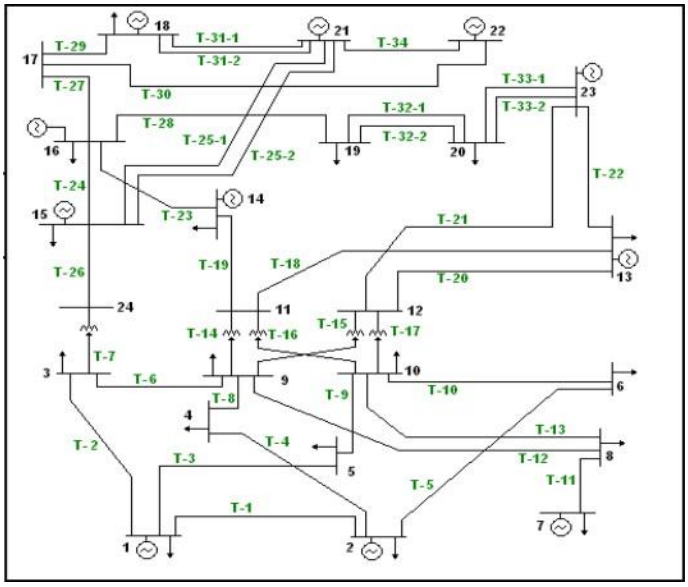


Figure 3. Single area IEEE RTS-96 grid (IEEE RTS-96, 1999).

Five different types of generating units are considered: oil/steam, coal/steam, oil/combustion turbine (CT), hydro and nuclear generating units. In Table 2, the A-HAZAN Table of a generating unit, i.e., the nuclear power plant sited at bus (ID)18, is exemplified. It is supposed to be located in a plain, near a river or a lake, possibly in a non-seismic area, far from urbanized areas and highways. In general, it can be supposed that the power plant does not lie under commercial air routes. This is a typical location for European nuclear power plants.

In the second column, the features of the power plant are summarized. Along with the generated power (size), i.e., 400 MW of active power and 137 MVAR of reactive power, and the percentage of the overall generated power, i.e., 12% of the entire power generated by the grid (level of supply), the type of generating unit, i.e., nuclear plant, and the total number of units of that type in the system, i.e., 1, are listed. A nuclear power plant has a high social criticality: the impact of an accident in the plant is high, both on public opinion and on public health. Due to its intrinsic dangerousness, a nuclear power plant is provided with maximum security measures and is classified as completely secure. The identification of environmental conditions is more suitable for a component than for an entire plant. For an entire plant, environmental conditions are replaced by the age of the infrastructure, or the characterizing ambiance: brackish air or degree of humidity. Maintenance accounts for all the joint operations regarding fuel supply and waste disposal, as well as components maintenance, e.g., turbines, thermal exchangers, or steam generators, as well as maintenance of the buildings and of the non-operational part of the plant, e.g., offices, air filtering systems.

In the third column, the specific area is detailed. In a general view, it could be assumed as a flat area.

Following the definitions given in Section 4, in the fifth column, threats are grouped under the label “sabotage” and “terroristic attacks”, where sabotage is meant as a deliberate action intended to ‘‘damage, disrupt, or subvert the organization’s operations for the personal purposes of the saboteur by

creating unfavorable publicity, embarrassment, delays in production, damage to property, the destruction of working relationships, or the harming of employees or customers'' (Crino, 1994). For example, in a workplace setting, sabotage is the conscious withdrawal of efficiency generally directed at causing some change in workplace conditions. On the other hand, terroristic attacks are actions intended to cause a strong psychological effect by means of disruption and death.

Task (Funtion)	**Compnent**		**Features**			**Geographic Location**	**Hazards**	
		ID		**MW**	**MVAR**			
USER	Load Bus	101	Size:	108	22	Area: 1 (zone11) Bus Name: Abel	INTENTIONAL ATTACKS Sabotage Terroristic attacks	
			Level of demand: 3,8%				RANDOM FAILURES from PRA/FMEA/FMECA	
			Level of protection: locked, non-complex barriers, fences				INTRA-System: lines outage : T-1, T-2, T-3	Degree
								Communities
								Centralities
			Social criticality: moderate				INTER-System	Physical: transport network (rail, road and air)
			Environmental conditions: aging and degradation of the plant, humidity, brackish ambience, temperature					
			Maintenance					Logical: economics, political
								Cyber: Scada System/ Power for switches
								Geographic: co-location
							EXTERNAL CAUSES Natural Hazards Human Activity	Diffuse: earthquakes, flooding, tornadoes, storms Local: stroke of lightning, landslides, snow, military maneuvers, aircraft crashes

Table 2. Qualitative part of the A-HAZAN Table for a nuclear generating unit in the RTS-96 (IEEE RTS-96, 1999).

Random failures are identified via FMEA/FMECA analysis on the power plant. Considering the function of the plant as a provider element, random failure rates are given in (IEEE RTS-96, 1999).

Intra-system features account for the connection of the component to other elements of the system, for example, lines T-29, T-31-1, T-31-2 outages which would prevent the power to flow from the provider to the users along the involved path.

Inter-system features encompass all interdependencies between the system and other infrastructures. We consider the possible interdependencies between the power plant and other infrastructure systems: water and fuel supply, transport network or the interdependencies between the communication system and the provider.

Natural hazards and human activities can be found under the definition of external causes. This is strictly tied with the location of the component and of the system in general. Natural hazards can beroughly divided as follows: seismic phenomena (i.e., earthquakes, landslides, etc.), volcano eruptions, meteorological phenomena (frequent meteorological events, i.e., wind, precipitations, snow pack, temperature, etc. and rare meteorological phenomena such as: tornadoes, storms, etc.) and flooding (i.e., hazards resulting from flooding of river, sea, lakes and semi-enclosed water bodies, groundwater, etc.). In Table 6, considering a frat terrain as the site, specific natural hazards are: flooding, storm and earthquakes.

The identification of potential sources of hazard due to unintentional human activities in the proximate areas are considered: processing or storing of potentially hazardous materials (such as explosive, flammable, corrosive, toxic or radioactive materials), aircraft crashes, railway rolling stocks and road traffic, military installations and future human activities in the planning stage, e.g., lands with potential for commercial development.

The external causes can be specified further in local and diffuse hazards, depending on the affected portion of the system. An example of local natural hazard is a stroke lightening on a transmission line. Diffuse hazards are mudslides due to heavy rains or flooding of rivers or lakes, for power plants and transmission towers.

In order to proceed in the analysis, a small portion of the RTS-96 (IEEE RTS-96, 1999) has been selected (Figure 4). It consists of a user (load bus (ID)101), a provider (generating unit located at bus (ID)1), a transmission line (line T-2) and a transmission line with a transformer (line T-7).

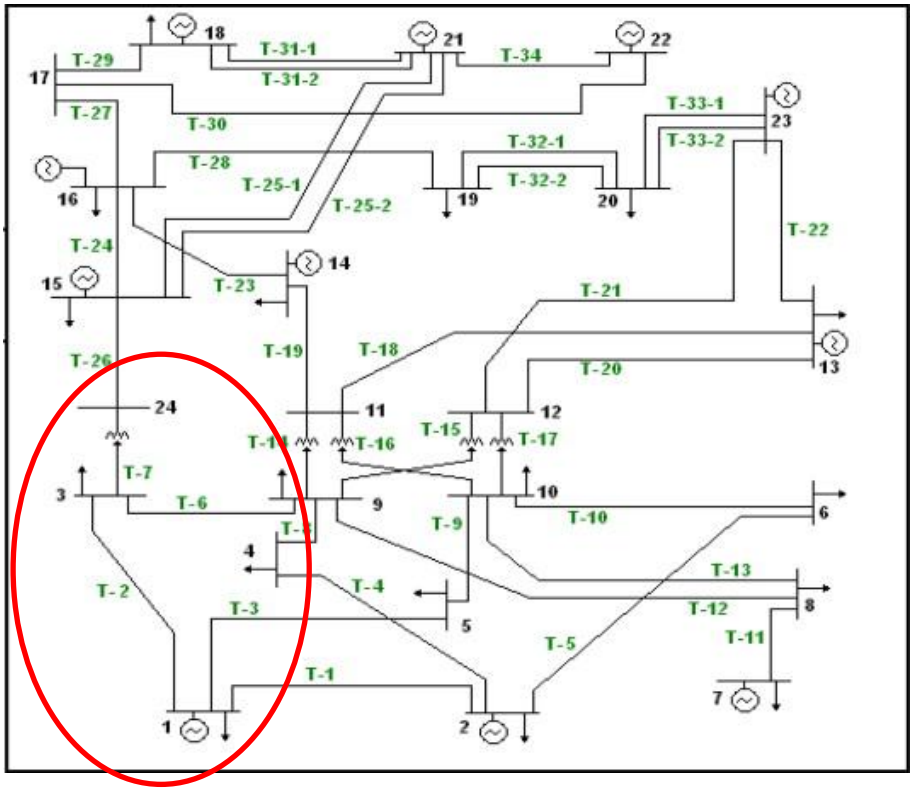


Figure 4. Portion of the RTS-96 (IEEE RTS-96, 1999) power grid described in Section 5 including bus (ID)101, generating unit (ID)1 and transmission lines T-2 and T-7.

The A-HAZAN Tables describing the characteristics of these components are here not reported, due to limitation of space. However, the salient aspects are described in the following.

In the A-HAZAN Table for the load bus, several features are highlighted, i.e., the absolute quantity of power requested (size), i.e., 108 MW of active power and 22 MVAR of reactive power needed by the bus (ID)101, and the percentage of power flow required by the particular bus with respect to the entire transmission network power requirements (level of demand), i.e., 3.8% of the total 2850 MW requested by the network, (IEEE RTS-96, 1999). Depending on the type of load bus, the level of protection is taken into account: a transmission grid load bus is typically located far from densely-populated areas and it is usually locked, surrounded by fences but no other complex barriers. The impact of the component role is assumed moderate, since it is expected that if a load bus cannot receive power, the generation of power is easily modulated, the power excess is eliminated, and the overall infrastructure still provides its service. The presence of special costumers on the load bus should be verified, e.g., hospitals, airports, energy-intensive factories. Special attention is devoted to the Environmental conditions, other than natural hazards, characterizing the component in a network. For example, the aging or the degradation of the components, due to specific conditions of the location site, i.e., the humidity of the air or the brackish ambience may cause corrosion or damage to the buildings. The temperature range is also an important issue to account for. An exceptionally hot summer or an extremely cold winter impose additional strain on the component.

In the Table, a description of the hazards is also given. Some of the items have been discussed in Section 4, and in reference to Table 2. For this particular component, the intra-system hazards are referred to failures that may occur when the connections (T-1, T-2 and T-3) between the component and the network are damaged.

A specific A-HAZAN contains the description of the generating unit sited at bus (ID)1. Additional features are, along with the generated power (size) and the percentage of generated power in the grid (level of supply), the type of generating unit and the number of units of the particular type of provider. The level of protection is also considered: the plant is isolated from the urbanized areas and it is usually guarded with security patrols, video surveillance of the entire power plant and alarms. A provider is assumed to have a high level of criticality, because in general its entire power supply cannot be readily replaced by alternative generation. Environmental conditions should also be taken into account; age and degradation affect the power plant (or its constituents) full functionality.

For this particular component, the intra-system hazards are referred to failures that may occur when the connections (T-1, T-2 and T-3) between the plant and the network are damaged.

Two tables are used to report the features pertaining to components referred to as transmitters because they are not the recipients of the electrical service but perform the propagation of the service. Transmitters include transmission lines and transformers. The highlighted physical parameters are: the direction of the lines, from bus 101 to bus 103, their capacities, their lengths, 55 miles, and the electrical characteristics of the transmitter, i.e., resistance ($R = 0.055\ \Omega$), reactance ($X = 0.211\ \Omega$) and susceptance ($B = 0.057$ S). Transmission towers are usually located in isolated sites, e.g., open country and they are not provided with any particular fence or barriers. Nor are they watched by patrol. Transmission lines environmental issues are salt pollution depositing on insulators on overhead lines and on substations, or floods and fires adjacent to electrical equipment, e.g., beneath overhead lines. The pruning of trees sited along transmission lines is also crucial: for example, the Italian 2003 blackout was triggered and accrued by two consecutive flashovers towards a tree of two overhead lines (Sforna and Delfanti, 2006).

Particular features of the transformer that connects bus 103 to bus 124: a zero miles length line, a very small resistance, $R=0.002\ \Omega$, a reactance $X= 0.084\ \Omega$ and no susceptance, $B = 0.000$ S.

6. Conclusions

A practical all-hazard analysis framework has been presented for merging two different perspectives on vulnerability of critical infrastructures. On one hand, it captures the vulnerabilities due to random failures and natural hazard; on the other hand, it includes vulnerabilities due to malevolent acts. In this sense, it extends common approaches of system hazard identification.

A general organization of the relevant information on the system components is offered on the basis of their tasks and of the features that characterize them as potential sources of vulnerability.

For the characterization, inter- and intra-dependencies are considered. The A-HAZAN framework is intended as a tool for managers, analysts and stakeholders of CIs to carry out the identification of the sources of vulnerability in an all-hazard perspective. It can serve as an entry point into the quantitative evaluation of the degree of exposure of CIs to hazards of different nature.

The future step of the analysis will be the development of a decision logic framework for evaluating the susceptibility to the hazards that loom over a CI, i.e., random failures, unintentional acts and natural hazards, but also malevolent acts.

References

(Apostolakis and Lemon, 2005) G. E. Apostolakis and D.M. Lemon, A screening methodology for the identification and ranking of infrastructure vulnerabilities due to terrorism, *Risk Analysis*, vol 25, no2, 2005.

(Crino, 1994) M. D. Crino, Employee Sabotage: A Random or Preventable Phenomenon?, *Journal of Managerial Issues*, Vol. 6, pp. 311-330, 1994.

(DHS, 2011) http:// www.dhs.gov/ xabout/ laws/ gc_1214508631313. shtm

(ECSS, 2004) ECSS European Cooperation for space standardization ECSS P-001B, 14 july 2004.

(Glass et al., 2003) Glass, R.J., Beyeler, W.E., Conrad, S.H., Brodsky, N.S., Kaplan, P.G., Brown, T.J., "Defining research and development directions for modeling and simulation of complex, interdependent adaptive infrastructures", *SNL paper SAND 2003-1778P*.

(IEEE RTS-96, 1999) Reliability test system task force of the application of probability methods subcommittee. The IEEE reliability test system – 1996. IEEE Trans Power Syst 1999; 14; 1010 – 20.

(ISO guide, 2009) ISO guide 73:2009. Risk management – Vocabulary. ISO Concept database. https://cdb.iso.org/

(Johansson and H. Hassel, 2010) J. Johansson and H. Hassel, An approach for modeling interdependent infrastructures in the context of vulnerability analysis, *Reliability Engineering and System Safety*, 95, 2010.

(Jönsson et al., 2008) J.,Jönsson, J., Johansson and H., Johansson, Identifying critical components in technical infrastructure networks, Proc. IMech E Vol. 222 part O: *J. Risk and Reliability*, 2008.

(Konce et al., 2008) A.M. Konce et al. [2008] A.M. Konce, Bulk power risk analysis: Ranking infrastructure elements according to their risk significance, *Electrical Power and Energy Systems*, 30, 2008.

(Kroger and Zio, 2011) W. Kroeger and E.Zio, *Vulnerable Systems*, Springer, 2011.

(MIL STD, 1993) MIL STD – 882C Military standard system safety program requirements, 19 January 1993.

(Morgan et al. 2000) Morgan, M., G., Florig, H., K., DeKay, M., L., Fischbeck, P., " Categorizing Risks for Risk Ranking", *Risk Analysis*, Vol. 20, No.1, 2000.

(OHS, 2009) Office of Homeland Security, National Infrastructures Protection Plan, 2009.

(Patterson and Apostolakis, 2007) S.A. Patterson and G. E. Apostolakis, Identification of critical locations across multiple infrastructures for terrorist actions, *Reliability Engineering and System Safety*, 92, 2007.

(Piwowar et al., 2009) J. Piwowar et al., An efficient process to reduce infrastructure vulnerabilities facing malevolence, *Reliability Engineering and System Safety*, 94, 2009.

(Pollet and Cummins, 2009) Pollet J. and Cummins, J., "All-Hazard approach for Assessing Readiness of Critical Infrastructure", *IEEE Conference on Technologies for Homeland Security*, 2009.

(Sforna and Delfanti, 2006) Sforna, M. and Delfanti, M., Overview of the events and causes of the 2003 Italian blackout, Proceedings of the PSCE '06. - Power Systems Conference and Exposition, Oct. 29 - Nov. 1, 2006, Atlanta, GA, IEEE Power Engineering Society, 2006.

(Turner et al, 2003) B.L. Turner et al., A framework for vulnerability analysis in sustainability science, *PNAS*, vol. 100, no.14, 2003.

(UNISDR, 2004) United Nations International Strategy for Disaster Reduction, *Living with Risk. A Global Review of Disaster Reduction Initiatives* – 2004 version, ISDR, Geneva, 2004.

(Waugh, 2004) Waugh, W. L., "Terrorism and the All-Hazard Model", *IDS Emergency Management On-Line Conference*, June 28-July 16, 2004.

(White, 1974) G.F. White, *Natural hazards: local, national, global*, new York: Oxford University Press, 1974.

Paper II

A Framework for Ranking the Attack Susceptibility of Components of Critical Infrastructures

Enrico Zio, R. Piccinelli, Giovanni Sansavini

Chemical Engineering Transactions, vol. 26, 309-314, 2012.

A Framework for Ranking the Attack Susceptibility of Components of Critical Infrastructures

Enrico Zio*[a], Roberta Piccinelli[b], Giovanni Sansavini[b]

[a] Chair on Systems Science and the Energetic challenge, European Foundation for New Energy-Electricite' de France Ecole Centrale Paris and Supelec, Grande Voie des Vignes, F92-295, Chatenay Malabry Cedex

[b]Politecnico di Milano, Energy Department, Nuclear Section, c/o Cesnef, via Ponzio 33/A , 20133, Milan, Italy

enrico.zio@ecp.fr

Abstarct

Outages or destruction of Critical Infrastructures (CIs) cause disruption of fundamental services for Society's well being and result in diverse consequences with economical and social implications. An analysis of all the occurrences that can cause such undesired events is necessary. This calls for an *all-hazard approach* capable of dealing with diverse events and conditions such as deterioration and failures, natural disasters, accidents and malevolent acts. In this paper, an *All-HAZard ANalysis* (A-HAZAN) framework is proposed. Starting from the identification of the task of each component in the infrastructure, we use a Table to organize the information on the susceptibility to attacks, and to single and cascading failures. We qualitatively assess the susceptibility to attacks as a function of size, level of demand, level of protection and social criticality. Then, we give a methodology to rank the components with respect to their susceptibility to attacks. The systematic process of analysis is here presented by way of its application to a case study of literature.

1. Introduction

Systems providing energy (electricity, oil & gas supply), transportation (rail, road, air, sea), information and telecommunication (e.g. the internet) are all large-scale, man-made, networked systems, called *infrastructures*. They are named *critical* as any incapacity or destruction can have a debilitating impact on Society's health, safety, security, economics and well being (Kröger and Zio, 2011), so determining their degree of exposure to hazards and intentional attacks has become a topic of great concern.

An *all-hazard approach* is required for identifying the causes of damage or disruption of services in CIs (Waugh, 2004), (Pollet and Cummins, 2009). In particular, the approach must handle also malevolent acts, which differ from natural or other man–made hazards and lacks of a well-established methodology for uncertainty assessment.

The all-hazard approach is intended to provide the basis for addressing unexpected events of any nature. The need is to capture the CI vulnerability sources and issues, given technical and physical features of the system, and the dependencies and interdependencies on other systems. This requires an evaluation of the susceptibility to different hazards, including threats of malevolent acts. Whereas the

susceptibility to hazards leading to random failures can be quantified in terms of probability, the susceptibility to intentional malevolent acts lacks a well-established framework of evaluation. We propose a framework for an *All-HAZard ANalysis* (A-HAZAN) which relies on tailored tabular procedures to organize the qualitative and quantitative features of the system relevant for revealing and highlighting its vulnerabilities. The A-HAZAN framework is intended as a tool for managers, analysts and stakeholders of CIs to carry out the identification of all the sources of vulnerability in an all-hazard perspective. In particular, a methodology to assess the susceptibility to intentional attacks is introduced. The paper is organized as follows. In Section 2, a framework of the A-HAZAN analysis is presented. In section 3, a methodology for assessing the susceptibility to attacks is explained. In Section 4, the A-HAZAN methodology is applied to a literature case study. Conclusions are drawn in Section 5.

2. Framework of analysis

The proposed tabular framework for the all-hazard vulnerability analysis is divided into two steps.

A preliminary qualitative step aims at identifying the relevant features, operating conditions and failure modes. This step highlights all the variables and states that impact on the component's role as a possible source of vulnerability and has been introduced in (Zio et al., 2011).

The second step aims at evaluating the components susceptibility to the various hazards: the need is to provide a measure of the likelihood that the specific hazard results in a mishap to the components and the CI. This evaluation requires different treatments and tools suitable to the nature of the specific hazard and the related uncertainties. The uncertainties associated with random failures and natural hazards can be treated within a probabilistic framework, while the uncertainties associated with threats might require an alternative treatment. With the framework, we aim at structuring the identification of the variables and states of the components that affect and contribute to their susceptibility, at pointing out the sources of uncertainties and the most suitable tools to treat them. For the assessment of threats due to malevolent acts, a qualitative evaluation of susceptibility is given.

3. Methodology

In the evaluation of the susceptibility to hazards of the system and its components, four categories of possible hazards have been identified: namely, attacks, random single failures, cascading failures induced by initiating failure within the CI, and failures induced by the coupling with other CIs.

- The susceptibility to *attacks* is characterized in terms of *attractivity* and *accessibility*. The term *attractivity* considers the appeal to intentional attacks, while the term *accessibility* considers that components have been designed for efficiency and convenience, yet access must be easy for maintenance staff but difficult for attacks.

- The susceptibility to *random failures* refers to the random failures of the components, i.e., permanent outages and transient outages, the effect of environmental conditions and temperature, or random failures due to maintenance operations.
- The susceptibility to *cascading failures* considers how the components are related in the network, i.e. the connectivity/topology of the network through which an initiating failure may propagate.
- The susceptibility to hazards that originate from *interdependent systems* whose input from external systems is required for the safe operations of the component.

The susceptibility to attacks is influenced by two variables; namely attractivity and accessibility. In turn, attractivity and accessibility are composed by four elements: size, level of demand, level of protection and level of social criticality:

- *size* (physical) of the component;
- *level of demand, transmission or supply*: the importance of the task of the component is measured both as the consumed, transmitted or supplied power of the component and as the fraction of the overall consumed, transmitted or produced power with respect to the whole system. From the point of view of accessibility, a component can be considered to be relevant in so far as it consumes, transmits or provides to the functioning of the system; for example, if the component is a generating unit and it provides a consistent amount of power to the network, then it should be accessible for maintenance or repairs;
- *level of protection*: the logical and physical barriers deployed to prevent or discourage malevolent acts;
- *social criticality:* given that the attack will be successfully accomplished, the impact on public opinion is influenced by the (conditional) effects caused by the achieved intentional act. The most relevant consequences here considered are in terms of human lives and geographic extension of the event.

We assume that the more protected a component, because of the security measures, the less accessible it is, but, the most attractive it is perceived, from a malevolent act point of view. This assumption relies on the idea that a malevolent will think that if a component deserves a good level of protection, then its damage would produce a large disruption. For example, nuclear power plants are strongly defended from the point of view of security; in this sense, they could be challenging and "attractive" to attacks. The combination of the above four elements yields different values of susceptibilities that are used in sorting the vulnerabilities of different components. Unfortunately, data and information concerning malevolent acts are incomplete, imprecise, ambiguous. To account for this, one may embrace various uncertainty representation approaches (Marseguerra et al., 2004).

