PREDICTABILITY LIMIT OF CASCADES IN INTERCONNECTED MODELS

By

Xin Lin

A Dissertation Submitted to the Graduate

Faculty of Rensselaer Polytechnic Institute

in Partial Fulfillment of the

Requirements for the Degree of

DOCTOR OF PHILOSOPHY

Major Subject: COMPUTER SCIENCE

Examining Committee:

_____

Boleslaw K. Szymanski, Dissertation Adviser

_____

Gyorgy Korniss, Dissertation Adviser

_____

Christopher Carothers, Member

_____

Malik Magdon-Ismail, Member

Rensselaer Polytechnic Institute
Troy, New York

April 2017
(For Graduation May 2017)

CONTENTS

LIST OF TABLES

LIST OF FIGURES

ACKNOWLEDGMENT

I am using this opportunity to acknowledge everyone who helped me during my graduate studies. Without their help, I cannot finish my thesis. First of all, I wish to express my most heartfelt gratitude to my advisor Prof. Boleslaw Szymanski. I appreciate the chance to work with him in the field of complex network. In the past few years, he gave me substantial valuable advice, teaching me to be a better researcher. At each stage of my studies, I always received his assistance to overcome challenges in my research and daily life. His rich experience in the academia, deep insight of the entire industry, innovative ideas in research design, and rigorous working style all left a deep impression on me and pointed out the correct direction for me to move forward.

Also, I am very grateful for my co-adviser Prof. Gyorgy Korniss. He gave me constructive opinions to improve my research and helped me solve research problems. In addition, I would like to thank the rest of my doctoral committee, Prof. Malik Magdon-Ismail, and Prof. Christopher Carothers. I appreciate their help and time for my thesis. I also need to thank my colleagues, Alaa Moussawi, Noemi Derzsy, Andrea Asztalos, and Sameet Sreenivasan. It is a very unforgettable and rewarding experience to work with them.

Last but not the least, I would like to express my deepest love and gratitude for my parents. Their generous supports encourage and inspire me greatly for a long time.

ABSTRACT

In recent years, cascading failures in the complex network, which cause enormous damages in the real world, have been studied intensively because most critical information and infrastructure systems belong to the complex network, such as the Internet, finance, and power grid systems. The spreading of cascading failures could be significantly affected by the topology, initial attacks, and evolution dynamics of an interconnected network. It is now evident that higher connectivity makes the whole system more vulnerable to disruptions. Hence, there is a strong desire to probe the properties of cascading processes and predict the damage in advance. In this thesis, we analytically formulated abstract networks consisting of conceptual objects, such as the global risk network and the fire propagation model. We also studied a real-world system, which is the European power grid.

In the abstract models, we proposed an innovative methodology to quantitatively analyze the cascading failures in a system consisting of multiple non-homogeneous Poisson processes. In this thesis, we formulated an Alternating Renewal Process (ARP) model of global risks in cross domains to capture the actual dynamics of the cascade propagation. In our methodology, we simulated the time series of discrete states with different lengths of time and fit the training data to estimate the parameters of our model. We formulated a practical methodology to simulate complicated processes with hidden variables. The recovery of the hidden and explicit parameters of the model enables the predictions of the activation of global risks. The critical infrastructures are sensitive to cascading failures, which cause a huge threat to the global stability. Hence, assessing the reliability of parameter recovery is important.

Next, to verify the parameter recovery using the method of maximum likelihood estimation (MLE), we compared the estimated parameters with ground-truth parameters in another similar model, which simulates the fire propagation in a city. In this model, we demonstrated that the convergence and asymptotic properties of the maximum likelihood estimation are consistent with the theoretical analysis.

This study provides a quantitative perspective of cascading evolutions and identifies the detrimental risks. In the interconnected network with discrete states, our methodology delivers a great estimation of evolution dynamics and makes accurate predictions of future activities.

In addition, we also studied the cascading failures in a real-world resistor network model, which considers the spatial constraints of the European power grid. In the cascades triggered by a single-node removal, we introduced the stochastic capacity allocation to mitigate the potential damages. Moreover, in the cascades triggered by multiple-node removals, we detected a bimodal distribution of cascading damages as a function of total load and degree of initiators. If the spatial constraints of initiators decline, the cascading failures exhibit more randomness patterns, and the complicated topology is more difficult to predict the cascades than it is in standard structures. Decision makers can benefit from our analysis to design efficient mitigation strategies to protect the vulnerable parts of the system.

CHAPTER 1
INTRODUCTION

The 2008 financial crisis emerged originally in the real-estate market and the banking system in the United States. However, this crisis was not limited to a single country. The successive impacts spread to other countries and caused heavy damage in many aspects of the society, such as the labor market, government credits, manufacturing industry, energy supply, and so on. Presumably, the high connectivity between various aspects of our society exacerbates the spreading of a crisis from a localized area to global-wide. This kind of behavior is called "cascading failures" since the initial breakdown of small components may cause a catastrophe to the complex network [1]–[3]. Therefore, the cascading failures drew much attention in the past decades, which have been studied in many specific domains, such as the standard networks [4]–[6]; bank systems [7], [8]; power grid [9]–[11]; the Internet [12], [13]; traffic systems [14], [15]; and interdependent networks [16]–[19].

In the beginning, we introduced many recent studies, which intensively discuss the abstract and real-world networks. Some studies focus on the cascading properties of particular networks, for example, detecting the relationship between the cascading damage and the initially failed node, or revealing the critical values for the number of initiators to collapse the entire network. Other studies emphasize the robustness of the network, predictability, and mitigation of cascading failures. Although the background and essential details of the previous studies are quite different, these rigorous studies create a broad foundation for this thesis.

Today, the improvement of technology and science makes the world smaller and increasingly connected. Although a more connected system possesses a higher efficiency, this interconnected system may have a greater vulnerability than before, and the materialization of a risk may trigger other connected risks and accelerate the spreading of cascading failures [20]. The globalism and heterogeneity of the entire network make the whole system more vulnerable to a catastrophe [21]. Thus, it is crucial to study these cross-domain systems. Inspired by the previous studies, we

built a quantitative model for global risks, and the detailed information is provided in following sections. In this thesis, we formulated an abstract cross-domain network consisting of 50 global risks from 5 domains: economic, political, societal, environmental, and technological. Experts from the World Economic Forum evaluated the likelihood, impact of future occurrence and connections of each risk [47]. We used the expert assessments to build the structure of the global risk network and detect the evolution dynamics of the model from historical occurrences of risks. Then, we demonstrated the activity level of the network (the number of active risks) at each time step of the stochastic process and detected the most persistent risks based on the simulation. When risks from different categories are connected, the coupled global risk network is a huge potential threat to the world. The primary purpose of this model is to accurately predict the future behaviors of global risks. Although not all risks and events can be precisely predicted, forecast plays a significant role in mitigation strategies.

After we introduced the methodology to analyze the cascading failures in the global risk network, a new question naturally comes up: how to evaluate the quality of our prediction? To answer this question, we applied our methodology to a fire propagation model in an artificial city to probe the limit of predictability. The city structure includes three types of houses: small, medium, and large; each type of house has unique properties. The quality of predictions relies on the approximation of control parameters in our model. Since there are discrete states for each house in the network, and we assumed the state transitions follow Poisson process, the maximum likelihood estimation (MLE) is suitable for parameter recovery from historical observations [22]–[25]. As the input dataset increases, the estimated parameters have a more accurate approximation to the ground-truth values, and the variance of multiple estimations shrinks quickly in a power-law decay, which is consistent with the asymptotic limits of MLE [22]–[24]. This study gives us a better understanding of our methodology, verifies the estimation precision of control parameters, and indicates the limit of our predictions of cascading failures.

Last but not the least, cascading failures not only exist in a global-wide network across multiple domains but also occur in a specific field, such as the electricity

system [9]–[11]. We analyzed the cascading failures in the European power grid system with geometrical constraints. The complicated topology significantly influences the damage. Different initiators selected from the same region could lead to opposite results. The cascading damage is either small or large in different scenarios. Simply increasing the capacity of the system does not save more components since the removals of the "fuse" nodes could block the spreading of cascading failures. We introduced a stochastic capacity allocation for each node to mitigate the damage efficiently. In the cascades triggered by regional attacks, the cascading size presents a bimodal behavior as a function of the total degree and load of the initiators. This study provides valuable analysis to manage potential risks and identify the dangerous parts of the system.

This thesis is organized as follows: Chapter 2 presents the literature reviews for interconnected networks and cascading failures analysis. Chapter 3 discusses the global risk network, such as the definition, dynamics, data fitting and cascade simulations. In Chapter 4, the predictability limits of the abstract model are studied to show the quality of prediction using the maximum likelihood estimation. In Chapter 5, we analyzed the cascading failures in a spatially embedded resistor power grid system and detected the relationship between cascading damage and topological properties. The last chapter contains the conclusion of this thesis.

CHAPTER 2
LITERATURE REVIEW

2.1 Cascade in abstract networks

2.1.1 Standard model

Watts studied the general features of a systemic cascade on an arbitrary random graph [4]. Two important features are derived from observations: first, a little disruption or shock could trigger a systemic cascade in extreme cases; second, such a systemic cascade occurs rarely. A normally working system repeatedly faces potential shocks or attacks. A robust system can endure the influence of such shocks. However, if the system cannot properly withstand the shock under the abnormal state, the subsequent collapse destroys the whole system. Hence, one interesting question arises: how severe is the damage of cascading failures? To answer this question, we need to understand the distribution of cascade damages or losses in a particular network. Generally, nodes and edges are critical to determining the features of cascades. The author pointed out two possible distributions for the cascading sizes. Network connectivity mitigates the severity of successive failures. When the network connectivity is dense, nodes have weak influence on each other. In scale-free networks [26], few nodes have an extremely large connections, which act as hubs of the network and are critical to the global cascades instead of the small-degree nodes. To mitigate the damage, stakeholders need to control these highly connected nodes. If the network connectivity is sparse, the cascading failures are determined by the stability of each node, and the distribution of cascade size approximates a bimodal distribution. In small-world networks [27], large part of nodes possess a similar value of connections. The failures of average-degree nodes cause catastrophic losses. In this case, the system resists most small triggers but fails to handle massive disruptions.

To verify this conclusion, *Watts* used a binary decision model [4]. In economic and societal information networks, individuals often make a binary choice based on their neighbors' decision. For instance, a person is more likely to choose

the restaurant that has high review scores and recommendations, specifically by friends. Hence, the word-of-mouth effect plays a significant role in influencing everyone's decision. Because of the incomplete information, each individual has to rely heavily on others' decision. This social behavior exists in many specific networks and demonstrates a threshold feature. If the threshold is exceeded, an individual abandons current selection for a more appealing one, such as when stock traders make transactions, and customers find restaurants with good reviews on Yelp. In this binary model, each node has a particular threshold ϕ, which is the fraction of neighbors with the same state. The values of the threshold and degree for each node are drawn from designated distributions. The main idea is to build a random network with a heterogeneous connectivity and thresholds. This model is different from other models for several reasons. First, the threshold effect just has a local influence, which is unlike the epidemic model threshold with a global impact. Second, the threshold is the fraction of neighbors rather than the frequency. Third, the degree distribution is heterogeneous, which is different from the uniform distribution in a regular lattice. By tuning the parameters of the network, it is possible to simulate various types of random networks: sparse or dense connections, high or low thresholds. The author pointed out a compromise for enlarging and declining the heterogeneity of the binary network. A heterogeneous threshold makes the system vulnerable, in contrast, a heterogeneous degree distribution makes the system stable.

Dynamics of many social networks exhibit the threshold effect, and one initially failed node may be not sufficient to trigger others. *Centola* indicated that increasing the structured randomness could reduce the probability for global cascades [5]. The authors built a similar model to previous research [4], which studies the threshold effect in the model dynamics. The difference is that active nodes will stay unchanged, and there is no recovery in this model. The simulation is implemented in a two-dimension lattice which has the small-world network features between regular and random structures [27]. To increase the randomness of the topology, the authors implemented two methods: rewire and swap. In the first method, for each edge of node i, there is a probability p for this edge to connect

with other nodes. In the second method, two edges swap ending node and reserve the degree with a probability of p. For example, edge e_{ij} and e_{kl} switch the ending node j and node l. Then, new edges e_{il} and e_{kj} are obtained in the network. In a small-world network, there are 10000 nodes with an average degree of 8. *Centola* varied the value of threshold T and probability p to detect in which cases a global cascade would occur. The results show that a small value of threshold T makes the network more vulnerable to a global cascade. As the value of threshold T increases, it takes a longer time for a cascading process to materialize every member in the network. Moreover, the severity of materialization in the small-world network is larger than that in a regular network eventually. As the value of p increases, more nodes update their connections, which makes the uniform degree distribution more heterogeneous. However, the additional randomness does not accelerate the propagation of cascading failures. Because after rewiring and swapping, fewer nodes share common neighbors, it is less likely for a node to exceed its threshold. This conclusion is consistent with previous study [4], which states that a heterogeneous degree distribution makes the network resistant to small disturbances.

In the small-world networks and random networks, most nodes have the number of connections close to the average value, and the degree distribution has a peak at the average value and decreases exponentially before or after the peak. Besides, another important category is the scale-free network, where the degree distribution exhibits a power-law behavior. Few nodes have extremely large connections acting as hubs in the network. This skewed distribution reflects that the scale-free network has more heterogeneity than the small-world network. *Albert et al.* investigated how initial failures impair the structure of networks [6]. The first network is the Erdős-Rényi (ER) network [28], [29], which has N nodes, average degree of $\langle k \rangle$, and a probability p to connect pairs of nodes. The second network is a scale-free (SF) network. The size and average degree of network are identical in these networks. The authors built this network iteratively. At every step, there is a new node adding to the network with a fixed degree of m. The probability to add a new edge is determined by the degree of old nodes. If an old-node has more connections, the new-node has a higher chance to connect with this node. Since every new node

starts with m edges, the degree distribution has the following expression [6]:

$$p(k) = 2m^2/k^3, \tag{2.1}$$

where the exponent value of the power law distribution is 3. *Albert et al.* used the diameter d to reflect the performance of the entire system, which is defined as the average length of all shortest paths in the network. A large d means two nodes are distant in the network. The growth of d could quantify the damage of the cascading process. If the cascading failures destroy numerous intermediate nodes, the value of d becomes larger, which makes a longer path for communications. The authors assumed two cases of initial attacks: intentional failures and random failures [6]. It is interesting to find how different initial attacks damage the network. There are four cases: the ER model with random failures, the ER model with intentional failures, the SF model with random failures, and the SF model with intentional failures. At first, the authors revealed the relationship between the network diameter d and the initial fraction f of removed nodes. The value of d increases gradually as the value of f becomes larger in the ER model, which is the same for both triggering strategies. Intentional and random removals do not differ much while the system changes. However, in the SF network, the diameter d in the case of intentional failures is much larger than that of random failures. The gap between diameter in these cases has a positive correlation with f. Regarding the relative size of the surviving giant component G, the ER model exhibits a monotonic decrease as a function of the initial-attack proportion f for both cases. When f exceeds the critical point, there is a phase transition for G, which drops dramatically and becomes very close to zero. This trend is similar to the SF model with intentional failures. However, in the case of the SF model with random failures, the value of S decreases slowly against f, without a phase transition behavior. This is because the SF network has a more skewed degree distribution. The intentionally attacked nodes act as hubs in the network, whose failures impact tremendous healthy nodes. However, the randomly attacked nodes are likely to have few connections, which make them less likely to trigger a global cascade. *Albert et al.* concluded that the SF network has a higher tolerance to random failures compared with the ER network. Moreover, the

intentional failures in the SF network are critical for a systemic cascade since the deliberately targeted attacks make the network more vulnerable, which is defined as a "robust-yet-fragile" feature [6], [9] in the SF networks. The heterogeneous topology makes the system stable for random disruptions, but very fragile for deliberate attacks.

2.1.2 Flow-based model

So far, we introduce studies focused on the standard network whose evolution dynamics is based on the influence of neighbors. In reality, cascades also exist in various flow-based networks, such as the transportation system (traffic flow) [15], the power grid (current flow) [9]–[11], and the Internet (information flow) [12], [13]. The allocation and spreading of flow realize the functionality of these networks.

Motter et al. analyzed the cascading failures in flow-based networks [9]. In the cascading process, at each time step, overloaded nodes and edges are removed from the system, and the load distribution is recalculated. The authors assumed that for any two nodes (a source-target pair), one unit of flow travels through the shortest path. After averaging on all possible source-target pairs, the authors counted the number of shortest paths including the node i, which is defined as the load L_i. The capacity of node i is defined as [9]:

$$C_i = (1 + \alpha)L_i, \tag{2.2}$$

where α is the tolerance parameter describing the excessive capacity of node i. In the end, the load distribution remains unchanged, and there are no further failures. However, once certain parts of the system are removed, the length of the shortest path enlarges leading to a raise of load. *Motter et al.* considered a single node removal and formulated three initial attack strategies to trigger the cascade. The principal purpose is to detect how cascading failures develop as a function of time due to different initial attacks. The authors compared results of a uniform network with a scale-free network. The former network consists 5000 nodes with the same degree of 3 for all nodes. The following network has an exponent value of 3 for the power-law distribution, the average degree of 3.1, and 5000 nodes [9]. The value

of G reflects a monotonic increase against the variable α. In distinct initial-attack scenarios, the relationship between G and α is different. *Motter et al.* concluded that intentional attacks, such as the high degree or load nodes, are more likely to cause a catastrophe in a heterogeneous network. The authors applied the same analysis to the Internet network and a power grid system. Similar behaviors are observed in these systems which endorse the authors' conclusion.

Crucitti et al. also proposed a study on the cascades in a flow-based network [10]. The flow dynamics is similar to the previous research [9]. There are three innovations in this model. First, the variable *efficiency* (E) is used to measure the size of a cascade [31]. Second, if node failures occur, there is no removal of the failed nodes or edges. The authors assumed that flow can avoid the overloaded nodes and choose alternative paths. Third, at time step t, each edge has an efficiency value $e_{ij}(t)$ in the range $[0, 1]$. When $e_{ij}(t)$ is 0, the flow cannot travel along the edge between these two nodes. If the value is 1, the edge load of $e_{ij}(t)$ does not have any loss. The network efficiency can be calculated using the following formula [31]:

$$E(G) = \frac{1}{N(N-1)} \sum_{i \neq j \in G} \frac{1}{d_{ij}} \tag{2.3}$$

where d_{ij} is the minimal distance between node i and node j, N is the size of the network. Based on the definition of E, *Crucitti et al.* considered all possible shortest paths. The higher efficiency means a faster communication velocity. The variable $L_i(t)$ has the same definition as the previous study [9]. At time t, the edge efficiency $e_{ij}(t)$ updates its value according to the following formula [10]:

$$e_{ij}(t+1) = \begin{cases} e_{ij}(0)\frac{C_i}{L_i(t)} & \text{if } L_i(t) > C_i \\ e_{ij}(0) & \text{if } L_i(t) \leq C_i \end{cases} \tag{2.4}$$

where C_i is the capacity for node i, $e_{ij}(0)$ is the initial edge efficiency before the cascade. When the load is larger than the link capacity, the edge efficiency is adjusted to a smaller value. Hence, the excessive flow will reallocate to other possible paths.

Crucitti et al. compared cascades in an Erdős-Rényi (ER) network [28], [29]

and a Barabási-Albert (BA) network [30]. The sizes of network and connections are the same for both networks. To trigger the cascading failure, the authors removed a node either randomly or intentionally. The authors showed how the efficiency of network E varies with the value of α. In the ER network, E is very sensitive to the increase of α. A small value of α can cause a dramatic jump for the value of E. This behavior is similar to the cascade triggered by a random node removal in the BA network. However, in the BA network with intentional removals of the highest load node, the cascade is persistent in the network when the degree distribution is extremely imbalanced. The authors also applied the same analysis in two real world networks. Similar behaviors are observed and consistent with the results in BA network. *Crucitti et al.* concluded that in the scale-free network, single node removal, such as the highest load or degree node, could cause a catastrophe to the entire system. Oppositely, the network withstands most random attacks very well. The heterogeneous topology determines the likelihood of cascades instead of the dynamics of model evolution.

Wu et al. studied the features of cascading process in a flow-based network with a community structure [11]. Individuals from the same community have similar behaviors. The connections inside one community are much stronger than those between communities. It is interesting to study how the cascading process propagates within one community or across multiple communities. The authors focused on the scale-free network (Barabási-Albert network) with a community structure, which is a dynamic network with the preferential attachment. If an existing node has more connections, the probability to connect with the new node is higher. The number of communities is denoted as c. At every step, one new node is connected to n nodes inside the community and $m - n$ nodes outside, totally m edges. The power-law degree distribution has an exponent of 3 as shown in [11]: $p(k) \propto k^{-3}$. The community is measured by the modularity Q, which is defined in [32]. A higher value of Q (approaching 1) means a stronger connected community. If Q approaches 0, it is hard to observe any communities in this network.

In this flow-based model, *Wu et al.* designed two methods to trigger the cascade [11]. The first method is to eliminate the edges inside a community with

the highest load. The second method is to remove the edge between communities with the highest load. After the removals of edges and nodes, the load distribution is recalculated. The value of load $L_{ij}(t)$ is proportional to the total number of the shortest paths that passing through edge e_{ij}. In this study, the authors varied the value of modularity Q and compared different network structures. Based on the simulation results, *Wu et al.* stated that a system with a high modularity possesses to an excellent stability and resistance against the cascades. As the entire network shrinks, the surviving giant component decreases gradually as a function of time. Interestingly, both initial removal strategies could lead to a global cascade in certain situations. To mitigate the cascading failures, decision-makers should increase the modularity of communities and the tolerance parameter of the network.

2.2 Cascade in real-world networks

In addition to the important studies on standard networks, various real world networks also draw substantial attention recently. Related studies concentrate on the features of cascades in power grid, the robustness of the system and mitigation strategies.

2.2.1 Power grid networks

Our daily life relies heavily on the infrastructure networks, such as transportation network, communication network, water system, energy supply and electric power [14]. Among these critical infrastructures, power grid system has been studied intensively because of its significance to our society.

Hines et al. analyzed the historical recordings of blackouts in North America during the year $1984 - 2006$ [33], [34]. The authors indicated that the blackout size and cascading frequency have a power-law relationship. Most blackouts just impact small-scale users, but few power outages could cause devastating damage to the entire system. The blackout damage is measured by the reduction of electricity supply and the number of impacted customers. The extreme weather and the equipment malfunction are the main reasons to cause blackouts. Moreover, the occurrence of blackouts is also affected by the demand of power during a day and

year. At the peak time of a year, such as the summer and the winter, the likelihood of a blackout is substantially larger, which means the power grid faces enormous disturbances proportional to the demand growth. Although the whole system tolerates small-scale disturbances well, suddenly occurred large-scale disturbances make the system very vulnerable to a wide-spreading blackout. However, the authors found that the blackout sizes does not have a high correlation with the duration time. The worst scenario could be a large blackout that lasts for a long time. Large-scale and lone-term blackouts could cause tremendous loss to our society.

Dobson et al. formulated a threshold model for cascading failures [35], [36]. In this model, the threshold value of capacity L_{fail} is the same for all nodes, and the initial load L_i of node i is randomly selected from a range $[L^{min}, L^{max}]$. During the cascading process, once the load of node i exceeds the threshold capacity, node i is failed, and a fixed amount of flow P is redistributed to the remaining parts of the system. Initially, a disturbance flow D is imposed to the network to start the cascade. The authors tried many scenarios with different parameter values. Intriguingly, by adjusting the average value of the initial load L_i, the probability of overload size exhibits either an exponential or a power-law decay. Most cascades just cause small-scale damages, but rare cases could destroy the whole system. *Dobson et al.* also showed the critical point for the average initial load to cause a phase transition. Beyond the critical point, the average number of overloaded nodes increases dramatically from a small value up to the system size.

2.2.2 Internet

Cohen et al. probed the stability of the Internet with respect to random and intensional attacks [12], [13]. The connectivity of the Internet shows a power-law distribution [12], [13]:

$$P(k) = ck^{-\alpha}, \quad k = m, m+1, ..., K \tag{2.5}$$

where k is the degree of a single node on the Internet, m is the lower bound for the connection, and K is the upper bound. The authors detected how the relative size of the surviving giant component G is affected by the fraction of the initial removal

nodes p. In the case of random removals and $\alpha \leq 3$, the Internet is very robust to the random disturbances, and the whole system stays connected [12]. However, in the case of removing nodes with high degrees, the network is more vulnerable than before. When the fraction of removed nodes p reaches the critical value p_c, the remaining system is divided into smaller components reflecting a phase transition. The critical value p_c decreases gradually as the exponent value α increases, which indicates that the robustness of the Internet has a better performance in a more skewed degree distribution. Also, the value of G shows a monotonic decrease when p increases, so the network with higher capacities may result in more severe damages. Hence, the Internet with a skewed degree distribution is more vulnerable to the intentional attacks rather than random attacks.

2.2.3 Financial networks

Extensive studies have been done in financial networks because of their ubiquitous influence on our everyday life. Among the various critical factors of economic systems, the network topology draws much attention. *Roukny et al.* analyzed how different factors affect the cascades in a banking system, and detected the importance of the topology of an interconnected network [7]. The authors concluded that a network with a divergent structure is either more stable or more vulnerable than a uniform network. The influence of the network topology enhances when the liquidity level deteriorates, and more strongly connected components become infected by their neighbors. The nodes in the network represent commercial banks, and directed edges represent the direction of investments. The authors defined the ratio between the net capital and total assets of bank i to be the robustness η_i of a bank [7]. When the invested institutions fail, the investor banks undergo a devaluation of their assets and robustness to some extent. If the total assets of one bank are less than the debt or liabilities, the bankruptcy occurs and leads to a lower liquidity level of the market since it is hard to find investors to buy assets so that the assets have to be sold under market price. Hence, the authors varied the key factors of the model to simulate the cascading failures in different scenarios. These key factors include the initial attack strategy, the correlation between the node degree and node robustness,

the connectivity, and the liquidity of the market.

Roukny et al. demonstrated how the cascading failures change with the average degree k and the average robustness m. They identified the phase boundary between the global cascade (damage more than 90% of nodes) and the local cascade as a function of k and m. The negative correlation between k and m shows that when the value of k is small, we need a large value of m to avoid the global cascade. As k increases, the critical value for m decreases gradually, which means a stronger connection is more stable than a sparser structure. This trend is the same for three different networks: the uniform, random and scale-free network. However, distinct topologies react differently against initial attacks. When nodes are randomly removed to trigger the cascade, the phase boundary of three topologies looks almost the same. When the highest degree node is removed intensionally, in the scale-free network, the average robustness m is larger than that of other topologies, which means the scale-free network is more vulnerable to initial shocks. The other two networks have a similar trend in the case of intensional attacks. The random network has a slightly higher robustness level to make a phase transition than the uniform network. When considering a situation of less liquidity and random attacks, the phase boundary between the global cascade and the local cascade as a function of k and m is non-monotonic. There is a turning point of the critical value of robustness. In this case, lower liquidity means a higher likelihood for a bank to fail. When the k increases, more connections indicate more contagions between failed nodes and healthy nodes, presumably, there are more paths for cascading propagation. Among three topologies, the scale-free network has the worst stability for both random and intensional attacks. In the case of intensional attacks, the gap with others is even larger. In the scenario of a positive correlation between individual robustness and degree, a node with a higher degree should have a higher value of robustness. The scale-free network performs better than other networks since the correlation increases the heterogeneity of the network, and makes the network more stable in this particular scenario. Therefore, a heterogeneous topology reflects an extreme behavior in different scenarios: either stable or vulnerable when the market liquidity is bad. In a healthy market with a high level of liquidity, various topologies have no

obvious distinctions [7].

Huang et al. built a bipartite banking system and analyzed how the cascade propagates between banks and assets [8]. This model provides useful guidances about harmful and fragile components. One part of this model is the failed banks during 2008 financial crisis, and the other part is the assets for these banks. Each bank has a unique asset portfolio such as mortgages, investments, residential properties, loans and so on. Each category of the assets is shared by multiple banks. Hence, the authors computed the components of bank assets and the market value of each asset [8]. The fraction of different assets owned by one bank is calculated using the formula: $w_{i,m} = B_{i,m}/B_i$. The bank i possesses $B_{i,m}$ amount of the asset m, and B_i is the total property of the bank i. The share of one asset is defined as $s_{i,m} = B_{i,m}/A_m$, where A_m is the total value for asset m in the market. To trigger the cascade in the banking system, the authors imposed an initial devaluation p on all assets. The updated portfolio value of the asset m after the initial shock is pA_m. The banks that possess this asset will also suffer from a reduction of the asset value: $B_{i,m}(1-p)$. Therefore, the authors can update the value of each asset for each bank iteratively. Once the properties of a bank are below its liabilities or debts L_i, this bank becomes bankrupt and is removed from the system. The assets owned by the failed bank experience a devaluation of market value $\alpha B_{i,m}$. Once triggered, the cascading failures propagate between banks and assets until reaching the steady state. The primary motivation of this study is to identify which assets determine the bankruptcy and predict the potential failures of banks. For example, the agricultural loans are critical for bankruptcy. Failed banks have a weaker survivability than good banks. Besides, the predictability of this model is measured using the ROC curve analysis showing that this model makes more true positive predictions than false positive predictions. In the end, the authors revealed the phase boundary between the stable and unstable states for banks as a function of p and α. *Huang et al.* indicated that banks should keep a balanced asset portfolio and increase the liquidity to resist market shocks to avoid bankruptcy.

2.2.4 Interdependent networks

As our world is connected tightly and strongly, the components of one network could be another network. Multiple networks with various functionality combine together to assemble a more sophisticated and integrated system. More interdependent systems exist in our society, which attracts more attentions than before.

Buldyrev et al. modeled the cascading failures on interdependent networks, which include a power grid and a controlling system in Italy [16]. The simulation demonstrates reflect that single node failure in the power system could trigger a catastrophic cascade in both networks, breaking the whole network into fragments. In the interdependent systems, two networks have the same size and all nodes are pair-wisely connected. The dependency is bi-directional, and nodes have different degrees. In addition, all edges connected with node a_i and b_i in both networks are also removed. In the remaining networks, only the mutually connected components are functional, in which all nodes are connected either inside one network or between two networks. As iteratively updating the interdependent networks, the whole system splits into clusters, and the surviving giant component diminishes gradually until reaching the steady state. *Buldyrev et al.* compared the cascading behaviors in three types of networks: random regular (RR), Erdős-Rényi (ER), and scale-free (SF) network [16]. All networks have the same size and average degree. The authors measured the probability of obtaining a mutually connected giant component and observed a phase transition of the probability when increasing the number of initial failures. In general, a more heterogeneous topology has a smaller critical value of removal fraction, which means it is less likely to survive from cascading failures. Since in a heterogeneous topology, such as a SF network with a high exponent value, the degree distribution is skewed, and the failure of a small-degree node could disconnect a large component in the interdependent network. For example, a small-degree node in one network could connect with a high-degree node in the other network. This extreme disparity could exaggerate the damage of cascading failure. When designing the interdependent network, decision makers should consider a stable structure.

Pinnaka et al. studied the cascades in an interdependent network consisting of the United States infrastructure systems [17]. Nodes of this network are from

different infrastructure networks, such as energy, nuclear, communication, food and agriculture and so on. Edges represent the dependency between nodes. If two nodes are mutually dependent, the edge is bi-directional. The functionality of the whole system is measured by the flow robustness defined in [37]. This variable is determined by the network connectivity and the total amount of flows of the network. When parts of the network are removed, the flow robustness is recomputed, and the reduction of the flow robustness reflects the damage of cascading failures. To trigger the cascade, the authors removed the nodes with the highest degree, closeness, and betweenness. The cascades can be triggered by distinct types and numbers of initial attacks. The flow robustness reflects a monotonic decreasing as a function of the severity of initial failures. This trend is similar to different initial attacks, which means the failures of the centrality nodes indeed have a huge impact on the system. Interestingly, when the authors intentionally removed the edges with a high centrality, the whole system tolerances to the initial attacks much better than the case of random removals, which indicates that the centrality is critical to a global cascade. Thus, *Pinnaka et al.* calculated the centrality of each subnetwork and pointed out the vulnerability of several significant infrastructure components, such as *energy, communication* and so on [17]. The ranking of these critical systems provides valuable guidance to reduce the risk. To increase the stability of the network, stakeholders should protect the vulnerable components of the interdependent network.

2.3 Predictability of cascading failures

In previous sections, we discuss the features in standard networks and real-world networks. It is important to understand the dynamics of cascading evolution in different cases. Our primary motivation is to predict the future cascades and protect the system in advance. Hence, we have a strong desire to find what is the accuracy of predictions. In this section, we introduce related studies.

Daqing et al. analyzed the spatial correlation of cascades in real world networks [15]. The spatial correlation is defined as [15]:

$$C(r) = \frac{1}{\sigma^2} \frac{\sum_{ij,i\in F} (x_i - \bar{x}) \left(x_j - \bar{x}\right) \delta \left(r_{ij} - r\right)}{\sum_{ij,i\in F} \delta \left(r_{ij} - r\right)} \tag{2.6}$$

where x_i represents the state of node i (value of 1 for overloaded and value of 0 for the normal state), $\bar{x}$ is the mean value of state for all nodes, F is the overloaded nodes, σ is the standard deviation of F, r_{ij} is the geometrical distance between node i and j, and δ function assigns the value of r_{ij} equal to r. A positive value of $C(r)$ means the overloaded nodes are spatially close to each other. *Daqing et al.* indicated that $C(r)$ in the traffic flow network of Beijing, China exhibits a power-law distribution in terms of r. During rush hours, the spatial correlation reflects a power-law decay. However, at off-peak hours, the decay is more like an exponential decay. Consequently, the time factor affects the cascading size, which means in rush hours, these values reach the peak. In a real-world power grid, the authors could also recognize similar behaviors. *Daqing et al.* indicated that the spatial correlation could be a good predictor for the future cascading failures and improve the mitigation strategies.

Shunkun et al. implemented machine learning algorithms to predict the overload rate for each node in a power grid system [38]. In a complex network, it is difficult to characterize the dynamics of nodes. The cascading process is a stochastic process since the initiators are uncertain. A group of nodes with same characteristics possesses a higher predictability than isolated nodes. The authors started with a two-dimension lattice with a size of 50×50. Each node represents a subnetwork or community. The weight of each edge is randomly drawn from a Gaussian distribution. The weight is guaranteed to be non-negative. the authors used the number of shortest paths including node i as the load L_i. Capacity is calculated by multiplying a factor of $1 + \alpha$ to the initial load. Initially, the center 2×2 square is removed to trigger the cascading failures. Since the lattice can be divided into multiple layers of squares from center to boundary, the distance between a large square and the central square is used to show far away for the nodes in the large square to the central square. In a 50×50 lattice, there are 23 possible values for the distance of a node from the central square. The authors implemented several machine learning algorithms to predict the cascade rate for different distances. The training data is

from multiple cascade simulations, and the prediction is from the regression results. The relative error is very close to zero, which reflects the prediction is accurate. As the tolerance parameter increases, the relative error decreases gradually, and the support vector regression has the smallest error. Machine learning algorithms are able to predict the future behaviors based on historical results, which is an indirect method, but works well when the dynamics of cascading failure is complicated.

Zhao et al. studied how cascading failures propagate through the network as a function of time and space [39]. In a geometrically concentrated network, the cascading propagation is constrained by the network topology. Two variables are defined to measure the propagation properties. First, $r_c(t)$ is defined as the average location of all overloaded nodes at time step t. Second, $F_r(t)$ is defined as the amount of overloads at time t, which indicates the severity of cascading failures at current step. The authors calculated these two variables in a two-dimension regular lattice with a size of $L \times L$. Initial failures are located at the center of the lattice. At first several steps $t < 10$, the value of $r_c(t)$ increases linearly, consequently, cascading failures propagate at a nearly constant velocity. When $t > 10$, the increase of $r_c(t)$ slows down. $r_c(t)$ becomes saturated when the cascading failures approach the network boundary. $r_c(t)$ has a positive correlation with L and a negative correlation with tolerance parameter α. Intuitively, in a network with a larger size and lower capacity, the cascading failures travel faster than in other cases. After taking normalization, different networks have a similar value of r_c/L. However, $F_c(t)$ experiences a bell-shaped behavior. The value of $F_c(t)$ increases quickly, reaches the peak when the time step is around 5, then $F_c(t)$ decreases gradually. $F_c(t)$ also has a positive correlation with L, and different networks have a similar normalized value: $F_c(t)/L^2$. In the network with a large tolerance parameter α, $F_c(t)$ decreases dramatically. Based on theoretical analysis and simulation results, *Zhao et al.* concluded that the normalized variables $r_c(t)/L$ and $F_c(t)/L^2$ could be used to predict the future cascading failures in a spatially constrained network. In general, a high capacity of the system is necessary to slow down the propagation of cascade process.

2.4 Robustness of complex network

Robustness of complex network reflects the ability to withstand disruptions. Based on the topology and properties of the network, we could measure the robustness in various ways. In this section, we discuss the related studies about the metrics of robustness.

Koç et al. proposed a robustness metric for the power grid [40]. The cascading failures are triggered by initial removals of targeted edges. The authors designed the edges to be directed and used the direct current (DC) approximation in the power grid. The robustness metric considers the topology and flows properties of the network. At first, the authors utilized the concept of the information entropy to define the nodal robustness [40]:

$$R_{n,i} = -\sum_{k=1}^{L} \alpha_k p_k \; log p_k \tag{2.7}$$

where n is the out-degree of node i, L is the neighbors of node i, α_k is the tolerance parameter of neighbor node k, p_k is the proportion of edge load f_k over the sum of all edge loads. The definition of p_k is shown in [40]:

$$p_k = \frac{f_k}{\sum_{j=1}^{L} f_j} \tag{2.8}$$

where f_k is the value of load on the edge e_{ik}. In addition, *Koç et al.* defined the nodal significance as shown in [40]:

$$\delta_i = \frac{P_i}{\sum_{j=1}^{N} P_j} \tag{2.9}$$

where P_i is the load of node i. This value reflects the proportion of a single-node load over the total load. In the end, the network robustness metric is a product of nodal robustness and significance [40]:

$$R_{CF} = \sum_{i=1}^{N} R_{n,i} \delta_i \tag{2.10}$$

Then, *Koç et al.* focused on the node with the highest significance value and re-

moved the edge with the highest load to trigger the cascade in a real-world power grid. The authors detected the correlation between the network robustness R_{CF} and the network survivability, which related to the link and capacity. Based on simulation results, network robustness is positively correlated with survivability, which is 0.76 between R_{CF} and link survivability, 0.75 between R_{CF} and Capacity survivability. *Koç et al.* concluded that the robustness metric R_{CF} is a good measurement of the stability of the power grid concerning cascading failures. Well-designed improvements focusing on network topology and dynamics can reduce the vulnerability of the whole system.

Albert et al. stated that North American power grid has a high robustness for most perturbations, but fails to withstand targeted attacks [41]. The nodes in the power grid has three identifications: generators, transmission nodes, and consumers. The cumulative probability function of the degree in this power grid follows an exponential distribution [41]:

$$P(k > K) \sim exp(-0.5K) \tag{2.11}$$

where k is the degree of one node. The load of node i has the same definition as in the previous study [42]. The load distribution at the initial state is as follows [41]:

$$P(l > L) \sim (2500 + L)^{-0.7} \tag{2.12}$$

where l is the load of one node. From these two distributions, it is evident that the load distribution is more skewed than the degree distribution. Few nodes have a significant load, and their failure could cause catastrophic damages. The robustness of a power grid considers not only the topological properties but also the flow dynamics inside the network. Hence, *Albert et al.* introduced the connectivity loss C_L to measure the network robustness [41]:

$$C_L = 1 - \frac{\left\langle N_g^i \right\rangle_i}{N_g} \tag{2.13}$$

where N_g represents total number of generators, N_g^i represents how many genera-

tors are connected with the consumer i, $\langle N_g^i \rangle_i$ is the average value of N_g^i over all consumer nodes. When the generators are removed initially to trigger the cascading failures, the connectivity loss C_L does not increase too much and stays close to the minimal value. However, when the transmission nodes fail initially, the cascading failures have a different behavior. Cascade triggered by the removals of transmission nodes with a high degree or load causes a significant connectivity loss. Hence, the power grid fragments into smaller components and the cascading failures profoundly affects the functionality of the system. As the fraction of initially attacked nodes increases, the connectivity loss grows quickly and saturates in the end. According to the simulation results, *Albert et al.* concluded that generators of the power grid are critical to cascading failures. In general, decision-makers need to build the generators close to consumers, reduce the distance of electricity transmission, and increase the capacity or connectivity of the system.

Parandehgheibi et al. proposed a robustness metric of the interdependent networks [18]. A good example of such networks is a power grid and a communication system, which are mutually dependent. In the power grid, there are two types of nodes: generators and substations. In the communication system, the control nodes dispatch operation instructions, and router nodes connect with substations. The power grid needs the control signal to work continuously, and the communication system relies on the electricity power. Meanwhile, all generators only connect to substations, and all control nodes only connect with router nodes. A single node is randomly removed to trigger the cascading failures. The authors introduced a robustness metric *minimal total failure removals* (MTFR) for the entire system collapse [18]. A larger value of the MTFR means a higher stability of the interdependent networks and vice versa. *Parandehgheibi et al.* proved that in the case of directed edges, it is NP-hard to find the MTFR value. However, regarding undirected edges, it just takes polynomial time to obtain the MTFR. The MTFR reflects a linear relationship with the size of interdependent networks. And the robustness of an undirected network is higher than that of a directed network. *Parandehgheibi et al.* computed the MTFR in the interdependent networks in Italy, which includes a power grid and a communication system. The whole system tolerates most of the

initial disruptions. However, parts of the system are very vulnerable. For example, even only three initially failed nodes could destroy the northen parts of the Italian power grid (one-third of the entire system). This study provides valuable guidance to build a robust structure of the interdependent networks.

2.5 Mitigation strategies against cascading failures

After we get familiar with the dynamics of cascading process, there is a strong desire to propose efficient strategies to reduce the cascading damages. Our motivation is to use the limited resource to powerfully protect the system. Due to the complex features of the system, it is difficult to find a general method to apply everywhere. Since the propagation of cascading failure is very fast through the network, it is costly for our society to suffer the damage passively without any reactions. An active and well-designed defense strategy can save a huge part of the network efficiently. One reasonable procedure is to start with simple theoretical models and extend the research to real world networks. Several related studies are briefly introduced in this section.

Motter et al. extended his research on cascading failures in a flow-based network, and introduced a defense strategy to reduce the cascade damages [43]. The structure of the network is the same as that in the previous study [9]. The main idea is to intentionally remove specific nodes or edges immediately after the initial attack to enable more nodes to survive eventually. The author concluded that removing the small-load nodes and the large-load edges during the cascading process pro-actively could dramatically mitigate the cascading damage [43]. Although the intentional removals enlarge the overloaded parts at each step of the cascade, these components act as "fuse" effectively blocking the propagation of failure to the remaining parts. To determine the initial attacks, the author focused on the actual load that travels through the selected node and its contribution load generated to the whole system. The actual load of one node L_k is defined as follows [43]:

$$L_k = \sum_{i,j} L_k^{(i,j)}, \tag{2.14}$$

where $L_k^{(i,j)}$ is the load of node k with a specific source and target pair (i, j). The flow is assumed to travel on the shortest path between the source and target nodes. After summing up all the possible pairs, the author got the actual load of node k in this network. The contribution load L_i^g is defined as [43]:

$$L_i^g = \left(\bar{D}_i + 1\right)(N - 1), \tag{2.15}$$

where $\bar{D}_i$ represents the averaged length of all shortest paths starting from node i and ending with other nodes, N is the size of the network. If the real load L_i is larger than contribution load L_i^g, node i have a higher ability to handle additional flow. If $L_i < L_i^g$, node i generates more load than the actual load, which means the excessive load is delivered to other parts of network. If there are numerous nodes, whose contribution load is higher than their actual load, other nodes have to consume the additional load and are more likely to overload during the redistribution of load. Hence, a natural idea is to remove the nodes with a significant contribution load but with a small actual load. In addition to the gap between two loads, *Motter et al.* also designed another three triggering strategies: removing the node with the largest contribution load, real load, and degree. The cascading damage is much smaller on the network with the intentional removal strategies than the network without any control actions. Since L_i and L_i^g is negatively correlated, a smaller L_i leads to a higher L_i^g. Thus, removing the node with the lowest load at each step of cascading process leads to the best performance of mitigation. Regarding removing edges, *Motter et al.* applied same analysis. The conclusion is to eliminate the edge with largest loads to optimize the defense.

Schneider et al. developed an edge-rewiring methodology to effectively mitigate the cascading damages [44]. The main idea is to rewire a small proportion of connections, preserve the degree of each node, and maintain the functionality of the network. A fraction of nodes q is removed initially to trigger the cascading failure. The critical value of this fraction q_c is popularly used to reflect the resilience of the system, at which the whole system fails entirely. However, in some cases, cascading failure causes enormous damages but fails to wreck the system. To capture more features of cascading dynamics, the authors defined a new variable R to represent

the robustness of the network [44]:

$$R = \frac{1}{N}\sum_{Q=1}^{N} s(Q) \tag{2.16}$$

where N is the total number of nodes, Q is the number of initial removals, $s(Q)$ is the relative size of the surviving giant component in terms of Q initially failed nodes. A greater value of R reflects a higher robustness of the network. Because of the limited resource, it is impractical to add too many connections or capacities to the system. Hence, *Schneider et al.* focused on changing current edges and assumed that the cost of changing edges is smaller than that of varying the node degree. The strategy is simple: randomly choose two edges e_{ij} and e_{kl}, swap the endpoints of both edges, and generate new edges e_{ik} and e_{jl}. If the new graph has a higher robustness R than before, this change is accepted, and the same procedure is repeated until no further changes can be made.

Schneider et al. tested this mitigation strategy in two real-world networks: the European power grid and the Internet. After applying the mitigation strategy, the relative size of the surviving giant component $s(Q)$ becomes larger than the original value, which means the robustness increases, and the severity of cascades decreases accordingly. The robustness R of both networks improved 45% and 55% respectively. The robustness R increases as more edges are rewired. Meanwhile, the functionality of both networks remains unchanged after the edge reconstruction. Also, *Schneider et al.* concluded that it is possible to find the most robust structure based on a particular degree distribution.

Parandehgheibi et al. proposed a mitigation strategy for interdependent networks including a power-grid system and a communication network [19]. To reveal the difference between the interdependent and isolated networks, the authors started with two standard network scenarios: first, a single Erdős-Rényi (ER) network, and second, two interdependent ER networks. It is evident that the interdependent networks suffer a greater damage, and the average size of the giant component G of the interdependent networks is always smaller than that of the single network. When the fraction of initial removals exceeds 0.5, the value of G drops dramatically and

approaches zero.

Next, *Parandehgheibi et al.* designed the interdependent networks with a power grid and a communication network. Both networks have 500 nodes, have the average degree of 4, and are mutually dependent. In the power grid $G_P = (V_P, E_P)$, a direct current (DC) power approximation is applied, and three types of nodes are considered: generators, consumers, and substations. If one node fails, the flow redistributes to other parts of the grid. In the communication network $G_C = (V_C, E_C)$, router nodes exchange signals, and control nodes dispatch instructions. Each generator in the power grid connects with a router node, and several router nodes connect with one control node. The communication network replies on the electricity from the power grid, and the power grid relies on the control instructions from the communication network. If a power node disconnects from the corresponding router node, this power node is out of control and is removed from the power grid. A fraction of power nodes p are removed initially to trigger the cascade. If there are no mitigations, even a small value of p could cause huge damages, and the average size of the surviving giant component G is close to 0, which means the interdependency amplifies the cascading damages. Hence, *Parandehgheibi et al.* formulated a two-phase mitigation strategy: first, the inevitably failed nodes are identified and removed from both networks; second, the supply and demand of power nodes are rebalanced in order to guarantee the normal operations in both networks. The cascading failures are mitigated but cannot be avoided. Reducing the supply of power nodes makes the overload in power gird less likely. However, increasing the supply of power nodes satisfies the minimal requirement for the communication network. After applying this mitigation strategy, the average reduction of power supply decreases dramatically compared with the non-controlled case, which means the mitigation strategy indeed protects the interdependent networks from the global cascades.