The variables upon which vulnerability depends may be described as follows:

- x_1, size: the generating units are measured by their production of megawatts. They are grouped into small plants (power production ≤ 100 MW), medium plants (100 MW ≤ power production ≤ 300

MW) and large plants (power production ≥ 300 MW). Since the physical dimensions of a power plant are proportional to the produced power, in this frame the size of the power plant and the level of the produced power are joined under the *size* label;

- x_2, level of protection: as proposed in (Koonce and Apostolakis, 2008), six levels of protection may be considered:

Table 1: Levels of protections of infrastructure elements (Koonce and Apostolakis, 2008)

Level	Description
6 – Extreme (E)	Completely secure
4 – Moderate (M)	Secure area
3 – Low (L)	Complex barriers, security patrols, video surveillance
2 – Very low (VL)	Unlocked, non-complex barriers
1 – Zero (Z)	Completely open, no control, no barriers

- x_3, social criticality: here, the psychological impact on Society is considered, and attention is focused on human losses and the geographical spread of the event.

Table 2: Levels of social criticality of infrastructure elements

Level	Description
4 – Severe (S)	Many victims, widely spread extension
3 – High (H)	Many victims, contained extension
2 – Moderate (M)	Few victims, widely spread extension
1 – Low (L)	Few victims or no victims, contained extension

The variables x_1 and x_3 influencing accessibility can be described as for attractivity; the variable x_2, level of protection, acts on accessibility in the opposite way as for attractivity. In other words, the level of accessibility of a component is in inverse proportion to the adopted security measures: the weaker the security measures, the more accessible, and therefore more prone to attacks is the component.

Table 3: Levels of accessibility (y_1) and attractivity (y_2)

Level	Description
Y_5 – Extreme	Extreme degree of attractivity or extreme ease of accessibility
Y_4 – High	Elevated level of attractivity or high ease of accessibility
Y_3 – Moderate	Average degree of attractivity or medium ease of accessibility
Y_2 – Low	Low level of attractivity or low ease of accessibility
Y_1 – Very low	Almost null degree of attractivity or very low ease of accessibility

Then, the six-level scheme of protection of (Koonce and Apostolakis, 2008) is proposed in reverse order, level 0 applying to completely secure and level 6 to completely open, no barriers.

The accessibility and attractivity variables (y_1, y_2) are characterized by five levels, as shown in Table 3.

Evaluation can be done by combining in an IF–THEN decision logic the levels of (x_1, x_2, x_3) for attractivity and (x_1, x_2, x_3) for accessibility. For example, the logic rules could be:

$$\text{IF } x_1 \text{ is } X_{1k} \text{ AND IF } x_2 \text{ is } X_{2m} \text{ AND IF } x_3 \text{ is } X_{3n} \text{ THEN } y_i \text{ is } Y_1, \qquad 3 \le k+m+n \le 5 \qquad (1)$$

$$\text{IF } x_1 \text{ is } X_{1k} \text{ AND IF } x_2 \text{ is } X_{2m} \text{ AND IF } x_3 \text{ is } X_{3n} \text{ THEN } y_i \text{ is } Y_5, \qquad 12 \le k+m+n \le 13 \qquad (2)$$

$$\text{IF } x_1 \text{ is } X_{1k} \text{ AND IF } x_2 \text{ is } X_{2m} \text{ AND IF } x_3 \text{ is } X_{3n} \text{ THEN } y_i \text{ is } \begin{cases} Y_2 \text{ if } k+m+n = 6 \text{ or } 7 \\ Y_3 \text{ if } k+m+n = 8 \text{ or } 9 \\ Y_4 \text{ if } k+m+n = 10 \text{ or } 1 \end{cases}, \quad 6 \le k+m+n \le 11 \qquad (3)$$

where the index k can vary in the range [1,2,3], m in the range [1,2,3,4,5,6] and n in the range [1,2,3,4]. The rules are set considering all the input variable levels as subsets labeled in linguistic terms as shown in Tables 1, 2 and 3. Each variable x_i is attributed an indexed function X_i representing the corresponding level and the logic rules (1), (2) and (3) are proposed. These rules have been chosen arbitrarily and will change depending on the characterization of the situation under analysis.

Table 4: rules for accessibility variable (y_1) when the size variable x_1 is (a) small, (b) medium and (c) large

	L	M	H	S
Z	Y_3	Y_3	Y_4	Y_4
VL	Y_2	Y_3	Y_3	Y_4
L	Y_2	Y_2	Y_3	Y_3
M	Y_1	Y_2	Y_2	Y_3
H	Y_1	Y_1	Y_2	Y_2
E	Y_1	Y_1	Y_1	Y_2

(a)

	L	M	H	S
Z	Y_3	Y_4	Y_4	Y_5
VL	Y_3	Y_3	Y_4	Y_4
L	Y_2	Y_3	Y_3	Y_4
M	Y_2	Y_2	Y_3	Y_3
H	Y_1	Y_2	Y_2	Y_3
E	Y_1	Y_1	Y_2	Y_2

(b)

	L	M	H	S
Z	Y_4	Y_4	Y_5	Y_5
VL	Y_3	Y_4	Y_4	Y_5
L	Y_3	Y_3	Y_4	Y_4
M	Y_2	Y_3	Y_3	Y_4
H	Y_2	Y_2	Y_3	Y_3
E	Y_1	Y_2	Y_2	Y_3

(c)

Table 5: rules for attractivity variable (y_2) when the size variable x_1 is (a) small, (b) medium and (c) large

	L	M	H	S
Z	Y_1	Y_1	Y_1	Y_2
VL	Y_1	Y_1	Y_2	Y_2
L	Y_1	Y_2	Y_2	Y_3
M	Y_2	Y_2	Y_3	Y_3
H	Y_2	Y_3	Y_3	Y_4
E	Y_3	Y_3	Y_4	Y_4

(a)

	L	M	H	S
Z	Y_1	Y_1	Y_2	Y_2
VL	Y_1	Y_2	Y_2	Y_3
L	Y_2	Y_2	Y_3	Y_3
M	Y_2	Y_3	Y_3	Y_4
H	Y_3	Y_3	Y_4	Y_4
E	Y_3	Y_4	Y_4	Y_5

(b)

	L	M	H	S
Z	Y_1	Y_2	Y_2	Y_3
VL	Y_2	Y_2	Y_3	Y_3
L	Y_2	Y_3	Y_3	Y_4
M	Y_3	Y_3	Y_4	Y_4
H	Y_3	Y_4	Y_4	Y_5
E	Y_4	Y_4	Y_5	Y_5

(c)

In Tables 4 and 5, the model rules are represented: they are shown with reference to input x_2 (level of protection) and x_3 (social criticality) and are parameterized with respect to the three different linguistic terms of variable x_1 (size). Then, accessibility (y_1) and attractivity (y_2) are combined to yield the susceptibility to attacks, y. The logic approach deployed to combine x_1, x_2 and x_3 into y_1 and y_2, is applied to y_1 and y_2 to yield the susceptibility, y. The logic rules that yield the values of y are described in the following. Again, the IF – THEN logic rule is proposed.

Five Threat Conditions are identified by means of Roman numerals. From lowest to highest, the levels are:

Low = **I**. This condition refers to a low susceptibility of terrorist attacks.
Guarded = **II**. This condition is declared when there is a general susceptibility of terrorist attacks.
Elevated = **III**. An elevated condition is declared when the susceptibility of attack is significant.
High = **IV**. A high condition is declared when there is a high susceptibility of malevolent attacks.
Severe = **V**. A severe condition reflects a severe susceptibility of terrorist attacks.

The higher the threat condition, the greater is the susceptibility of an intentional attack, going from **I**, low susceptibility to attacks, to **V**, severe condition of malevolent acts. To evaluate susceptibility to attacks, the set of rules that relate the two input variables, attractivity, y_1 and accessibility, y_2, to the output, susceptibility, y, are assigned in Table 6.

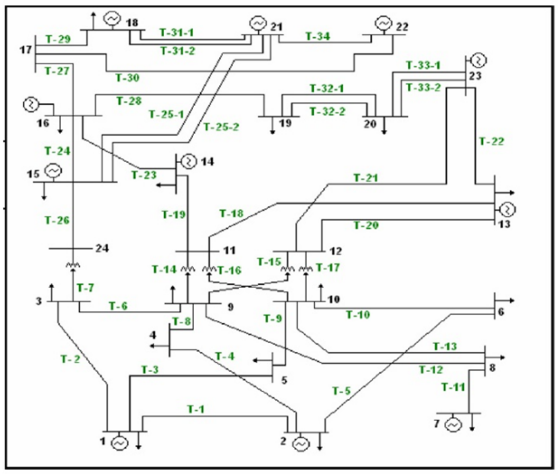
Table 6 shows how the levels of attractivity and accessibility yield the different levels of susceptibility to attacks. The numeral numbers distribute along east-west diagonal lines. The susceptibility to attack increases from the upper left corner where the susceptibility to intentional attacks is low, **I** level, to the lower right corner where the threat of attacks is severe, **V** level.

Table 6: Rules linking the linguistic variables attractivity (y_1) and accessibility (y_2) to the linguistic variable susceptibility y. The Roman numerals refer to the five levels: I, II, III, IV and V

y_2 / y_1	Y_{21}=I	Y_{22}=II	Y_{23}=III	Y_{24}=IV	Y_{25}=V
Y_{11}=I	**I**	**I**	**I**	**II**	**III**
Y_{12}=II	**I**	**I**	**II**	**III**	**IV**
Y_{13}=III	**I**	**II**	**III**	**IV**	**V**
Y_{14}=IV	**II**	**III**	**IV**	**V**	**V**
Y_{15}=V	**III**	**IV**	**V**	**V**	**V**

4. Application

For illustration purposes, the rules have been applied to a study case of literature: the IEEE RTS-96 (IEEE RTS-96, 1999) power grid showed in Figure 1.



Provider ID	Attractivity	Accessibility	Susceptibility to attacks
1	4	2	III
2	4	2	III
7	4	2	III
13	4	3	IV
14	3	2	II
15	4	2	III
16	4	2	III
18	5	3	V
21	5	3	V
22	4	3	IV
23	4	3	IV

Figure 1: Single area IEEE RTS-96 grid (IEEE RTS-96, 1999) and list of the Susceptibility levels of the providers of IEEE RTS-96 grid

It consists of 24 load buses (users), 11 of these are generating units (providers), and 38 transmission lines (transmitters). In order to calculate the susceptibility of the elements of the network, each element has been assigned a level of protection and a level of social criticality.

Concerning the users, the level of protection is taken into account depending on the type of load bus. For example, a transmission grid load bus is typically located far from densely-populated areas and it is usually locked, surrounded by fences but no other complex barriers (level 3, Table 1). The impact of the component role is assumed moderate (level 2, Table 2) since it is expected that if a load bus cannot receive power, the generation of power can be modulated, the power excess eliminated, and the overall infrastructure put in conditions to still provide its service. On the other hand, a provider is assumed to have a high level of criticality (level 4, Table 2), because in general its lost power supply may not be readily replaced by alternative generation. From the level of protection point of view, it can be assumed that the plant is isolated from the urbanized areas and it is usually guarded with security patrols, video surveillance of the entire power plant and alarms (level 5 or 6, Table 1). Finally, transmission towers are usually located in isolated sites, e.g., open country and they are not provided with any particular fence or barriers, nor are they watched by patrol (level 1 or 2, Table 1). The impact of their role may be assumed as low (level 1, Table 2). By these considerations, all the components of the network have been assigned attractivity and accessibility levels. The susceptibility to attacks can then be derived by the rules in Section 3 and the levels of susceptibility for providers are reported in Figure 1. The Tables of the levels of susceptibility for users and transmitters are here not reported, due to limitation of space. However, the salient aspects are described in the following. As can be seen from Figure 1, the susceptibility to attacks turns out to be strongest for generators sited in the upper part of the grid where the highest percentage of the generation is provided. The levels of susceptibility to

attacks are low both for users and transmitters: however it is worthwhile noting that the transmission lines that have the highest susceptibility to attacks are placed in the central part of the network (for example, lines connecting bus 11 to bus 13 and bus 12 to bus 23). In the performed analysis the components of the network with the highest level of susceptibility turn out to be the providers: an attack to them will cause a strong effect in terms of disruption, hence the highest success from a terroristic perspective.

5. Conclusions

An approach for identifying the vulnerabilities of critical infrastructures has been presented within an all-hazard analysis framework which allows merging two different perspectives on vulnerability: on the one hand, there is the demand to encompass the vulnerabilities due to random failures and to natural hazards; on the other hand, there is the need to include vulnerabilities due to malevolent acts.

A structured organization of the relevant information on the system components is made on the basis of their tasks and of the features that influence them in the potential role as source of vulnerability. Then, an evaluation is made of the degree of exposure, i.e., the susceptibility of the components to malevolent acts. The future step of the analysis will be the development of a quantitative decision logic method for evaluating the susceptibility encompassing the whole set of hazards, i.e., random failures unintentional acts and natural hazards, but also malevolent acts, while accounting for the related uncertainties.

References

IEEE RTS-96, 1999, Reliability test system task force of the application of probability methods subcommittee. The IEEE reliability test system – 1996. IEEE Trans Power Syst; 14, 1010 – 20.

Koonce A.M., Apostolakis G.E., 2008, Bulk power risk analysis: ranking infrastruture elements according to their risk significance, Electrical Power and Energy Systems, 30, 169-183.

Kroeger W., Zio E., 2011, Vulnerable Systems, Springer, London. ISBN 978-0-85729-654-2.

Marseguerra M., Zio E., Bianchi M., 2004, A fuzzy modeling approach to road transport with application to a case of spent nuclear fuel transport, Nuclear Technology, 146 (3), 290-302.

Pollet J., Cummins, J., 2009, All-Hazard approach for Assessing Readiness of Critical Infrastructure, HST'09 IEEE Conference on Technologies for Homeland Security, 366-372.

Waugh, W. L., 2005, Terrorism and the All-Hazard Model, Journal of Emergency Management, 3(2), 8-10.

Zio E., Piccinelli R., Sansavini G., 2011, An All-Hazard Approach for the Vulnerability Analysis of Critical Infrastructures, Proceedings of the Annual Conference ESREL, 18-22 september 2011, Troyes, France, 2451-2458.

Paper III

Randomized flow model and centrality measure for electrical power transmission network analysis

Enrico Zio, Roberta Piccinelli

Reliability Engineering and System Safety 95, 379 – 385, 2010.

Randomized flow model and centrality measure for electrical power transmission network analysis

Enrico Zio, Roberta Piccinelli

Politecnico di Milano, Dipartimento di Energia, Via Ponzio 34/3, I-20133 Milano, Italy

Abstract. Commonly used centrality measures identify the most important elements in networks of components, based on the assumption that flow occurs in the network only along the shortest paths. This is not so in real networks, where different operational rules drive the flow. For this reason, a different model of flow in a network is here considered: rather than along shortest paths only, it is assumed that contributions come essentially from all paths between nodes, as simulated by random walks. Centrality measures can then be coherently defined. An example of application to an electrical power transmission system is presented.

Acknowledgements

This work has been partially funded by the Foundation pour une Culture de Securité Industrielle of Toulouse, France, under the research contract AO2006-01.

1. Introduction

Modern society is witnessing a continuous growth in the complexity of the infrastructure networks which it relies upon. Reliable electric power supply, for example, is crucial for many of the services that are taken for granted today; disturbances in the power supply have the potential of severely disrupting these indispensable services. This raises significant concern about reliability and resilience to disturbances and failures of various types of infrastructure systems, and a corresponding demand for methods capable of analyzing the vulnerabilities of these systems [1].

The developments contained in this paper are motivated by the interest in the analysis of electric power networks. In this context, the network analysis paradigm set up to study the dynamics of the relations in social networks has been previously utilized to analyze the vulnerability of electric power infrastructure systems [2, 3]. The focus of these types of studies is typically on analyzing the structural properties of the system from a topological point of view, i.e., considering only the connectivity properties of the network and not the actual physical flow through it [4, 5]. Three drawbacks associated with the related measures of network performance are that they are based on:

- binary links among network nodes (or components), thus neglecting the strength of the connections (or links or arcs or edges); this has been pointed at as a limitation both in social networks, where the strength and depth of interpersonal relationships is of relevance [6, 7, 8]

and in engineered network infrastructures, where the capacities of the arcs connecting the components limit the flow among them [5];

- a simplified modeling scheme which assumes that flow (communication, in the social case) between a pair of components (persons, in the social case) in the network takes place only along the shortest path linking them [8, 9]; this has been considered a limitation in many cases, because the flow from one node of a network to another is typically a global phenomenon which does not depend only on the links on the direct and shortest paths, since it is quite possible that information will take a more circuitous route; this is true both in social networks, where information may travel by random communication or be intentionally channeled through intermediaries, and in network infrastructures, where flow is channeled through selected routes, following the specific operative rules and constraints which apply to the system;
- a simplified modeling scheme which neglects the possibility of failures in the interconnections between pairs of linked components; this is particularly relevant for the engineered infrastructure networks made of fallible hardware and software, operated by (unfortunately) not error-free human operators.

In synthesis, when looking at the safety, reliability and vulnerability characteristics of an infrastructure such as the electric power transmission network, one should take into account the capacities of the transmission elements and their probability of failure, and examine the different transmission routes available to the flow. This would entail undertaking a complex and detailed mechanistic modeling effort of the entire network system, which is in practice often unfeasible, both with respect to its development and its computation. For this reason, a framework of analysis has been proposed to integrate models at different levels of detail, in a problem-driven approach to solution; complementation of network analysis, for performing an initial screening of the vulnerabilities of a critical infrastructure with object-oriented modeling, to further deepen the vulnerability assessment of the screened scenarios has been investigated as a feasible way to proceed in such direction [10].
To improve the physical description of the network characteristics within a network analysis for preliminary screening, a model based on random walks is here introduced as an extension of the model in [8] giving proper consideration to the following facts:

- each link connecting two nodes is characterized by a transmission capacity which cannot be exceeded;
- the capacities of the network lines are assumed to stochastically vary, to account for the inherent uncertainties;

- not only the links on the direct and shortest paths are considered in the analysis of the transmission of flow; this is achieved by a randomization of the direction of the flow in output from a node; the randomization is driven by the capacities of the outgoing links, with the highest capacity links most probably channeling the flow;
- the network interconnecting links are assumed fallible, with given probabilities;
- source generation and load demands are assumed to vary stochastically, to account for the fluctuations inherent in the network behavior and operation.

From the analysis of the network characteristics and behavior, it is also important to gain an understanding of the role that the elements of the infrastructure network play in determining the flow through it, as this can be of great practical aid to network designers and operators in providing indications for network protection. From a topological viewpoint, various measures of the importance of a network node, can be introduced. These so-called centrality measures, take into account the different ways in which a node interacts/communicates with the rest of the network. Classical topological centrality measures are the degree centrality [11, 12], the closeness centrality [12; 13; 14], the betweenness centrality [12] and the information centrality [15]. The major drawback of these measures is that to assess the node importance they rely only on topological information based on the three previously mentioned model simplifications. Then, based on the model proposed in this paper an extension of the betweenness centrality measure of [16] is computed, to more realistically capture the importance of the role played by the different components in determining the flow through the network.

An application of the proposed approach is illustrated with reference to a power transmission network system of literature [17].

The paper is organized as follows. In Section 2, a description of the random walk flow propagation model is provided. In Section 3, the topological concept of betweenness centrality measure is recalled and then extended to its randomized flow definition. The results obtained on a case study of literature are discussed in Section 4. Conclusions are drawn in Section 5.

2. Randomized flow model of a power transmission network infrastructure

The topological interconnection of a power transmission system can be modeled as a network consisting of N nodes (also called vertexes) and K edges (also called arcs or lines): the buses of the electric grid are represented as nodes interconnected by undirected edges representing the transmission lines; N_S nodes are power sources (generators), N_T nodes are targets (loads) and the rest are transmission nodes. The N×N adjacency matrix $\{a_{ij}\}$ defines the topological structure of the network, i.e., the pattern of connectivity among its nodes, with the matrix entry a_{ij} being equal to 1 if there is an

edge linking i and j and 0 otherwise; the entries on the diagonal elements, a_{ii}, are undefined and for convenience they are set equal to 0.

The matrix $\{q_{ij}\}$ defines the probabilities of failure of the links.

The capacities of the links are assumed to vary stochastically, to account for the uncertainties inherent in their behavior and operation; then, to each capacity value w_{ij} is associated a probability distribution $\pi(w_{ij})$ of the possible values.

The underlying strategy to model the flow in the network is to choose a source node, follow at random one of the departing links to one of its neighbors, take this as the source and iterate this process until the required target is reached. The random choice of the arc to follow is based on the actual capacity of each arc outgoing from the node: higher capacity arcs have larger probability to be selected as flow carriers.

Accordingly, the algorithm to evaluate the service reliability performance characteristics of the network, and its related vulnerabilities, consists of three nested cycles of randomization; the steps are as follows:

1. Sample the fault configuration of the network on the basis of the failure probabilities of each element (node or arc) of the system.
2. Sample the production from the sources, the demand at the targets and the capacity of the arcs.
3. Build the discrete cumulative distribution function of the capacities of the arcs leaving the source node and sample the flow direction from it.
4. Develop the flow propagation cycle, for each source:
    - 4.1 the random walk of flow follows the arc sampled on the basis of the actual capacities of the arcs departing from the successive nodes traversed by the flow;
    - 4.2 if the flow goes into an isolated node with no departing connections, the cycle ends;
    - 4.3 the flow between a pair of nodes is accounted once (repeated flows between the same pair of nodes are neglected);
    - 4.4 once the flow arrives at a target node, the capacities of the incoming arcs are checked: if their sum is larger than the maximum capacity of the node, an overload is recorded;

4.5 if the flow does not reach the target, a new source of random walk is sampled. If no flow arrives at any of the targets, then a blackout is recorded.

3. Randomized betweenness centrality measure

Determining the critical elements of large-scale network infrastructures is an important issue for the reliability and the protection of the network. From a topological point of view, a number of centrality indices have been introduced as measures of the importance of the nodes in a network [18]. These indices take into account the different ways in which a node interacts and communicates with the rest

of the network and have proved of value in the analysis and understanding of the role played by the elements in the network.

A classical topological centrality measure is the betweenness centrality [12]. This measure is based on the idea that a node is central if it lies between many other nodes, in the sense that it is traversed by many of the shortest paths connecting pairs of nodes. The topological betweenness centrality C_i^B of a given node *i* in a network G(N, K), where N is the number of nodes and K is the number of links connecting them, is quantitatively defined as:

$$C_i^B = \frac{1}{(N-1)(N-2)} \sum_{j,k \in G, j \neq k \neq i} \frac{n_{jk}(i)}{n_{jk}} \tag{3.1}$$

where n_{jk} is the number of topological shortest paths between nodes j and k, and $n_{jk}(i)$ is the number of topological shortest paths between nodes j and k which contain node *i*. C_i^B assumes values between 0 and 1 and reaches its maximum when node *i* falls on all geodesics (paths of minimal length between pairs of nodes).

From the definition, betweenness centrality can be regarded as a measure of the influence a node has on the spread of the flow through the network, of the extent to which a node has control over the flow between other nodes. In a network in which flow is entirely or at least mostly distributed along geodesic paths, the betweenness of a vertex measures how much flow will pass through that particular vertex.

In most networks, however, flow does not occur only along geodesic paths; in some cases, flow may not follow the ideal route to get from one place to another, and "wander around" in a random-like fashion or as directed by the system operative rules and constraints. In most cases, a realistic betweenness measure should include non-geodesic paths in addition to geodesic ones [19].

To account for this issue, a betweenness centrality measure based on the concept of network flow has been suggested [8]. The edges of a network are considered as channels of communication linking pairs of nodes; the value of the connection of two nodes i and j determines the capacity w_{ij} of the channel linking them, or the amount of information that can pass between them. Information is assumed to flow along these channels. If f_{ij} is the amount of information passing on a channel linking node i directly to node j then $f_{ij} \leq w_{ij}$, i.e., the amount of information flowing along a channel that links directly connected vertices cannot exceed the capacity of that channel. What is relevant here is not just the direct flow between connected nodes, but the overall flow between pairs of nodes along all the paths that connect them: if a node i is chosen as an information source, or transmitter, and another node j as an information sink, or receiver, information from i may reach j along an edge linking i directly to j or along any and all indirect paths that begin at i, pass through one or more intermediate nodes and end at j.