CHAPTER 3
FAILURE DYNAMICS OF THE GLOBAL RISK NETWORK

In modern society, there are various functional systems connected with each other in an explicit or implicit manner. Because of the development of technology and science, the efficiency of a system is increasing quickly. However, this trend may worsen the vulnerability of the systems since potential risks heavily threaten the world. Unlike the traditional risks analysis, which only considers one particular domain, such as the power grid [36], Internet [12], [13], and transportation systems [45], various risks from diverse domains interconnect and form an interdependent network, where one risk can be not only triggered internally but also by others. Not many studies focus on the cross-domain risks. The materialization of a single risk may trigger the cascading failures of other risks, cause a catastrophe in the entire system, and lead to the social unrest in the end. Hence, there is a strong desire to study the combined effects of global risks [4], [20]. It is difficult to distinguish the combined effects of risk materialization in such a complex network. Moreover, the dynamics of the risk propagation is complicated. Fortunately, crowd-sourcing assessments could solve these problems. Taking advantage of expert assessments, we developed a quantitative model analyzing the cascading dynamics and predicting the materialization of global risks [46]. A better understanding of the underlying connections between sub-systems provides us more valuable guidances to mitigate the damage of failures of the global risks. Stakeholders and decision makers can benefit from our studies to fight against potential threats and disruptions [47].
The main steps of designing the global risk network in this thesis are as follows:

1. Use the expert assessments from the World Economic Forum Global Risk 2013 Report [47] to build the structure of the global risk network. A survey

Portions of this chapter previously appeared as: B. K. Szymanski, X. Lin, A. Asztalos, and S. Sreenivasan, "Failure dynamics of the global risk network," *Sci. Rep.*, vol. 5, no. 10998, Jun. 2015. Accessed on: Apr. 21, 2017. [Online]. Available: https://www.nature.com/articles/srep10998.

in the report provides us with the likelihood of materialization, the impact of failure and the connectivity among 50 global risks in 5 domains: economic, environmental, geopolitical, societal and technological.

2. Collect the historical status of each risk in each month during the period from 2000 to 2012. We searched academic articles, on-line encyclopedias, question-and-answer sites, news websites, magazines, books and other resources to collect training data for parameter recovery.

3. Use a Markov chain to simulate the evolution of risks. There are only two discrete states: normal and materialized. We assume the state transitions follow a Poisson process. Three independent state transitions determine the next state of each risk. Each state transition has a control parameter.

4. Use maximum likelihood estimation to find the optimal value of control parameters [22]–[25]. Simulate the cascades in various scenarios. In addition, detect the precision of predictions and compare with other models. Identify detrimental risks, predict future activity of each risk, and mitigate the damage according to simulation results.

3.1 Model definition

3.1.1 Crowd-sourcing assessment

Crowd-sourcing is defined as an efficient method to accumulate contributions for solutions, ideas, and services from a group of people (professional or non-professional) and although used for centuries, its use accelerated in the recent decade [48]–[52]. Crowd-sourcing has been implemented massively in business organizations to seek technical solutions. What's more, this trend becomes more popular in non-commercial areas.

Since global risks have drawn much attention in the past decades, we focus in our study in this popular area. We utilize the crowd-sourcing assessment coming from the World Economic Forum (WEF) Global Risk Report 2013 [47]. The main part of this report is based on an annual survey by over 1000 experts from various

Table 3.1: List of global risks.

ID	Risk	ID	Risk
1	Chronic fiscal imbalances	2	Chronic labor market imbalances
3	Extreme volatility in energy prices	4	Hard landing of an emerging economy
5	Major systemic financial failure	6	Prolonged infrastructure neglect
7	Recurring liquidity crises	8	Severe income disparity
9	Unforeseen negative consequences	10	Unmanageable inflation or deflation
11	Antibiotic-resistant bacteria	12	Failure of climate change adaptation
13	Irremediable pollution	14	Land and waterway use mismanagement
15	Mismanaged urbanization	16	Persistent extreme weather
17	Rising greenhouse gas emissions	18	Species overexploitation
19	Unprecedented geophysical destruction	20	Vulnerability to geomagnetic storms
21	Critical fragile states	22	Diffusion of weapons of mass destruction
23	Entrenched organized crime	24	Failure of diplomatic conflict resolution
25	Global governance failure	26	Militarization of space
27	Pervasive entrenched corruption	28	Terrorism
29	Unilateral resource nationalization	30	Widespread illicit trade
31	Backlash against globalization	32	Food shortage crises
33	Ineffective illicit drug policies	34	Mismanagement of population aging
35	Rising rates of chronic disease	36	Rising religious fanaticism
37	Unmanaged migration	38	Unsustainable population growth
39	Vulnerability to pandemics	40	Water supply crises
41	Critical systems failure	42	Cyber attacks
43	Failure of intellectual property regime	44	Massive digital misinformation
45	Massive incident of data fraud/theft	46	Mineralresource supply vulnerability
47	Proliferation of orbital debris	48	Unforeseen consequences of climate change
49	Unforeseen consequences of nanotechnology	50	Unforeseen consequences of new life science

fields such as academia, industry, government and so on. In addition, respondents also have different organizations, residences, expertises, genders, and ages.

There are $N = 50$ global risks in 5 categories: economic, environmental, geopolitical, societal and technological in this report [47]. The respondents focused on these 50 risks and evaluated the likelihood of their occurrence in the next 10 years, the severity of the impact if they occur and the connectivity among them. The WEF provided analysis based on this survey and also showed how the background of respondents influences the evaluation. With the help of experts, we can have a clearer understanding of the opaque behaviors among various risks. For example, each survey respondent assessed the likelihood using a score from 1 to 5, where lower score represents a lower likelihood. The score is either an integer or a midpoint between integers. In the end, the average value of all respondents is used as the final score of the likelihood. This process is the same for evaluating impact. For connectivity, the respondents can choose from three up to ten strongest edges. The

experts provided 515 different connections. Besides the average value of evaluations, the standard deviation and margin of error are also shown in the results. Due to the high quality of assessment, these margin of error is about 2%, which is a very low level.

Table 3.1 shows a list of 50 risks in our model [47], which are grouped into five categories and listed in the following order of categories: economic, environmental, geopolitical, societal and technological. Some risks such as "Chronic fiscal imbalances" and "Chronic labor market imbalances" may share an overlapped definition. They could be triggered by the similar reasons. Hence, some real world events can be labeled by multiple risks which show a complicated feature of these events. On the other hand, some other risks have very clear definition and boundary, for example, natural disasters and water supply crisis. They concentrate on one aspect of our society and are very easy to identify and recognize their materialization. However, the materialization of these risks could cause other risks to fail quickly if they are out of control. One example is that a severe food shortage and water supply crisis will make critical states fragile and lead to a global governance failure eventually. Hence, one small risk may trigger cross-domain cascade and trigger some risks which may seem to be unrelated to the initial risk. Because of the density of possible paths for risk to propagation, we need to analyze the whole network systematically and consider all the potential influences at the same time. It is a very complex scenario, and our study is to capture the underlying features inside this network and understand them better.

As described above, the intelligent opinions have been transformed to a quantitative scale by professional agents resulting in a higher quality scale that could have been expected from non-experts. It is possible to find the most impactful and dangerous risks in an explicit manner. According to the WEF Global Risk 2013 Report [47], the following tables show examples of the survey results:

As defined in the report [47] and shown in Table 3.2, "severe income disparity" has the highest likelihood of materialization of 50 risks. In Table 3.3, the top five highest impact risks are not the same as the highest likelihood risks [47]. It is interesting that this risk also has high connections. Hence, decision makers

Table 3.2: Top five risks with the highest likelihood.

No.	Global Risk	Likelihood
1	Severe income disparity	4.22
2	Chronic fiscal imbalances	3.97
3	Rising greenhouse emissions	3.94
4	Water supply crisis	3.85
5	Mismanagement of population aging	3.83

Table 3.3: Top five risks with the highest impact.

No.	Global Risk	Impact
1	Major systemic financial failure	4.04
2	Water supply crisis	3.98
3	Chronic fiscal imbalances	3.97
4	Diffusion of weapons of mass destruction	3.92
5	Failure of climate change adaptation	3.90

Table 3.4: Top ten highest degree risks.

No.	Global Risk	Connections
1	Global governance failure	44
2	Severe income disparity	41
3	Critical fragile states	40
4	Food shortage crises	33
5	Mismanaged urbanization	33
6	Pervasive entrenched corruption	32
7	Extreme volatility in energy & agriculture prices	30
8	Failure of climate change adaptation	30
9	Unsustainable population growth	30
10	Chronic fiscal imbalances	29

need to pay much attention to this risk because of the high likelihood and degree. There might be an underlying correlation between the likelihood and connections in Table 3.4 [47]. However, this is not the same for all risks. Some risks have high likelihood or impact, but few connections. Since there is no obvious metric to identify the important risks, it is necessary to develop a quantitative model by simulating the propagation dynamics among global risks and predict their future behavior.

3.1.2 Global risks network

Helbing studied the enhanced vulnerability among systemic risks by forming a network of networks [20]. In this model, an initial failure can destroy the whole system very quickly. The complexity of the connected networks makes it very hard to track the propagation path of the cascade process.

Inspired by the previous study, we take advantage of expert assessment to build an interconnected global risk network. The damage or influence of one risk is no longer local but global. We want to find how the failure propagates across multiple domains.

The likelihood given by expert assessment in the range between 1 and 5 is denoted as L_i for risk i. In later steps, we normalize it to be within 0 and 1. Although the experts give us a weighted evaluation on the connection intensity for 515 pairs with an average degree of 20.6, we just use a binary state to represent these connections denoted as $b_{i,j}$. The $b_{i,j}$ is the adjacent matrix element defining connection between risk i and j. Value 1 means there exists a connection between risk i and j; value 0 means no edge between them. The reason is that we get the same value of maximum expectation from historical data using this simplified information as the case using full information that includes how many times the experts listed a given connection.

In addition to the likelihood and connections, we assume in this network, the state of each risk is binary at each time step: each risk is in either inactive or active state. In our model, we use one time step to represent one month of time in real world. Risks either manifest their failures or not, regardless of impact. The average degree of each node is 20.6. The probabilities of state transitions in this model are expressed in terms of L_i and b_{ij}.

We use a Stochastic Block Model [53] to represent the structure of global risks. Since there are five domains inside this network, each domain is treated as a group denoted as g. The probability of a connection between two nodes in the same group p_g varies from the probability of connection between nodes in two groups. The probability of a connection between two nodes from different groups is defined as $p_{g1,g2}$. These two values can reflect how strong the connectivity is within one group or

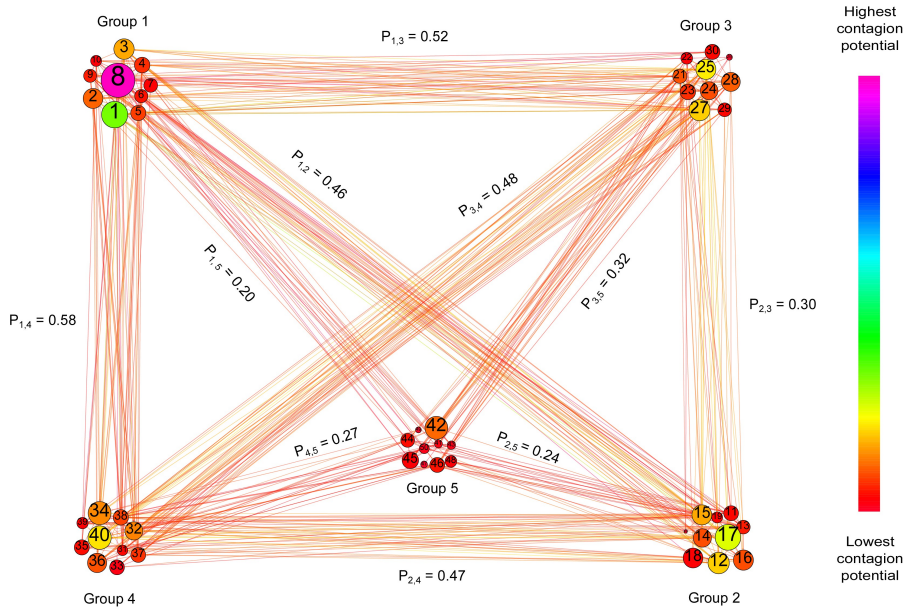


Figure 3.1: Connectivity of risk groups.

between two groups. Figure 3.1 visualizes the inter-group connectivity where node's color corresponds to its total connectivity and the number of lines connecting the groups indicate strength of the inter-group connectivity. Groups 1 (economic risks), 2 (environmental risks), and 3 (geopolitical risks) are the best connected, so risks from these groups dominate the list of most persistent nodes. The remaining two groups: 4 (societal risks) and 5 (technological risks), have fewer connections to other groups. The inter-edges are labeled with probability of inter-group connections.

3.1.3 Historical dataset of global risks

The expert assessment is used to build the structure of the global risk network. A time series of binary states is used to represent how each risk changes its state with time. Moreover, the time unit represents one month. As we mentioned above, to find the optimal values of control parameters in our model, we require such a time series in a contiguous period as input data. So we collected the historical events for each risk over the period 2000 – 2012 so of 13 years and formulated a table of time series comprising binary states. We searched academic articles, online encyclopedias, question-and-answer sites, news websites, magazines, books

and other resources systematically about the historical risk materialization events. This collected dataset is called *historical* dataset, and it includes 7800 points of data for parameter training.

Table 3.5: Examples of historical events.

Risk	Event	Impact	Start date	End data
Chronic Fiscal Imbalances	European sovereign debt crisis	Global	2008.10	2012.12
	2008 financial crisis in US	Global	2009.01	2012.12
	Early 2000s recession	Regional	2000.08	2003.05
	Dubai financial crisis	Regional	2009.11	2010.01
Global Governance Failure	Crisis in Syria	Global	2011.05	2012.12
	Crisis in Libya	Regional	2011.02	2012.12

In Table 3.5, we show some examples of events in the historical dataset. Each historical event is recorded with the following information: name, impact (global or regional), starting date and ending date. Only the events with global impact are represented in the history of the corresponding risk. The time series is formulated based on one-month time unit.

3.2 Model dynamics

3.2.1 Discrete time model

The failure dynamics of the global risk network can be modeled using Alternating Renewal Processes (ARP) [54], which were initially used for engineered systems, but recently have been applied to science and economic problems [72]. Nodes under ARP alternate between the normal state and the failure state, and the corresponding events of failure activation and recovery are responsible for the state transitions. Typically, these events are assumed to be triggered by homogeneous (time invariant intensities) or the non-homogeneous Poisson Processes [54]. In our systems, use of the Poisson distribution is justified because each failure represents a systemic risk that can be triggered by many elementary events distributed all over the globe. Such triggering distorts any local temporary patterns of events (such as periodic weather related local disasters in some regions of the globe). Moreover, the time-dependence of the intensity of risk activation is the results of influence that active risks exert on passive risks connected to them. Accordingly, each risk at time t is either in state 1

(materialized or *active*) or state 0 (not materialized or *inactive*).

In a traditional ARP, there are two directly observable processes, one of risk activation and the other of recovery from the active risk. However, in our model, we introduce two latent processes that together represent the risk activation process. As explained later, we use the maximum likelihood estimation Algorithm [22]–[25] to find model parameters that make the model optimally matching historical data. Consequently, we assume that changes in the state of each risk result from events generated by three types of Poisson processes.

First, for a risk i, given that it is in state 0, its *spontaneous* or *internal* materialization is a Poisson process with intensity λ_i^{int}. Similarly, given the risk in state 1, its recovery from this state, and therefore transition to the state 0, is a Poisson process with intensity λ_i^{rec}. Finally, given that risks i is in state 0, and j in state 1 are connected, the materialization of risk i due to the *external* influence of risk j is a Poisson process with intensity $\lambda_{ji}^{\text{ext}}$. We assume that each of these processes is independent of each other. We also evaluated models in which recovery is represented by two latent processes, one of internal recovery and the other of recovery induced by either the connected passive or active risks. In both cases, the optimal intensity of the externally induced recovery was 0. Thus, the simpler model with just internal recovery is used as it yields the same results as the more complex models using latent processes for recovery.

For nearly all events that we consider here, it is difficult to assign precise starting and ending times for their periods of materialization. Thus, it is more proper to consider a Bernoulli process in which the time unit (and also time step of the model evolution) is one calendar month (we ignore the minor numerical imprecision arising from the fact that calendar months have different numbers of days). Consequently, all events starting in the same month are considered to be starting simultaneously. Hence, at each time step t, each risk i is associated with a binary state variable $S_i(t) \in \{0, 1\}$. The state of the entire set of risks at time t can, therefore, be represented by a state vector $\vec{S}(t)$. Thus, the dynamics progresses by assuming that at each time step $t > 0$:

1. a risk i that was inactive at time $t - 1$ materializes internally with probability

$p_i^{\text{int}} = 1 - e^{-\lambda_i^{\text{int}}}$.

2. a risk j that was active at time $t-1$ causes a connected to it risk i that was inactive at time $t-1$ to materialize with probability $p_{ji}^{\text{ext}} = 1 - e^{-\lambda_{ji}^{\text{ext}}}$.

3. a risk i that was active at time $t-1$ continues its materialization with probability $p_i^{\text{cont}} = e^{\lambda_i^{\text{rec}}} = 1 - p_i^{\text{rec}}$.

It is easy to show that for real time t (months in our case), the Poisson process assumption for events results in a probability $1 - e^{-\lambda\lceil t\rceil}$ of an event happening in at most $\lceil t\rceil$ time units, which are identical to the assumed Bernoulli process. The advantage of the latter process is that in each step the probability of an event is known, simplifying maximum likelihood evaluation of the model parameters. Finally, the dynamics described above imply that the state of the system at time t depends only on its state at time $t-1$, and therefore the evolution of the state vector $\vec{S}(t)$ is Markovian.

Given the probabilities of internal materialization, external influence and internal continuation, that is just 1 minus the probability of recovery, the probability of a transition in a risk's state between consecutive time steps can be written in terms of these probabilities:

$$\begin{aligned}
\mathcal{P}_i(t)^{0\to1} &= 1 - e^{-\lambda_i^{\text{int}} - \sum_{j\in A(t-1)} \lambda_{ji}^{\text{ext}}} \\
\mathcal{P}_i(t)^{0\to0} &= 1 = \mathcal{P}_i(t)^{0\to1} \\
\mathcal{P}_i(t)^{1\to0} &= 1 - e^{-\lambda_i^{\text{rec}}} \\
\mathcal{P}_i(t)^{1\to1} &= 1 - \mathcal{P}_i(t)^{1\to0}
\end{aligned} \tag{3.1}$$

where $A(t)$ represents the risks that are active at time t, and $\mathcal{P}_i(t)^{x\to y}$ is the probability that risk i transitions between time $t-1$ and t from state x to state y, or in other words $S_i(t-1) = x$ and $S_i(t) = y$.

3.2.2 Continuous time model

The model is similar to a model of house fires in a city, where some houses burn alone from a self-started fire, but others are ignited by the burning neighboring

houses. Yet, the recovery, in this case rebuilding of a burnt house, is independent of the state of its neighboring houses.

Denoting by $s_i(t)$ the probability at time t that the state of a risk i at that time is 1, we can express the expected number of risks materialized at time t as the sum of all $s_i(t)$'s, each of which is defined by the following Ordinary Differential Equation (ODE), as stated in [46]:

$$\frac{ds_i(t)}{dt} = \lambda_i^{\text{int}}(1 - s_i(t)) - \lambda_i^{\text{rec}} s_i(t) + \lambda_i^{\text{ext}}(1 - s_i(t)) \sum_{j=1, j\neq i}^{N} a_{i,j} s_j(t) \tag{3.2}$$

Checking stability, we conclude that this system of non-linear ODEs has only one unique stable point in the feasible range $0 \leq s_i(t) \leq 1$, which can easily be found numerically. Moreover, this system of ODEs for a fully connected graph when the intensities $\lambda_s, \lambda_r, \lambda_e$ of the three Poisson processes are independent of the node on which they operates, and for all nodes starting in the same initial condition $s(0)$ has the analytic solution of the form, as stated in [46]:

$$s_i(t) = \frac{a + b * \tanh\left(b * t/2 + \operatorname{arctanh}\left(\left(2\lambda_E * s(0) + a\right)/b\right)\right)}{2\lambda_E}, \tag{3.3}$$

where $\lambda_E = (n-1)\lambda_e$, $a = \lambda_s + \lambda_r - \lambda_E$, and $b = \sqrt{a^2 - 4\lambda_s\lambda_E}$. This solution tends asymptotically to $\frac{2\lambda_s}{\lambda_s+\lambda_r-\lambda_E-\sqrt{(\lambda_s+\lambda_r-\lambda_E)^2-4\lambda_s\lambda_E}}$.

The mapping of the Poisson process intensities into the expert assessments is described in next section. Each of the probabilities of Bernoulli processes is mapped onto the probability obtained from expert assessment of likelihood of risk failure by single-parameter formula. We find the values of the model parameters that optimize the model match with the historical data, while we use expert assessments to individualize probabilities of Bernoulli processes for each risk. In essence, the expert assessments are defining those probabilities for each risk in relations to probabilities for other risks, while model parameters map performance of all risks onto historical data. By distinguishing between internal and external materialization factors, the mapping of parameters onto historical data enables us also to decompose risk

materializations into these two categories. Once the mapping is done, the model is complete and can be used to evaluate global risk dynamics.

From several alternative models discussed in following section, we discuss below the best performing network model which uses all three parameters, and the independent model which sets the value of probability of influence of a risk materialization on any other risk to zero, effectively isolating risk materializations from each other.

3.2.3 Methods

The first step to define the model is to relate the Poisson process intensities that determine the event probabilities in the model, to quantities provided by the expert assessments, namely, the likelihoods L_i of internal materializations of risks over a decade, and the influence that a given risk's materialization has on other risks.

3.2.4 Mapping expert assessments to Poisson process intensities

We first normalize the likelihood values to probabilities in their natural range of $[0, 1]$ by a simple linear transformation as stated in [46]:

$$p_i = (L_i - 1)/4 \tag{3.4}$$

This normalized likelihood value p_i is in direct proportion to the expert assessment L_i, and for our purposes captures the risk's vulnerability to failure. Next, we assume that the relationship between the probability, p_i^{int}, that a risk i materializes internally in a time unit (one calendar month) and this risk normalized likelihood value obeys the following polynomial form, with a parameter α defining the exact mapping stated in [46]:

$$p_i^{\text{int}} = 1 - (1 - p_i)^{\alpha} \tag{3.5}$$

Thus, the probability of failing within a time period increases as the vulnerability p_i increases, and Equation 3.5 coupled with our earlier assumption that the

internal risk materialization is a Poisson process with intensity λ_i^{int}, yields in [46]:

$$\lambda_i^{\text{int}} = -\alpha \ln (1 - p_i) \tag{3.6}$$

An advantage of Equation 3.5 and Equation 3.6 is that their forms remain invariant under changes of the time-scale under consideration. Indeed, multiplying the original value of time unit by factor f simply changes the Poisson process intensity and the value of α by the same factor f. For example, for the time unit of the expert materialization likelihood assessment set to a decade, the corresponding value of α is 120 times larger than the value obtained when the time unit is set to a month. Moreover, the ratio of intensities is defined entirely by the ratio of the corresponding model parameters and is independent of the risk for which the corresponding Poisson processes generate events. So model parameters define the same ratio of intensities of all risks, while likelihood assessments define individual values of these intensities for each risk.

Another advantage of the form of Equation 3.5 is that it can represent convex (for $\alpha > 1$), linear (with $\alpha = 1$), or concave (for $\alpha < 1$) function, with the parameter α defining the shape that best matches a given set of historical data.

We adopt a similar reasoning to the mapping between the probability of continuation in a time unit p_i^{con} and the normalized likelihood values p_i. We start with the assumption that the probability of a materialized risk continuing over a time unit is $1 - (1 - p_i)^{\gamma}$, where parameter γ defines the mapping from likelihood to probability. This dependence captures the increasing likelihood of continuation as the vulnerability p_i increases and leads to the following equation which are stated in [46]:

$$\lambda_i^{\text{con}} = -\gamma \ln (1 - p_i) \tag{3.7}$$

Finally, following similar arguments as above, the intensity $\lambda_{ji}^{\text{ext}}$ of the Poisson process that enables a materialized risk j to influence the materialization of risk i is a function of parameter β defined as, as stated in [46]:

$$\lambda_{ji}^{\text{ext}} = -\beta b_{ji} \ln (1 - p_i) \tag{3.8}$$

The factor b_{ji} on the right hand side merely serves to capture the fact that the risks i, j must be perceived by the experts as having an influence on each other, in order for the probability of influence to be non-zero.

The forms provided in Equation 3.6, Equation 3.7, and Equation 3.8 define the model completely, and all that remains is to fit the parameters α, β, and γ optimally to the historical data capturing the risk materialization events over the last 13 years. In the historical dataset, each risk is assigned a state per month (the fundamental time unit) over the period of $2000 - 2012$. Thus, the likelihood of observing this particular sequence of risk materialization events through the dynamics generated by our model can be written as in [46]:

$$\mathcal{L}\left(\vec{S}(1), \vec{S}(2) \cdots, \vec{S}(T)\right) \equiv \prod_{t=2}^{T} \prod_{i=1}^{N} \mathcal{P}_i(t)^{S_i(t-1) \to S_i(t)} \tag{3.9}$$

where $T = 156$ is the number of time units in the historical dataset and $N = 50$ is the number of risks. Consequently, the logarithm of the likelihood of observing the sequence is, as stated in [46]:

$$\ln \mathcal{L}\left(\vec{S}(1), \vec{S}(2) \cdots, \vec{S}(T)\right) \equiv \sum_{t=2}^{T} \sum_{i=1}^{N} \ln\left(\mathcal{P}_i(t)^{S_i(t-1) \to S_i(t)}\right) \tag{3.10}$$

Following the well-known process of maximum likelihood estimation [22]–[25], we find the arguments that maximize the log-likelihood to optimize the model fitness. For a given set of values of parameters α, β, and γ, one can compute the log-likelihood of observing the given time-series of risk materialization using Equation 4.14 and Equation 4.15. Thus, by scanning different combinations of α, β, and γ over their respective feasible ranges, and by computing the resulting log-likelihoods, one can find with the desired precision the values of α, β, and γ that maximize the likelihood of observing the data. The likelihood function is smooth (see the plot in Figure 3.2) with a unique maximum that guarantees that found parameter values are indeed globally optimal for the model considered. With the time unit of a decade, these optimal values (marked by * superscript) are $\alpha^* = 0.365 \approx 4/11, \beta^* = 0.14 \approx 1/7, \gamma^* = 427$, and the log-likelihood of observing

the data given these parameters is -415.6. We refer to so-defined model as *network model*.

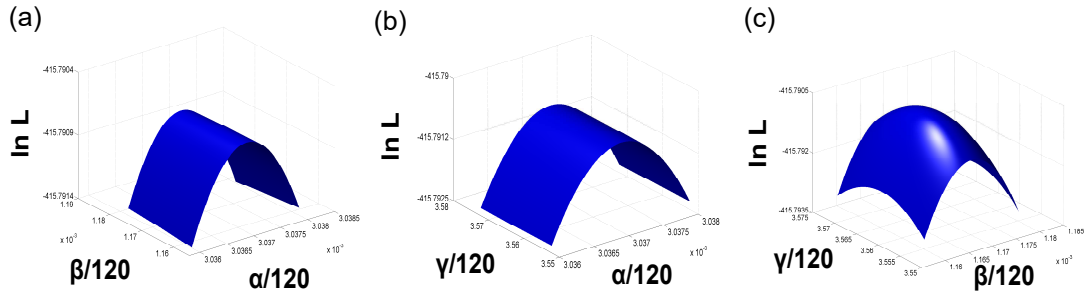


Figure 3.2: Log-likelihood of data as a function of model parameters.

3.2.5 Establishing model properties

Next, we measure how vulnerable our model is to noise in the expert data. To do this, we randomly perturb each average likelihood value provided by the experts to a value within one standard deviation from the average, and create 20 sets of such randomly perturbed likelihood data. Next, we compute 20 parameter sets that maximize log-likelihood of observing the historical data. Then, we run 20 models, termed noisy data models, defined by the obtained parameter sets. For each noisy model, we compute its monthly activity level, which is the number of risks active in each month averaged over 10^6 runs of this model. Finally, we compute the maximum differences between parameters and monthly activity levels at each month of the network model and all 20 noisy data models.

From Figure 3.3, the bars represent activity level and the blue curve represents relative error for each random likelihood result compared with basic case. Since the fluctuation of results is small, the robustness of the model is verified. The value of activity level arbitrary varies, but the sum of relative errors is very close to zero. The max relative error is no more than 1.5%. Random likelihood does not have an obvious impact on the result of risk propagation. The optimal parameters of noisy data models were within $\pm 1.4\%$ of those values for network model. Finally, the maximum log-likelihoods of noisy data models are within $\pm 1\%$ of this value for network model. Since each set of activity level is very close to the base case, the

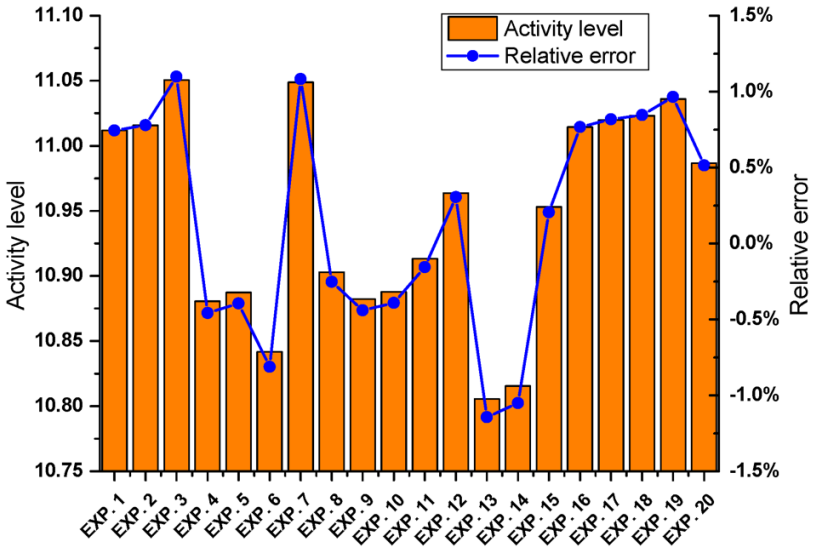


Figure 3.3: Random likelihood experiment.

robustness of the network is good enough to handle the error of likelihood. Hence, the background noise of crowd-sourcing data can be ignored.

Alternative models

We also measure the importance of network effects by comparing the maximum log-likelihood obtained above to the corresponding maximum log-likelihood value for a model which is oblivious to network effects, so has $\beta = 0$. We refer to this model as the *independent model*. With the time unit of a decade (which experts used in their likelihood assessment), the two optimal parameters are $\alpha^d = 0.91$, $\gamma^d = \gamma^* = 427$ and maximum log-likelihood is 420.1. Using the likelihood ratio (LR) test [56], we conclude the network model outperforms the independent model at a significance level of 0.01. This result demonstrates that to fully uncover the value of expert data requires accounting for network effects, as we have done in our network model.

Setting $\alpha = 1.0$ creates a model that we refer to as *expert data based model* which yields maximum log-likelihood of 420.1 that is only slightly higher (by 0.02%) than for the optimal independent model. More importantly, it results in particularly simple form for one-decade risk i materialization probability: $p_i^{\text{int}} = p_i$. This linear

mapping demonstrates that the averages of experts' assessments of risk materialization likelihoods are in fact excellent estimates of probabilities of risk failures in the ten-year period. It also attests to validity of our historical data and of expert assessments, since any mistake in those two datasets would make a mapping from expert data to probabilities a complex function. Similar high quality expert forecast in strategic intelligence was discussed in [57]. Yet, this results uncovers the limit of expert assessments, as they capture the aggregate probability of failures resulting from internal and external risk materializations without ability to distinguish between them. Since external materialization depends on which risks are active, any change in the states of the risks changes such aggregate probability. Our model, through parameter mapping onto the historical data, is able to separate external and internal materialization probabilities and therefore is valid regardless of the current or future states of the risks.

We also evaluate the value of experts' assessments of risks susceptibility to failures and their influence on each other for modeling risks. An alternative model with individual parameters for each risk susceptibility and influence would have too many degrees of freedom to be well-defined. However, the network model applied to risks with uniform likelihood and influence, a model to which we refer to as *uniform model*, and which is therefore agnostic to expert data, yields a maximum log-likelihood of 437.1, far from what the independent and expert data based models deliver. According to the LR test [56], the independent and expert data based models cannot be distinguished from each other with any reasonable significance level. However, the same test allows us to conclude that these two models outperform a simple uniform model based only on historical data with a significance level of 0.001.

Summary of models discussed here is provided in Table 3.6. Parameters for the models mentioned in the text, and the data utilized to estimate the respective parameters. Parameters α, β, and γ govern the Poisson process intensities for internal materializations, pairwise influence, and continuation, as expressed in Equation 3.6, Equation 3.8, and Equation 3.7 respectively. L_i represents the likelihood score provided for risk i by the WEF report [47], and b_{ji} is a binary variable that adopts

Table 3.6: Summary of models on likelihood-ratio test.

Model	Parameters	Data used
network model	α, β, γ	L_i, b_{ji}, historical data
independent model	α, γ ($\beta = 0$)	L_i, historical data
expert data based model	γ ($\alpha = 1, \beta = 0$)	L_i, historical data
uniform model	$\lambda^{\text{int}}, \lambda^{\text{con}}, \lambda^{\text{ext}}$	historical data

a value of 1 if the materialization of risk j is deemed to have an influence on the materialization of risk i in at least one of the experts' opinion. The expert data based model is the independent model in which value of parameter α is restricted to 1.0. The uniform model uses two parameters for the Poisson process intensities for internal materialization and materialization continuation (Equation 3.6, Equation 3.7) which are assumed to be identical for all risks. It uses the third parameter to define the influence probability between risks (Equation 3.8) which is assumed to be the same for all risk pairs. The network model outperforms all other models in explaining the observed data.

Parameters α, β, and γ govern the Poisson process intensities for internal materializations, pairwise influence, and continuation, as expressed in Equation 3.6, Equation 3.8, and Equation 3.7 respectively. L_i represents the likelihood score provided for risk i by the WEF report, and b_{ji} is a binary variable that adopts a value of 1 if the materialization of risk j is deemed to have an influence on the materialization of risk i in at least one of the experts' opinion. The expert data based model is the independent model in which value of parameter α is restricted to 1.0. The uniform model uses two parameters for the Poisson process intensities for internal materialization and materialization continuation (Equation 3.6, Equation 3.7) which are assumed to be identical for all risks. It uses the third parameter to define the influence probability between risks (Equation 3.8) which is assumed to be the same for all risk pairs. The network model outperforms all other models in explaining the observed data.

3.3 Contagion potentials of risks

Here, we investigate the relative importance of different risks.First, in analogy with epidemic studies, we calculate the *contagion potential* of individual risks, i.e., the mean number of materializations that a risk induces given that it alone has materialized. For risk i, the exact expression for this quantity is, as stated in [46]:

$$C_i = \sum_{j=1, j\neq i}^{N} \frac{p_i^{\text{con}} p_{ij}^{\text{ext}}}{1 - p_i^{\text{con}} + p_i^{\text{con}} p_{ij}^{\text{ext}}} \tag{3.11}$$

where N refers to the total number of risks. This expression assumes that risks other than i can only be activated through the influence of risk i and not internally.

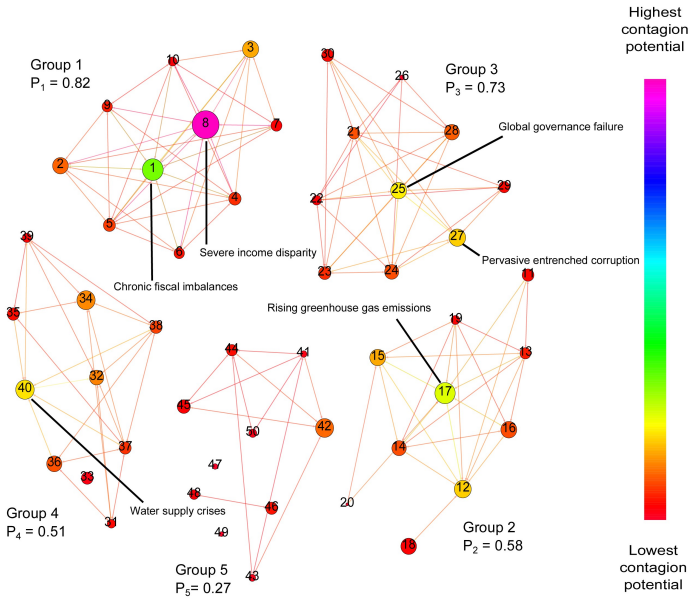


Figure 3.4: Global risk network intra-group connectivity and node congestion potentials.

Figure 3.4 shows a visualization of the network capturing the contagion potentials as well as the internal failure probabilities in the network model. As illustrated, the internal failure probability does not strictly show a positive correlation with contagion potential. Hence, a frequently materializing risk does not necessarily inflict the most harm to the system as a result of its influence on other risks. For example,

although risk 42 - "Cyber attacks" has a relatively high probability of internal materialization, its contagion potential is low. In contrast, risk 25 - "Global governance failure" has both a high probability of internal materialization and a high contagion potential. However, most striking is the fact that risk 8 - "Severe income disparity" has both the highest internal materialization probability and the highest contagion potential.

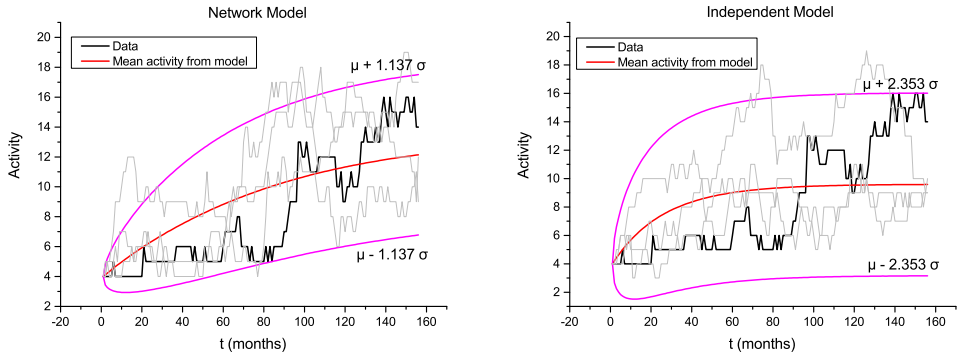
The five risks with the highest contagion potentials are: 8 - "Severe income disparity", 1 - "Chronic fiscal imbalances", 17 - "Rising greenhouse gas emissions", 40 - "Water supply crises", and 12 - "Failure of climate change adaptation". When ranked purely by raw likelihood values L_i (or equivalently by the internal failure probabilities p_i^{int}), the only change is on the fifth position, where risk 12 is replaces by risk 34 - "Mismanagement of population aging" moves up from eleventh position to fifth.

3.4 Network activity level and risk persistence

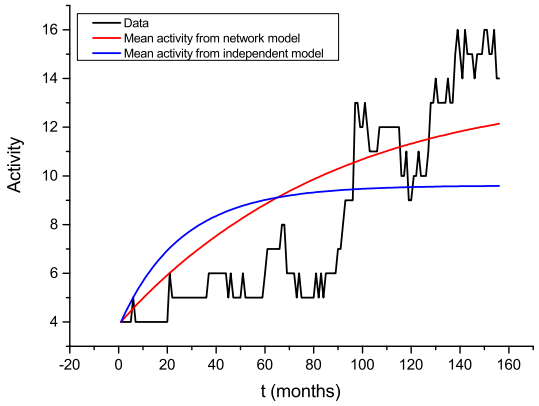
Next, we perform Monte-Carlo simulations of both the network model and the independent model.

As shown in Figure 3.5a, simulations of the network model with optimal values of all three parameters, produces a mean activity level that is commensurate with the historical data. The activity levels observed in the historical data for each month lie within 1.32 standard deviations of the mean activity level obtained from 10^6 simulations of the network model. In comparison, the most extreme activity levels observed in the historical data lie about 2.35 standard deviations away from this mean in the case of the independent model Figure 3.5b. This large difference further corroborates the fact that network effects are indeed important in reproducing the observed data. Figure 3.5c shows explicitly the comparison between mean activities produced by the two models.

Figure 3.6 shows the fraction of time steps over 10^6 simulations, each consisting 2200 time steps, that a given risk was active (the initial transient consisting of 200 steps was ignored). We call this fraction the *persistence of the risk*. Each simulation was initiated with the same active risks that are present in the first



(a) Network Model **(b) independent Model**



(c) Compare Two Models

Figure 3.5: Activity level measured as a function of time.

month of historical data (i.e. January 2000). The most persistent risk was 8 active 90% of the time, followed by risk 1, active 68% of the time, risk 17, active 64% of the time, risk 40, active 56% of the time, and risk 12, active 51% of the time.

Another interesting aspect is the distribution of the number of active risks obtained in the simulation. The 10^{th} percentile value of the number of active risks is below 8, while the 90^{th} percentile value of the number of active risks is over 19, implying that about 80% of the time, the number of active risks will lie between these two values.

The steady state (long-time limit) activity levels indicate that the *carrying capacity* of the global risk network at the present time is 27% of the size of the network, i.e., about 13.8 risks are active all the time. The top seven risks observed

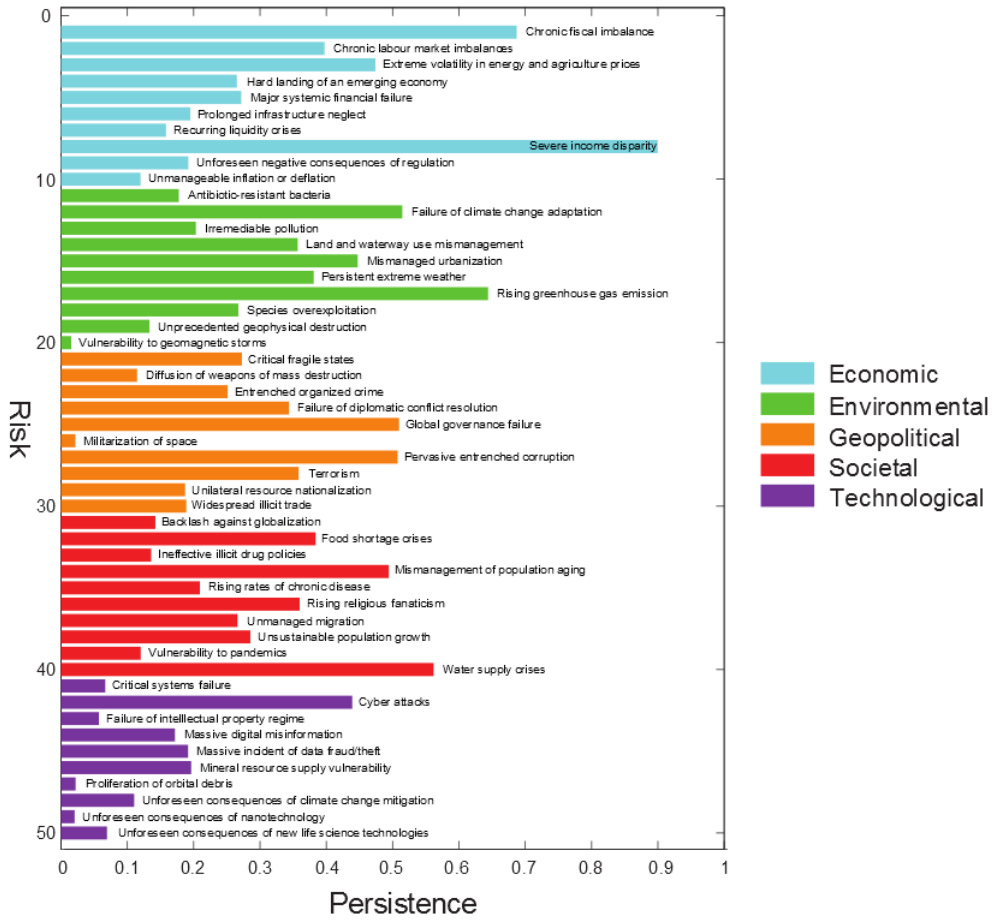


Figure 3.6: Persistence of risks in model simulation.

to be active most frequently in simulations are $8, 1, 17, 40, 12, 25$, and 27. These seven risks contribute on average 4.3 members to the total activity level at any month. This implies that other 43 risks together contribute on average the remaining 9.4 active risks, thus their activity level per risk is nearly three times lower than it is for the top seven risks.

3.5 Cascades due to single risk materializations

We further study the effect of risk interconnectivity by investigating the survival probability of a failure cascade initiated by a particular risk's materialization. Specifically, we perform 10^6 Monte-Carlo simulations of the model, each running for

50,000 time steps, starting with a given single risk active and setting the internal failure probabilities of all risks to zero. Thus, all subsequent risk materializations (after the initial failure) are caused purely by the *cascade* propagating within the network. Note that this is different from the true activity dynamics within the network discussed previously. These simulations are carried out to demonstrate the extent to which the connectivity between risks facilitates secondary activations. Shown in Figure 3.7 are the survival probabilities for cascades initiated by five highly contagious risks ranked in descending order of contagion potential. The linear nature of the curves on the linear-logarithmic scale indicates that survival probabilities decay exponentially with time. Despite that, even the cascade initiated by the *least* contagious risk in the displayed data, risk 40 "Water supply crisis", has a greater than 1% chance of continuing beyond 10,000 months, i.e., over eight centuries. These long cascade lifetimes, even in the absence of internal failures, demonstrates the profound disadvantage of interconnectivity of global risks.

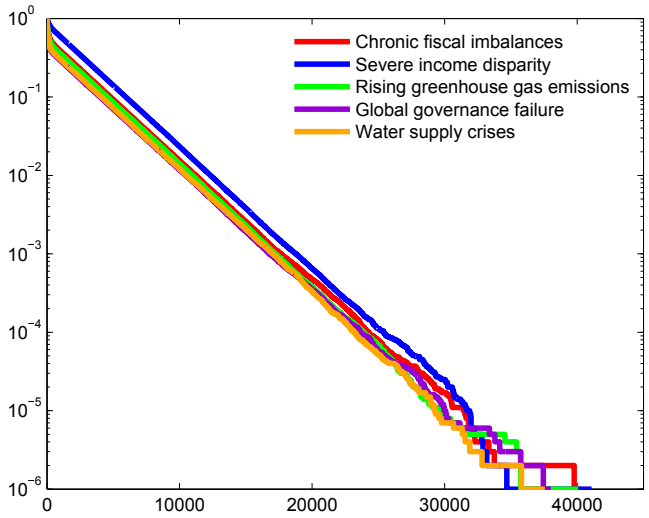


Figure 3.7: Survival probability of risks in cascades.

Next, we investigate which risks are predominantly responsible for the cascades persisting for such long time-scales. Figure 3.8 shows the expected fraction of the lifetime of a cascade for which a particular risk is active, in ranked order. The bar graph shows the fraction of the total lifetime of a cascade that a given risk

is expected to be active, as obtained from 10^6 simulations for each of 15 different initiators, where initiators are chosen from sets of risks with high contagion potential, medium contagion potential and low contagion potential. The specific risks chosen as initiators were risks $1, 8, 9, 12, 16, 20, 23, 25, 26, 27, 31, 33, 42, 47, 49$. The top five highest active risks are 8, active for 83% of the cascade lifetime, 1, active for 53% of the lifetime, 17, active for 46% of the lifetime, 40, active for 39% of the lifetime, and 12, active for 35% of the lifetime. Interestingly, the lists of top five most persistent risks observed in the cascades and seen in the full dynamics of activation (when all nodes undergo both internal and external activation Poisson processes) are identical.

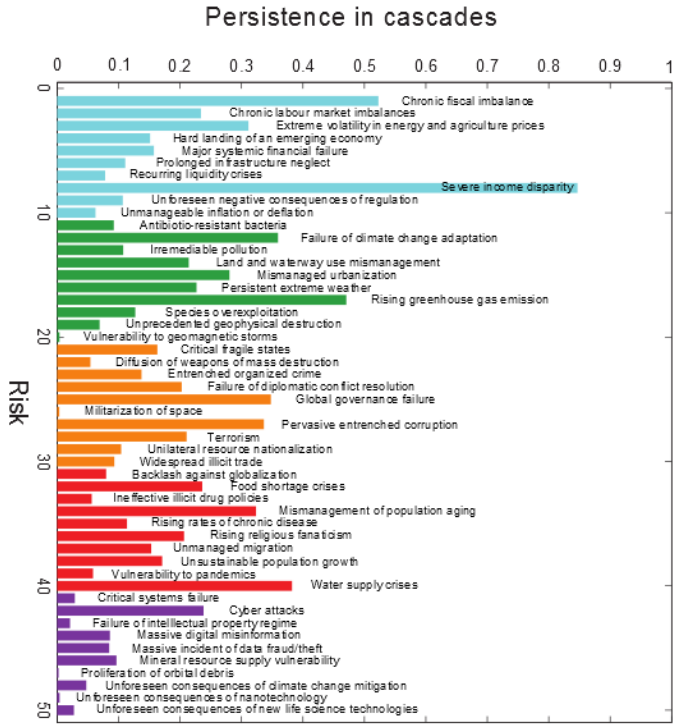


Figure 3.8: Persistence in cascades.