Thus, the flow between two nodes is a global phenomenon: it depends, not just on the capacity of the channel linking two nodes directly, but on the capacities of all the channels on all the paths – both direct and indirect – that connect the two.

Ford and Fulkerson [16] introduced a model to determine the maximum flow from any source i to any sink j. Let m_{jk} be the maximum flow from a node i to a node k and let $m_{jk}(i)$ be the maximum flow from j to k that passes through node i; then, the flow betweenness centrality measure may be quantitatively defined as [8]:

$$FC_i^B = \frac{\sum_{j=1}^{N}\sum_{k=j}^{N} m_{ik}\ (i)}{\sum_{j=1}^{N}\sum_{k=j}^{N} m_{jk}} \tag{3.2}$$

Flow betweenness is based on the idea of maximum flow; the flow betweenness of a node i is defined as the amount of flow through i when the maximum flow is transmitted from source (s) to target (t), averaged over all s and t. Each edge in a network can be thought of as a transmission line carrying a flow of current. In general, more than a single unit of current can be carried between s and t by making simultaneous use of several different paths through the network.

In practical terms, the flow betweenness measures the betweenness of nodes in a network in which a maximal amount of information is continuously pumped between all sources and targets. Necessarily that flow still needs to "know" the ideal route (or one of the ideal routes) from each source to each target in order to realize the maximum flow. This still seems an unrealistic definition, in that it is often the case that flow does not take any sort of ideal path from source to target. To model this, a new betweenness centrality measure has then been introduced, the *random walk betweenness* [20]. Roughly speaking, the random walk betweenness of a node i is equal to the number of times that a random walk starting at s and ending at t passes through i along the way, averaged over all s and t. This measure is appropriate to a network in which information wanders about essentially at random until it finds its target, and it includes contributions from many paths that are not optimal in any sense. Let I_i^{st} be the current flowing from s to t, through node i. Quantitatively the random betweenness centrality measure is defined as:

$$RWC_i^B = \frac{\sum_{i=1,s<t}^{N} I_i^{st}}{\frac{1}{2}N(N-1)} \tag{3.3}$$

This measure seems an intuitively reasonable one to describe the fact that current will flow along all paths from source to target, and nodes that lie on no path from source to target get a betweenness of zero.

In this paper, an attempt to evaluate random betweenness centrality measures is made considering the physical characteristics of the transmission network in terms of length, capacity and failure probability of each transmission line, and types of nodes.

4. Application

The artificial transmission network system IEEE 14 BUS [17] is taken as reference case study. The network represents a portion of the American Electric Power System and consist of 14 bus locations connected by 20 lines and transformers, as shown in Figure 1. The transmission lines operate at two different voltage levels, 132 kV and 230 kV. The system working at 230 kV is represented in the upper half of Figure 1, with 230/132 kV tie stations at Buses 4, 5 and 7. The system is also provided with voltage corrective devices in correspondence of Buses 3, 6 and 8 (synchronous condensers). Buses 1 and 2 are the generating units.

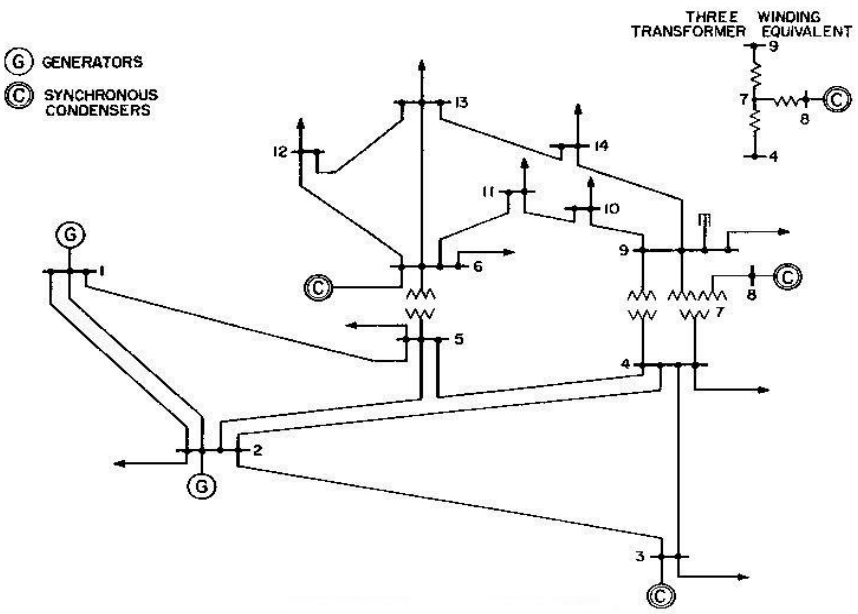


Figure 1. Transmission network IEEE 14 BUS [17].

To carry out the analysis, each network component is transposed into a node or edge of the representative topological network, as it is shown in Figure 2. Three different physical types of nodes are considered: source nodes (where the electricity is fed into the network), load nodes (where customers are connected) and transfer or transmission nodes (without customers or source).

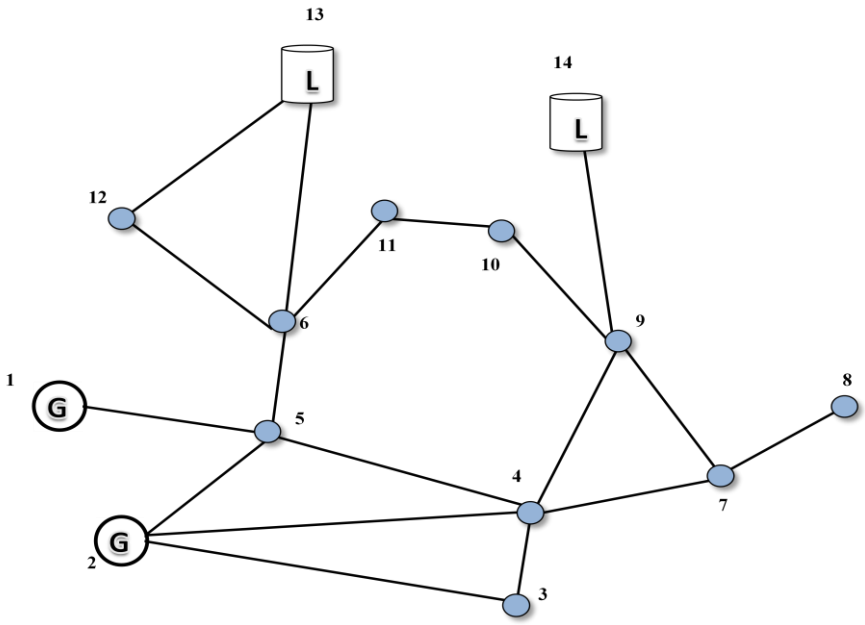


Figure 2. The IEEE 14 BUS transmission network's graph representation [21]. The white circles labeled with **G** represent the generator nodes, the colored circle nodes represent the transmission nodes and the white cylinders labeled with **L** represent the load nodes.

Table I provides the adjacency matrix that defines the topological structure of the network.

$$\{a_{ij}\} = \begin{vmatrix} 0 & 1 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 1 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 1 & 0 & 1 & 0 & 1 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 1 & 1 & 0 & 1 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 1 & 1 & 1 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & 1 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 1 & 0 & 0 & 1 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 1 & 0 \end{vmatrix}$$

Table I. The IEEE 14 BUS adjacency matrix

Table II reports the failure rate values of the components of the transmission network, as inferred from literature data [22].

From BUS	To BUS	Failure rate/yr/100km	Failure rate/yr	Equipment
1	2	1.0858		132 kV transmission line
1	5	1.0858		132 kV transmission line
2	3	1.0858		132 kV transmission line
2	4	1.0858		132 kV transmission line
2	5	1.0858		132 kV transmission line
3	4	1.0858		132 kV transmission line
4	5	1.0858		132 kV transmission line
4	7		0.01045	132/230 kV transformer
4	9		0.01045	132/230 kV transformer
5	6		0.01045	132/230 kV transformer
6	11	0.5429		230 kV transmission line
6	12	0.5429		230 kV transmission line
6	13	0.5429		230 kV transmission line
7	8		0.01045	132/230 kV transformer
7	9		0.01045	132/230 kV transformer
9	10	0.5429		230 kV transmission line
9	14	0.5429		230 kV transmission line
10	11	0.5429		230 kV transmission line
12	13	0.5429		230 kV transmission line
13	14	0.5429		230 kV transmission line

Table II. Failure data of the arcs of the IEEE 14 BUS transmission network

The failure probability of edge ij is defined as:

$$q_{ij} = 1 - e^{-\lambda_{ij}T} \tag{4.1}$$

where λ_{ij} is the constant failure rate per unit time of the edge ij linking nodes i and j (column 4 in Table II) and T is the reference time for the analysis, here chosen equal to 1 year.

Because the failure rate data are usually given as function of the length of each transmission line (column 3 in Table II), in order to compute the failure probability (4.1) transmission line lengths must be known. In this work, these have been inferred from the available data as follows. The total impedance Z_{ij} of a transmission line ij is dependent on the length of the line l_{ij}:

$$Z_{ij} = (r_{ij} + jx_{ij})l_{ij} \tag{4.2}$$

where r_{ij} is the resistance per unit length of arc ij and x_{ij} is its reactance per unit length. While the resistance of the line depends both on the length and on the thickness of the wire, the reactance of the line depends only on the length [23]. Based on relation (4.2), the lengths of the transmission lines in the IEEE 14 BUS system have been obtained from literature data by taking a system power base of 100 MVA and a conversion factor of 0.48 Ω/km [24]; lines containing transformers are considered to be zero-length lines (Table III).

From BUS i	To BUS j	Length (km) l_{ij}	Failure Probability q_{ij}
1	2	22	0.2125
1	5	81	0.5768
2	3	72	0.5338
2	4	64	0.4932
2	5	63	0.4884
3	4	62	0.4828
4	5	15	0.1498
4	7	-	0.0104
4	9	-	0.0104
5	6	-	0.0104
6	11	220	0.6970
6	12	283	0.7847
6	13	144	0.5425
7	8	-	0.0104
7	9	-	0.0104
9	10	93	0.3978
9	14	299	0.8027
10	11	212	0.6843
12	13	221	0.6999
13	14	385	0.8762

Table III. Length and failure probabilities of the transmission lines

The shortest lines are concentrated in the lower half of the network, which contains the generating units, while the longest lines belong to the upper half of the network, which contains the loads. The largest failure probabilities are concentrated on the edges directly connected with sources and with loads (edges 1 – 5, 2 – 3, 2 – 4 and 12 – 13, 9 – 14, 13 – 14).

Source generation is sampled from a normal distribution with a mean value of 30 and a variance of 100, in arbitrary units (a.u.). The values of the capacities of the network links are assumed all distributed according to a normal distribution of mean value 100 a.u. and a standard deviation of 10 a.u. The direction of flow is sampled on the actual capacities of the arcs. Once the flow arrives at a target node, the capacities of the incoming arcs are checked: if their sum is larger than the maximum capacity value of the node, an overload is recorded. The targets are absorbing nodes: the flow stops and the received flow is recorded for evaluating the network lost load and the network service efficiency; if no flow reaches any target, a service blackout is recorded.

The network performance characteristics computed on the basis of the above data are reported in Table IV:

Blackout (%)	0.82
Overload (%)	0.00
Network service efficiency	0.18
Network demanded load (a.u.)	59.93
Network received load (a.u.)	10.92
Network lost load (a.u.)	49.01

Table IV. Network performance indicators

where:

- blackouts and overloads are evaluated considering the average value of the flow that does not reach the targets or that exceeds the capacities of the transmission lines, respectively;
- the network demanded load is the average sum of the power generated from all the sources s_i, i=1, 2,…, N_S:

$$NDL = \sum_{i=1}^{N_S} s_i \tag{4.3}$$

- the network received load is the average sum of the flow reaching the targets t_i, i=1, 2, …, N_T:

$$NRL = \sum_{i=1}^{N_T} t_i \tag{4.4}$$

- the network lost load is obtained as the difference between demanded and received loads:

$$NLL = NDL - NRL \tag{4.5}$$

- the network service efficiency is obtained as the ratio between received and demanded loads:

$$NSE = \frac{NRL}{NDL} \tag{4.6}$$

In the artificial case considered, the network service efficiency is low and the blackout probability is high: very little of the generated power is received from the loads. This is caused by the high probability of failure of the transmission lines, due to their relatively high failure rates and large lengths. To see the effect of the lines lengths, a second computation has been made using average transmission line lengths inferred from literature [24]. Two line lengths of 48 and 50 km have been considered (Table V).

From BUS i	To BUS j	Length (km) l_{ij}	Failure probability q_{ij}
1	2	48	0.4079
1	5	48	0.4079
2	3	48	0.4079
2	4	48	0.4079
2	5	48	0.4079
3	4	48	0.4079
4	5	48	0.4079
4	7	-	0.0104
4	9	-	0.0104
5	6	-	0.0104
6	11	50	0.2372
6	12	50	0.2372
6	13	50	0.2372
7	8	-	0.0104
7	9	-	0.0104
9	10	50	0.2372
9	14	50	0.2372
10	11	50	0.2372
12	13	50	0.2372
13	14	50	0.2372

Table V. Length and failure probabilities of the arcs

In this case, the lower half of the network has the largest failure probability. The network performance characteristics are reported in Table VI.

Blackout (%)	0.44
Overload (%)	$3.33 \cdot 10^{-4}$
Network service efficiency	0.60
Network demanded load (a.u.)	59.93
Network received load (a.u)	36.24
Network lost load (a.u)	23.70

Table VI. Network performance indicators

With respect to the previous case, targets now receive a larger load, the service efficiency is higher and the blackout probability is lower. This is due to the values of the transmission lines failure probabilities, which are smaller than in the previous case because of the smaller line lengths.

Finally, the centrality betweenness based on the proposed random walk model has been computed for the two examples. The results are shown in Figure 3.

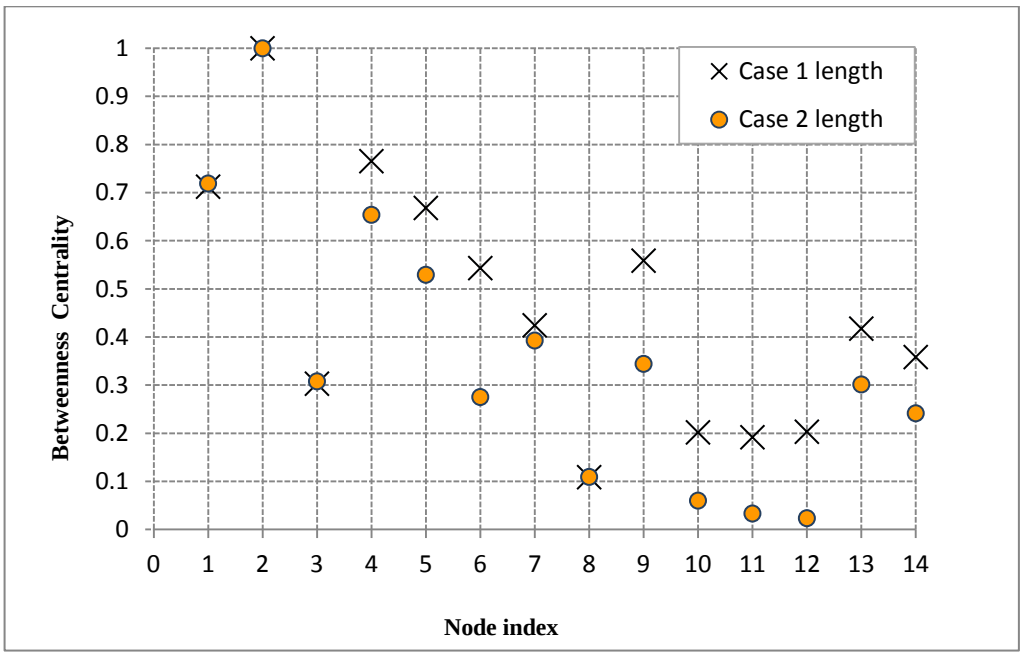


Figure 3. Betweenness centrality according to the random flow model proposed.

Betweenness values evaluated for case 1 (×, Figure 3) using higher lengths computed from eq. (4.2), are equal or higher than the values obtained for case 2 using smaller average lengths (○, Figure 3), due to the associated larger failure probabilities. Equal values are obtained for nodes 1 and 2, i.e., the

source nodes, for node 3 and for node 8 which is an isolated node (Figure 1). The lower half of the network (nodes 1, 2, 3, 4 and 5), which contains the generating units (nodes 1 and 2), has higher values of betweenness than the upper half, which contains the load nodes 13 and 14. Nodes 10, 11 and 12 act as a tie for the flow. Node 2, which represents a source node and a transfer node as well, is the most important from the betweenness point of view, in both studies.

5. Conclusions

In previous works, network analysis has been shown suitable for a preliminary analysis of complex infrastructures aimed at identifying structural criticalities, e.g. the most connected nodes, shortest path lengths of connection, most vulnerable nodes, etc. Limitations of the analysis relate to the neglecting of the actual capacities of the links, their probabilities of failures and the fact that flow among network nodes is typically a global phenomenon, not restricted to only direct, shortest paths as typically assumed.

To overcome some of these limitations, in this paper a model of random flow propagation has been introduced and the topological concept of betweenness centrality has been accordingly extended to account for the random flow propagation across the network. The randomization of the flow out of a node is driven by the capacity values of its outgoing links and allows non-geodesic paths to be travelled by the flow. Variability in the behavior and operation of the links, source and target nodes is also accounted for by varying stochastically the capacities, the productions and demands, respectively.

The modeling approach has been applied to the artificial transmission network system of the IEEE 14 BUS and indications derived from the betweenness centrality measure values have been analyzed for different transmission lines lengths. Each equipment of the system has been transposed into a node or edge of the representative network and the length of the arcs has been calculated for two cases: case 1, in which the lengths have been obtained from the impedances of the lines and case 2, in which average line lengths have been considered from literature data. The network performance characteristics and the random walk betweenness centrality measures have highlighted the weaknesses of the network structure, for the failure data used.

Acknowledgments:
The authors wish to thank Prof. Maurizio Delfanti and Dr.Mauro Pozzi of the Department of Energy of the Politecnico di Milano, for their contribution to the work.

References

[1] Albert, R., Albert, I., Nakarado, G.L., “Structural vulnerability of the North American power grid”, *Phys. Rev E* **69**, 025103 (R), (2004).

[2] Holmgren, A.J., “Using Graph Models to Analyze the Vulnerability of Electric Power Networks”, *Risk Analysis*, Vol. 26, No. 4 (2006).
[3] Zio, E., Petrescu, C.A., Sansavini, G., “Vulnerability analysis of a power transmission network”, Proceedings of PSAM9 - International Probabilistic Safety Assessment and Management Conference, Hong Kong , China, 18-23 May (2008).
[4] Cadini, F., Zio, E., Petrescu, C.A., ”Using centrality measures to rank the importance of the components of a complex network infrastructure”, CRITIS’08 - International Workshop on Critical Information Infrastructures Security, Frascati (Rome), Italy, 13-15 October (2008).
[5] Hines, P., Blumsack, S., “A Centrality Measure for Electrical Networks”, Proceedings of the 41st Hawaii International Conference on System Sciences (2008).
[6] Festinger, L., “The analysis of sociograms using matrix algebra”, *Human Relations* 10, pp. 153 – 158 (1949).
[7] Yan, X., “A fuzzy set analysis of sociometric structure.” *Journal of Mathematical Sociology* 7 pp 159 – 180 (1988).
[8] Freeman, L.C., et al., “Centrality in valued graph: a measure of betweenness based on network flow”, *Social Networks* **13**, 141-154, (1991).
[9] Stephenson, K., Zelen, M., “Rethinking centrality: Methods and examples.” , *Social Networks* 11, pp. 1 – 37 (1989).
[10] Eusgeld I., Kröger W., Sansavini G., Schläpfer M., Zio E., “The role of network theory and object-oriented modeling within a framework for the vulnerability analysis of critical infrastructures”, *Journal Reliability Engineering & Systems Safety*, Vol. 94, No 5, pp.954-963, 2009.
[11] Nieminen J., On Centrality in a Graph, *Scandinavian Journal of Psychology*, n.15, 1974.
[12] Freeman, L.C., “Centrality in Social Networks: Conceptual Clarification”, *Social Networks*, **1**, pp.215 – 239, (1979).
[13] Sabidussi G., The Centrality Index of a Graph, *Psychometrika*, n.31.1966.
[14] Wasserman S., Faust K., Social Networks Analysis, Cambridge U.P., Cambridge, UK.
[15] Latora V. and Marchiori M, A Measure of Centrality Based on the Network Efficiency, *New Journal of Physics* 9, 188, (2007).
[16] Ford, L.R. and Fulkerson, D.R.., “Maximal flow through a network”, *Canada Journal of Mathematics* **8**, 399 404, (1956).
[17] The IEEE 14 BUS data can be found on: http://www.ee.washington.edu/research/pstca/.
[18] Zio, E., “From Complexity Science to Reliability Efficiency: A new Way of Looking at Complex Network Systems and Critical Infrastructures”, *Int. J. Critical infrastructures*, Vol. 3, Nos. ¾, pp.488 – 508, (2007).
[19] Borgatti, S.P., ”Centrality and Network Flow”, *Social Networks* **27**, 55 -71, (2005).
[20] Newman, M.E., “A measure of betweenness centrality based on random walks”, Preprint cond-mat/0309045, (2003).
[21] Pajek, Program for large network analysis : http://vlado.fmf.uni-lj.si/pub/network/pajek/.
[22] Billington, R., Li W., Reliability Assessment of electric power system using Monte Carlo Methods, pp 19 – 20, (1994).
[23] Grisby, L.L., Power Systems, CRC Press, (2007).
[24] IEEE Transaction on Power systems, Vol.14, No. 3, August 1999.

Paper IV

A comparison of the load flow and random flow models for the analysis of power transmission networks

Enrico Zio, Roberta Piccinelli, Maurizio Delfanti, Valeria Olivieri, Mauro Pozzi

Reliability Engineering and System Safety 103, 102 – 109, 2012.

Application of the load flow and random flow models for the analysis of power transmission networks

Enrico Zio[1, 2*], Roberta Piccinelli[2], Maurizio Delfanti[2], Valeria Olivieri[2], Mauro Pozzi[2]

[1]*Chair on Systems Science and the Energetic challenge, European Foundation for New Energy-Electricite' de France, Ecole Centrale Paris and Supelec*

[2]*Politecnico di Milano, Milano, Italy*

*corresponding author:

enrico.zio@ecp.fr, enrico.zio@supelec.fr, enrico.zio@polimi.it

Abstract.
In this paper, the classical load flow model and the random flow model are considered for analyzing the performance of power transmission networks. The analysis concerns both the system performance and the importance of the different system elements; this latter is computed by power flow and random walk betweenness centrality measures. A network system from the literature is analyzed, representing a simple electrical power transmission network. The results obtained highlight the differences between the LF "global approach" to flow dispatch and the RF local approach of randomized node-to-node load transfer. Furthermore, computationally the LF model is less consuming than the RF model but problems of convergence may arise in the LF calculation.

Acknowledgements

This work has been partially funded by the Foundation pour une Culture de Securité Industrielle of Toulouse, France, under the research contract AO2006-01.

1. Introduction

A fundamental objective of power transmission systems design is to contain economic expenses as much as possible, while ensuring the required reliability levels. Power flow studies, often referred to as load flow (LF) analyses, are the backbone of power systems analysis and design. They are necessary for planning, operation, economic scheduling and exchange of power between utilities; they are also required for contingency studies.