We also compute the probability that the cascade resulting from the materialization of a given initiator risk would result in the materialization of a selected risk as shown in Figure 3.9. The bar graphs show the materialization probabilities of four labeled risks, as a function of the initiator of the cascade. Each experiment ended when either the selected risk was infected, or all risks became inactive. Specifically, we consider the probability of materialization of the four risks, 8, 1, 17 and

25, observed to be among the top five risks most frequently active in simulations. Risk 8 is the initiator that yields the highest materialization probability for risks 1, 17, and 40, while it itself materializes with highest probability when the initiator risk is 1 and is followed by risks 17, 25, and 40. The risks 8 and 1 materialized with highest probability for initiators 17 and 25.

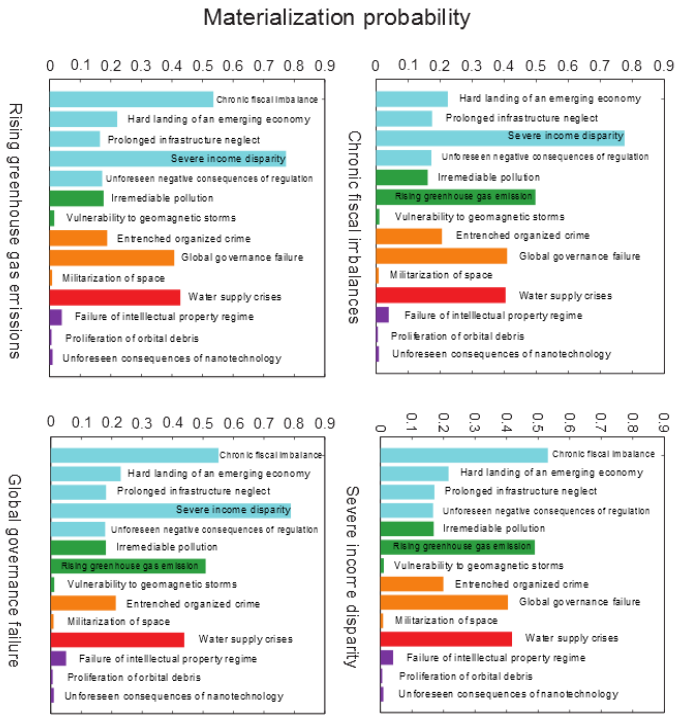


Figure 3.9: Materialization probabilities of risks in cascades.

3.6 Predicting risk materializations

In addition, we quantified the ability of our network model to predict future risk materialization events and compared its prediction errors to those of alternative models. Specifically, for evaluating prediction errors incurred by the network model, we do the following. We split the 156 months of historical data into two sets. Data from the first 132 months are used as the training set, and model parameters (α^*,β^* and γ^*) are derived using this set. The second set contains data for months $133-156$, which we used to evaluate the predictive ability of the model. Thus, this set constitutes the test set. We start simulations of 10^6 experiments at month 133

using a vector of historical data for the month 132. Then, month after month, we use the currently simulated month result to obtain subsequent month result, which emulates the process of obtaining a two-year prediction of the global risk network. In the end, for each simulated month, we compute the average frequency of risk i being active in experiments in this month and use the result as the predicted probability of risk i being active in this month.

To evaluate the quality of the model, we use a conventional measure [58] of prediction error known as cross entropy. For the case where a prediction constitutes the probabilities, $p_i^m, i = 1 \cdots N, m = 1 \cdots M$ that each of $N \times M$ distinct binary variables will adopt value 1, the cross-entropy error is given by:

$$CE = -\frac{1}{NM}\left[\sum_{i=1}^{N}\sum_{m=1}^{M} t_i^m \log\left(p_i^m\right) + \left(1 - t_i^m\right)\log\left(1 - p_i^m\right)\right] \tag{3.12}$$

where $t_i^m \in 0, 1, i = 1 \cdots N, m = 1 \cdots M$ are the realized values of the binary variables. In the present context, each binary variable is set to 1 if the specific risk i is active in the month m and 0 otherwise, so N corresponds to the number of risks and M to the number of months. The lower the CE value, the better are the predictions made by the model. In Figure 3.10a, we compare the CE for predictions generated by the network model and the independent model. (Predictions for the independent model are made using the same procedure as described for the network model, except that the value of the parameter β is zero by definition.) For the m^{th} month in the test set, the $N = 50m$ predictions made up to and including that month are utilized in the computation of CE. Although the differences are small, predictions of the network model consistently outperform those generated by the independent model, and the CE error of the latter at the end of the test period is 6.0% higher than that of the network model.

Additionally, we also evaluated the predictive abilities of two simpler models that depend only on the first 132 months of historical data and which ignore the data provided by the experts. For the first of these models, each risk i is assumed to be active at any time with probability p_i^1 which is the fraction of months in the training set for which the risk i was active. Thus, the predicted probability that a

risk is active is independent of time, according to this model. We call this model the *activity model*. The difficulty with this model arises for risks that either have not been active for the period covered in the training set, or were active for the entire period. A risk materialization for the former case, or risk recovery for the latter case in the test set will cause the prediction error for these risks to be infinite. To avoid that, we assume that materialization (recovery) will appear in the testing set if we double it on each of its sides. Accordingly, the activity probability is set to either $1/(3d+1)$ or $3d/(3d+1)$ respectively in these two cases, where d denotes the length of training data.

The second model computes the transition probabilities $\mathcal{P}_i(t)^{0\to1}, \mathcal{P}_i(t)^{1\to0}$ for each risk i directly from the training set. The first probability is for the transition from the inactive state to the active state, which is the risk materialization probability. The second probability is for a transition from the active to the inactive state which corresponds to the recovery probability. We call this model the *switching state model*, or *switching model*, in short. The same problem as in the *activity model* arises here for risks which are in the same state over the entire testing period. However, the same solution applies, which is adding needed events in the extended training period. An additional problem arises if a single month of activity or inactivity appears in the training data. For example, with only a single active month in the training set, the estimated probability of recovery, i.e., $\mathcal{P}_i(t)^{1\to0}$ would be 1 and the complement of recovery probability $\mathcal{P}_i(t)^{1\to1}$ would be 0, resulting in an error of infinity if there is any transition in the test set between two active states. To prevent the prediction error from diverging to infinity in such cases, we set the complement of recovery probability in the first case, and the materialization probability in the second case, to the frequency of risk materialization in the training period. Figure 3.10b shows how the values of CE for these two models that are agnostic to expert assessments, compares to the network model that utilizes expert data. The advantage yielded by exploiting the expert data is clearly demonstrated by the markedly lower error of the network model's predictions. We also checked that our conclusions regarding the relative predictive abilities of the various models are not dependent on the specific error measure used.

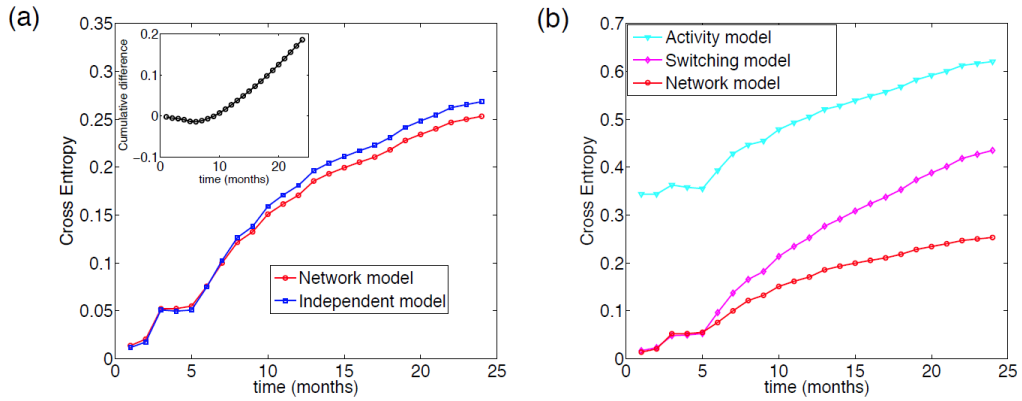


Figure 3.10: Predictability of the network model measured by cross entropy.

In addition to evaluating the predictive ability of models using cross entropy, we also used the Brier score which measures the predicted probabilities of occurrence of N events as follows [59]:

$$B = \frac{1}{NM} \sum_{i=1}^{i=N} \sum_{m=1}^{M} (p_i^m - t_i^m)^2 \tag{3.13}$$

where p_i^m is the predicted probability that the event i occurs in month m, and t_i^m is the realized outcome i.e. 1 if the event i occurs in month m and 0 otherwise. Figure 3.11a shows the comparison between the Brier metrics for the network model and the independent model. As was the case where prediction error was measured by mean cross entropy, in this case too, the network model demonstratively outperforms the independent model. Figure 3.11b shows that even using the Brier score, the prediction errors produced by the network model are far smaller than those incurred by the activity model and the switching model.

3.7 Mitigation of cascading failures

In previous sections, we demonstrate the persistent features and cascading failures of most concerned risks. In addition, we also introduce the details of model dynamics and how to estimate the parameters from the historical dataset. Besides these achievements, another important application is to mitigate the damage of

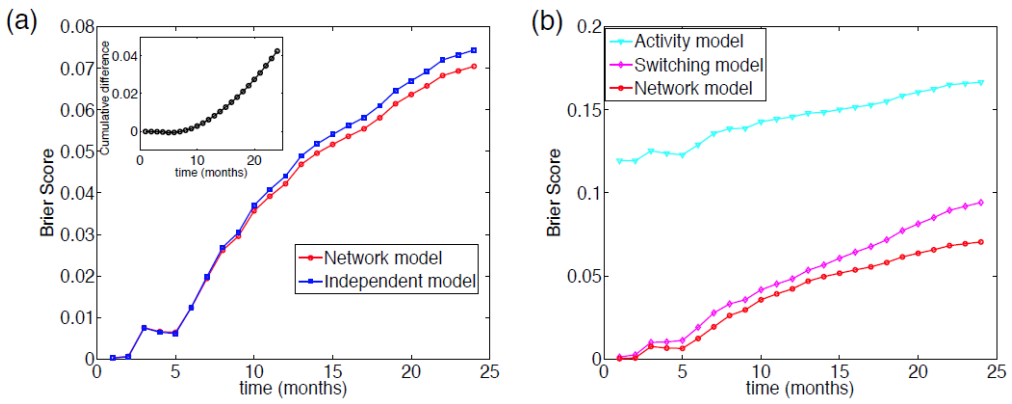


Figure 3.11: Predictability of models measured by Brier score.

cascading damages. As we know, there could be many efficient mitigation strategies. What's more, since the resource of our society is limited, we cannot afford to control every part of the global risk network. We have to find a good balance between the cost and benefit. Hence, we extend our study to develop an efficient method to control the network and reduce the damage.

3.7.1 Reducing the likelihood and connections

Which part of the network should we control? To answer this question, we have two choices: risks and connections. By controlling risks, we reduce the likelihood of each risk to some extent. The value of likelihood defines directly the ability to trigger the risk and indirectly the activity level for one risk. Since the controlled risks will be more passive, fewer materializations make the whole system more stable. By controlling connections, we remove some edges of targeted risks to decrease the network connectivity. It is intuitive that the number of possible paths for cascade propagation is lower in a sparse network so that the vulnerability decreases accordingly. It seems hard to determine which one is a better choice. We design the following experiment to detect the effects of these two factors. We assume three kinds of intervention strategies:

(1) I1: halving the likelihood of top ten risks

(2) I2: halving edges of top ten risks

(3) I3: combining interventions I1 and I2

The reason for controlling top 10 risks is due to the limits of available resources. It would be prohibitively expensive to control all the risks together. As we mentioned before, several persistent risks contribute most to the activity level of the whole system. Choosing top 10 risks can be a good start to probe the difference between likelihood and connections.

Next question is how to determine the top 10 risks. There are various metrics for ranking the risks. And the top 10 risks will change according to different rankings. Based on the available information about global risks, we formulate the following 9 different ways to rank the risks:

(1) R1: by internal probability: p_i^{int} for risk i.

(2) R2: by the product of internal probability and expert-assessed impact: $p_i^{\text{int}} \times impact_i$ for risk i.

(3) R3: by the product of internal probability and exponential value of impact: $p_i^{\text{int}} \times e^{impact_i}$ for risk i.

(4) R4: by persistence: P_i

(5) R5: by the product of persistence and expert-assessed impact: $P_i \times impact_i$ for risk i.

(6) R6: by the product of persistence and exponential value of impact: $P_i \times e^{impact_i}$ for risk i.

(7) R7: by contagion potential: C_i for risk i.

(8) R8: by the product of contagion potential and expert-assessed impact: $C_i \times impact_i$ for risk i.

(9) R9: by the product of contagion potential and exponential value of impact: $C_i \times e^{impact_i}$ for risk i.

Internal probability p_i^{int}: a risk i materializes due to internal factors with probability p_i^{int}. This value is computed, as described in Section 3.2.1 of this thesis, based on the expert assessment given in the WEF 2013 report [47].

Persistence of the risk P_i: this value is defined in previous experiment (see Section 3.4 of this thesis): the fraction of time over 10^6 simulations, each consisting 2200 time steps, that a given risk was active (the initial transient consisting of 200 steps was ignored). A risk with a higher persistence will stay active longer than other risks once it materializes. For example as shown in Figure 3.6, risk 8 - "Severe income disparity" stays in materialized state for more than 90% of time. Risk 1 - "Chronic fiscal imbalance" and risk 17 - "Rising greenhouse gas emission" both have a high persistence because they are materialized for more than 60% of time.

Contagion potential C_i: the mean number of materializations that a risk induces given that it alone has materialized. For risk i, the exact expression for this quantity is, as stated in [46] and in Section 3.3 of this thesis:

$$C_i = \sum_{j=1, j\neq i}^{N} \frac{(1 - p_i^{\text{rec}})\, p_{ij}^{\text{out}}}{p_i^{\text{rec}} + p_{ij}^{\text{out}} - p_i^{\text{rec}} p_{ij}^{\text{out}}} \tag{3.14}$$

As shown in Figure 3.4, the color of nodes demonstrates the value of contagion potential. Only a few risks have a high contagion potential since this value combines the internally triggering ability together with connections. Most risks enjoy a low value of contagion potential.

1. ***Expert-assessed impact*** $impact_i$: the impact of risk if it occurs. The experts give the evaluation in the report [47]. These values reflect the risk impact perceived from the global economy point of view. We combine these 9 different rankings with 3 interventions to detect which strategy reduces the damage of cascading most. Totally there are 27 combinations. To measure the damage, we design three different metrics:

2. ***Activity level***: the sum of probabilities of each risk being active.

3. ***Linear impact***: the sum of products of each risk probability of being active and its impact value.

4. ***Exponential impact***: the sum of products of each risk probability of being active and its exponential impact value.

Activity level only considers the expected number of materializations by summing up probabilities of each risk. The "Linear impact" and "Exponential impact" consider the effect of evaluated impact. In all 27 cases, the initial conditions for simulations are the same: using the first month of historical data and simulating 1000 time steps. The fraction of time steps that each risk is active is recorded in the simulation. We finish 10^6 independent realizations and average the fraction of the active time. In the end, we compute the three metrics defined above to show the mitigation results.

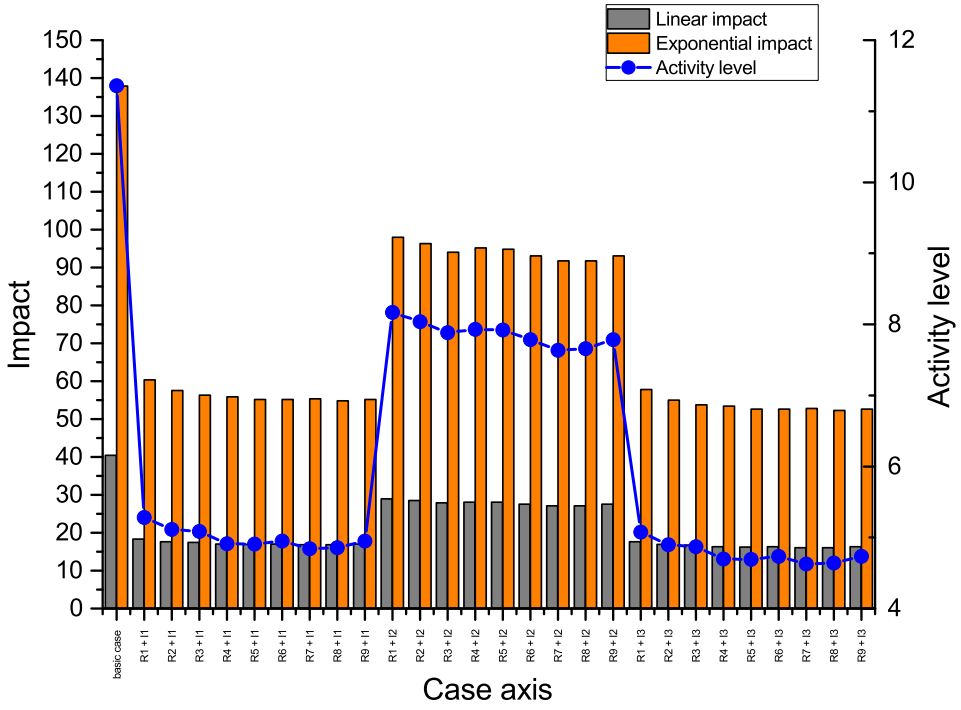


Figure 3.12: Mitigation strategies results.

As shown in the Figure 3.12, among all the cases, the basic case has the highest activity level, linear impact, and exponential impact since there is no intervention involved. It is evident from the plots that interventions on the likelihood of risks perform much better than those eliminating edges. When looking at the activity

level, the basic case has a value about 11 expected materializations in a steady stage. In the cases applying I2 intervention (reducing the connections), no matter which ranking to use, the activity level drops to around 9 which is a 18% improvement. In the cases applying likelihood mitigation (I1 and I3), the activity level becomes much lower than other cases, which is a 40% improvement. These results are valid under the assumption that the cost of reducing half of the connection is the same as the cost of halving the likelihood which may depend on the number of connections the node have. The performance of all I3 cases is slightly better than I1 cases, which is not surprising since I3 is more costly than either I1 or I2. Moreover, it is hard to distinguish the difference between all rankings. The ranking of risks is not critical for interventions. Since nine ranking categories have almost the same results, it is reasonable to conclude that ranking is not as important as controlling the likelihood. This trend is similar when we focus on another two metrics: linear impact and exponential impact. Based on this figure, the improvement comes from mitigation of likelihood instead of edges.

3.7.2 Mitigation cost for various number of risks

As discussed in the previous section, controlling likelihood is a better choice, then controlling connectivity, so next step is to find how much we should change it. One simple idea is to manage all risks and reduce their likelihood to some extent. However, it is not feasible to do so since in reality, because of the prohibitively high cost of mitigation of every risk in the network. We need to find the optimal outcome based on the available resources. Due to the limited budget, the more risks we control, the weaker our mitigation can be and vice versa. Hence, there should be a relationship between the number of mitigated risks and the scale of the mitigation. It is very complicated to find a precise formula to represent this relationship. So we start with a simple equation defined as follows:

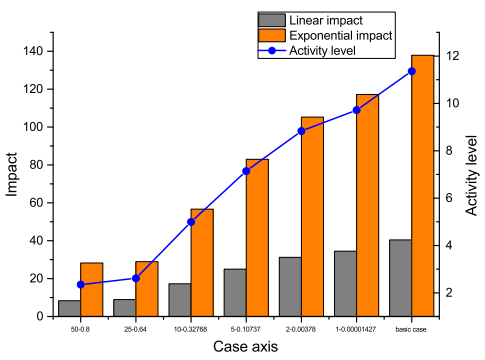
$$d_k = d_N^{N/k} \tag{3.15}$$

where d_N is the reduction level for all $N = 50$ risks, d_k is the reduction level for $k < 50$ risks. The new likelihood is $L_i \times d_k$. Given the value of d_N and k, we can

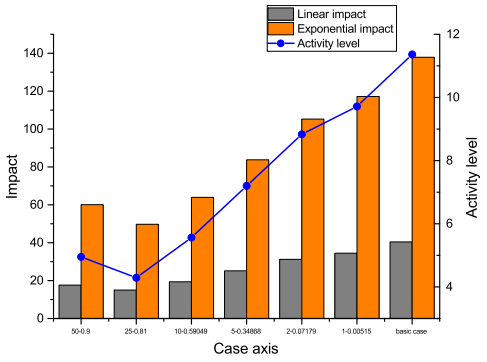
compute the corresponding value for d_k. We assume an exponential relationship between the number of controlled risks and the value of reduction level instead of a linear one.

In this experiment, we set the value of d_N to be $0.8, 0.9, 0.95, 0.99$ and 0.999. The initial condition is the first month status in historical data and simulation time length is 1000 steps. We run 10^6 realizations. The definition of three metrics is the same as in the previous experiment. The mitigation focuses only on the likelihood of risks: we just control different numbers of risks and keep the connections unchanged. In the following experiment, we vary the number of controlled risks (k) to be $1, 2, 5, 10, 25$ and 50.

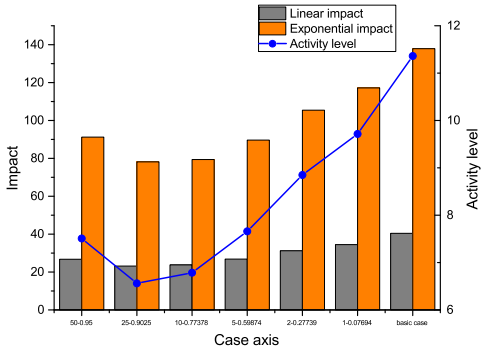
In Figure 3.13, we demonstrate the mitigation results for 5 different reduction levels d_N. For each value of reduction level, we vary the number of controlled risks: $1, 2, 5, 10, 25$ and 50. The more risks we control, the lower the reduction level is. For example, if the reduction level for 50 risks is 0.8, then the corresponding reduction level for 25 risks is 0.64, for 10 risks it is 0.32768, for one risk is 0.00001427. The controlled risks are selected according to internal probability (R1). In the case $d_N = 0.8$ as shown in Figure 3.13a, the mitigation results change monotonically. As the number of controlled risk increases, the mitigation improves. This trend is the same for activity level, linear impact, and exponential impact. Hence if we control on half of the nodes and for all nodes, the improvement is much better than controlling on a smaller proportion of the nodes. The improved cascading impact for controlling on all nodes is only 20% of the basic case. However, in the case of $d_N = 0.9, 0.95$ as shown in Figure 3.13b, Figure 3.13d and Figure 3.13e, the curve for activity level, bars for linear impact and exponential impact show a non-monotonic behavior. The best performance is achieved when controlling 25 risks instead of 50 risks. In these scenarios, we should focus on half of the nodes. In the last case as shown in Figure 3.13e, the results of impact remain almost unchanged since the d_N is very close to 1 so the mitigation effect is not obvious. In general, according to different values of reduction level d_N, the optimal strategy may not be controlling all the nodes or a small part of them. Based on the performance curve, we can design an optimal mitigation strategy.



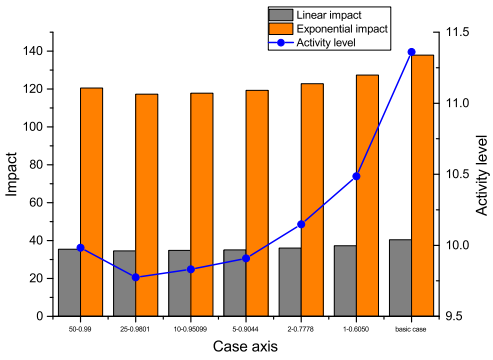
(a) $d_N = 0.8$



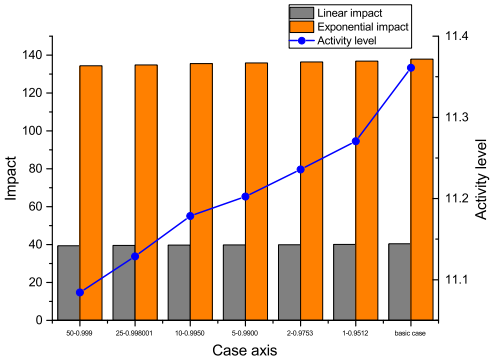
(b) $d_N = 0.9$



(c) $d_N = 0.95$

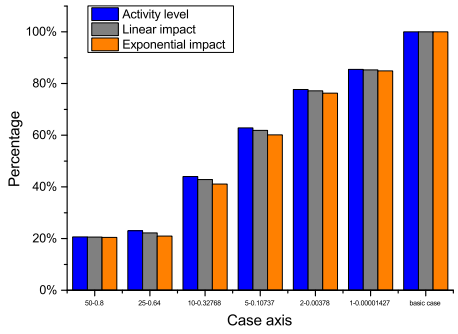


(d) $d_N = 0.99$

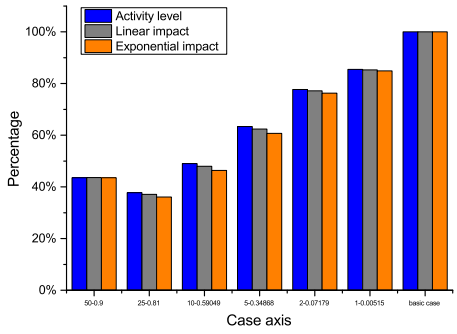


(e) $d_N = 0.999$

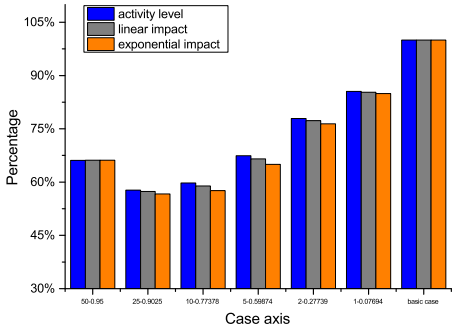
Figure 3.13: Mitigation with different reduction levels.



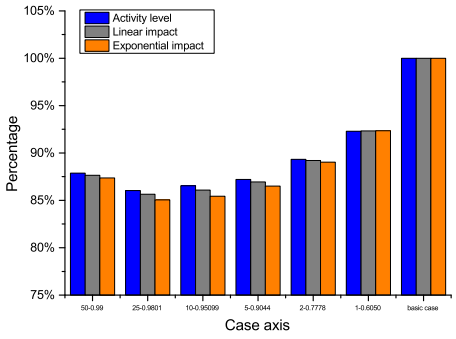
(a) $d_N = 0.8$



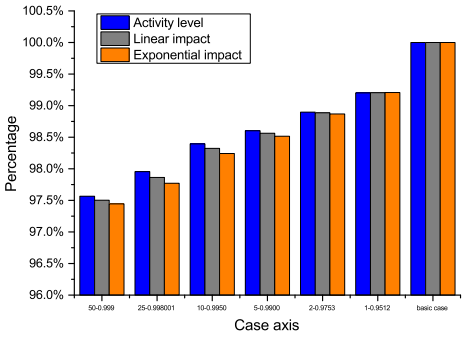
(b) $d_N = 0.9$



(c) $d_N = 0.95$



(d) $d_N = 0.99$



(e) $d_N = 0.999$

Figure 3.14: Normalized mitigation results with different reduction levels.

Figure 3.14 show the simulation results normalized by the basic case. The smaller proportion means a better mitigation improvement. In the case when d_N is equal to $0.9, 0.95$ and 0.99, the normalized results behave in a similar non-monotonic way regarding the number of controlled risks and mitigation performance. In these cases, the best performance arises around with controlling 25 risks.

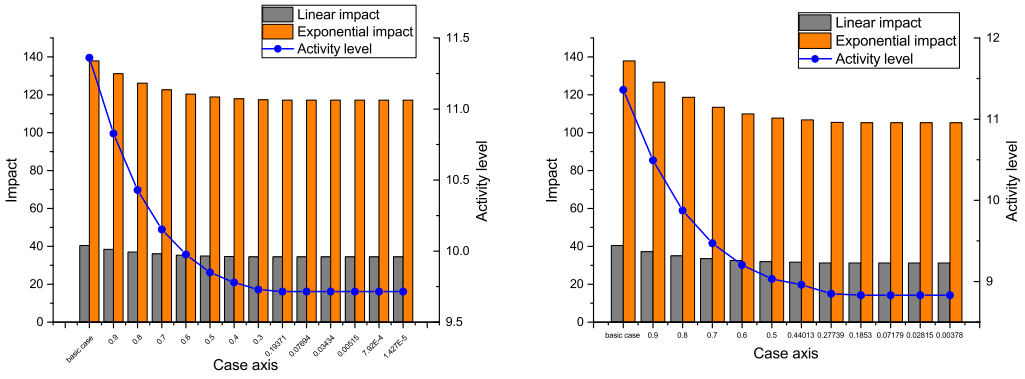
Figure 3.15 compares 6 different numbers of controlled risks: $1, 2, 5, 10, 25$ and 50. X-axis enumerates different levels of reduction factor d_N. The left y-axis measures the impact while the right y-axis shows the activity levels. In each case, we vary the value of reduction levels. All results are normalized by the value in basic case. Simulation time steps are 1000 and the number of realization is 10^6. In the cases with a small number of risks, the difference between various reduction levels is not evident. With a large number of risks, activity level and impact both drop dramatically as the reduction level increases. These figures provide some guidances how to choose the optimal value of controlled risks.

In the Figure 3.16, the x-axis shows the case index as a pair of numbers, such as $1-0.9$, $50-0.1$ and so on. The left y-axis is for impact value, and the right y-axis is for activity level. The first digit is the number of changed nodes and the second number is the reduction factor. In this figure, the bars are sorted from largest to smallest, and the curve becomes lower gradually but not linearly. As we expected, controlling on a large part of nodes increases reduction dramatically. If we have to configure a strategy, it is necessary to influence as many risks as possible.

3.8 Conclusion

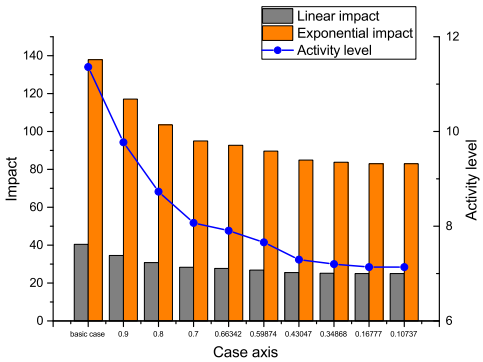
To summarize, in this study we have presented a method of obtaining a quantitative insights into the global risk network, starting from the qualitative observations provided by 1000 WEF experts. We assume a three parameter network model for the propagation of risk materialization (representing the corresponding network node failures), and obtain maximum likelihood values for the parameters using historical data on risk materialization.

Our model was built upon expert assessments available in the WEF report which enabled us the construction of a detailed and heterogeneous weighted net-

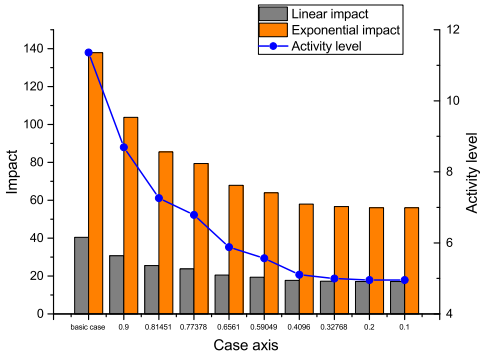


(a) 1 **risk**

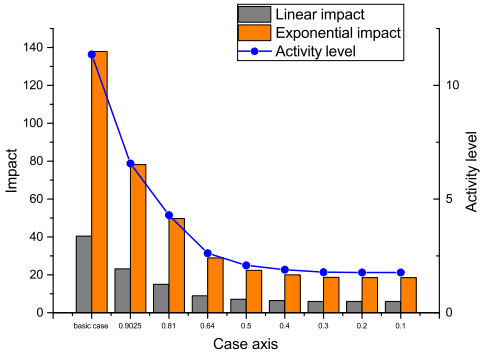
(b) 2 **risks**



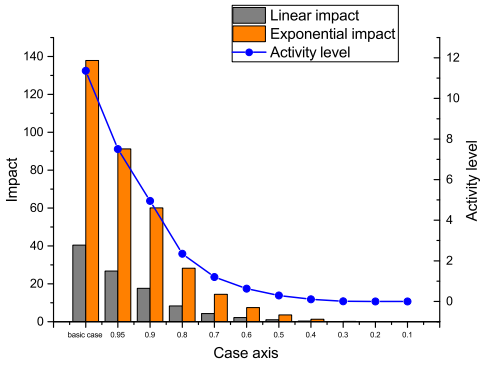
(c) 5 **risks**



(d) 10 **risks**



(e) 25 **risks**



(f) 50 **risks**

Figure 3.15: Mitigation for various controlled risks.

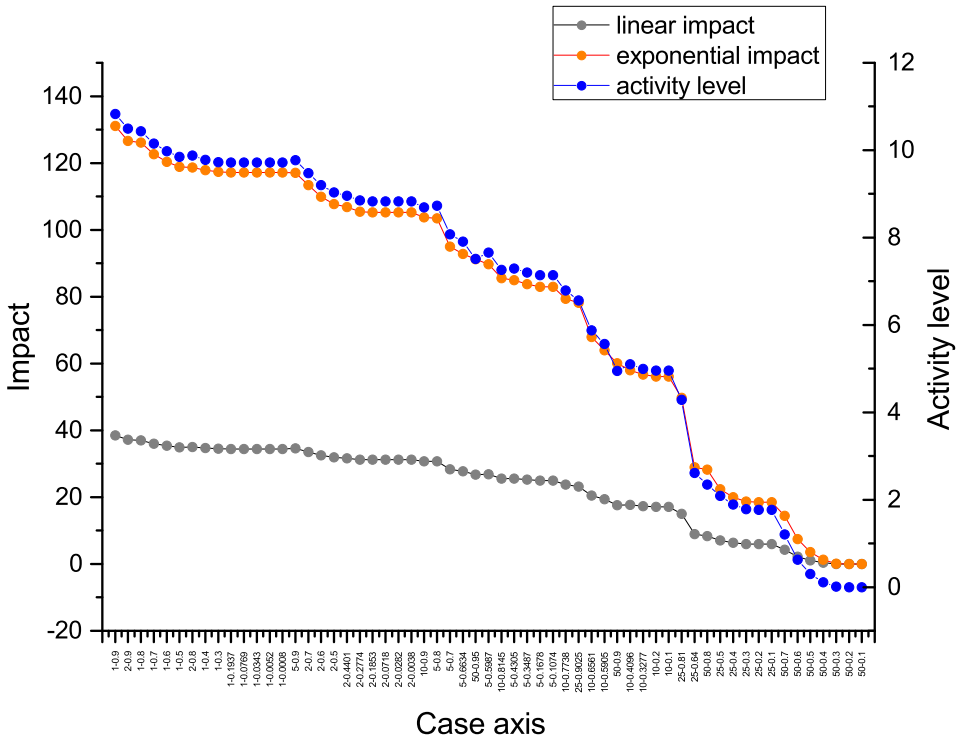


Figure 3.16: Mitigation comparison for all cases.

work of risks. As we show, ignoring network effects (i.e. the independent model) or ignoring specific heterogeneities in the failure likelihoods and influence (i.e. the uniform model) yielded poor results in comparison to the network model. This underscores the importance of the expert assessments in building a model capable of matching the available activity data well, and therefore yielding reliable insights. We have also found the greatest strength of expert assessments, which is nearly perfect forecast of aggregated failure probabilities of different risks, but also those assessments greatest weakness, which is inability to separate external risk materialization probabilities form internal ones. We have developed an approach in which by selecting proper model parameters and using maximum likelihood estimation to find optimal model parameters, we are able to do such separation.

We have uncovered the global risk network dynamics and measured its resilience, stability, and risks contagion potential, persistence, and roles in cascades of failures. We have identified risks most detrimental to system stability and measured

the adverse effects of risk interdependence and the materialization of risks in the network. According to these studies, the most detrimental is risk 8 -"Severe income disparity". Other risks that play a dominant role due to either their contagion potential or their persistent materialization are: 1 - "Chronic fiscal imbalances", 25 - "Global governance failure", 27 - "Pervasive entrenched corruption", 12 - "Failure of climate change adaptation", 17 - "Rising greenhouse gas emissions", and 40 - "Water supply crises".

Utilizing the complete network model generated using the WEF data provides a much more detailed picture of the threat posed by different risks than the one obtained by simply relaying only on their failure-likelihood L_i values and using the independent model. Additionally, our analysis demonstrates that the carrying capacity of the network i.e. the typical activity expected in the network given the current parameters, is about 13.7 risks or 27% of the total number of network nodes, of which four are persistently chosen from a subset of seven risks (see Figure 3.6). Aiming to reduce this overall carrying capacity could potentially be an overarching goal of global risk minimization.

We have compared the global risk network with several simpler models which ignore the connections or use uniform expert assessments. The network model utilizing crowd-sourcing evaluation has the best performance based on likelihood ratio test and cross entropy error score. In addition, we also detected the mitigation strategies in global risk network. We found that mitigation of the likelihood of risks leads to a better reduction of damage than mitigation on connections. In order to optimize the benefit of mitigation expense, we assumed an exponential relationship between the number of controlled risks and reduction level. Interestingly, there is a non-monotonic behavior for the mitigation improvement against the controlled risks in some cases. Hence the most efficient strategy is not to control the entire network, sometimes, controlling on half of the network may benefit the network most. It is complicated to design an optimal mitigation strategy due to the complex topology and imperfect expert assessments.

There are several prospects for extending the model that we presented here and its further analysis. First, obtaining more robust historical estimates of risk

materialization may help us improve the fitting of the model. Secondly, it will be beneficial to account for slow evolution of network parameters in time. This change in network characterization will be captured by a model through expert data provided in yearly WEF reports, resulting in time dependent L_is and b_{ij}s. Furthermore, the accuracy of the model could possibly be improved by assuming the existence of different dynamics for chronic risks as compared to sporadic risks.

From a larger perspective, our attempt here has been to utilize data crowd-sourced from experts towards gaining a quantitative picture of the network of global risks, which in turn has yielded some actionable insights. The network by definition has risks of varying complexity, which arguably makes the risk mitigation process more involved for some risks than for others. In such a scenario, our quantification of the relative impacts of different risks could provide a valuable guidance to any cost-benefit analysis involved in the design of policies or strategies aimed at global risk minimization.

CHAPTER 4
LIMITS OF PREDICTABILITY IN A CASCADING ALTERNATING RENEWAL PROCESS MODEL

Most risk analysis models systematically underestimate the probability and impact of catastrophic events (e.g., economic crises, natural disasters, and terrorism) by not taking into account interconnectivity and interdependence of risks. To address this weakness, we propose the Cascading Alternating Renewal Process (CARP) model to forecast interconnected global risks. However, assessments of the models prediction precision are limited by lack of sufficient ground truth data. Here, we establish the prediction precision using alternative long historical data generated by simulations of the CARP model with known parameters. We illustrate the approach on a model of fires in artificial cities assembled from basic city blocks with diverse housing. The results confirm that parameter recovery variance exhibits power law decay as a function of the length of available ground truth data. Using CARP model, we also demonstrate the estimation the real-world prediction precision for the global risk model based on the World Economic Forum Global Risk 2013 Report [46], [47]. We conclude that the CARP model is an efficient method for predicting catastrophic cascading events with potential applications to emerging local and global interconnected risks.

4.1 Introduction

A generalized Alternating Renewal Process model referred to as Cascading Alternating Renewal Process (CARP), has been recently proposed for dynamically modeling a global risk network represented as a set of Poisson processes [46]. This

Portions of this chapter previously appeared as: B. K. Szymanski, X. Lin, A. Asztalos, and S. Sreenivasan, "Failure dynamics of the global risk network," *Sci. Rep.*, vol. 5, no. 10998, Jun. 2015. Accessed on: Apr. 21, 2017. [Online]. Available: https://www.nature.com/articles/srep10998.

Portions of this chapter have been submitted to: X. Lin, A. Moussawi, G. Korniss, J. Bakdash and B. K. Szymanski, "Limits of risk predictability in a cascading alternating renewal process model," submitted for publication.

approach raises questions about the reliability of such recovery and the limit of the prediction precision in different scenarios. Since most of the world's critical infrastructures, including the global economy, form a complex network that is prone to cascading failures [47], this question is important but difficult to answer given the lack of ground truth data. Here, we discuss how the limits of prediction precision can be established by simulating the CARP model with known parameters to generate many alternative ground truth datasets of arbitrary length. We develop a model of fire propagation in an artificial city with different sizes of houses and time periods. Next, we use maximum likelihood estimation (MLE) [22]–[25] to recover the parameters and calculate the error of parameter estimation from the ground-truth values. We also measure the recovered parameter precision as a function of (i) the complexity of the system (in our case the size of the cities) and (ii) the number of events present in the historical data. Finally, using real-world data with the developed methodology, we assess the precision of parameter recovery in the World Economic Forum model [46].

4.1.1 Risk modeling

Most quantitative risk analysis models (e.g., Value at Risk and Probability-Impact models) systematically underestimate the probability and impact of worst-case scenarios (i.e., maximum loss for a given confidence level, typically "tail" outcomes) for catastrophic events such as economic crashes, natural disasters, and terrorist attacks [61]. Underestimation in such models is due to speciously assuming the sequences of random variables in probability distributions are normal, independent, and identically distributed (IID) [62], [63] thus discounting the potential of interdependencies and interconnections between events.

Few quantitative risk models capture the non-IID properties of risk factors and their impacts. The Havlin model uses branching to predict cascading failures (e.g., power grids, communication networks) [16]. A small fraction of initial failures could cause catastrophic damage in mutually dependent systems. The authors used the percolation theory and detected a phase transition for the robustness and functionality of the interdependent networks. The Ganin model of resilience provides

an analytical definition for determining critically to design more resilient technical systems [64]. Moreover, even when interconnections are included in risk modeling, the traditional method for determining the implied correlations (among assets) in financial models misestimates their values, resulting in potentially massive underestimation of both the probabilities and impacts of the decrease in value [65].

While above techniques go beyond simple risk models, which assume risks exhibit independent probabilities and impacts, neither aims to quantify the limits of predictability for interconnected risks. In contrast, the CARP model is a novel method in which interdependencies and interconnections are explicitly represented and the model parameters are recovered from historical data using maximum likelihood estimation. Moreover, the CARP model offers a quantitative risk assessment for interconnected conceptual models such as Reason's Swiss cheese model of failure [66]. In Reason's model, defenses preventing failure are layers (of Swiss cheese) and when the holes of the layers align a risk may materialize. Previously, Reason's model has been formalized using percolation theory [67], but this formalization has not been validated.

4.1.2 Alternating renewal process

The definition of Alternating Renewal Process (ARP) originated in renewal probability theory. A simple renewal process alternates between two states: the normal state representing its operational time and the abnormal (failure or repair) state representing its holding (non-operational) time. An independent Poisson process governs each state [68], [69]. The Alternating Renewal Process can be used to find the best strategy for replacing worn-out machinery [70], [71]. A real-world example is the electrical devices. The inter-arrival time between equipment failures follows an exponential distribution. For each device, the evolution of states is a stochastic process as a function of time. And we can find the asymptotic limit of the proportion time of each state [72]. Once we get the expected length of the normal and abnormal state, the probability of the system in the normal state is just the proportion of time in the normal state over total time. Here, we introduce recent studies on the features and applications of ARP model.

Roldán et al. studied the distribution of precipitation in the United States and compared results from two methods: a first-order Markov chain and the Alternating Renewal Process [73]. The stochastic process of precipitation has observations of binary states (wet and dry days), which come from five National Weather Service daily stations in the United States. The period of observations is more than 20 years. To consider the seasonal variation, the authors either assigned constant parameters or used finite Fourier series. *Woolhiser et al.* developed a model using first-order Markov chain to simulate the precipitate process and calculate the coefficients using maximum likelihood estimation [74]. The likelihood-ratio test verifies the quality of results. Also, the ARP model simulates stochastic process consisting observations of discrete states. The intervals of a dry day are assumed to be independent of each other, so a conditional probability p_{ij} represents the likelihood of state transitions [73].

Given the probabilities of state transitions and historical observations, the authors used maximum likelihood estimation to find the optimal values of parameters. *Roldán et al.* concluded that the ARP method requires longer running time and higher computing ability to finish the optimization process and gives us similar results as the method of first-order Markov Chain. The methodology in this thesis utilizes a similar procedure. We start with the ARP model to determine the distribution details; then we use maximum likelihood estimation to approximate the parameters of the proposed distribution.

Besides the application in weather services, the ARP model can also be used in manufacturing industry. In the coating process, there is a strong desire to predict the distribution of coating mass on particles or items. *Freireich et al.* formulated an ARP model to approximate this distribution based on the observations of coating mass per visit and the cycle time of the process [76]. The ARP model has the advantage of easy accomplishment and compact expression. Since there are small fluctuations during the coating process caused by many factors, the coating mass of particles follows an unknown distribution. It is critical to detect the details of this distribution to control the quality of coating operation. A well-known metric to reflect the degree of coating uniformity is the coefficient of variation (CoV), which

is defined as the ratio between standard derivation of coating mass and the mean value of coating mass [76]. Since the coating mass is proportional to the time of spray step, the authors focused on the distribution of coating time. *Freireich et al.* also introduced several alternative models. First is the Bernoulli trial model, which focuses on the fraction of time for spraying one particle in the spraying step. The expression is defined as [76]:

$$CoV = \sqrt{\frac{1}{\alpha}\frac{\triangle t}{t}} \tag{4.1}$$

where α is the mean value of time fraction for particles staying in the coating spray step, $\triangle t$ is the actual time in spray step in one trail, and t is the total time spent in the coating process. Other factors, besides the time spending in spraying step, could affect the coating mass distribution, such as location, waiting time, and velocity. During each simulation, the values of these factors are selected randomly from a designated distribution. In the end, statistical conclusion is summarized from many realizations. In the population balance model, the coating mass is a function of time and we can use partial differential equations to reflect the dynamics [76]:

$$\frac{dm}{dt} = JA\alpha(d) \tag{4.2}$$

where m is the mass of coating and J is the average value of spray flow at the unit of time. A is the area of particle projection, d is the size of the particle, and α is the mean value of time fraction for particles staying in the coating spray step. As the spray process goes on, the value of mass m, projected area A, and particle size d all increase gradually. In the ARP model, *Mann* firstly introduced the use of renewal theory to represent the distribution of coating mass [77]. The main difference of this model is that the coating mass and coating time are two independent variables. The coefficient of variation is defined as [76]:

$$CoV = \sqrt{\frac{\mu_C}{t}\left[\left(\frac{\sigma_W}{\mu_W}\right)^2 + \left(\frac{\sigma_C}{\mu_C}\right)^2\right]} \tag{4.3}$$

where σ_W is the standard deviation of coating mass during the spray process and σ_C

is the standard deviation of the coating time. μ is the mean value of these variables. In this ARP model, coating mass and spray time both play important roles in determining the quality of coating procedure. This method also demonstrates an asymptotic CoV expression for a long term running time. *Freireich et al.* concluded that this ARP model with a simple expression is an excellent complement to other models.

In addition to above studies, there are many other applications of the renewal theory. *Samuels* stated that the superposition of two ARPs with a positive inter-arrival time is also an ARP with a positive inter-arrival time [78]. *Ferreira* extended this conclusion to the ARPs with the inter-arrival time close to zero [79]. The renewal processes can follow either Poisson or binomial distributions. These conclusions inspire us since the state transitions in these studies follow a Poisson distribution.