The traditional, deterministic load flow analysis finds nodal voltages and line flows under a specific operating condition. On the other hand, the information sustaining load flow calculations are stochastic in nature (e.g. bus powers and configurations): hence, the input quantities for the calculations should be treated as random variables. As outcomes, probabilistic load flow calculations assess adequacy indices such as the probability of a line flow being greater than its thermal rating and the probability of a bus voltage being outside its operational constraints. These indices are typically obtained under steady state conditions.

Some effort has been devoted to solve the load flow problem probabilistically [1, 2, 3], but the great majority of methods only account for load and power generation data uncertainties, whereas the network configuration is considered fixed. In [4], methods are presented for obtaining a probabilistic load flow solution when network outages are modeled as random variables.

So far, the effects of the configuration uncertainties have not been considered in full. Any change in the network of links of the power transmission system will alter the configuration and, consequently, the set of functions relating inputs and outputs. Three basic sources of variation can be identified: the first is the variation in the actual parameters defining loads and power generating units; the second is the variation of lines capacity, transformers' or other components' parameters and the third source of variation is associated with the availability or unavailability of components such as transmission lines, transformers, etc, as they are subject to outages due to faults and maintenance.

In this paper, the classical load flow model, enriched with the above mentioned sources of variation, and the random flow model [5] are considered for the analysis of a power transmission network. Two kinds of outputs are of interest in the analysis: the system performance and the importance of its elements. While system performance indicators encompass the static characteristics of the whole system and the dynamic evolution of its flow therein, the importance of the different elements (nodes) in the system (network) can be seen from the point of view of their individual connectivity efficiency and/or their contribution to the propagation of failures through the system network of connections [8-17]. In this paper, betweenness centrality is used [9]. In its original formulation, this measure describes the idea that a node is central if it lies between many other nodes, in the sense that it is traversed by many of the shortest paths connecting pairs of nodes. However, in our case the metrics of betweenness from a pure topological perspective fails to capture some basic and important features such as electrical distance or line flow limit. For this reason, the betweenness centrality measure is re-defined within the two particular paradigms of flow modeling here considered, the load and random walk models.

The IEEE 14 BUS network system [6] is considered as reference case study of the analysis, in order to critically interpret the differences identified in the results.

The paper is organized as follows. For completeness of the paper, in Section 2 short descriptions are provided of the load flow and random flow models. In Section 3, the system performance indicators are introduced and the classical topological concept of betweenness centrality measure is recalled, and then extended to its power flow and random flow definitions. The presented modeling frameworks are applied to the IEEE 14 BUS network system and the results obtained are discussed in Section 4. Conclusions are drawn in Section 5.

2 Load flow and random flow models of a power transmission network system

2.1 Load flow model (LF)[7]

From the physical point of view, the interconnection of the different elements of a power transmission system provide the basis for the development of an overall load flow model for computational simulation of the system performance under a wide variety of projected operating conditions. Successful power system operation under normal conditions requires the following:

-generation supplies the demand (load) plus losses;

-bus voltage magnitudes remain close to their rated values;

-generators operate within specified power limits;

-transmission lines and transformers are not overloaded.

The power flow problem (also known as the load flow problem) addressed by the computational model can be stated as follows: “for a given power network, with known power loads and some set of specifications or restrictions on power generations and voltages, solve for any unknown bus voltages and for the power flow in the network components” [7].

The power network description may be given in the form of a system map and accompanying data tables specifying the components characteristics. These include the failure probabilities of components and transmission lines and the values of generating units, loads and line parameters.

The steps to be followed by the Direct Current (DC) power flow computation here of interest can be summarized as follows:

1. sample the fault configuration of the network on the basis of the failure probabilities of each element (node or arc) of the system;
2. sample the production from the sources, the demand at the targets and the capacity of the arcs;
3. develop the mathematical model describing the power flow in the network, under the sampled conditions;
4. solve for the voltage profile of the network;
5. solve for the power flows in the network;
6. check for constraint violations: the capacities of the transmission lines are checked; if their values are larger than the maximum capacities of the node to which they are connected, overloads are recorded; if the load exceeds the generation supplies, then a blackout is recorded.
7. evaluate the system performance indicators and the power flow betweenness centrality measure defined in Section 3.

2.2 Random flow model (RF)[5]

From a topological point of view, a power transmission system can be modeled as a network consisting of N nodes (also called vertexes) and K edges (also called arcs or links): the buses of the electric grid are represented as nodes interconnected by (un)directed edges representing the transmission lines; N_S nodes are power sources (generators), N_T nodes are targets (loads) and the rest are transmission nodes. The N×N adjacency matrix $\{a_{ij}\}$ defines the topological structure of the network, i.e., the pattern of connectivity among its nodes, with the matrix entry a_{ij} being equal to 1 if there is an edge linking nodes i and j, and 0 otherwise; the entries on the diagonal elements, a_{ii}, are undefined and for convenience they are set equal to 0. The matrix $\{q_{ij}\}$ defines the probabilities of failure of the links over the period of interest.

The model based on random walks gives consideration to the following facts:

- each link connecting two nodes is characterized by a transmission capacity which cannot be exceeded;
- the capacities of the links are assumed to vary stochastically, to account for the uncertainties inherent in their behavior and operation, e.g., losses (due to the resistance and the conductance, or thermal and voltage drop limits that can restrict currents on transmission lines): then, to each capacity value w_{ij} is associated a probability distribution $\pi(w_{ij})$ of the possible values.
- the direction of the flow in output from a node is driven by the capacities of the outgoing links: the highest capacity links most probably channel the flow;
- the network interconnecting links are assumed fallible, with given probabilities;
- source generation and load demands are assumed to vary stochastically, to account for the fluctuations inherent in the network behavior and operation.

The underlying strategy to model the flow in the network is to choose a source node, follow at random one of the departing links to one of its neighbors, take this as the source and iterate this process until the required target is reached. The random choice of the arc to follow is based on the actual capacity of each arc outgoing from the node: higher capacity arcs have larger probability to be selected as flow carriers.

Accordingly, the developed algorithm to evaluate the system performance indicators and elements importances, consists of three nested cycles of randomization; the steps are as follows:

5. sample the fault configuration of the network on the basis of the failure probabilities of each element (node or arc) of the system;
6. sample the production from the sources, the demand at the targets and the capacity of the arcs;

7. build the discrete cumulative distribution function of the capacities of the arcs leaving the source node and sample the flow direction from it;
8. for each source, develop the flow propagation cycle:
    8.1 the random walk of flow follows the arc sampled on the basis of the actual capacities of the arcs departing from the successive nodes traversed by the flow;
    8.2 if the flow goes into an isolated node with no departing connections, the cycle ends;
    8.3 the flow between a pair of nodes is accounted once (repeated flows between the same pair of nodes are neglected);
    8.4 once the flow arrives at a target node, the capacities of the incoming arcs are checked: if their sum is larger than the maximum capacity of the node, an overload is recorded;
    8.5 if the flow does not reach the target, a new source of random walk is sampled. If no flow arrives at any of the targets, then a blackout is recorded.

5. evaluate the system performance indicators and the random flow betweenness centrality measure defined in Section 3.

3. System performance indicators and element importance measures

The performance of the system is evaluated with respect to the following indicators:

- blackouts and overloads are evaluated considering the average value of the flow that does not reach the targets or that exceeds the capacities of the transmission lines, respectively;
- the network demanded load is the average sum of the power generated from all the sources s_i, i=1, 2,…, N_S:

$$NDL = \sum_{i=1}^{N_S} s_i \tag{3.4}$$

- the network received load is the average sum of the flow reaching the targets t_i, i=1, 2, …, N_T:

$$NRL = \sum_{i=1}^{N_T} t_i \tag{3.5}$$

- the network lost load is obtained as the difference between demanded and received loads:

$$NLL = NDL - NRL \tag{3.6}$$

- the network service efficiency is obtained as the ratio between received and demanded loads:

$$NSE = \frac{NRL}{NDL} \tag{3.7}$$

The importance of the elements in the system is evaluated with respect to measures of centrality [8-12]. In general, determining the critical elements of a large-scale network infrastructure is an important issue for the reliability and the protection of the network. In particular, there exists a close relation between topological structure and physical properties of power system. From a topological point of view, a number of centrality indices have been introduced as measures of the importance of the nodes in a network [8]. These indices take into account the different ways in which a node interacts and communicates with the rest of the network and have proved of value in the analysis and understanding of the role played by the elements in the network.

Among them, the betweenness centrality plays a key role in the identification of critical components of complex networks [9]. This measure is based on the idea that a node is central if it lies between many other nodes, in the sense that it is traversed by many of the shortest paths (geodesics) connecting pairs of nodes. The topological betweenness centrality C_i^B of a given node *i* in a network G(N, K), where N is the number of nodes and K is the number of links connecting them, is quantitatively defined as [9]:

$$C_i^B = \frac{1}{(N-1)(N-2)} \sum_{j,k \in G, j \neq k \equiv i} \frac{n_{jk}(i)}{n_{jk}} \tag{3.1}$$

where n_{jk} is the number of topological shortest paths (geodesics) between pairs of nodes j and k, and $n_{jk}(i)$ is the number of topological shortest paths (geodesics) between pairs of nodes j and k which contain node *i*. The centrality betweenness may be normalized by dividing by the number of pairs of vertices not including *i*, which is $(N - 1)(N - 2)$ for directed graphs, so that C_i^B assumes values between 0 and 1 and reaches its maximum when node *i* falls on all paths of minimal length between all pairs of nodes in the network.

Note that the measure of betweenness is based on the assumption that information is passed only along the shortest paths and can be regarded as a measure of the influence a node has on the spread of the flow through the network, of the extent to which a node has control over the flow between other nodes. In a network in which flow is entirely or at least mostly distributed along geodesic paths, the betweenness of a vertex measures how much flow will pass through that particular vertex.

In most networks, however, flow does not occur only along geodesic paths; in some cases, flow may not follow the shortest or the direct route to get from one place to another, and “wander around” in a random-like fashion or as directed by the system operative rules and constraints. In most cases, a realistic betweenness measure should include non-geodesic paths in addition to geodesic ones [10].

To account for this issue, a modified betweenness centrality measure based on the concept of network flow has been suggested [11] to include contributions of maximum flow. The edges of a network are considered as channels of communication linking pairs of nodes; the value of the connection of two nodes *i* and *j* determines the capacity w_{ij} of the channel linking them, or the amount of information that

can pass between them. Information is assumed to flow along these channels. If f_{ij} is the amount of information passing on a channel linking node i directly to node j then $f_{ij} \leq w_{ij}$, i.e., the amount of information flowing along a channel that links directly connected vertices cannot exceed the capacity of that channel. What is relevant here is not just the direct flow between connected nodes, but the overall flow between pairs of nodes along all the paths that connect them: if a node i is chosen as an information source, or transmitter, and another node j as an information sink, or receiver, information from i may reach j along an edge linking i directly to j or along any and all indirect paths that begin at i, pass through one or more intermediate nodes and end at j.

Thus, the flow between two nodes is a global phenomenon: it depends, not just on the capacity of the channel linking two nodes directly, but on the capacities of all the channels on all the paths – both direct and indirect – that connect the two.

To model this for the power transmission network whose flow is described by the LF model, a power flow betweenness centrality measure (PFC_i^B) is here originally introduced. Let PF_i^{st} be the power flow from s to t, through node i; quantitatively the power flow betweenness centrality measure is defined as:

$$PFC_i^B = \frac{\sum_{i=1,s<t}^{N} PF_i^{st}}{\frac{1}{2}N(N-1)} \tag{3.2}$$

The calculation considers the load that must be dispatched to the targets and, consequently, calculates the flow through all the transmission lines of the network in order to serve them [8]. The PFC_i^B measure describes the fact that power flows along all paths (transmission lines) from source to target not only the shortest ones, and nodes that lie on no path from source to target get a betweenness of zero.

Similarly, a random flow betweenness centrality (RFC_i^B) has been introduced [12] based on the random flow model. Roughly speaking, the random flow betweenness of a node i is equal to the number of times that a random walk starting at s and ending at t passes through i along the way, averaged over all s and t. This measure is appropriate to a network in which information wanders about essentially at random until it finds its target, and it includes contributions from many paths that are not optimal in any sense. Letting I_i^{st} be the current flowing from s to t, through node i, the random flow betweenness centrality measure (RFC_i^B) is defined quantitatively as [12]:

$$RFC_i^B = \frac{\sum_{i=1,s<t}^{N} I_i^{st}}{\frac{1}{2}N(N-1)} \tag{3.3}$$

In this paper, power flow and random flow betweenness centrality measures are evaluated considering the physical characteristics of the transmission network in terms of length, capacity and failure probability of each transmission line, and types of nodes.

4. Application

The artificial transmission network system IEEE 14 BUS [6] is taken as reference case study. The network represents a portion of the American Electric Power System and consists of 14 bus locations connected by 20 lines and transformers, as shown in Figure 1. The transmission lines operate at two different voltage levels, 132 kV and 230 kV. The system working at 230 kV is represented in the upper half of Figure 1, with 230/132 kV tie stations at Buses 4, 5 and 7. The system is also provided with voltage corrective devices in correspondence of Buses 3, 6 and 8 (synchronous condensers). Buses 1 and 2 are the generating units.

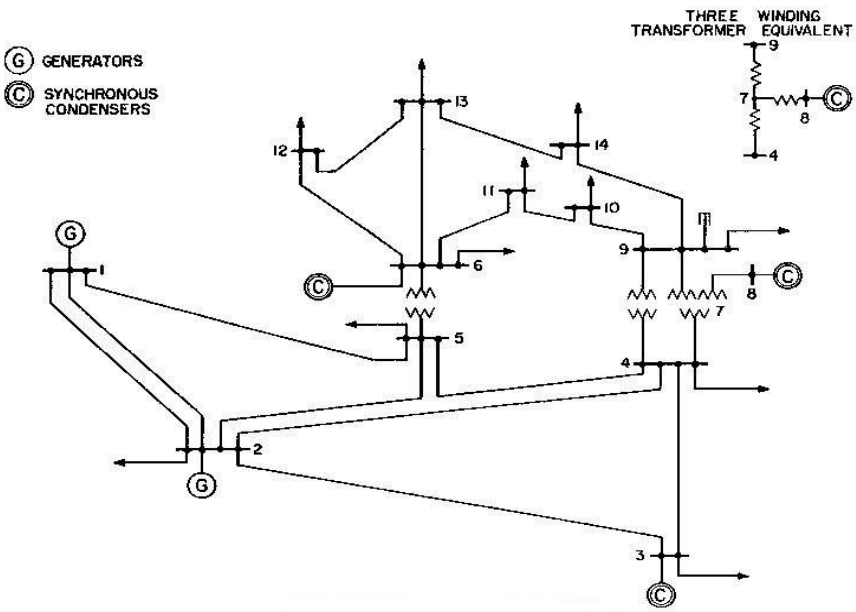


Figure 3. Transmission network IEEE 14 BUS [6].

To carry out the analysis, each network component is transposed into a node or edge of the representative topological network, as shown in Figure 2. Three different physical types of nodes are considered: source nodes (where the electricity is fed into the network), load nodes (where customers are connected) and transfer or transmission nodes (without customers or source).

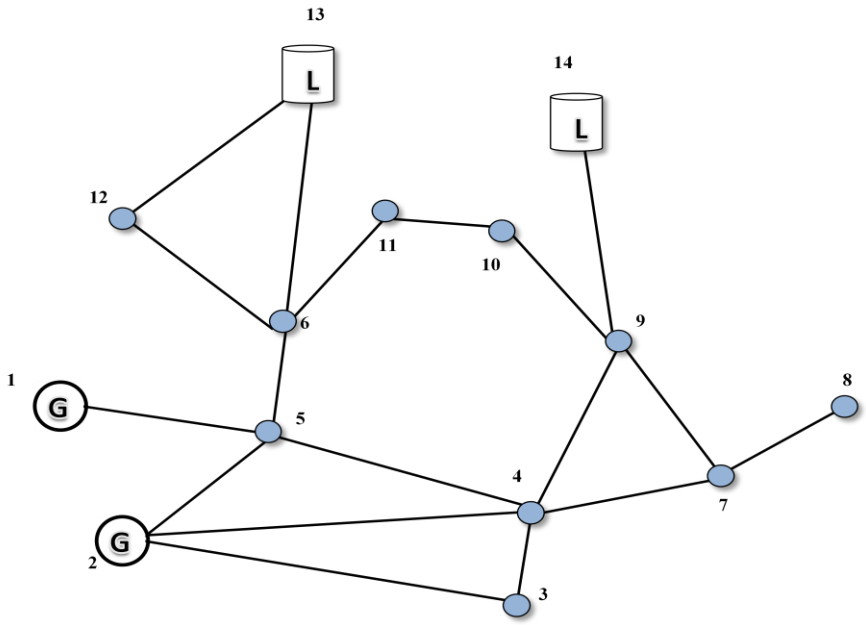


Figure 4. The IEEE 14 BUS transmission network's graph representation [18]. The white circles labeled with **G** represent the generator nodes, the colored circle nodes represent the transmission nodes and the white cylinders labeled with **L** represent the load nodes.

Table I reports the failure rate values of the components of the transmission network, as inferred from literature data [6].

From BUS	To BUS	Failure rate/yr/100km	Failure rate/yr	Equipment
1	2	1.0858		132 kV transmission line
1	5	1.0858		132 kV transmission line
2	3	1.0858		132 kV transmission line
2	4	1.0858		132 kV transmission line
2	5	1.0858		132 kV transmission line
3	4	1.0858		132 kV transmission line
4	5	1.0858		132 kV transmission line
4	7		0.01045	132/230 kV transformer
4	9		0.01045	132/230 kV transformer
5	6		0.01045	132/230 kV transformer
6	11	0.5429		230 kV transmission line
6	12	0.5429		230 kV transmission line
6	13	0.5429		230 kV transmission line
7	8		0.01045	132/230 kV transformer
7	9		0.01045	132/230 kV transformer
9	10	0.5429		230 kV transmission line
9	14	0.5429		230 kV transmission line
10	11	0.5429		230 kV transmission line
12	13	0.5429		230 kV transmission line
13	14	0.5429		230 kV transmission line

Table I. Failure data of the arcs of the IEEE 14 BUS transmission network

The failure probability of edge ij is defined as:

$$q_{ij} = 1 - e^{-\lambda_{ij} T} \tag{4.1}$$

where λ_{ij} is the constant failure rate per unit time of the edge ij linking nodes i and j (column 4 in Table I) and T is the reference time for the analysis, here chosen equal to 1 year.

Because the failure rate data are usually given as functions of the lengths of the transmission lines (column 3 in Table I), in order to compute the failure probability (4.1) transmission line lengths have been obtained from impedance data, as reported in [5].

The lengths of the transmission lines in the IEEE 14 BUS are listed in Table II. Lines containing transformers are considered to be zero-length lines.

From BUS i	**To BUS j**	**Length (km) l_{ij}**	**Failure Probability q_{ij}**
1	2	22	0.2125
1	5	81	0.5768
2	3	72	0.5338
2	4	64	0.4932
2	5	63	0.4884
3	4	62	0.4828
4	5	15	0.1498
4	7	-	0.0104
4	9	-	0.0104
5	6	-	0.0104
6	11	220	0.6970
6	12	283	0.7847
6	13	144	0.5425
7	8	-	0.0104
7	9	-	0.0104
9	10	93	0.3978
9	14	299	0.8027
10	11	212	0.6843
12	13	221	0.6999
13	14	385	0.8762

Table II. Case 1: length and failure probabilities of the transmission lines

Source generation is sampled from a normal distribution with a mean value of 30 and a variance of 100, in arbitrary units (a.u.). The values of the capacities of the network links are assumed all distributed according to a normal distribution of mean value 100 a.u. and a standard deviation of 10 a.u.

The application of the general model follows the procedural flow of Figure 3.

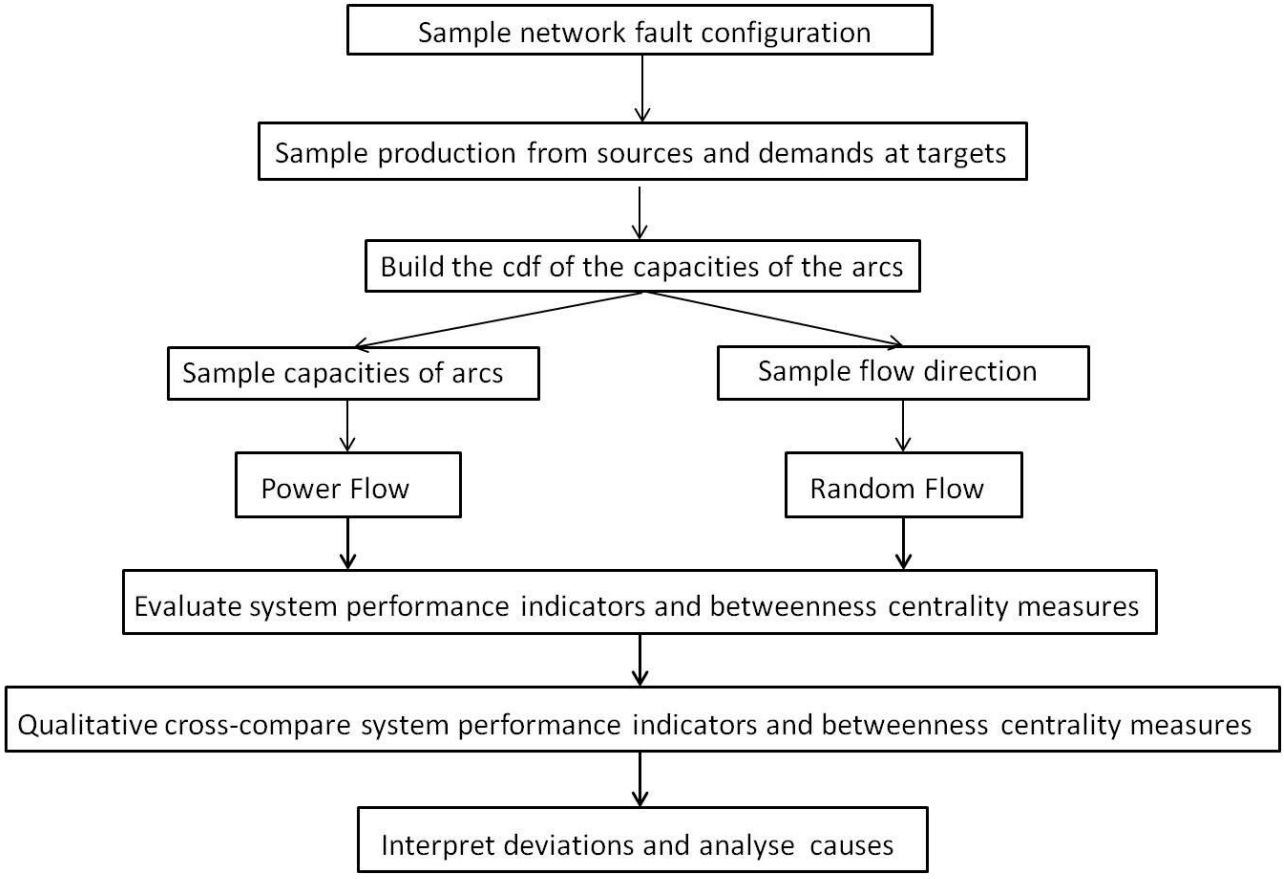


Figure 3. Procedural steps of the general evaluation model.

The network performance characteristics, computed on the basis of the LF model are reported in Table III. For comparison, the values obtained in [5] by the RF model are also reported.

Quantity	LF	RF[5]
Blackout(%)	0.77	0.82
Overload(%)	0.04	0
NDL(%)	100.0	100.0
NRL(%)	15.88	18.22
NLL(%)	84.12	81.78
NSE	0.16	0.18

Table III. Case 1: network performance indicators

To see the effect of the lines lengths, a second computation has been made using average transmission line lengths taken from literature [19]. Two line lengths of 48 km and 50 km have been considered (Table IV). In this case, the lower half of the network has the largest failure probability, while in the previous case the highest failure probabilities concern the upper half of the network .