Another application of the ARP model is in the health insurance industry. *Silvestrov et al.* indicated that a homogeneous ARP model could predict how many claims the customers will report in the future and the amount of payment for the illness or temporarily disabled clients [80]. An insured person only has binary status: healthy or ill, which are independent in most cases. Hence, the ARP model has a good applicability in this scenario. *Young et al.* demonstrated the ability of the ARP model to simulate the traffic flow in high-speed communication networks [81]. There are also binary states in this scenario. When the state is on, data flow is traveling from the source node to the target node along all possible paths in the system. In contrast, in the off state, there is no data transmitting. By tunning appropriate parameters, the ARP model is an efficient method to capture the dynamic features of data flow in a communication network. In addition, ARP model can also be used in neural science. *Steele et al.* designed an ARP model to represent the buildup function which simulates the switch of perception between two states: integration and segregation [82]. Buildup function is widely used in a behavioral experiment to indicate the ability of adaptation. The principal purpose is to capture model dynamics from short-term observations and estimate long-term behaviors of perceptual groups. *Steele et al.* assumed the waiting-time interval

for each event is independent, and the analytical solution agrees with the Monte Carlo simulation results. The authors concluded that the ARP model captures the transient dynamics of the perceptual organization very well and make good predictions for long-term behaviors.

The CARP model expands ARP in several ways. The most important extension is to allow for defining risks as a network of risks with the given set of weighted edges representing interdependencies and probabilities of state transitions associated with each edge. Since state transitions are often observed without distinction for the inner or external causes or triggers, the latter representing, edge induced transitions, both are treated as hidden variables, since only their joined effect is observed directly. Another generalization is not restricting the number of states to two. Finally, we associate with CARP an MLE procedure for recovery of CARP parameters from historical data of the system evolution over time.

4.1.3 Model structures

The city modeled is diverse. A small fraction of houses is large and placed sparsely with low fire propagation probabilities and high recovery rates. A larger fraction of houses with a medium size have a higher spatial density and fire propagation probabilities. Finally, the most densely packed small houses have the highest fire propagation probabilities, and lowest recovery rates.

Houses are grouped into blocks stitched together into a continuous city as shown in Figure 4.1. A sample city consists of four basic blocks. Three types of houses are represented as nodes. Red circles show the range of external fire spread. All houses in this circle may catch fire from the burning center node. Large and small houses sit on the West and East boundaries of the city, respectively. The North and South boundaries border all three types of houses. The houses are placed on rectangular lots whose size is commensurate with the size of the house. A large house occupies a square lot of one by one unit size. There are four such squares horizontally laid out with no space in between. Each medium house has a rectangular lot of half unit length vertically by one unit length horizontally. Each block holds 8 medium houses. Finally, repeating a similar pattern, small houses sit on a rectangular lot of

one by a quarter unit size. One block contains 16 small houses. Distances between two adjacent houses vary from 1 for large houses to 1/2 for medium and 1/4 for small houses. A basic block holds 28 houses on 12 square units of land. A city of arbitrary large size can be built by repeatedly adding blocks to it. Figure 4.2 shows the degree of each house, which is defined as the number of houses within a fixed distance. As shown in Figure 4.1, all houses within a fixed distance may catch fire from the burning center house. Hence, the small houses have a relative higher degree because of the high density of nearby houses. The surrounding houses on the boundary of the city and on the border of different communities experience a slightly smaller degree compared with other same-size houses.

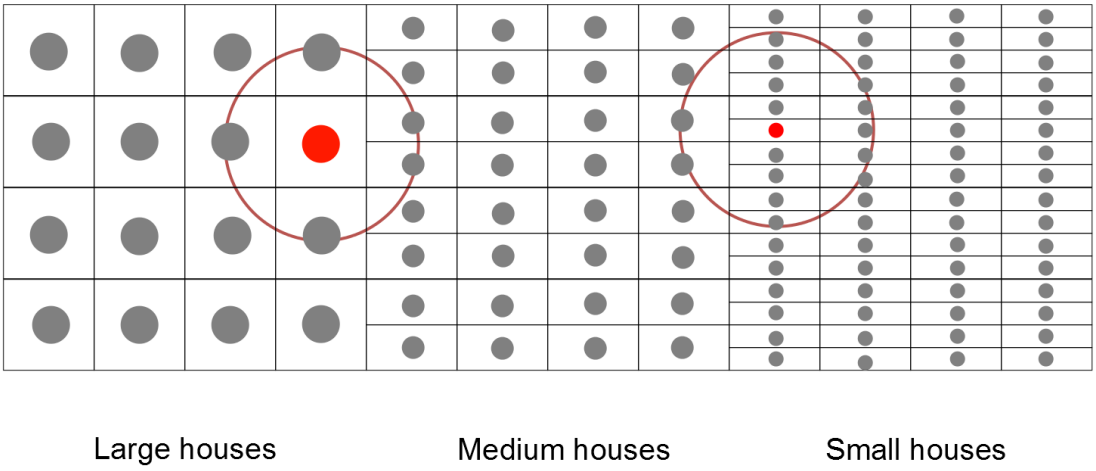


Figure 4.1: Basic block of the city structure.

A city with four basic blocks is shown in Figure 4.1 with the number of houses $N = 28 \times 4 = 112$. We assume each on-fire house can only ignite its neighbors within an igniting circle. The size of circle is fixed to be the same for all houses. Various sizes of circle show different external triggering ability inside this model. It is obvious that fire will stay active longer in a denser community.

Figure 4.2 shows the degree of houses in the city. The degree of houses decreases generally from small houses to large houses. The small houses (in red color) with the highest degree among three types indicates a high likelihood to ignite fires.

Intuitively, the likelihood of a house catching fire is determined by two factors: the materials the house is made of and the density of its neighbors. Let $N_{1,i}$ be the

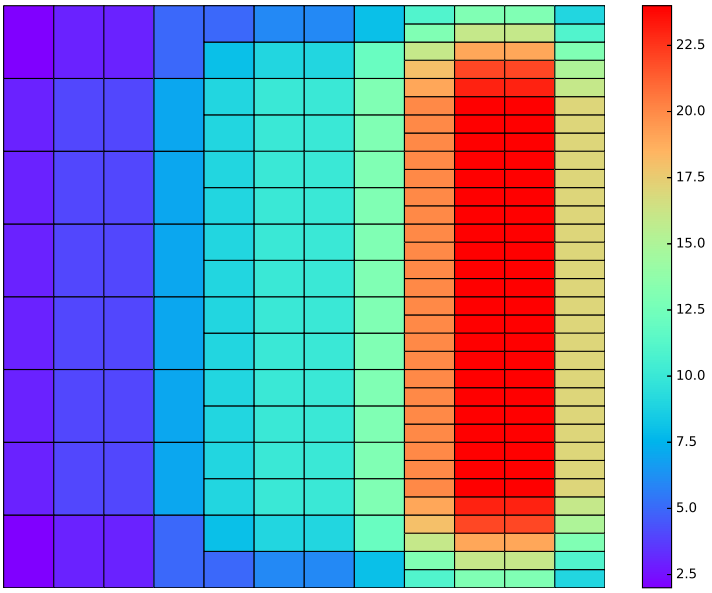


Figure 4.2: House degrees in a city of 224 houses.

likelihood of house i to catch fire (internally or externally). This value is assumed to be uniform for houses of the same category and should be positively correlated with the size of the house. On the other hand, larger houses have access to more resources, making the recovery process faster. Hence, we use another parameter $N_{2,i}$ for the recovery likelihood of house i, and the value of $N_{2,i}$ is negatively correlated with the size of the house.

4.2 Methods

4.2.1 Discrete model

Based on the structure of the fire-propagation model, we can simulate the fire cascades throughout the entire network using CARP. At time t, each house is either in state -1 (recovery), state 0 (fully operational) or state 1 (on-fire). Houses in the recovery state are under reconstruction and are immune to fire. Hence only fully operational houses are susceptible to fire. The burning (on-fire) state with certain probability switches to a state of recovery. Each house alternates between these three states.

The state transition is invoked by four types of Poisson processes. A house

i transitions from state 0 to 1 (on-fire) for internal reasons according to a Poisson process with intensity λ_i^{int}. In this event, a fire starts inside of the house, caused by events such as overheating of electrical appliances, unattended stoves or other possible accidents. This house can make this transition if its neighbor j ignites the fire externally through a Poisson process with intensity $\lambda_{ji}^{\text{ext}}$. A transition from state 1 (on-fire) to -1 (recovery) also follows a Poisson process, with intensity λ_i^{etg} (extinguishing the fire). Finally, transition from state -1 to 0 is caused by a Poisson process with intensity λ_i^{rec} (completing the recovery process).

For all the events discussed above the exact time of occurrence is not known, hence we use a discrete time step to accommodate this uncertainty and round up the event time to the nearest integer step. Hence, the evolution of the system can be viewed as a discrete-time series of stochastic processes with three states. For convenience, we assume here that each time unit represents one day of the real world. Hence, there are multiple state-transition events in a single day. As shown in [46], at one time unit (day), each Poisson process (state transitions) is actually the same as the corresponding Bernoulli trials.

Here, like in [46], we assume that experts provide assessments of each house's fire resistance. The value of $N_{1,i}$ represents the likelihood of house i to catch fire internally or externally, and $N_{2,i}$ represents the likelihood that the house fire is extinguished and house rebuilt over large period of time. Here, we set a control parameter for each state transition process. The internal risk materialization is controlled by parameter α. The external risk materialization is controlled by parameter β. However, only the state transition from fully operational to on-fire is observed, without knowing the actual reason for it. So, impact of an individual parameter α or β is hidden from direct observation. In contrast, parameters γ, controlling recovery and δ controlling fire extinguishing are independent of each other, so impact of each of the two can be observed in the changes in the evolution of corresponding underlying process.

We list events and parameters in the order compatible to the way they were ordered in the World Economic Forum model [46] in which only three events (int, ext, rec) and parameters (α, β, γ) exist. We set $N_{1,i} = 0.4$ for large, 0.3 for medium

and 0.2 for small houses and $N_{2,i} = 0.2$ for large, 0.3 for medium houses and 0.4 for small houses. We assign target values to our parameters $\alpha = 0.08$, $\beta = 0.012$, $\gamma = 0.016$ and $\delta = 0.032$.

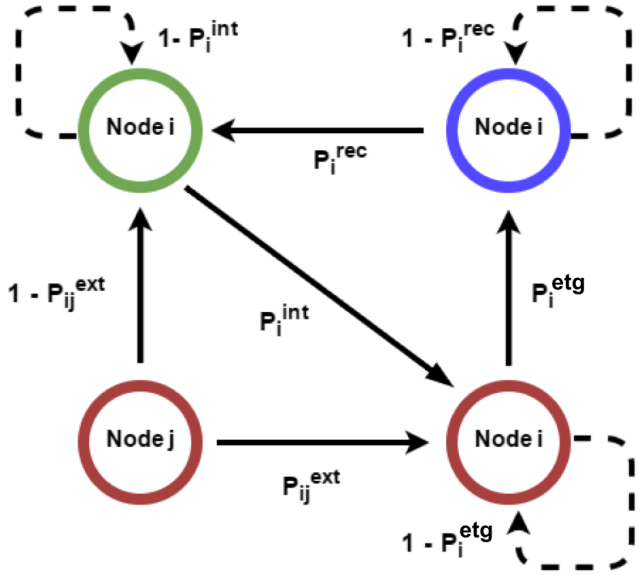


Figure 4.3: Fire propagation dynamics.

The dynamics process is shown in Figure 4.3. The nodes represent a fully operational house (green), a burning house (red) and a recovering house (blue). Houses in normal state can be ignited to on-fire state (red color) due to external and internal triggering. The probability of internal triggering of house i is denoted as p_i^{int}. The probability of external triggering by on-fire house j is denoted as p_{ji}^{ext}. The state transition from operational to on-fire state is governed by these two processes. For an on-fire house, the probability of extinguishing fire is p_i^{etg}. Finally, a house in recovery state (blue color) becomes operational again with a probability of p_i^{rec}.

To summarize, the dynamics progress in discrete steps $t = 1, \ldots, T$ and the probability of transition in each step is defined by the intensity of the corresponding Poisson process as shown in Table 4.1. The dynamics described above and shown in Figure 4.3 imply that the state of the system at time t depends only on its state at time $t-1$, and therefore the evolution of the state vector $\vec{S}(t)$ is Markovian. The definitions of parameters and equations mapping them into intensities of Poisson

processes are listed in Table 4.1.

Table 4.1: Intensities of Poisson processes and state transition probabilities.

House type	λ_{int}	λ_{ext}	λ_{rec}	λ_{etg}	p_{int}	p_{ext}	p_{rec}	p_{etg}
Large	0.0073	0.0109	0.0257	0.0515	0.00730	0.01094	0.02542	0.05020
Medium	0.0096	0.0144	0.0192	0.0385	0.00959	0.01434	0.01908	0.03779
Small	0.0129	0.0193	0.0146	0.0293	0.01279	0.01913	0.01455	0.02889

Thus, the dynamics progresses at each time step $t > 0$:

1. House i that was fully operational at time $t-1$ gets on-fire internally with probability $p_i^{\text{int}} = 1 - e^{-\lambda_i^{\text{int}}}$.

2. House j that was on-fire at time t ignites a fire of a neighboring house i that was fully operational at time t with probability $p_{ji}^{\text{ext}} = 1 - e^{-\lambda_{ji}^{\text{ext}}}$.

3. House i that was on-fire at time $t-1$ switches to the state of recovery at time t with probability of $p_i^{\text{etg}} = 1 - e^{\lambda_i^{\text{etg}}}$.

4. House i that was in recovery state at time $t-1$ becomes susceptible at time t with probability $p_i^{\text{rec}} = 1 - e^{\lambda_i^{\text{rec}}}$.

$$
\begin{aligned}
\lambda_i^{\text{int}} &= -\alpha \ln(N_{1i}) \\
\lambda_{ji}^{\text{ext}} &= -\beta \ln(N_{1i}) \\
\lambda_i^{\text{rec}} &= -\gamma \ln(N_{2i}) \\
\lambda_i^{\text{etg}} &= -\delta \ln(N_{2i}) \\
p_i^{\text{int}} &= 1 - {N_{1i}}^{\alpha} \\
p_{ji}^{\text{ext}} &= 1 - {N_{1i}}^{\beta} \\
p_i^{\text{rec}} &= 1 - {N_{2i}}^{\gamma} \\
p_i^{\text{etg}} &= 1 - {N_{2i}}^{\delta}
\end{aligned}
\tag{4.4}
$$

Given the value of N_{1i}, N_{2i} and control parameters, we can compute the probability of state transition explicitly. For a real city, the values of N_{1i} and N_{2i} would be assessed and provided by experts and the control parameters would be learned

from the historical records of city fires. In our artificial city case, the evolution of the system can be simulated as a discrete time stochastic process with three states for the desired period of times to provide historical data for parameter recovery. Moreover, repeating simulations with different random number generator seeds, we can obtain alternative historical data, which is of course impossible in the real life.

At each time step t, we can calculate the proportion of houses on-fire for the simulated set of parameters. The probabilities of events are derived from expert assessment (assumed by us the case of the artificial city) as the input of the model and used to compute probabilities using Equation 4.4. Like in the case of global risk network, experts should provide estimation of the likelihood of catching fire and of rebuilding, as well as of the number of neighbors that can ignite a fire for each house. Here, we established them to have them perfectly corresponding to the definition of Poisson processes used in simulation. The complete algorithm in pseudo-code is given in display Algorithm 1:

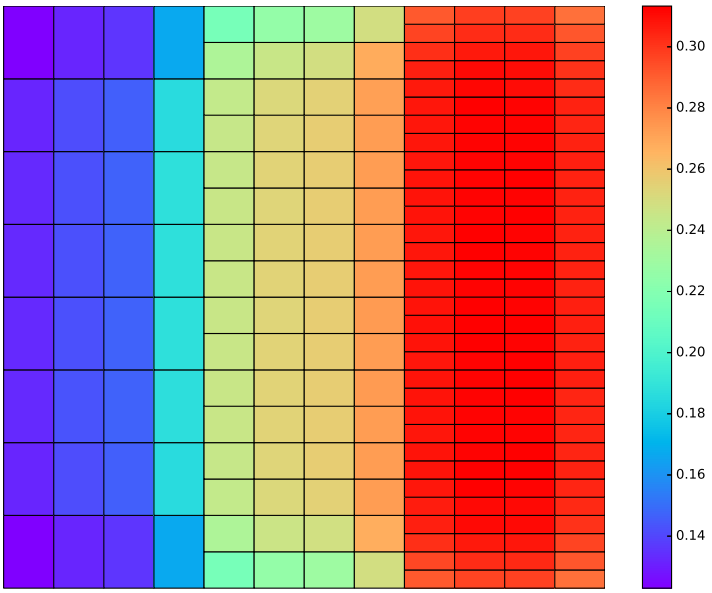


Figure 4.4: Fraction of on-fire time of houses.

With the assumed expert assessments and the created ground truth parameter values for the model, we simulate the evolution of house states for a particular period and record the frequency of emerging fires. In Figure 4.4, we show the fraction of

Algorithm 1 Fire propagation algorithm

Inputs:
  $\vec{S} = s_i$
Initialize:
  $s_i^0 \leftarrow 0,\ i = 1, \ldots, N$
  $\vec{S}_0 \leftarrow \vec{S}$
for $t = 1$ *to* T **do**
  $\vec{S}_t \leftarrow \vec{S}_{t-1}$
  for $i = 1$ *to* N **do**
    if $s_i^{t-1} = 0$ **then**
      if random number $r < p_i^{\text{int}}$ **then**
        $s_i^t = 1$;
      else
        $s_i^t = 0$;
      end if
    else if $s_i^{t-1} = 1$ **then**
      if random number $r < p_i^{\text{etg}}$ **then**
        $s_i^t = -1$;
      else
        $s_i^t = 1$;
      end if
    else
      if random number $r < p_i^{\text{rec}}$ **then**
        $s_i^t = 0$;
      else
        $s_i^t = -1$;
      end if
    end if
  end for
  for $i = 1$ *to* N **do**
    if $s_i^{t-1} == 1$ **then**
      $N_i \leftarrow$ neighbors of i
      for $j \in N_i$ **do**
        if $s_j^t == 0$ and $s_j^{t-1} == 0$ and random number $r < p_j^{\text{ext}}$ **then**
          $s_j^t = 1$;
        else
          $s_j^t = 0$;
        end if
      end for
    end if
  end for
end for

time each house is on-fire during one million time steps. Results are averaged over 100 independent realizations. Each simulation runs 10^6 time units. The houses are fully operational initially. The simulations used parameter values listed in Table 4.1. The range of this on-fire fraction varies from 0.12 to 0.32. The fraction increases when the size or degree of the house increases, but areas of higher density housing suffer higher on-fire fractions than indicated by their degree.

4.2.2 Continuous model

In addition to the time series with discrete states, we can also describe the dynamics in a continuous formulation. In each time step, we use a probability of each state denoted as $s_i(t), f_i(t)$ and $r_i(t)$ to represent the expected value of occurrence instead of the discrete state. $s_i(t)$ defines the probability of susceptible at time t for house i. Similarly, $f_i(t)$ is the probability for the state of on-fire and $r_i(t)$ is the probability for the state of recovery. The state transitions can be expressed by Ordinary Differential Equation (ODE):

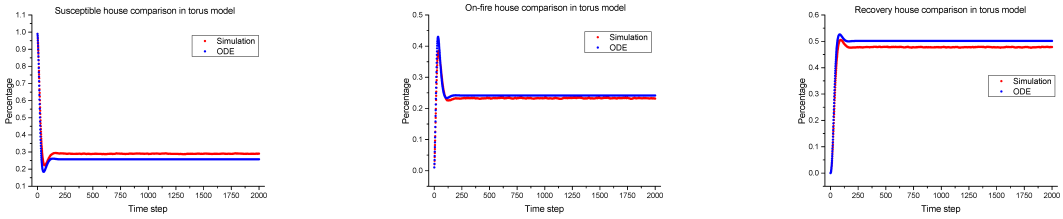
$$\frac{ds_i(t)}{dt} = -\lambda_i^{\text{int}} s_i(t) - \lambda_i^{\text{ext}} s_i(t) \sum_{j=1, j\neq i}^{N} a_{i,j} f_j(t) + \lambda_i^{\text{rec}} r_i(t) \tag{4.5}$$

$$\frac{df_i(t)}{dt} = \lambda_i^{\text{int}} s_i(t) + \lambda_i^{\text{ext}} s_i(t) \sum_{j=1, j\neq i}^{N} a_{i,j} f_j(t) - \lambda_i^{\text{etg}} f_i(t) \tag{4.6}$$

$$\frac{dr_i(t)}{dt} = \lambda_i^{\text{etg}} f_i(t) - \lambda_i^{\text{rec}} r_i(t) \tag{4.7}$$

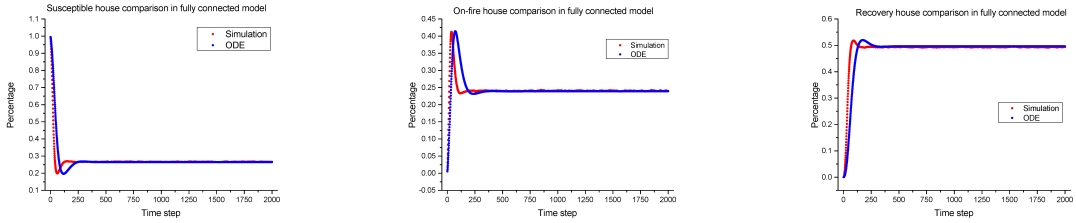
Summing up all three equations, we get $\frac{ds_i(t)}{dt} + \frac{df_i(t)}{dt} + \frac{dr_i(t)}{dt} = 0$ which means the total size of the system does not change along with time. The non-linear ODEs have stable points in the range of $[0, 1]$.

To verify the simulated time series, we compare the simulation results with the numerical solution of ODEs in two simple models. One is a torus model with 8 neighbors for each house; the other is a fully connected model. Both models have the same number of identical houses which is 224. The parameters for each state transition are also the same. The difference between these models is the number of neighbors of each house. We calculate the proportion of each state (susceptible,



(a) Susceptible **(b) On-fire** **(c) Recovery**

Figure 4.5: Comparison in the torus model.



(a) Susceptible **(b) On-fire** **(c) Recovery**

Figure 4.6: Comparison in the fully connected model.

on-fire and recovery) for a time steps of 2000. From the Figure 4.5 and Figure 4.6, the gap between simulation and numerical results is larger in the torus model than the fully connected model, which means the higher model connectivity, the better match between discrete and ODE results. The red curve represents the simulation result and blue curve is for ODEs results. The x axis is the time step and y axis is the proportion of each state. It is intuitive that a higher degree graph has smaller topology structures. In the fully connected model, there is only one possible topology structure. For the torus model with 8 neighbors, it is only one specific structure for the network with a uniform degree of 8.

4.2.3 Precision limit of maximum likelihood estimation

In our approach, we use maximum likelihood estimation (MLE) to recover parameters from ground truth data. The main reason for this choice is that state transitions are governed by independent inhomogeneous probability distributions for which MLE delivers consistency and asymptotic normality with sufficient amounts of observed data [22]–[25]. The historical data represents the combined effects of four Bernoulli processes. Our purpose is to recover the unknown parameters α, β, γ and

δ mapping the expert assessments into event probabilities for Bernoulli processes of our model. We denote $\hat{\alpha}, \hat{\beta}, \hat{\gamma}$ and $\hat{\delta}$ to be the recovered (estimated) values of each parameter.