From BUS	To BUS	Length (km)	Failure probability
1	2	48	0.4079
1	5	48	0.4079
2	3	48	0.4079
2	4	48	0.4079
2	5	48	0.4079
3	4	48	0.4079
4	5	48	0.4079
4	7	-	0.0104
4	9	-	0.0104
5	6	-	0.0104
6	11	50	0.2372
6	12	50	0.2372
6	13	50	0.2372
7	8	-	0.0104
7	9	-	0.0104
9	10	50	0.2372
9	14	50	0.2372
10	11	50	0.2372
12	13	50	0.2372
13	14	50	0.2372

Table IV. Case 2: length and failure probabilities of the arcs

The network performance characteristics are reported in Table V, together with the values obtained by the RF model [5].

Quantity	LF	RF[5]
Blackout(%)	0.04	0.44
Overload(%)	0.76	$3.33*10^{-4}$
NDL(%)	100.0	100.0
NRL(%)	19.87	60.46
NLL(%)	80.13	39.54
NSE	0.2	0.6

Table V. Case 2: network performance indicators

Cross-comparing the results of Tables III and V for the LF and RF models, one sees larger blackout and lower of overload percentages in case 1 than in case 2, both for the LF and RF models. Moreover, in case 1 the LF and RF models show nearly the same service efficiency. On the contrary, differences appear in case 2 with respect to the network service efficiency and the received and lost loads computed by the two models; for example, the LF model has an average lost load higher than the RF model. Indeed, in case 2 the two models behave differently than in case 1. This is due to the fact that in case 2 the failure probabilities of the upper part of the grid, $q_{ij} = 0.2372$, are lower than those of the lower part, $q_{ij} = 0.4079$ (Table IV): compared to case 1, there is then more probability for the flow to reach the target, and hence, the blackout probability is lower and the overload probability increases.

The differences in the values obtained by the LF and RF models in case 2 are consistent with their underlying principles. In the LF model, the load distribution leads the flow towards the targets: since the lower part of the network has lower failure rates than the upper part, there is a larger probability for the flow to reach the targets (global approach to network flow distribution). On the contrary, in the RF model the flow moves locally from node to node, wandering around towards the targets: since the lower half of the network has higher failure rates than the upper half, the random walk of the flow has more difficulty to reach the targets (local approach to network flow distribution).

The difference in the analysis results obtained by the two methods calls for a word of caution to the decision makers who would need to inform their decisions based on results from different methods taking different views in the analysis of the system.

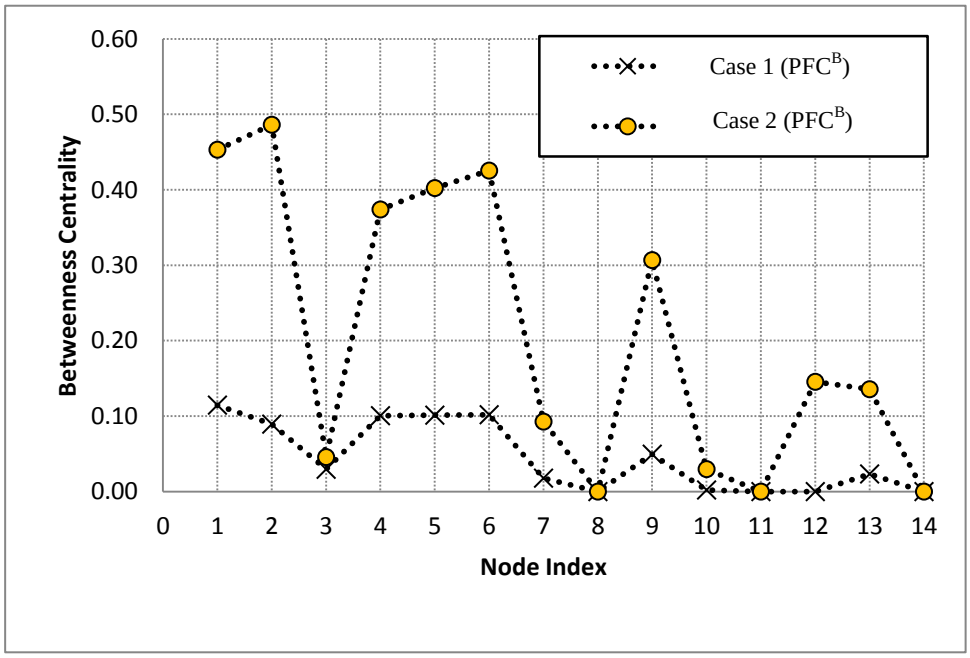


Figure 4. Power flow betweenness centrality (PFC^B) according to the LF model proposed.

Figure 4 reports the values of the power flow betweenness centrality measures (PFC_i^B) as inferred from the LF model calculations, both for the case of the average lengths and the case of the computed lengths. It can be observed that the two cases (with different arc failure probability values) show a similar trend although the magnitudes of the betweenness centrality are different.

Case 1 (×, Figure 4) refers to the case of transmission lines lengths inferred from impedance data: a very small amount of the generated flow is available to the network (Table III) and consequently, the power flow centrality betweenness values are low. In this case, the blackout percentage is high: a very little average flow is available to the network. Case 2 (○, Figure 4) refers to the case of transmission lines lengths. In both cases, nodes 1 and 2 show the highest values: they are source nodes. Nodes 8 and 11 are zero-betweenness centrality nodes; the arcs that connect these nodes to the network undergo failure most frequently. In the lower part of the network, betwenness centrality values are higher than in the upper part. Moreover, node 14 does not receive any flow.

The LF-computed betweenness centralities (PFC^Bs) can be compared to those obtained by the RF model (RFC^Bs) [5]. Case 1 results are shown in Figure 5. Although different in magnitudes, the trends of the LF and RF models betweenness centralities are similar.

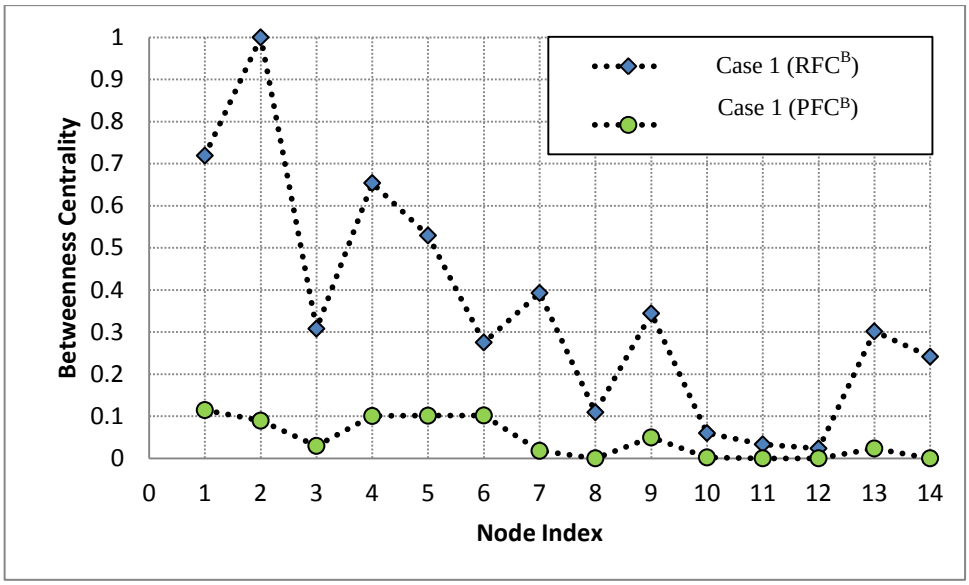


Figure 5. Betweenness centrality: the circles refer to PFC^B values, the diamonds refer to RFC^B.

Case 2 is shown in Figure 6. The PFC^B measure presents the same trend as the RFC^B one, although some differences can be seen for nodes 4, 5 and 6: the RF model suggests a decrease in betweenness centrality going from node 4 to node 6, while the LF model sees an increase in the betweenness centrality for the same nodes. As defined by Eq. (3.2), the RFC^B accounts for the contribution to the flow from every path directed into the node and not only from the geodesic ones: in Figure 2 it can be seen that nodes 4 and 5 connect the two halves of the network. Moreover, node 4 is connected to the

upper half through two links (edge connecting nodes 4 and 7 and edge connecting nodes 4 and 9): these links make a loop and the flow, in the random flow model, is likely to wander around and more flow will pass through them. On the other hand, node 5 connects the lower half of the network directly to the upper one through node 6: in this case, there may be less flow passing through the link connecting nodes 5 and 6.

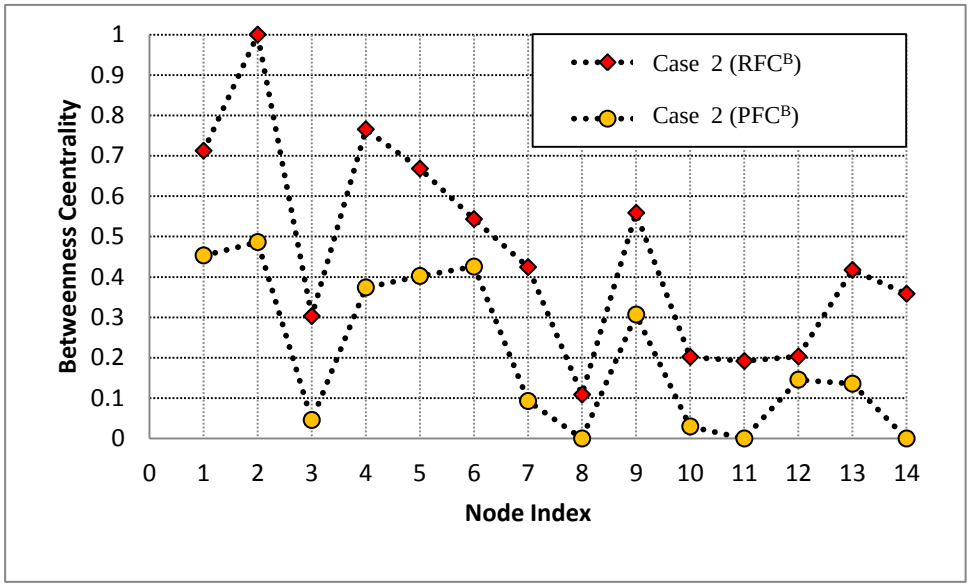


Figure 6. Betweenness centrality: the circles refer to PFC^B values, the diamonds refer to RFC^B.

As before, the differences in the results can be explained in terms of the principles underlying the flow calculation in the two models. In the RF model, the flow starts from the source, follows at random one of the departing links to one of its neighbors, takes this as the source and iterate this process until the target is reached. The random choice of the arc to follow is based on the actual capacity of each arc outgoing from the node: higher capacity arcs have larger probability to be selected as flow carriers.

On the other hand, the LF model calculation considers the load that must be dispatched to the targets and, consequently, calculates the flow through all the transmission lines of the network in order to serve them [20]. Nodes 4, 5 and 6 connect the lower half of the network with the upper half: a reduced flow passing through them prevents the flow coming from the sources (lower half) from joining the target nodes (upper half), especially in the LF model considering that nodes 10 and 11 show nearly no betweenness centrality (Figure 5), that is there is nearly no flow passing from the upper left part (nodes 6, 9 and 13) to the right part (nodes 9 and 14) of the network (Figure 2).

The results are summarized in the following Tables: Table VI qualitatively compares the performance indicators of Case 1 and Case 2 for the LF model and the RF model. Table VII compares the two models for each case .

Performance Indicators	Load Flow (LF)		Random Flow (RF)	
	Case 1 (computed lengths)	Case 2 (average lengths)	Case 1 (computed lengths)	Case 2 (average lengths)
Blackout	Case 1 has an higher percentage of blackout than Case 2		Case 1 has an higher percentage of blackout than Case 2	
Overload	Case 1 has lower percentage of overload than Case 2		Case 1 has lower percentage of overload than Case 2	
NDL	100%	100%	100%	100%
NRL	Low	Low	Low	High
NLL	High	High	High	Low
NSE	Very Low	Very Low	Low	High
PCF^B	Low amount of generated flow available, so PCF^B values are low both for Case 1 and Case 2		–	
RFB^B	–		The amount of generated load is higher in case 2 than in Case 1, so RFB^B values are higher in Case 2 where more flow is available to the network	

Table VI. Performance indicators' overview of Case1 and Case 2 of the LF and RF models.

Performance indicators	Case 1 (computed lengths)	Case 2 (average lengths)
Blackout	The RF model shows a slightly higher blackout percentage than the LF model	The RF model shows an higher blackout percentage than the LF model
Overload	The LF model shows a small overload percentage while the RF model shows no overload percentage	The LF model shows an higher overload percentage than the RF model
NDL	100%	100%
NRL	The RF model shows an higher received load than the LF model	The RF model shows an higher received load than the LF model
NLL	The LF model shows an higher lost load than the RF model	The LF model shows an higher lost load than the RF model
NSE	The LF and RF models show nearly the same service efficiency	The RF model shows an higher service efficiency than the LF model
Betweenness centralities ($PCF^B$and RFB^B)	Similar trends of betweenness centralities, but different in magnitude(RFB^B measures are higher than PCF^B)	The betweenness measures have the same trend, and PCF^B and RFB^B have close values.

Table VII. Performance indicators' comparison of Case1 and Case 2 for the LF and RF models.

5. Conclusions

The load flow and random flow models have been applied to the analysis of power transmission networks with respect to a number of performance indicators. One relevant output derived from the analysis is the betweenness centrality of the network elements, which allows identifying the importance of the different components in determining the flow through the network. In this regard, an original formulation of the betweenness centrality measure has been here introduced, tailored to the power flow description provided by the load flow model.

The LF and RF models have been applied to an electrical power transmission network of literature. For the analysis, each equipment of the system has been transposed into a node or edge of the representative network. Two different cases of transmission lines lengths have been considered: in case 1, the lengths have been obtained from the impedances of the lines; in case 2, line lengths have been taken from literature data.

The analysis results show a good agreement between the two models in the computation of the percentage of blackout and overload for case 1 and a difference in case 2; this is due to the fact that the LF model takes a "global approach" to the flow dispatch while the RF model has a local approach of randomized node-to-node load transfer. The betweenness centrality measures computed by the two models show similar trends, although differences exist in the absolute values. From the computational point of view, the LF model is not as time consuming as the RF model (in the case study presented, the ratio of the computation time is 0.82) ; however, problems of convergence in the calculation process may arise in the LF model when dealing with the stochastic nature of the network: indeed, isolated nodes in a faulty configurations may prevent the algorithm from reaching convergence. Instead, the RF model does not suffer from these convergence problems and has proved to be an effective option to gain relative insights into the complexity of the network system, particularly in terms of the importance of its elements, when local transfer is the governing flow distribution regime.

These considerations would need to be confirmed in the scaling of the methods to larger systems, which is object of further research.

Acknowledgements: the authors are indebted to the anonymous reviewers for their suggestions which have helped improving the paper.

References

[1] Borkowska, B. "Probabilistic Load flow", *IEEE Trans.*, PAS-93, pp. 752-759, 1974.

[2] Dopazo, J.F., Klitin, O.A. and Sasson, A.M., "Stochastic load flow method", *IEEE Trans*, PAS-94, pp. 1551-1556, 1975.

[3] Sobjerajski, M., "A method of stochastic load flow calculation", Archiv für Elektrotecnik , pp. 37-40, 1978.

[4] Leite da Silva, A.M. et. al., "Probabilistic load flow considering network outages", IEE Proceedings, Vol 132, No.3, 1985.

[5] Zio, E., Piccinelli, R., "Randomized flow model and centrality measure for electrical power transmission network analysis" *Reliability Engineering and System Safety*, **95**, pp. 379 – 385, 2010.

[6] The IEEE 14 BUS data can be found on: http://www.ee.washington.edu/research/pstca/.

[7] Grisby, L.L., Power Systems, CRC Press, Boca Raton, 2007.

[8] Zio, E., "From Complexity Science to Reliability Efficiency: A new Way of Looking at Complex Network Systems and Critical Infrastructures", *Int. J. Critical infrastructures*, Vol. 3, Nos. ¾, pp. 488 – 508, 2007.

[9] Freeman, L.C., "Centrality in Social Networks: Conceptual Clarification", *Social Networks*, **1**, pp. 215 – 239, 1979.

[10] Borgatti, S.P., "Centrality and Network Flow", *Social Networks* **27**, pp. 55 -71, 2005.

[11] Freeman, L.C., et al., "Centrality in valued graph: a measure of betweenness based on network flow", *Social Networks* **13**, pp. 141-154, 1991.

[12] Newman, M.E., "A measure of betweenness centrality based on random walks", *Social Networks*, 27, pp. 39–54, 2005.

[13] Eusgeld, I., Koger, W., Sansavini, G., Schapfer, M., Zio, E., "The role of network theory and object-oriented modeling within a framework for the vulnerability analysis of critical infrastructures", *Reliability Engineering & Systems Safety*, **94(5)**, pp 954–963, 2009.

[14] E. Zio and L.R. Golea, "Analyzing the topological, electrical and reliability characteristics of a power transmission system for identifying its critical elements", *Reliability Engineering and System Safety*, **99**, pp. 172-177, 2012.

[15] Lin, Y., Yeh, C., "Maximal network reliability for a stochastic power transmission network", *Reliability Engineering and System Safety*, **96** (10), pp. 1332 – 1339, 2011.

[16] Johansson, J., Hassel, H., "An approach for modeling interdependent infrastructures in the context of vulnerability analysis", *Reliability Engineering and System Safety*, **95**, pp. 1335-1344, 2010.

[17] Rocco, C., M., Ramirez-marquez, J., E., "Vulnerability metrics and analysis for communities in complex networks", *Reliability Engineering and System Safety*, **96**, pp. 1360-1366, 2011.

[18] Pajek, Program for large network analysis: http://vlado.fmf.uni-lj.si/pub/network/pajek/.

[19] Billington, R., Li W., Reliability Assessment of electric power system using Monte Carlo Methods, Plenum Press, New York, pp. 19 – 20, 1994.

[20] IEEE RTS Task Force of APM Subcommittee, IEEE Reliability Test System, *IEEE Transaction on Power systems*, Vol.14, No. 3, August 1999.

Paper V

Unrcertainty propagation in power transmission networks

Enrico Zio, Roberta Piccinelli, Lucia R. Golea, Giovanni Sanavini

Submitted for publication

Uncertainty propagation in electric power transmission systems

(E. Zio[1] , R. Piccinelli[2], L. Golea[2], G. Sansavini[2])

[1]*Chair on Systems Science and the Energetic challenge, European Foundation for New Energy-Electricite' de France, Ecole Centrale Paris and Supelec*

[2]*Politecnico di Milano, Milano, Italy*

*corresponding author:

enrico.zio@ecp.fr, enrico.zio@supelec.fr, enrico.zio@polimi.it

Abstract

The purpose of this work is the characterization of uncertainties in electric transmission networks and of the impact that the propagation of the identified uncertainties has on the reliability of the electric infrastructure. To this aim, we developed a stochastic model that simulates the operations and the line disconnection and reconnection events of the electric network due to overloads beyond the rated capacity. We represent and propagate the uncertainties related to consumption variability, ambient temperature variability, wind speed variability and wind power generation variability. The model is here applied to a case study of literature. Conclusion are drawn on the impact that different sources of variability have on the reliability of the network and on the seamless electric power supply. Finally, we point out a possible system state in terms of a combination of power request and supply that is critical for network vulnerability, and may induce a cascade of line disconnection and massive network blackout.

1. Introduction

Systems of electric power generation, supply, and transmission play an extremely important role in modern societies. They depend on seamless electric power supply as an essential resource for communication, transportation, heating and cooling systems, lighting, and the powering of computers and electronics. Consumers have grown to expect electricity to be available instantaneously with a flick of a switch.

At the same time, providing electricity in a reliable fashion is a complicated and technically challenging task. It involves real-time assessment, coordination, and control of thousands of generating units, the transfer of electric power over networks of transmission lines and, finally, the delivery of electric power to the consumers. The high degree and inter- and intra-connectedness of networks for energy supply makes them vulnerable to global disruptions, when exposed to hazards of

various nature, from random/mechanical/physical/material failures to natural events, intentional malevolent attacks, and human errors.

The assessment of power system reliability is generally divided into two aspects: system adequacy and system security [Billinton and Li, 1994]. System adequacy deals with steady-state operation and planning of the power system, i.e., it gauges the ability of a power system to supply and deliver electric energy to satisfy customer demand. System security gauges the ability of a power system to respond to sudden changes or disturbances such as the loss of generators or transmission lines. Power system security involves two aspects. The first is related to the ability of the system to withstand internal failures and sudden natural disturbances, including network overload, voltage problems, and instability problems. The second aspect is related to the ability of the system to avoid external interference, attack, or coordinated physical assault on the system. Traditionally system planners dealt only with the former security aspect, i.e., problems arising from system operation, random failures of system equipment and natural disturbances [Ivey et al., 1999].

The overarching goal of electric resource planning is therefore to ensure that sufficient resources, delivery capacity, and reliability characteristics exist to meet future demand requirements in a reliable and economic manner [NERC, 2009]. All resource planners allow some percentage reserve margin of capacity above their demand requirements to ensure reliability following unexpected system conditions and to meet state regulatory and regional requirements. Reserve margins are determined by calculating the capacity of supply resources, discounted to reflect the potential unavailability of the resource at high risk times [Ivey et al., 1999]

The analytical processes used by resource planners range from relatively simple calculations of planning reserve margins to rigorous reliability simulations that quantify system Loss of Load Expectation (LOLE) or Loss of Load Probability (LOLP) values. In the latter case, planners periodically check resource adequacy indicated by the evaluated reserve margins through detailed reliability simulations that compare expected demand profiles against forced outage rates of generating units and maintenance schedules to yield LOLE or LOLP values [Billington and Li, 1994]. Moreover, the reliability simulations typically include probabilistic production cost simulations for meeting a specified demand (or chronological) curve from a specified generation fleet while incorporating the forced and unforced outage rates over the simulation period.

Deterministic approaches to power system security usually consider the worst-case scenario. The result of the analysis is most often qualitative and therefore difficult to use in a decision-making process. Deterministic methods also impose a hard limit on system operations. As a result, systems are often designed, planned or operated to withstand severe problems that have a low probability of occurrence.

Deterministic methods alone cannot adequately address the various transmission challenges such as the available transfer capability (ATC), long transmission and related voltage/reactive and security (stability) problems, transmission project ranking, transmission congestion alleviation, uncertainty of

weather, environmental constraints and the competitive environment, uncertainty of customer load demand, uncertainty of equipment failure and operation.

Probabilistic approaches consider factors that may affect the performance of the system and provide a quantified risk assessment using performance indices such as probability and frequency of occurrence of an unacceptable event, its duration and its severity. These performance indices are sensitive to factors that affect system reliability. Quantified descriptions of the system performance, together with external relevant factors such as environment impact, social and economic benefits etc., enter into the decision-making process and have an impact on operations, short-term planning, and long-range planning [Ivey et al., 1999].

In this paper, we address the reliability and safety of a power transmission network due to the uncertainties related to consumption variability, ambient temperature, wind speed variability and to the integration of large shares of electricity produced by wind energy. To this aim, we develop a stochastic framework that combines classical DC power flow, sequential Monte Carlo sampling of the possible uncertain system parameters, and a model for power line disconnection and reconnection due to line overloading beyond the rated capacity. Within this simulation frame of work, we identify and quantitatively propagate the uncertainties related to consumption, wind power generation, ambient temperature and wind cooling of lines, aiming at assess their impact on level II system adequacy [Billinton and Li, 1994] in terms of expected energy not served (EENS) and expected demand not supplied (EDNS). Based on the aforementioned modeling assumptions, we anticipate that the parameters whose uncertainty affects power system operations to a large extent, will have to be paid a special attention during the design and the management of power systems. Finally, we point out a critical system state related to the combination of wind power generation and load request that may lead to a cascade of line disconnection and a large system blackout.

The paper is organized as follows. In Section 2, we detail the power grid DC load flow, the wind turbine and the line overload models which enter the stochastic simulation frame of work. In Section 3, we identify and quantitatively represent the uncertainties related to power systems operations. In Section 4, we exemplify the developed framework with respect to a modified version of the IEEE RTS 96 test network system [Grigg, 1999], and provide reliability and vulnerability considerations about network operations. Conclusion are drawn in section 5.

2. Stochastic Framework of the Power Transmission System

In order to study the effects of the uncertainties (related to the load and the renewable generation forecasts on one side, and to weather parameters on the other side), that introduce disturbances in the grid and may line outages due to overloads, we developed a event-based stochastic framework which simulates the operations of the electric network. Our approach, inspired by the model introduced in

[Dobson et al., 2001; Anghel et al., 2007; Giorsetto and Utsurogi, 1983], combines: 1) a DC load flow algorithm that computes the distribution of power flows using a linear load flow approximation, 2) the contribution of wind generation power in a transmission power grid, 3) a strategy for generation dispatch in order to balance the power production and consumption throughout the network, 4) the dynamics of line temperatures as function of the power flow and environmental conditions (wind speed and ambient temperature), 5) the event of automatic line disconnection when the rated line temperature is reached, and 6) the event of line reconnection.

The evolution of cascading events in their slow initiating stages is described by means of transmission lines failures, caused by line overheating due to excessive power flows. To describe this effect, we monitor the evolution of the line temperature, and its dependence on electric flow redistributions, using the model of heat conduction in rods of small cross section in which an electric current of constant intensity flows [Anghel et al., 2007]. Further contributing to the evolution of cascades is a line restoration model which prevents a damaged line to be put back in service before a fixed restoration time has passed.

The model of transmission line failure due to loading over their transmission capacity and following restoration, is part of the developed event-based stochastic framework which has also the ability to represent daily hourly changes in power requests at customer side of the system, room temperature and wind speed variations.

The stochastic framework is based on sequential Monte Carlo Simulation (MCS) in which the combination of load requests, room temperature, wind power generation, wind speed and network topology is a system realization. Due to the yearly periodicity of the load request and the room temperature average values, each year is considered to be statistically equivalent to one another and the results are provided on the basis of yearly averages. The simulation begins by establishing the load demand, the room temperature and wind speed values. If no line disconnection due to excessive heating occurs, the next event corresponds to the occurrence of the next hourly time step ("next hour" event) with updated load demand, room temperature and wind speed conditions. If the temperature of a line exceeds the critical temperature set for that line, the "line disconnection" event may occur before the scheduled "next hour" event. A DC load flow is performed following the occurrence of each event. The "line reconnection" event occurs after a time chosen a priori for each line that is disconnected.

After each event, we solve the DC load flow equations in order to determine the line temperatures and the type of the next event. The change time to the next event is computed as the minimum between the time to the next hour change, the minimum failure time among all lines, and the minimum time to reconnection of all lines.

A. Formulation of the DC power flow

We assume that the electrical transmission system operates through steady-state conditions, also during the evolution of major disturbances in the system. This approximation does not hold during the

late stages of major disturbance events, and it can be relaxed if voltage dependent phenomena are modeled during these events.

In order to determine the steady-state operating conditions of the power grid, we should solve the full nonlinear power flow equations that provide information about the voltage magnitudes and phases and the active and reactive power flows along each transmission line. Unfortunately, since our simulations involve numerous power flow solutions for a power grid system that evolves in time, solving repeatedly the full non-linear power flow equations becomes computationally prohibitive. Moreover, the full nonlinear equations pose very difficult nonlinear optimization problems. We have therefore chosen to linearize the power-flow equations and to solve instead the so-called DC power flow equations that connect the flow of real power to the voltage phases of the system's buses, which results in a completely linear, non iterative, power flow algorithm [Wood and Wollenberg, 1996].

The DC power flow can only calculate real (MW) flows on transmission lines but it gives no answers to what happens to voltage magnitudes or reactive (MVAR) flows. Assuming that all bus voltages phasors are 1.0 per unit in magnitude, and defining the matrix B by $B_{ij} = -b_{ij}$ and $B_{ii} = \sum_j b_{ij}$, where b_{ij} is the susceptance of the transmission line joining buses *i* and *j* and the summation is over all nodes *j* connected to node *i*, the voltage phases θ_i are the solution of the linear power flow equation $P = B\Theta$. Here, P is the vector whose $N - 1$ components are the real powers injected at each node, except a reference node (slack node) for which the injected real power is computed from the power balance between total generation and total load. The vector Θ is the vector whose components are the voltage phases at each node in the network except the slack node which has phase zero. After solving the power flow equation for the vector Θ, the flow of real power along each transmission line is computed from $P_{ij} = b_{ij}\left(\theta_i - \theta_j\right)$ [Grigsby, 2007].

B. Line temperature model and overloaded-line failure

In order to model the failure of transmission lines due to loading beyond the rated transmission capacity, we consider the problem of heat conduction in rods of small cross-section [Carslaw and Jaeger, 1959] in which an electric current of constant intensity flows. We assume for simplicity that the transmission line is so thin that the temperature at all points of its cross-section is uniform. We suppose that the transmission line has constant cross-section area ω, perimeter *p*, thermal conductivity *K*, electrical conductivity σ, density ρ, specific heat *c*, diffusivity κ. We further assume that the heat flux across the surface of the line is proportional to the temperature difference between the surface and the surrounding medium and is given by $H(T-T_0)$, where T is the temperature of the line, T_0 is the temperature of the medium, and H is the surface conductance. The problem of heat conduction then becomes one of linear heat flow in which the temperature is specified by the time t and the distance x measured along the transmission line. Balancing the total heat generation in an element of volume

bounded by the cross-sections at x and $x+dx$ and the heat in flow across the surface minus the heat loss at the surface, we write the following heat equation,

$$\frac{\partial T(x,t)}{\partial t} = \kappa \frac{\partial^2 T(x,t)}{\partial x^2} + \alpha I^2 - \nu\left(T(x,t) - T_0\right) \tag{2.3}$$

where $\nu = Hp/(\rho c\omega)$, $\alpha = 0.239/\left(\rho c\omega^2\sigma\right)$, $\kappa = K/(\rho c)$ and $I = P/V$ is the current in the line measured in Amperes [Carslaw and Jaeger, 1959]. In order to estimate the surface conductance H, we will assume that the loss of heat across the surface of the line is due to forced convection. When fluid (gas or liquid) at temperature T_0 is forced rapidly past the surface of the line, it is found experimentally that the rate of loss of heat from the surface is given by $H(T - T_0)$, with a value of the coefficient H which depends on the velocity and the nature of the fluid and the shape of the surface [Carslaw and Jaeger, 1959]. For turbulent flow of air with velocity u perpendicular to a circular cylinder of diameter d, $H = 8 \times 10{-}5(u/d)^{1/2}\ cal/(cm^2 secK)$.

Assuming that fluctuations in power flows along the transmission lines propagates much faster than any heat flow transients, and since the heat source is equally distributed along the line, we can neglect the spatial variation in temperature along the line in order to get this simple equation describing the time evolution in the temperature of the line with the time evolution of the power flowing through the line:

$$\frac{\partial T(x,t)}{\partial t} = \alpha I^2 - \nu\left(T - T_0\right) \tag{2.4}$$

If the line is initially at temperature $T(0)$ and the power flowing through the line has the constant value P, the line temperature evolves according to this simple equation:

$$T(t) = e^{-\nu t}\left(T(0) - T_e(P)\right) + T_e(P), \tag{2.5}$$

where

$$T_e(P) = \frac{\alpha}{\nu}\frac{P^2}{I^2} + T_0, \tag{2.6}$$

is the equilibrium temperature that the line will reach when $t \to \infty$. If at some moment the power flow changes, we reset the clock and the initial temperature and use the same equation to describe the evolution of line temperature starting from this moment on.

A transmission line failure due to excessive heating, followed by line sagging and tripping, will happen if the present power flow through the line exceeds the maximum line rating. For each line l, we denote by T_{cl} the equilibrium temperature corresponding to a constant power flow equal to the line rating $P_l^{\max}$, i. e. $T_{cl} = T_e(P_l^{\max})$. When the power flow through the line changes, such that the present power flow P_l exceeds $P_l^{\max}$, the line will start heating toward the equilibrium temperature

corresponding to the new power flow. Since this equilibrium temperature exceeds T_{cl}, at some time t_{cl} the line temperature will reach T_{cl} and the line will fail. The failure time, t_{cl}, measured from the moment when the grid topology and the line flow has changed can be easily deduced from Eq. (2.5) and is given by

$$t_{cl} = \frac{1}{\nu} \ln \frac{T_{cl} - T_e(P_l)}{T(0) - T_e(P_l)} \tag{2.7}$$

Finally, in order to keep the heat equation linear, we have omitted on the right hand side of Eq. (2.3) a cooling term that takes into account that each element of the surface of the road loses heat by radiation to the surrounding medium — and provides cooling when the wind is absent.

Figure 1 shows the dynamics of the power flows (thin curve) and temperatures (thick curve) for one transmission lines during the operations of a power grid. We can see that line temperature reaches its threshold value at about t = 8467 hours due to excessive line flow (beyond scale in figure 1). When the threshold temperature is reached, the line is isolated and the electric flow rapidly drops to 0 MW, followed by line temperature decay to the room temperature by heat convection through the line surface. The line is assumed to be put back in service after 10 hours and the electric flow is restored.

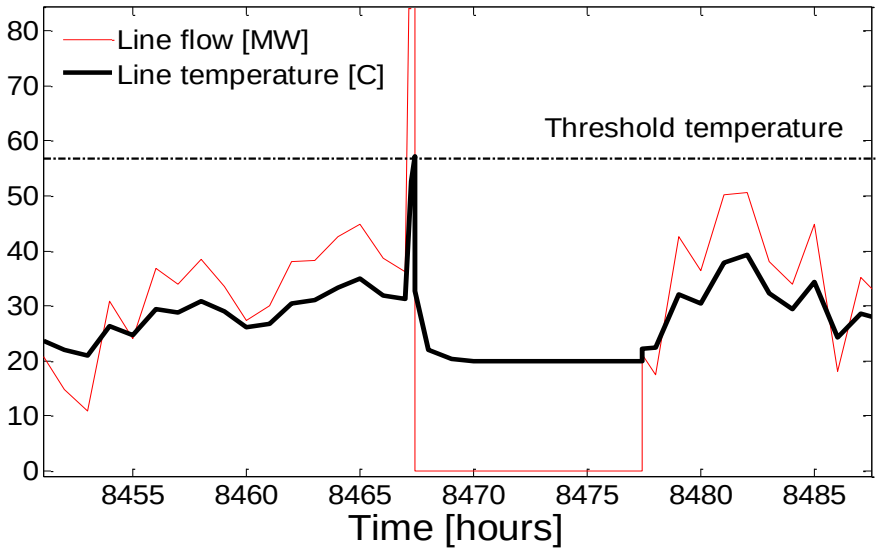


Figure 1. Dynamics of power flow and temperature for a transmission line during network operations

C. Wind turbine Model

The power output from a wind turbine generator (WTG) is determined using the functional relationships linking the characteristics of a WTG and the wind speed field [Giorsetto, 1983]. This function is described by the operational parameters of the WTG. The parameters commonly used are

the cut-in wind speed V_{ci} (at which the WTG starts to generate power), the rated wind speed V_r (at which the WTG generates its rated power) and the cut-out wind speed V_{co} (at which the WTG is shut down for safety reasons). Equation (2.1) [Giorsetto and Utsurogi, 1983] is used to obtain the power output of a WTG from wind speed (SW_t):

$$P(SW_t) = \begin{cases} 0 & 0 \le SW_t \le V_{ci} \\ \left(A + B * SW_t + C * SW_t^2\right) * P_r & V_{ci} \le SW \le V_r \\ P_r & V_r \le SW_t \le V_{co} \\ 0 & SW_t \ge V_{co} \end{cases} \tag{2.1}$$

The constants A, B, C depend on V_{ci}, V_r and V_{co} as expressed in eq. 2.2:

$$\begin{aligned} A &= \frac{1}{(V_{ci} - V_r)^2}\left\{V_{ci}(V_{ci} + V_r) - 4V_{ci}V_r\left[\frac{V_{ci} + V_r}{2V_r}\right]^3\right\} \\ B &= \frac{1}{(V_{ci} - V_r)^2}\left\{4(V_{ci} + V_r)\left[\frac{V_{ci} + V_r}{2V_r}\right]^3 - (3V_{ci} + V_r)\right\} \\ C &= \frac{1}{(V_{ci} - V_r)^2}\left\{2 - 4\left[\frac{V_{ci} + V_r}{2V_r}\right]^3\right\} \end{aligned} \tag{2.2}$$

Figure 1 presents the output power of a WTG:

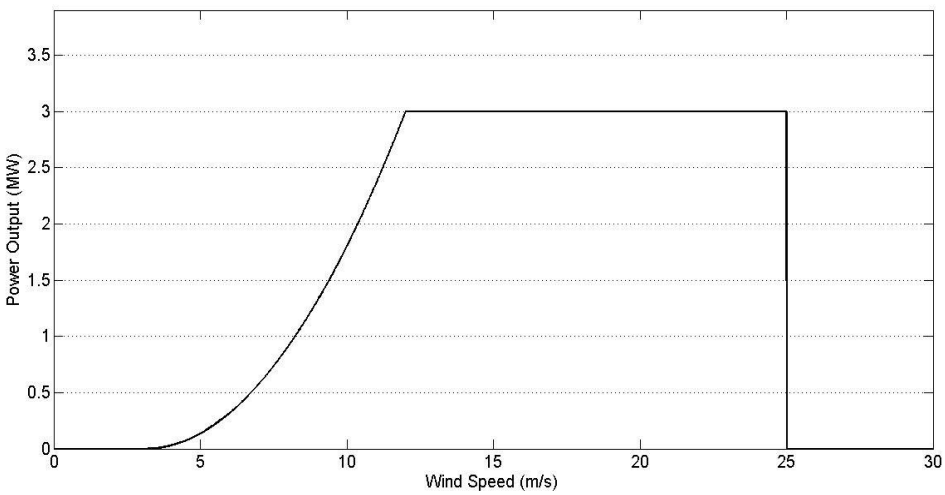


Figure 2. Power curve of a WTG with the following parameters: rated power P_r of 3 MW, cut-in speed, V_{ci}, of 3ms^{-1}, rated speed, V_r, of 12 ms^{-1} and cut-out wind speed, V_{co}, of 25 ms^{-1}

The wind turbine generating unit operates in four phases: a first standby phase in which wind speed is lower than V_{ci} = 3 ms^{-1} and there is no power production; a second phase in which power production increases with a nonlinear trend in the wind speed range from V_{ci} = 3 ms^{-1} to V_r = 12ms^{-1}. Wind turbines usually reach the rated power at a wind speed between 12 ms^{-1} and 16 ms^{-1}, depending on the design of the individual turbine. Finally, when the wind speed exceeds the rated wind speed, the wind generator is disconnected for protection purposes and the power production stops (cut-off phase). Hence, a wind turbine produces its maximum power, i.e. the rated power, within a certain interval that has its upper limit at the cut-out wind speed. Typical values of the cut-out wind speed range between 20 ms^{-1} and 25 ms^{-1}[Ackermann, 2005].

3. Identifying and Classifying Uncertainties in Power Transmission Systems

One of the main purposes of a power system is to satisfy the demands of customer loads in a reliable and economical manner. Failing to properly address planning problems and constraints will eventually yield operation problems and, therefore, will affect power system reliability. For example, failing to incorporate uncertainties in system planning may lead to an overestimation of safety margins and of system capabilities to maintain acceptable levels of reliability.

The appropriate incorporation of the uncertainty and the presentation of its implications are widely recognized as fundamental components in the analyses of complex electric systems [Billington and Huang, 2008]. There are two different forms of uncertainty in power system reliability assessment [Hoffman and Hammonds, 1994]. On the one hand, aleatory uncertainty arises because the system can potentially behave in many different ways. Components' failures and repair processes are random and are sources of aleatory uncertainty. On the other hand, uncertainty enters the system reliability assessment also due to the incomplete knowledge and information on the system and its related phenomena which leads to imprecision in the model representation of the system and in the evaluation of the system parameters. This latter type of uncertainty is often referred to as subjective, epistemic state-of-knowledge [Apostolakis, 1990]. In the field of power system research, epistemic uncertainty has been dealt within the fuzzy power flow analysis [Matos and Gouveia, 2008; Gouveia and Matos, 2009] where the power injections of all loads and generations are regarded as fuzzy variables.

Uncertainty in demand, transmission and generation parameters, line ratings, extreme weather, and other environmental factors, introduce uncertainty in operation and planning of electric networks. In general, the degree of uncertainty increases significantly from a shorter time frame in system planning to a longer time frame in system operation. Therefore, it is of paramount importance to identify and quantify these sources of uncertainty during the design phase of electric networks. Table I summarizes the uncertainties identified in the electrical transmission system. In this study, we represent and propagate the uncertainties related to (I) consumption variability, (II) ambient temperature variability, (III) wind speed variability and (IV) wind power generation.

Element	*Parameter*		*Source of uncertainty*	*Type of available information*	*Uncertainty representation*
Load bus	**Load value**		Consumption variability	Historical data	**Probabilistic** (Normal pdf)
Wind generating unit	Output power	**Wind speed**	Wind speed variability	Historical data	**Probabilistic** (Weibull pdf)
		Operation parameters	Wind power variability	Experimental data	**Probabilistic** (Normal pdf)
Weather	**Wind speed**		Wind speed variability	Historical data	**Probabilistic** (Weibull pdf)
	Ambient temperature		Temperature variability	Historical data	**Probabilistic** (Normal pdf)
Transmission Line	Line temperature		Material properties incomplete knowledge	Experts' judgment	Possibilistic

Table I. Uncertainties sources and their representation in the electrical transmission system

When the uncertainty in the variables is mainly due to their inherent randomness (aleatory uncertainty) and there is sufficient information to assign probability distributions and estimate their parameters, probabilistic modeling is embraced. The model output is represented by a function of *n* random variables, $Y = f(X_1,...,X_i,...,X_n)$, where X_i denotes the *i*-th probabilistic input variable with PDF $p_{X_i}(x)$. The probabilistic model defines the probability distribution of the output random variable *Y* as a function of the probability distribution of the inputs. Such distribution is evaluated analytically in simple cases, or by MCS for more realistic settings.

In power system studies, the MCS is typically embraced, given the large number of variables involved and their complex relationships, which make analytical models difficult or even impossible to derive [Billinton and Gao, 2008; Karki et al. 2010]. The operative procedure of MCS calls for a large number *m* of iterations: at each *e*-th iteration, an input vector of values $(x_1^e, x_2^e,...,x_n^e)$ is sampled from the PDFs of the input variables and a realization of the output value y^e is computed solving the system model. After *m* repetitions, an empirical estimate of the distribution of the system output is obtained.

3.1 Uncertainty representation of the power demand at load buses

The average hourly peak power demand follows the hourly load curve based on data from [Grigg et al. 1999]. The curve accounts for customer power need variations from day to night and from season to season. An example of a daily load peak curve, in different days and seasons, is given in Figure 3.

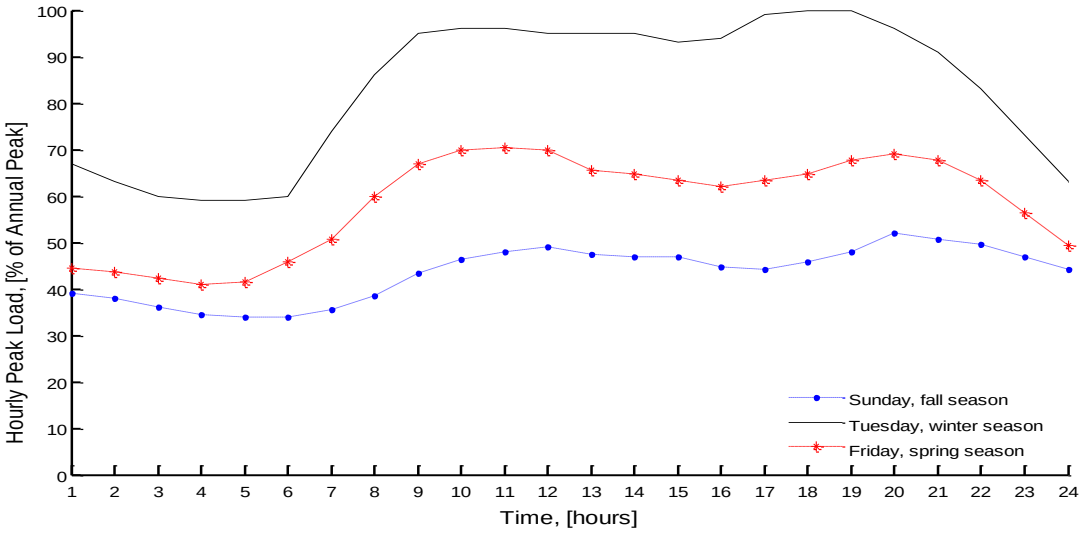


Figure 3. 24-hours load curve (Grigg et al. 1999). First hour corresponds to 12 a.m. – 1 a.m. interval of each day

Uncertainties in the hourly peak power demand arise from because power consumption by users is not exactly uniform and simultaneous, i.e. we assume that the power needed at the customers side may experience stochastic fluctuation from the average hourly peak power demand. Following [Billinton and Li, 1994], it is assumed that load uncertainty is well described by a normal distribution. Therefore, the load hourly values are sampled from normal distribution $N(\mu,\sigma^2)$ with mean μ equal to the hourly peak load considered in the deterministic case (Fig.3) and standard deviation σ assigned according to the perceived load forecast uncertainty, such as 10% of the mean value, $\sigma = 0.1\mu$ [Billinton and Li, 1994].

3.2 Uncertainty representation of the ambient temperature

In order to compute the annual ambient temperature curve (Fig.4), the daily minimum and maximum values during one year in a specific location of the United States were collected and analyzed. A linear variation of the temperature values between the daily minimum and maximum values is assumed, with the minimum and maximum peak registered at 5 a.m. and 4 p.m., respectively.

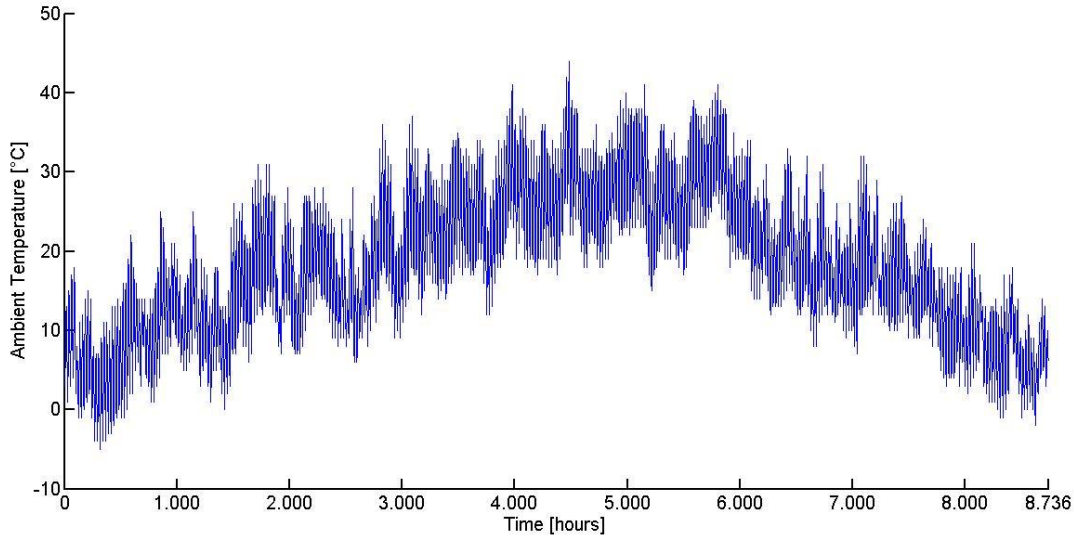


Figure 4 Ambient temperature curve used in the deterministic case. The maximum and minimum data values have been collected at the location of Bakersfield, CA, USA.

We assume that the uncertainty associated with the ambient temperature is well described by a normal distribution. Therefore, the temperature hourly values are sampled from normal distribution $N(\mu,\sigma^2)$ with mean μ equal to the hourly value from the annual ambient temperature curve (Fig. 4) considered in the deterministic case and standard deviation σ equal to 5% of the mean value, $\sigma = 0.05\mu$. This value has been identified by computing the standard deviations of the minimum of the median and of the maximum temperature values that are recorded within each month of the year by choosing the maximum among them.

3.3 Uncertainty representation of the wind speed

In order to compute the annual wind speed curve, the hourly values during a year in a specific location were collected and analyzed. In the deterministic case it is assumed, for each day of the year, that the wind speed is constant throughout the day and it is equal to the daily average value.

Following [Johnson, 1985], the Weibull distribution has been used to represent the wind speed randomness within a yearly time frame. It is shown that data collected at many locations around the world can be reasonably well described by the Weibull probability density function if the collection time frame is not too short, i.e. longer than several weeks. Figure 5 shows the distribution of the hourly wind speed collected at Bakersfield, CA, USA, and the Weibull distribution whose parameters are calculated by maximum likely estimation [Johnson, 1985] based on the hourly values. From the collected data, we notice that either the used anemometer has a lower bound of measuring at about roughly 6 ms^{-1}, or a wind speed below 6 ms^{-1} is an unlikely event at the considered location.

Moreover, we notice that the Weibull approximations holds beyond the maximum of the distribution while it is a rough approximation for lower speed values. Therefore, we expected that the wind speed values sampled from the Weibull distribution will be biased towards lower values if compared to the collected data. Nonetheless, this bias does not affect the modeling framework that is the main objective of the work.

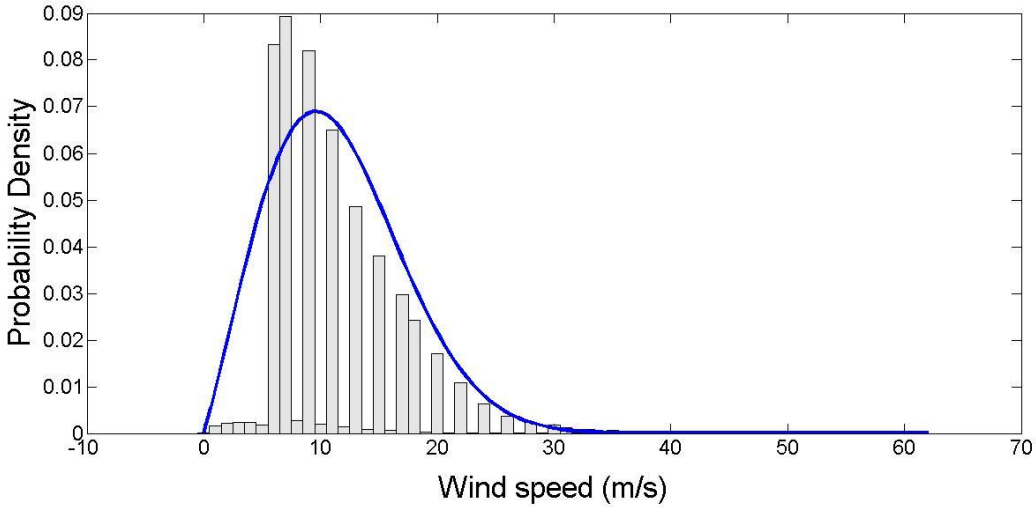


Figure 5. Distribution of the wind speed values collected at Bakersfield, USA, and the corresponding Weibull distribution evaluated through maximum likelihood estimation of the distribution parameters.

3.3 Uncertainty representation of the wind power generation

Finally, the variability of wind speed propagates to the power output of wind generators. The power output of a WTG depends strongly on the wind regime as well as on the performance characteristics and the efficiency of the generator. A fundamental assumption is made when considering the deterministic power curve (figure 2): the relationship between the wind speed and the output power is fixed, given the same type of WTG systems. In other words, the output power of the WTG is always the same at a specific wind speed. In reality, the output power for a fleet WTG of the same type always exhibits considerable variations even when they are operating at the same wind speed [Slootweg and Kling, 2002]. Moreover, Thiringer and Linders [Thiringer and Linders, 1993] analyzed the relationship between the wind speed and the output power based on a group of wind turbines. They found that the powers generated from individual wind turbines of the same type actually vary even at the same wind speed. These research findings suggest that a probabilistic model incorporating the power variations may be more appropriate to characterize the relationship between the wind speed and the actual output powers. Following [Jin and Tian, 2010], the actual output power P_d is proposed as a random variable which is characterized by the mean power output and its standard deviation:

$$P_d(x) = P(x) + \varepsilon \qquad (2.8)$$

where $P_d(x)$ represents the actual WTG power output, $P(x)$ represents the deterministic output governed by the equation (2.1) and ε represents the variation of the power output with $\varepsilon \sim N(0, \sigma_\varepsilon^2)$. Following [Jin and Tian, 2010], we assume $\sigma_\varepsilon = 0.1 \cdot P_r$, i.e. 10% of the rated power output. Since we also considered the uncertainty in wind speed, the function for power curve actually contains two random parameters: the wind speed $x \sim Weibull$ and the variation of the power output $\varepsilon \sim N(0, \sigma_\varepsilon^2)$.

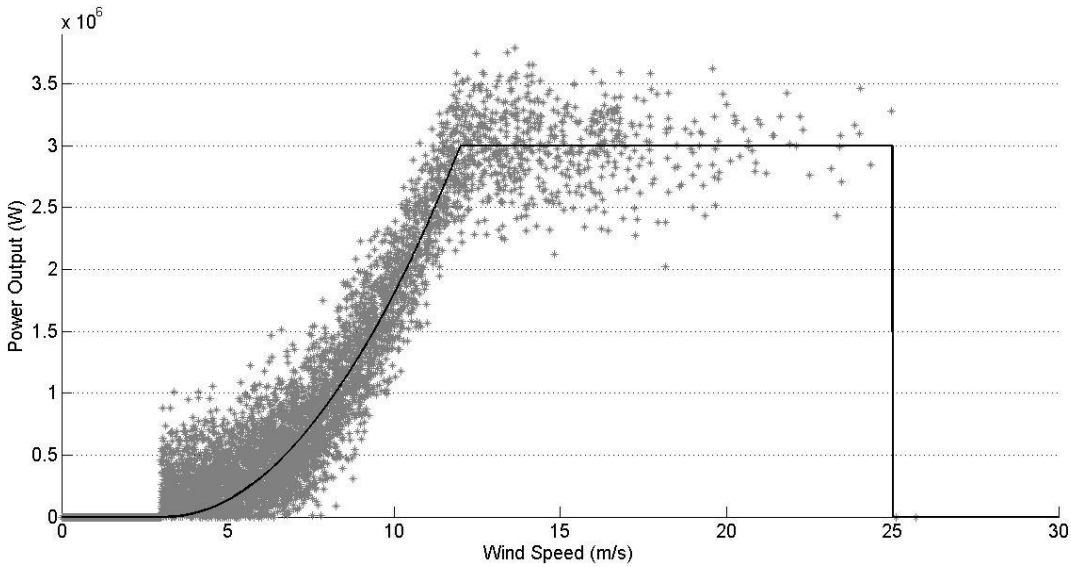


Figure 6. Power output realizations for the wind regime described by the Weibull distribution of figure 5 and the performance characteristic and efficiency of the generator.

Figure 6 shows the power output realizations (grey points) of the WTG of figure 2 for the wind regime described by the Weibull distribution of figure 5 and the performance characteristics and the efficiency of the generator described by $N \sim (0, \sigma_\varepsilon^2)$ [Jin and Tian, 2010]. Wind turbine starts producing power when wind speed equals the cut-in speed of 3 ms^{-1}: the majority of wind power generation concentrates during the nonlinear part of the output curve. This effect is consistent with the uncertainty in wind speed distribution.

4. Application to a modified version of IEEE RTS 96 [Grigg, 1999]

The composite test system IEEE-RTS (Fig. 7) was modified to exemplify the stochastic framework in section 2 with respect to a test system that reproduces the general conditions that exist in actual power systems. The original RTS has a very strong transmission network and a weak generation system. Following [Billinton and al. 2009], in this paper, the original RTS is modified to create a more practical system with a relatively weaker transmission network and a relatively stronger generation system with respect to the original IEEE-RTS test system.

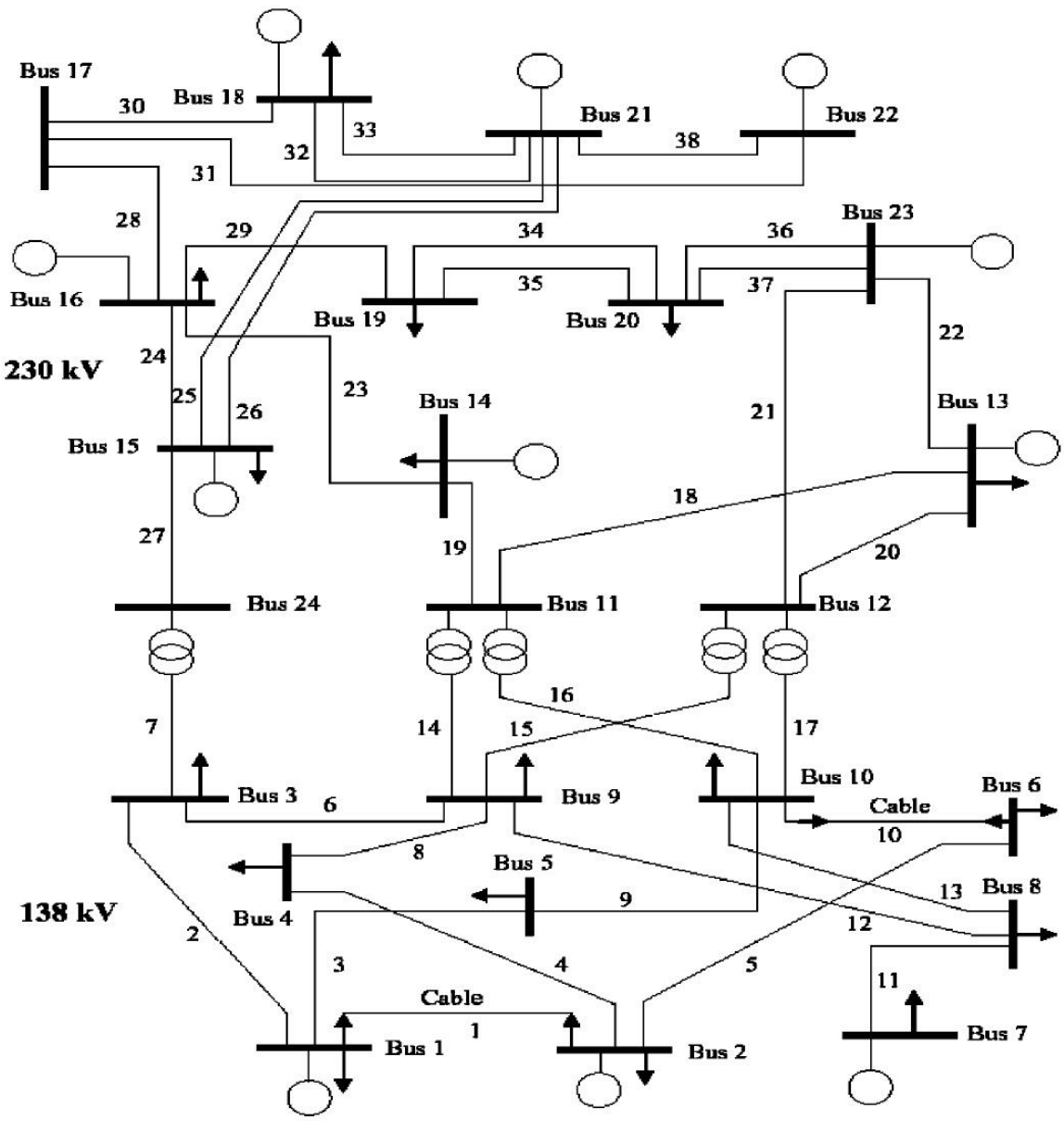


Figure 7. The single line diagram of the IEEE RTS 96/MRTS scheme [Griggs, 1999; Billinton, 2009].

The total installed capacity in the original RTS is 3405 MW in 32 generating units and the peak load is 2850 MW. In the modified version of RTS, henceforth designated as the MRTS, the lengths of all the 138-kV lines (lower part of the system in figure 7) are doubled except for line 10 which is 25.6-km cable. The 230-kV (the upper part of the system in figure 7) lines are extended as follows: the lengths of lines L21, L22, L31, L38 are increased by a factor of three; the lengths of lines L18 to L20, L23, L25 to L27 are increased by a factor of four; the lengths of lines L24, L28 to L30, and L32 to L37 are increased by a factor of six. To increase the utilization of the transmission network, the load levels at all delivery points were increased from 1.3 p.u. to 1.5 p.u. of the original values. The increase of the load levels is balanced by doubling the generating systems capabilities at Buses 16, 18, 21, 22 and 23. The total number of generating units is now 44. The total system capacity is 5320 MW and the peak load for the different load levels is given in Table II:

Load level (p.u.)	Peak load [MW]
1.3	3705
1.4	3990
1.5	4275

Table II. Load levels and corresponding peak loads for the modified power grid MRTS.

The single line diagram of the MRTS is the same as that of the unmodified version of RTS shown in Figure 7. Following the aforementioned modifications, the transmission utilization in the MRTS is significant as a considerable amount of power is transferred from the northern to the southern portion of the system. The modified system is used as a test bed to examine also the effects of uncertainties introduced by adding wind energy conversion systems (WECS) in two points of the transmission network: two additional 300 MW WECS are added through transmission lines at Buses 1 and 3 in the southern portion (138 kV) of the MRTS (Fig. 8).

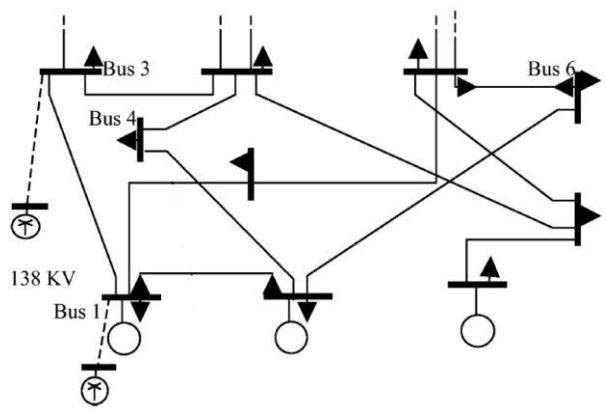


Figure 8. The two 300 MW WECs at Buses 1 and 3 in the MRTS (Billinton, 2009)

For each load level in Table II, four different scenarios are assessed in order to observe the effects of uncertainties and compared to the deterministic base case scenario in which uncertainty is neglected. In order to keep the comparisons consistent, the deterministic base case scenario is characterized by hourly variability of load, hourly variable ambient temperature, and a mean wind speed value which follows the hourly average values, with no associated uncertainty; the additional generators of 300 MW each, are supposed to have no uncertainty associated to their power production.

Following the uncertainties models described in Section 3, the first three scenarios deal separately once uncertainty at a time, with uncertainties in load demand (scenario I), uncertainty in ambient temperature (scenario II), and uncertainties in wind speed and power generation (scenario III). The fourth scenario combines the effects of all the uncertainties (scenario IV). The sensitivity of the annual energy loss with respect to costumers power requests, ambient temperature, wind speed, and wind power generation were quantified and reported in Table III.

Load Level = 1.3 (p.u.)	*Annual Energy Loss EENS[MWh]*	
	Mean	*Standard deviation*
Base Case	-	-
Uncertainty in load demand	85.217	11.523
Uncertainty in ambient temperature	61.574	5.779
Uncertainty in wind speed and power	393.50	19.57
Uncertainty	438.99	17.52

Load Level = 1.4 (p.u.)	*Annual Energy Loss EENS[MWh]*	
	Mean	*Standard deviation*
Base Case	2.0872e+003	-
Uncertainty in load demand	2.6895e+004	3.7602e+003
Uncertainty in ambient temperature	3.6079e+003	3.9286e+001
Uncertainty in wind speed and power	1.0116e+004	8.6805e+001
Uncertainty	9.8776e+003	7.4779+001

Load Level = 1.5 (p.u.)	*Annual Energy Loss EENS [MWh]*	
	Mean	*Standard deviation*
Base Case	2.6769e+004	-
Uncertainty in load demand	1.4636e+005	1.5323e+004
Uncertainty in ambient temperature	4.8056e+004	2.2806e+003
Uncertainty in wind speed and power	6.5395e+004	1.9002e+002
Uncertainty	6.6566e+004	6.7159e+002

Table III. Mean and standard deviation of the EENS for the deterministic base case and the four uncertain scenarios for the three load levels in Table II.

The annual energy loss is quantified by the Expected Energy Not Supplied, i.e. the average EENS index. The EENS index is an adequacy index for the transmission level of the electric infrastructure [Billington and Li, 1994]. It quantifies the annual electric energy that could have been provided by the generating system but that could not reach the customers due to bottlenecks in the transmission network.

Table III shows that the increase of the load level produces an increase in the system annual energy losses both for the deterministic base case scenario, and for scenarios that include uncertainties. When the load level is less than or equal to 1.3 p.u., the system experiences no power loss in the deterministic base case scenario. In particular, when the system assumes the lowest load level (1.3 p.u.), the major contribution to the annual power losses is determined by uncertainties in wind speed and wind power production. For larger load level values, i.e. 1.4 p.u and 1.5 p.u., uncertainties in the load demand cause the largest losses. Since we are considering stochastic simulations over 100 years, it can be expected that the size of the energy losses that are registered in each year may vary consistently. Nonetheless, in all the scenarios, the standard deviation value of the EENS is one to two orders of magnitude smaller than its mean value. It is worth noting that scenario IV that combines the effects of all the uncertainties, experiences an average EENS smaller than the average EENS of scenarios in which the different uncertainty parameters are propagated separately. This is an effect of compensation of the uncertainties that lowers the average energy not served to the customers when considering the simultaneous impact of all the different sources of uncertainty.

In the following analysis, we focus our attention on the largest load level value (1.5 p.u) because it is a good paradigm for systems with a high degree of utilization which operate in stress conditions, and represents the worst-case scenario in terms of energy not supplied. Figure 9 shows the impact that all the identified uncertainties have on the power grid when load level is 1.5 p.u., quantified in terms of power loss. These estimates are average values based on 100 samples, i.e. we simulated 100 years of system operations for each scenario that includes uncertainties. The power loss is quantified by the Expected Demand Not Supplied, i.e. the average EDNS index for each hour of the year. The EDNS index is an adequacy index for the transmission level of the electric infrastructure [Billington and Li, 1994]. It quantified the power not supplied to the network and it is a suitable index when dealing with events. From Figure 9, we notice that power losses occur in a burst fashion during the year with two loss peaks occurring at the very beginning of the year, and in summer; the highest peak load occurs in the week prior to the end of the year. In order to understand this behavior, we compare the EDNS and the hourly peak load curve (inset in Figure 9), which represents the seasonal load profile of system users. In our simulation, we assume that the first hours in the hourly peak load curve correspond to the first hours of the calendar year. We notice that the highest values in EDNS corresponds to the periods in the year where power requests reach the peak value.

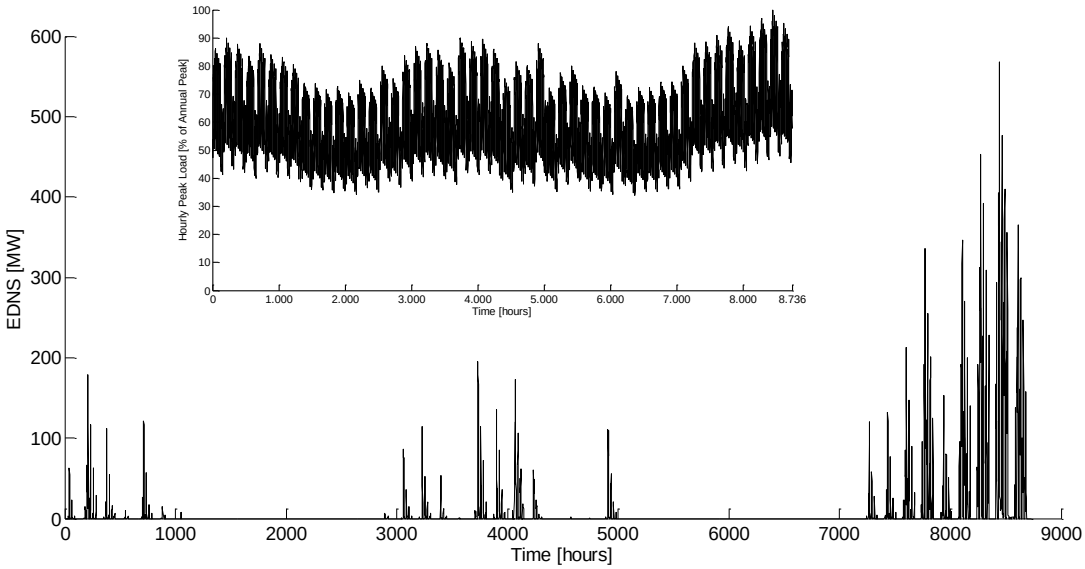


Figure 9. Representation of the hourly average loss of power due to uncertainties for the load level = 1.5. The average is considered over a period of 100 years. The inset represents the hourly peak load curve: the highest losses corresponds to the highest demand from the consumers.

In order to understand the causes for the demand not supplied in figure 9 and to devise operational safety margins which could prevent the occurrences of these losses, we investigate the global system parameter that guide the flow pattern in the system, i.e., the total load requested at all the buses. In figure 10, the EDNS index is expressed as function of the total load requested at all the buses. It shows that there are small power losses, on average, until the demanded load reaches a threshold value of 3600 MW. Above the threshold value, the system loses power proportionally to the increasing load demand (Figure 10).

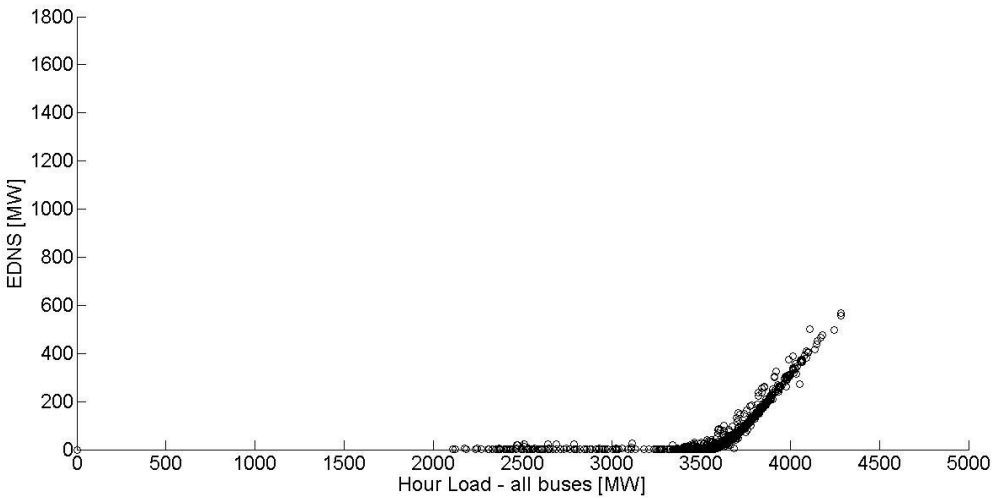


Figure 10. Average power loss as a function of the overall demanded load at all the buses. Each point is an average value evaluated in 100 simulated years

The EDNS index captures the average power losses. It may be expected that there will be years with very few losses and years where losses are significant; rare events, such as cascades, may pass unnoticed in an average analysis. Therefore, we considered the contribution to power losses of every hour load of each year of simulation (Fig. 11). The obtained Demand Not Supplied reveals that the majority of losses take place when hour loads values are around or exceed the threshold value of 3600 MW (bottom right of Figure 11).

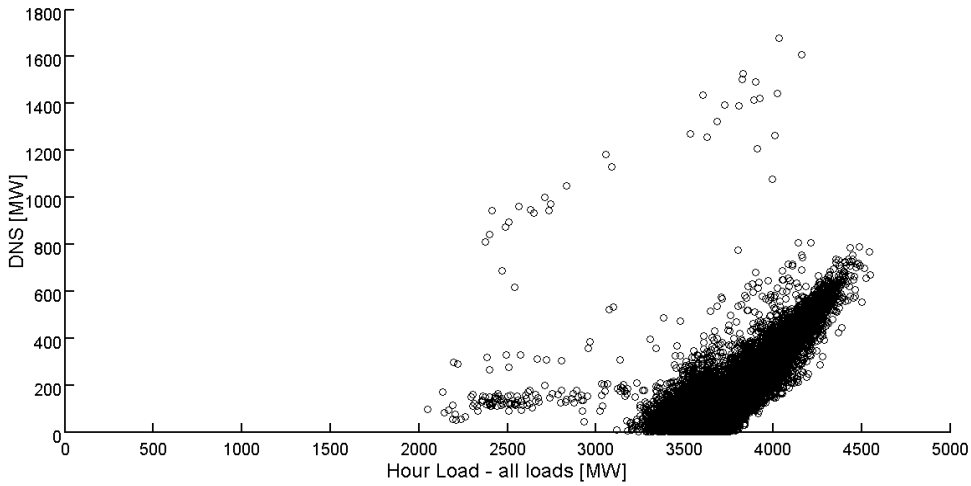


Figure 11. Hourly power loss as a function of the overall demanded load at all the buses

Nevertheless, we notice that the overall load requests alone cannot explain the magnitude of the demand not served. Indeed, some load demands may cause huge power losses as can be seen in Figure 11 (points at the top left and right), some other parameter must also influence the magnitude of the losses. For clarity, we refer to the four quadrants in which different power losses and power requests seem to subdivide the plane (Table IV):

	DNS [MW]	Hour Load [MW]
Quadrant 1	> 600	< 3200
Quadrant 2	< 600	< 3200
Quadrant 3	< 800	> 3200
Quadrant 4	> 800	> 3200

Table IV. Subdivision of the Cartesian plane in figure 11 into 4 quadrants

Quadrants 1 and 2 encompass the same hour load range, but different power loss range and so do quadrants 3 and 4. Quadrants 1 and 4 register the highest losses.

In order to understand which uncertain parameters affect the power losses in the transmission network, the cumulative distribution functions for ambient temperature, wind speed and wind power output in the four quadrants are represented in Figures 12-14.

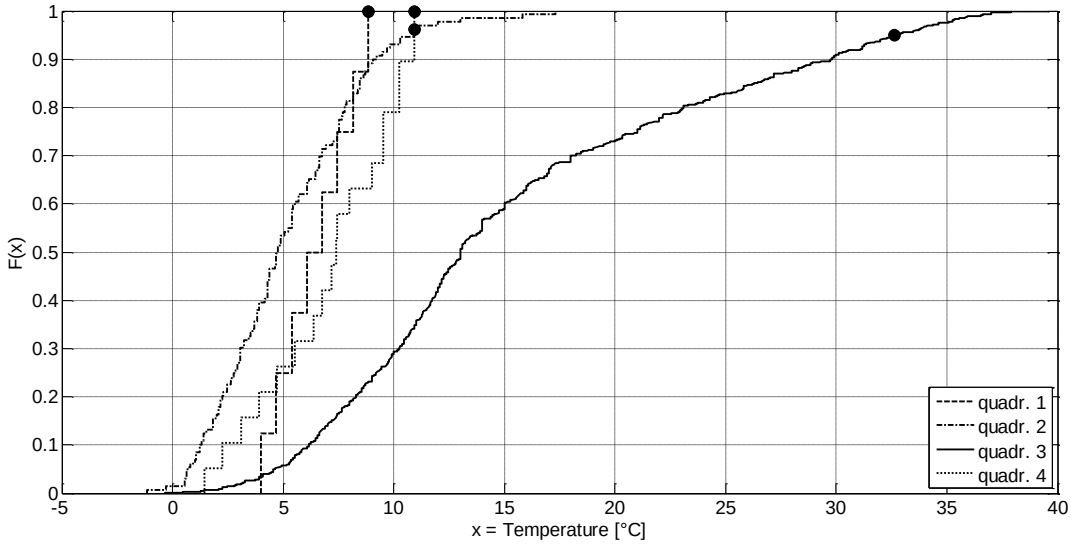


Figure 12. Cdf of ambient temperature distributions in the 4 quadrants. The black points represent the 95th percentile of each cdf.

In figure 12, the 95th percentile of each cdf is represented by a black point. Three of the four cumulative distributions of ambient temperature concentrate at low temperatures values, between 8°C and 11°C. These cdfs correspond to quadrants 1, 2 and 4 (Table IV): the large power losses (quadrant 1 and 4) happen during winter season, when also small losses occur (quadrant 2). It seems reasonable to exclude that hot ambient temperature contributes to the overheating of the transmission lines. Indeed, the cdf of quadrant 3 includes the majority of small losses which occur throughout the year: the temperature values for the cdf range from typically winter ambient temperature values, next to 0°C to summer hot ambient temperature values above 35°C.

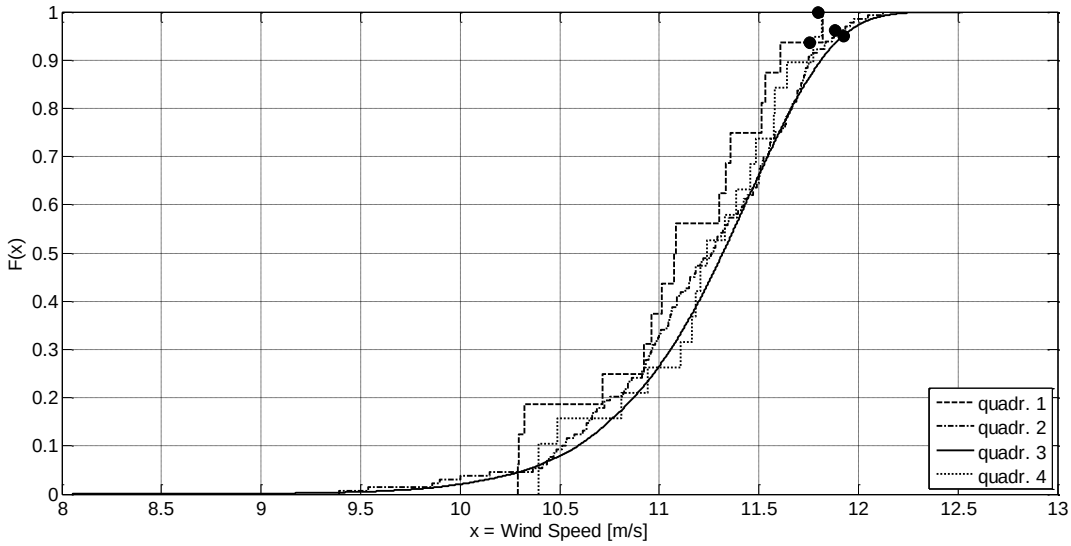


Figure 13. Cdf of wind speed distributions in the four quadrants. The black points represent the 95th percentile of each cdf.

Figure 13 represents the cumulative wind speed distributions. The 95th percentile values concentrate in a very small range of wind speed values, between 11.7 ms^{-1} and 11.9 ms^{-1}. The uncertainty associated to the cooling of the overhead line does not contribute to the power losses.

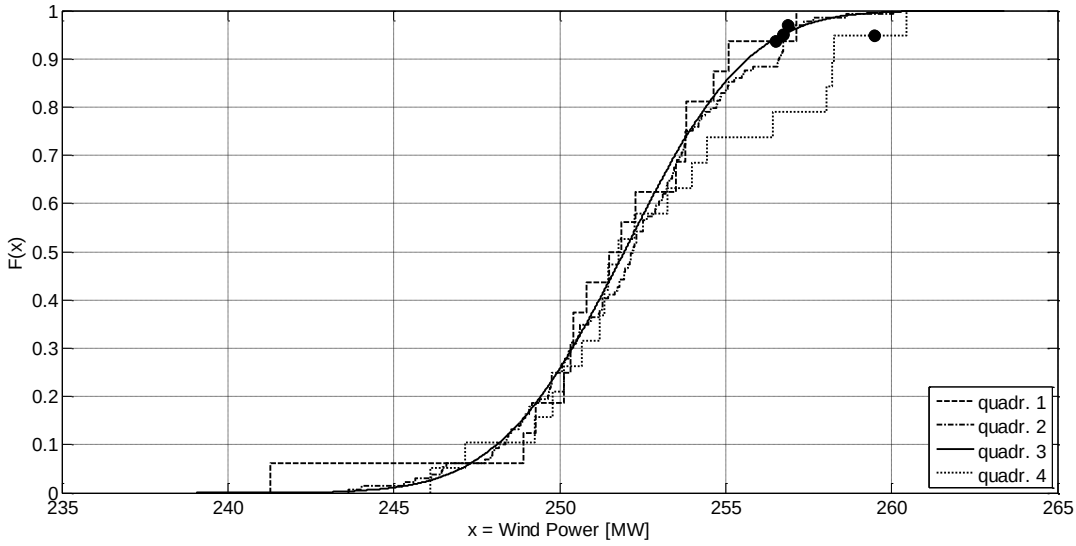


Figure 14. Cdf of wind power generation distributions in the four quadrants. The vertical lines represent the 95th percentile of each cdf.

Figure 14 shows the cumulative distribution for the wind power generation. The cumulative distributions related to wind speed and wind power distribution show similar trends in the four quadrants, although the 95th percentile value corresponding to quadrant 4 is larger if compared to the other three 95th percentiles.

From Figures 12-14, we conclude that no single parameter affects the magnitude of the power losses. Therefore, we investigate the combined effects of multiple system parameters on the power losses.

Cascades in the grid are triggered by overloaded lines: once a transmission line reaches the limit temperature (T_{cl}), the line is disconnected from the grid for a fixed amount of time during which its temperature drops under the T_{cl} value. The transmission lines that disconnect with larger frequency are line 25, line 26 and line 28; as can be seen in figure 7, these lines connect the upper four buses, i.e. bus 17, 18, 21, 22 with rest of the network.

If we define the unavailability of a transmission line as the percentage of time during the year that the line is disconnected from the system, we can see from Table V that, on average, line disconnections last longer when load levels are higher.

	Unavailability (Load level = 1.3)	Unavailability (Load level = 1.4)	Unavailability (Load level =1.5)
Line 25	0.0795	0.2109	0.3452
Line 26	0.2557	0.4188	0.5664
Line 28	0.0790	0.1969	0.3262

Table V. Unavailability of the disconnected lines for the different load levels of demand.

The disconnection frequency analysis reveals that the modified system with load level = 1.5 p.u. is highly stressed and it experiences a random sequence of disconnections of lines 25, 26, 27. When they are simultaneously disconnected, the system divides in two parts: an upper island including buses 17, 18, 21, 23 and a lower island composed by the remaining 20 buses (Fig. 7).

In the first island, composed by the upper part of the network, there are two generating buses, i.e. bus 21 and bus 22, that supply the only left load bus, i.e. bus 18: no losses or further disconnections are registered and the generation dispatch guarantee a balanced power supply in the island.

On the other hand, the lower part of the network has many load buses and a generating system that is no fit to supply enough power in response to the variable load demand. When requested loads in the sub-system reach large values, the demand cannot always be supplied by the generating units and some power demand cannot be served.

In 100 years of simulation, the behavior of the system has been analyzed, and attention has been focused on the same hours of the years when cascades occur only in year 30 and year 98, and on the same hours of years when no contingencies take place, years 2, 14, 31, 42, 54 were taken as examples.

The starting event for the cascades is the disconnection of line 18 always occurring at hour 8443 of the year; then, lines 20, 21, and 29 disconnect in sequence subdividing the lower part of MRTS system in two smaller islands where other disconnections occur propagating the cascade all through the lower

sub-system. Figure 15 shows that load demand in the hour during which the cascade occur, i.e. hour 8443, is large in year 30 and year 98,when the cascade occurs, but the power request is large also for the same hour in year 14 and year 42 when no cascade takes place.

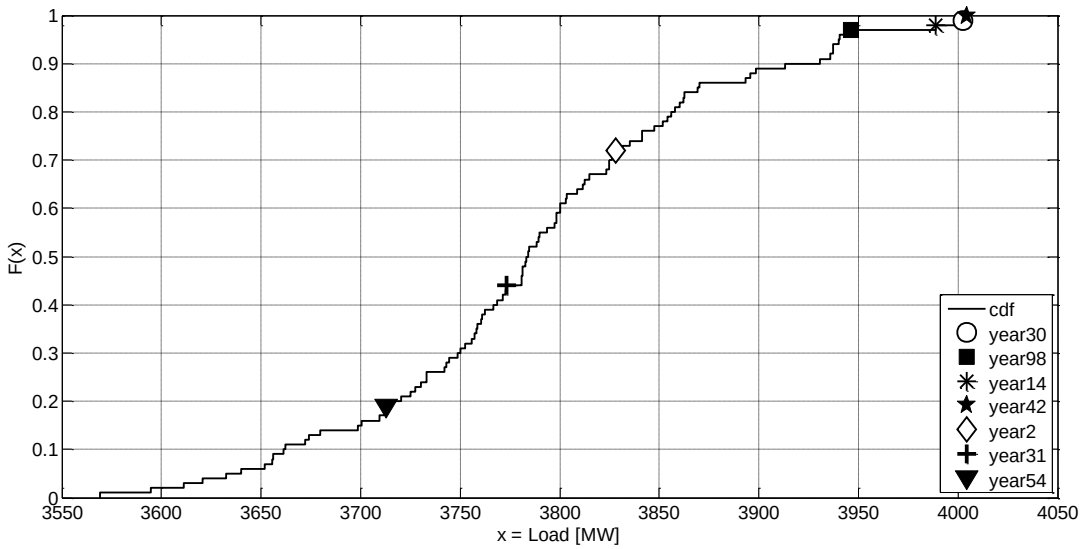


Figure 15 Cumulative distribution function of load demand at hour 8443 .

Yet, from figure 16, we notice that in correspondence of years 14 and 42, wind power generation at buses 1 and 3 is larger than it is in years 30 and 98, when cascades occur.

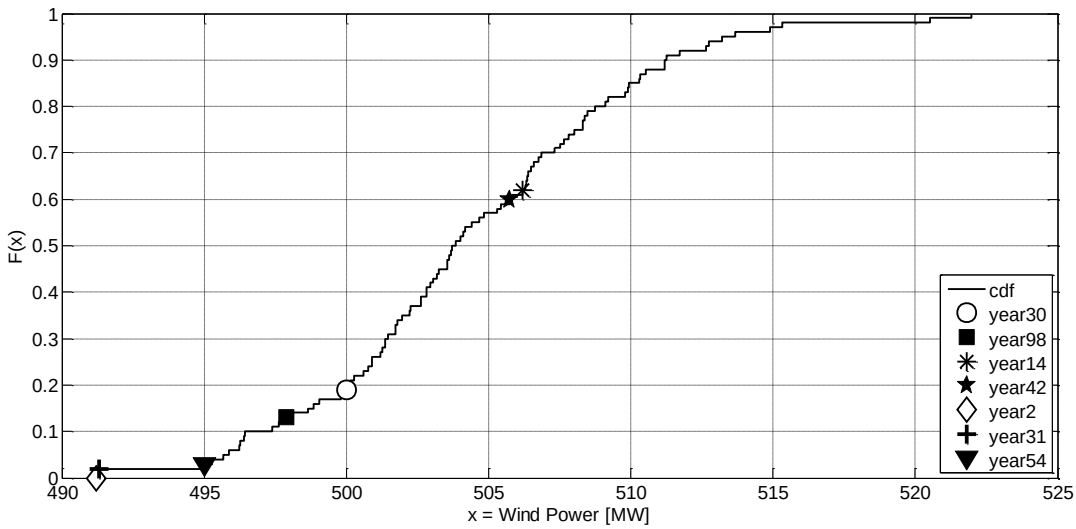


Figure 16. Cumulative distribution function for wind power output at hour 8443.

By the same token, years 2, 31 and 54 record a lower wind generation output than years 30 and 98 (Figure 16) but no contingencies occur because the corresponding load demand is small (Figure 15).

Therefore, uncertainties in the wind conversion system may prevent the transmission in the power grid causing cascades. The DNS values in quadrant 2 and 3 are due to the lack of adequacy of the lower island, i.e. the generation cannot match the demanded power in hours of large power requests. Conversely, the DNS values in quadrant 1 and 4 are due to the incapability of the generated power in reaching the load buses due to a cascade of line disconnections that isolates almost completely every load bus in the lower part of the network. The cascade of line disconnections is triggered by spatial unbalance between power requests and power generation, namely the wind power generation at buses 1 and 3 is not capable of meeting the local power request, and power has to flow from the upper right part of the lower island causing line disconnection.

The interplay between wind power generation and overall power request in the lower island constitutes a probabilistic safety margin which has to be monitored in order to avoid line disconnection cascade propagation and to limit the demand not supplied by the customers.

5. Conclusions

This paper presents the characterization of uncertainties related to composite generation and transmission networks incorporating large-scale wind energy facilities considering wind speed variability, wind power variability, ambient temperature variability and load variability. Transmission deficient environment was created in the MRTS by increasing the system load level and the generating capacity to represent general conditions that exist in actual power systems. The impact that the propagation of the identified uncertainties has on the reliability of the electric infrastructure has been quantified: variability in load and in wind power generation have the biggest impact on the system.

It is worth noting that the combined, simultaneous effect of all the different sources of uncertainty has a smaller impact on system safety in terms of expected energy not supplied, than scenarios in which the different uncertainty parameters are propagated separately. This is an effect of compensation of the uncertainties that lowers the average energy not served to the customers when considering the simultaneous impact of all the different sources of uncertainty.

The increase in the transmission system utilization can lead to cascade events if wind power output cannot sustain the load demand. Indeed, large power requests in the system cause the disconnections of the lines that link the upper part of the network, in which massive generating units are concentrated, and the lower part of the network, thus preventing the power transfer to the area with the highest concentration of buses. When the system tears apart, power request must be supplied by local power generation. In this situations, wind power generation may not be capable of satisfying the local power request causing cascade events of power line disconnections.

The interaction between wind power generation and the overall power request serves as paradigm for the assessment of the safety margins of the system, because no single parameter affects the magnitude of the power losses.

References

[Ackermann, 2005] Ackermann Thomas, Wind Power in Power Systems, Jonn Wiley and Sons Ltd, West Sussex, England, 2005. ISBN 0-470-85508-8

[Anghel et al., 2007] Angel, M., Werley, A.-K. and Motter A. E., Stochastic Model for Power Grid Dynamics, *Proceedings of the 40th Hawaii International Conference on System Sciences*, 2007.

[Apostolakis, 1990] Apostolakis G. E., The concept of probability in safety assessments of technological systems. *Science*, 250(4986), 1359–1364, 1990.

[Billinton et al. 2009] Billinton, R., Gao, Y., Karki R., Composite System Adequacy Assessment Incorporating Large Scale Wind Energy Conversion Systems, Considering Wind Speed Correlation, *IEEE Transaction on Power Systems*, Vol. 24, No.3, August 2009.

[Billinton and Gao, 2008] Billinton, R., Gao, Y., Adequacy assessment of composite power generation and transmission systems with wind energy, *International journal of Reliability and Safety*, vol. 2, No. 1, 79-98, 2008.

[Billington and Huang, 2008] Billington, R., and Huang D. Aleatory and Epistemic Uncertainty Considerations in Power System Reliability Evaluation, Probabilistic Methods Applied to Power Systems. PMAPS '08. *Proceedings of the 10th International Conference on*, 25-29 May 2008.

[Billington and Li, 1994] Billinton, R. and Li, W., *Reliability Assessment of Electric Power Systems Using Monte Carlo Methods*, New York: Plenum Press, pp.229-308, 1994.

[Carslaw and Jaeger, 1959] H. S. Carslaw and J. C. Jaeger, *Conduction of Heat in Solids*, 2nd ed. Clarendon Press, Oxford, 1959.

[Dobson et al., 2001] Carreras Lynch Newmann An initial model for complex dynamics in electric power systems blackouts, *Hawaii international Conference on System Sciences*, January, 3-6, Maui, Hawaii, 2001.

[Giorsetto and Utsurogi, 1983] Giorsetto, P., Utsurogi K.F., Development of a new procedure for reliability modeling of wind turbine generators, *IEEE transaction on Power Apparatus and Systems*, Vol. PAS-102, No.1, January, 1983.

[Gouveia and Matos, 2009] Gouveia E. M., Matos M. A., Symmetric ac fuzzy power flow model. *European Journal of Operational Research*, 197(3), 1012–1018, 2009.

[Grigg, C., P. Wong, et al, ., 1996] Grigg, C., P. Wong, et al., The IEEE Reliability Test System-1996. A report prepared by theReliability Test System Task Force of the Application of , -

ProbabilityMethods Subcommittee, *IEEE Transactions on Power Systems*, Volume 14, Issue 3, age(s):1010–1020, 1999.

[Grigsby, 2007] Grigsby, L., L., *Power System*, CRC press, Taylor and Francis Group LLC, 2007.

[Hoffman and Hammonds, 1994] Hoffman F.O. and Hammonds J.S., *Propagation of Uncertainty in Risk Assessments: The Need to Distinguish Between Uncertainty Due to Lack of Knowledge and Uncertainty Due to Variability*, Risk Analysis, Vol. 14, No. 5, pp. 707-712, 1994.

[Ivey et al., 1999] Ivey M., Akhil A., Robinson D., Stamber K., Stamp J., Consortium for Electric Reliability Technology Solutions Grid of the Future White Paper on Accommodating Uncertainty in Planning and Operations, Prepared for the Transmission Reliability Program Office of Power Technologies Assistant Secretary for Energy Efficiency and Renewable Energy U.S. Department of Energy, 1999.

[Jin and Tian, 2010] T. Jin and Z. Tian, "Uncertainty Analysis for Wind Energy Production with Dynamic Power Curves", , *IEEE 11th International Conference on Probabilistic Methods Applied to Power Systems (PMAPS)*, San Marcos, TX, USA, 14 – 17 june 2010.

[Johnson, 1985] Johnson., G.K., *Wind Energy systems*, Electronic edition, 2006.

[Karki et al. 2010] , Karki, R., Billinton R., Incorporating wind power in generating system reliability evaluation, *Int Journal of system assurance Engi. Manag*, 2, 120-128, 2010

[Thiringer and Linders, 1993]T. Thiringer, J. Linders, "Control by variable rotor speed of a fixed-pitch wind turbine operating in a wide speed range," IEEE Transaction on Energy Conversion, vol. 8, no. 3, 1993, pp. 520-526.

[Matos and Gouveia, 2008] Matos, M.A., Gouveia, E.M., The Fuzzy Power Flow Revisited. *IEEE Transactions on Power Systems*, 23(1), 213 – 218, 2008.

[NERC, 2009] NERC North American Electric Reliability Corporation, *Special Report: Accommodating High Levels of Variable Generation*, April 2009.

[Slootweg and Kling, 2002] J.G. Slootweg, W.L. Kling, "Modeling of large wind farms in power system simulations", IEEE Proc.-c, Power Engineering, PP.503-508, 2002.

[Wood and Wollenberg, 1996] A. J. Wood and B. F. Wollenberg, *Power Generation, Operation, and Control*, 2nd ed. John Wiley & Sons, New York, 1996.