The use of MLE to find the values of hidden parameters from observed events has not been studied, yet [22]–[24] indicate that it is feasible. In our approach, we split one of the parameters of directly observable events into a pair of hidden (and tangled) parameters of two processes and recover these two parameters from observed events. We denote the unknown parameters as θ. Given the n observations $x_1, x_2, ...x_n$, the likelihood function of this set of observations is defined as [24]:

$$\mathcal{L}(\theta) = f\left(x_1, x_2, ..., x_n|\theta\right) \tag{4.8}$$

When the distribution is discrete, f is a frequency function, and the likelihood function $L(\theta)$ shows the probability of observing the given data. The maximum likelihood method [24], [25] finds the values of parameters that yield the maximal probability of observing the given data. Logarithms are monotonic and therefore the likelihood and its logarithm have the maximum at the same argument. Since the observed data comes from independent distributions, the logarithm of likelihood function can be written as [24]:

$$\ln \mathcal{L}(\theta) = \sum_{i=1}^{n} \ln\left(f\left(x_i|\theta\right)\right) \tag{4.9}$$

For continuous and smooth likelihood functions, which is the case here, we can scan the parameter space in order to find the maximum point for $\ln \mathcal{L}(\theta)$. The historical data size is limited, so we need to study how this limitation affects the precision of results. In this section, we derive an estimation of the optimal parameters for large sample sizes. Under appropriate smoothness conditions, the estimate is consistent with large data sets and obeys the asymptotic normality. Detailed description of these conditions can be found in [22]. These conditions can be summarized as: the first three derivatives of the $\frac{\partial \log(f(x_i|\theta))}{\partial \theta}$ are continuous and finite for all values of x_i and θ; the expectation of the first two derivatives of $\frac{\partial \log(f(x_i|\theta))}{\partial \theta}$ can be obtained and

the following integral is finite and positive [22]:

$$\int_{-\infty}^{\infty} \left(\frac{\partial \log f}{\partial \theta} \right)^2 f dx \tag{4.10}$$

Let θ_0 denote the true value of the parameter θ and $\hat{\theta}$ to be its recovered value. In the asymptotic case for MLE with large amount of data size n, once the smoothness conditions are met, the recovered value $\hat{\theta}$ converges to the true value θ_0. If we normalize $\hat{\theta}$, we obtain an approximation from a normal distribution if the variance of MLE σ_{MLE}^2 exists [24]:

$$\lim_{n \to \infty} \left(\hat{\theta} - \theta_0 \right) = N \left(0, \sigma_{MLE}^2 \right) \tag{4.11}$$

Given the definition of Fisher information $I(\theta)$ shown in [24]:

$$I\left(\theta\right) = E \left[\frac{\partial}{\partial \theta} \log f \left(x | \theta \right) \right]^2 \tag{4.12}$$

The asymptotic normality of MLE can be written as [24]:

$$\lim_{n \to \infty} \left(\hat{\theta} - \theta_0 \right) = N \left(0, \frac{1}{nI\left(\theta_0\right)} \right) \tag{4.13}$$

The variance of the normalized estimate decreases as Fisher Information $I(\theta_0)$ and the amount of training data increases. Intuitively, higher information leads to a smaller variation and a lower uncertainty level. More observed data provides us more evidence to estimate true parameters. This asymptotic variance to some extent measures the quality of MLE. Although it is hard to compute the variance analytically in our model, we know there exists an estimation limit and the performance of MLE becomes better as the volume of given data increases. In the next section, we demonstrate that the variance indeed decreases as the size of the training dataset increases and the mean values of $\hat{\theta}$ approaches θ_0 with the error approaching 0.

4.2.4 Parameter estimation in fire-propagation model

Given the historical data about each house transition in a finite period, we use maximum likelihood estimation to find optimal values of model parameters which produce the maximum likelihood of the historical observations. State transitions are independent functions with unknown parameter values. Since the historical data is generated from the discrete stochastic process, the likelihood function of particular observations of risk materializations can be written as:

$$\mathcal{L}\left(\vec{S}(1), \vec{S}(2) \cdots, \vec{S}(T)\right) \equiv \prod_{t=2}^{T} \prod_{i=1}^{N} P_i(t)^{S_i(t-1) \to S_i(t)} \tag{4.14}$$

where T is the number of time steps, N is the number of houses in the model, $S_i(t)$ is the current state of house i at time t, and $P_i(t)^{S_i(t-1) \to S_i(t)}$ is the state transition probability of house i from time $t-1$ to time t. $\vec{S}(t)$ is the vector of states for each house in the network at time t. The logarithm of this likelihood is:

$$\ln \mathcal{L}\left(\vec{S}(1), \vec{S}(2) \cdots, \vec{S}(T)\right) \equiv \sum_{t=2}^{T} \sum_{i=1}^{N} \ln\left(P_i(t)^{S_i(t-1) \to S_i(t)}\right) \tag{4.15}$$

In the training process, we compute the probability of state transitions using p_i^{int} and p_{ji}^{ext} for transition into fire, p_i^{etg} for transition into recovery and p_i^{rec} for transition into the fully operational state. Correspondingly, there are three cases of state transitions in the historical data.

A transition from the fully operational state to the on-fire state ($0 \to 1$) happens when a house catches fire due to internal or external reasons. The probability of internal ignition is p_i^{int}. For external influence, we compute first the probability that none of the neighbors ignited this house, and then take the complement of this value. On-fire neighbor j fails to ignite house i with probability $1 - p_{ji}^{ext}$. The product of those over all on-fire neighbors defines the probability of house i not ignited by external fire:

$$prod_i^{0 \to 0} = \prod_{j \in A_i} \left(1 - p_{ji}^{ext}\right) \tag{4.16}$$

where A_i is the set of all on-fire neighbors of house i. The complement of this

product defines the probability that at least one neighbor ignites house i. Adding the probability of such external ignition to the probability of internal ignition, we obtain the total probability of a house catching fire:

$$P_i^{0\to 1} = p_i^{int} + \left(1 - p_i^{int}\right)\left(1 - \prod_{j\in A_i}\left(1 - p_{ji}^{ext}\right)\right) \tag{4.17}$$

Since internal and external ignition are mutually exclusive we include a factor of $(1 - p_i^{int})$ in the external ignition probability. Accordingly, the probability of not catching on fire is:

$$P_i^{0\to 0} = \left(1 - p_i^{int}\right)\prod_{j\in A_i}\left(1 - p_{ji}^{ext}\right) \tag{4.18}$$

A transition from being on fire to recovery state happens when the fire is extinguished and rebuilding process starts. This probability is defined as:

$$P_i^{1\to -1} = p_i^{etg} \tag{4.19}$$

Transition from recovering to fully operational state ($-1 \to 0$) happens when a house is completely rebuilt and becomes fully operational. The corresponding probability is:

$$P_i^{-1\to 0} = p_i^{rec} \tag{4.20}$$

The maximum likelihood parameters are obtained by summing the logarithms of corresponding probabilities. After scanning the potential ranges of the model parameters, we find the globally optimal values that maximize the likelihood of the historical data. The closeness of the recovered parameters to their values set in the simulations measures how precisely our model captures the dynamics of the system.

Unlike real-life, in simulations, we can arbitrarily vary the length of time over which we collect historical data and produce many variants of such data to measure the prediction precision of our model. We start with a mixed model with 8 large houses, 16 medium houses and 32 small houses, for a total of 56 houses. The parameter recovery is applied at seven different intervals: 100, 200, 400, 800, 1600, 3200 and 6400 time steps. The parameter recover is run on 50 versions of historical

data created by simulations running with different seeds for the random number generator to account for the randomness of the stochastic processes. To quantify the accuracy of parameter recovery, we compute the relative error between the recovered values and the target values used for creating ground truth data.

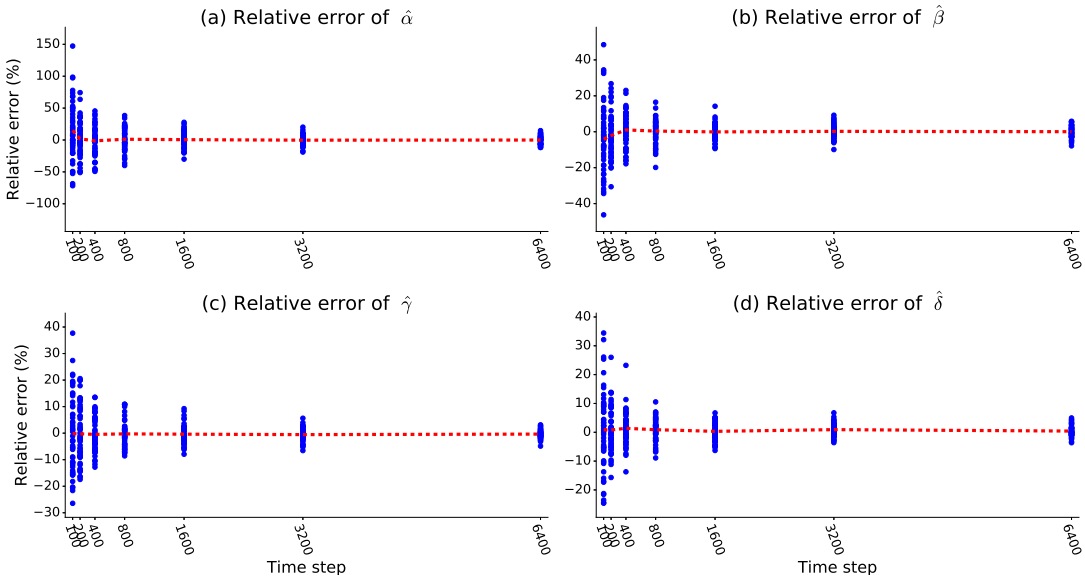


Figure 4.7: Parameter recovery in the fire-propagation model.

4.3 Results

As shown in Figure 4.7, the x-axis includes seven time intervals: 100, 200, 400, 800, 1600, 3200 and 6400 time steps. The y axis shows the relative error for the recovered parameters. Using parameter values shown in Table 4.1, parameter recovery was run on 50 different historical datasets generated by simulations with different seeds for the random number generator; the results of these runs are represented by blue dots. The red dashed curves show the average values of the relative error. The visible trend is that the average of relative errors tends asymptotically to zero and the variance exhibits power law decrease as the number of time steps increases, which means more training data improves the performance of parameter recovery. The blue vertical dots represent the relative error of parameter recovery in one realization. There are 50 blue dots for each particular length of training

dataset. The scattering of blues dots represents the variation of parameter recovery accuracy. The red curve is the average value among these 50 realizations, which is very close to zero. Obviously, variance is a better measurement for the precision since variance considers the positive and negative errors instead of canceling them out. The average of relative errors tends asymptotically to zero and the variance shrinks according to the power law with an increase in sample size which is the length of historical data series in this case. This trend is consistent with the asymptotic behavior of the MLE method [22], [24]. When the sample size is very large, the relative error of realization follows a normal distribution with 0 mean value and a finite yet small variance. More data decreases relative error, and therefore it is useful to find a balance between run time and prediction accuracy.

The relative errors of $\hat{\alpha}$ and $\hat{\beta}$ are larger than that of other parameters. This is due to the combined effects of two Poisson processes causing the same transition. As defined, α and β represent the intensity of internal and external fire ignition processes respectively. In real life, it is often hard to determine the actual reason. During the parameter recovery, these two parameters influence chances of each other to start a fire, which impacts the computation of likelihood function. This nonlinear effect tangles the errors of $\hat{\alpha}$ and $\hat{\beta}$ together as larger value of $\hat{\alpha}$ can be compensated by smaller value of $\hat{\beta}$ and vice versa.

We also compute the standard deviation of relative error of recovered parameters. Figure 4.8a shows that the distribution of the standard deviation follows a power law. We use seven simulation time intervals: 100, 200, 400, 800, 1600, 3200 and 6400 time steps. There are 56 houses. Using parameter values shown in Table 4.1, parameter recovery was run on 50 different historical datasets generated by simulations with different seeds for random number generator. The standard deviation decreases very quickly and then slowly as historical data size increases. The double-logarithmic plot in Figure 4.8b has a slope close to -0.5, which shows that the standard deviation decreases in a power-law fashion as the training data size increases.

For real-world case processes, it is impossible to get multiple historical datasets for parameter recovery, yet recovery error based on single dataset may be different

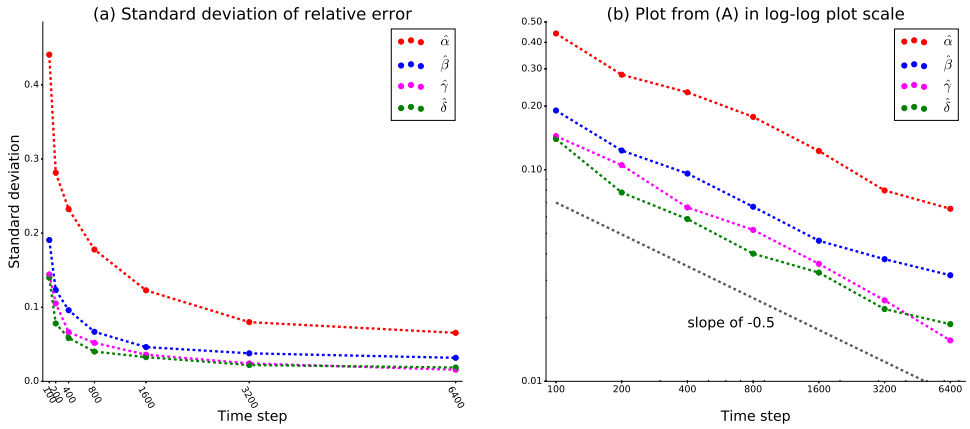


Figure 4.8: Standard deviation of relative error of parameter recovery.

from the one based on the average of such recovered values on multiple datasets. Hence, we study how sensitive our methodology is to imperfect input datasets. In order to test this sensitivity, we compare four cases of simulations in which we record the average length of time in each state and the number of emerging fires during five simulation periods: $400, 800, 1600, 3200$ and 6400. The first case is using the target values of parameters. The second case is using the averaged value of parameters recovered from 50 independent realizations. The third case employs adding σ to the average value of recovered parameters. The last case is subtracting σ from the average value of recovered parameters. The four sets of parameters employed in our simulation are listed in Table 4.2:

Table 4.2: Parameter values for simulation.

Parameters	α	β	γ	δ
Target value	0.00800	0.01200	0.01600	0.03200
Recovered value	0.00799	0.01199	0.01596	0.03204
Recovered value $+\sigma$	0.00853	0.01162	0.01622	0.03274
Recovered value $-\sigma$	0.00747	0.01237	0.01570	0.03134

The average values of recovered parameters come from 50 independent realizations with 6400 time steps of training data. The standard deviations are: $\sigma_\alpha = 0.00053, \sigma_\beta = 0.000372, \sigma_\gamma = 0.00026, \sigma_\delta = 0.0007$. When adding one standard deviation σ to average value of $\hat{\alpha}$, we subtract σ from $\hat{\beta}$ and vice versa. We assume complementary recovery errors on α and β since both of them contribute to the

fire triggering process so the maximum value of one is likely to be reached at the minimum of the other. This corresponds to the assumption that the unique historical dataset produces parameter values within one σ of their average values, more stringent assumption might require using broader interval around average values of parameters. In simulation, we record how long each house stays fully operational, on-fire and in recovery and record the number of new fires during the simulation. The number of houses is 56. Using parameter values shown in Table 4.2, parameter recovery was run on 50 different historical datasets generated by simulations with different seeds for random number generator. We gather results upon reaching five simulation periods: $400, 800, 1600, 3200$ and 6400. We complete 50 independent realizations to summarize the statistical conclusion of the predictability of our methodology. There is a trade-off between the running time and precision. To save computation effort, we do not finish too many realizations and choose an empirical value of 50. Figure 4.9 shows that all four parameter estimations have similar precision. In Figure 4.9a, the average duration of the fully operational state is almost the same for four cases since two parameters α and β have opposite influences on fully operational houses. In Figure 4.9b and Figure 4.9c, the gap between the cases of estimated parameters $\pm\sigma$ is very small. In Figure 4.9d, the number of emerging fires for all cases are increasing linearly as a function of time. The largest relative error is small in predicting the length of fire, which is just below 2%. Because of the specific target values of parameters, the duration of each state is in the same order, and the frequency of emerging fires is larger than that of a real-world scenario. We intentionally set large and comparable target values in order to generate sufficient state transitions and test the quality of parameter recovery in an arbitrary case. Table 4.2 shows that the estimated parameters approach the target values very well, even if the target values are in the same order.

Additionally, we design another method to find the $\pm\sigma$ boundary of the performance of recovered parameters where σ is the standard deviation. Initially, we complete 125 sets of parameter recovery for a period of 6400 time steps. Then, 5 different periods of time steps: 400, 800, 1600, 3200 and 6400 are simulated for each set of recovered parameters. In each case, 20 independent realizations are finished

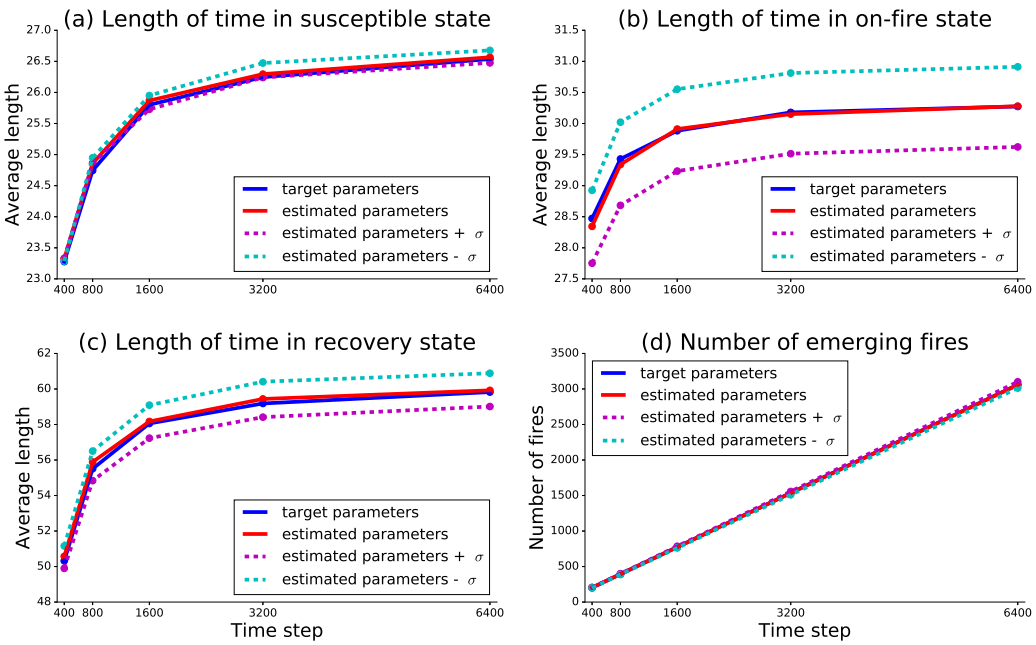


Figure 4.9: Impact of imperfect recovered parameters.

and the average value of four features is recorded for the performance: (i) the length of time in normal state, (ii) the length of time in on-fire state, (iii) the length of time in recovery state and (iv) the number of emerging fires. These values can be used as metrics to detect whether the performance of the estimated parameters is consistent within $\pm\sigma$ standard deviation. Then, we compute the relative error of these features between each set of estimated parameters and the ground truth parameters. For each feature, based on the absolute value of relative error and following Kolmogorov-Smirnov test [83], [84], we remove 39 results with largest error from the original 125 results. In this way, the 31.2% worst performance has been removed and the remaining 68.8% result showing a range of $\pm\sigma$ performance. For each set of recovered parameters, there are four features and each feature has two boundary values for the case of $+\sigma$ and $-\sigma$.

Figure 4.10 shows the performance of the $\pm\sigma$ case for five different periods. The gap between the cases of estimated parameters $\pm\sigma$ has a comparable relative error in all three cases in Figure 4.10, despite the fact that start of fire depends on two parameters. However the two parameters α and β have opposite influences on fully operational houses. In Figure 4.10d, the number of emerging fires for

all cases are increasing linearly as a function of time. The curve of ground truth parameter is located in the center between the curves of $+\sigma$ and $-\sigma$ cases and the distance between the results from $\pm\sigma$ boundary and from the ground truth parameters stabilizes as the simulation time steps increase. The absolute relative error between results from $\pm\sigma$ cases and from the ground truth parameters is low, about 2%. Additionally, the curve of the averaged results from estimated parameters (red dash line) almost overlaps with the corresponding results from the ground truth case which indicates a small bias overall. Using this method, we find an interval of 68% confidence level showing the variance of performance of estimated parameters.

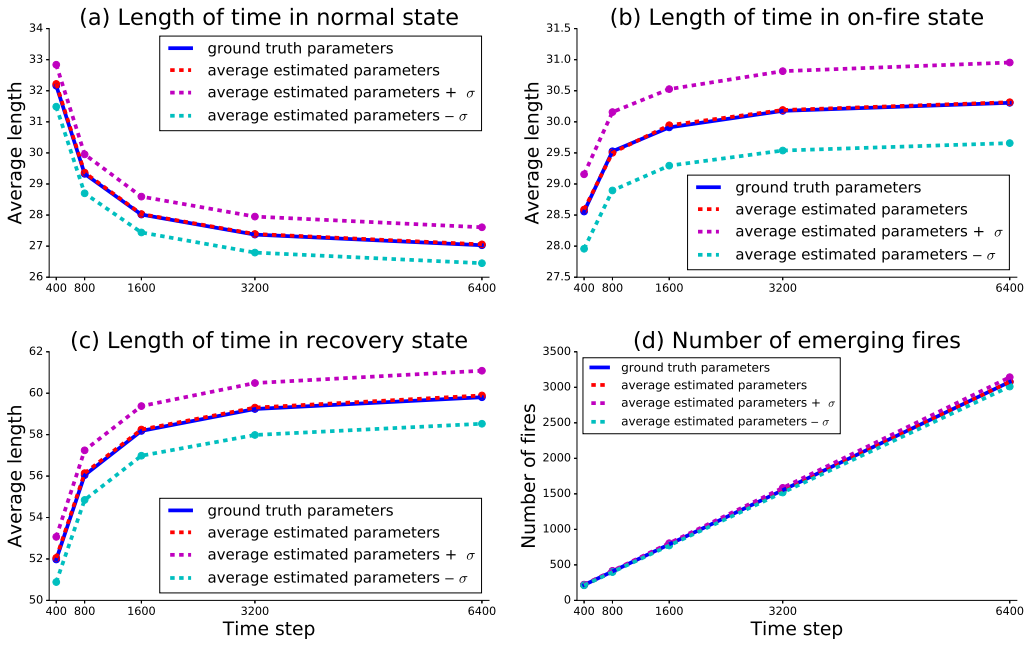


Figure 4.10: Performance of model prediction.

In addition to the parameter recovery from different lengths of historical datasets, we study how the precision of parameter recovery varies against the complexity of the system and the standard deviation of number of fires starting each day. Figure 4.11a and Figure 4.11b show the relative errors of recovered parameters for various city sizes (number of houses): 28, 56, 112, 224 and 448. Here we compare two cases. In the first case, we keep three types of houses and use the parameter values shown in Table 4.2. The blue dots in Figure 4.11a and Figure 4.11b represent the results of the first case. In the second case, we initialize all houses with the same

value of $N_{1,i} = 0.3$ and $N_{2,i} = 0.3$, which are equal to the values for middle houses. The red dots in Figure 4.11a and Figure 4.11b represent the results from this case. Therefore, we remove the influence of house types on the results. In both cases, the length of historical data is 1600 time steps. The dashed curve shows the mean values of relative error over 20 realizations and each dot represents one realization. Only two parameters ($\hat{\alpha}$ and $\hat{\beta}$) are shown in this figure since they are involved with the fire igniting process and another two parameters have similar trend. We find that the parameter recovery in a larger city has a smaller mean value and variance of relative errors. As the city's size increases, there are more state transitions within specific periods, leading to a more precise recovery of our parameters. To study the impact of intensity of emerging fires on recovery precision, we change the values of $N_{1,i}$, which define the intensities of fires. We compare three cases: $\sqrt{N_{1,i}}$ in red, $N_{1,i}$ in blue and $N_{1,i}^2$ in cyan. Figure 4.11c and Figure 4.11d show the relative error of recovered $\hat{\alpha}$ and $\hat{\beta}$. Red dots represent the case of $\sqrt{N_{1,i}}$, blue represents the case of $N_{1,i}$, and cyan represents the case of $N_{1,i}^2$. The results come from 20 independent realizations and 6 different lengths of historical dataset: 100, 400, 800, 1600, 3200 and 6400. $N_{1,i}$ is 0.4 for large 0.3 for medium and 0.2 for small houses. The model with more emerging fires at each day enjoys a smaller variance of relative errors. The reason is similar to that shown above: more state transitions in the historical dataset and higher precision of the recovered parameters.

4.3.1 Sensitivity analysis

In real world case, the input training data including some error which may influence the accuracy of the output. In order to study how the bias of input effects the results, we conduct the sensitivity analysis in the same model. We intend to add an uniformly distributed error to the target value of each parameters.

$$\begin{aligned} \alpha_i^* &= \alpha \times (1 + \epsilon_i) \\ \beta_i^* &= \beta \times (1 + \epsilon_i) \\ \gamma_i^* &= \gamma \times (1 + \epsilon_i) \\ \delta_i^* &= \delta \times (1 + \epsilon_i) \end{aligned} \tag{4.21}$$

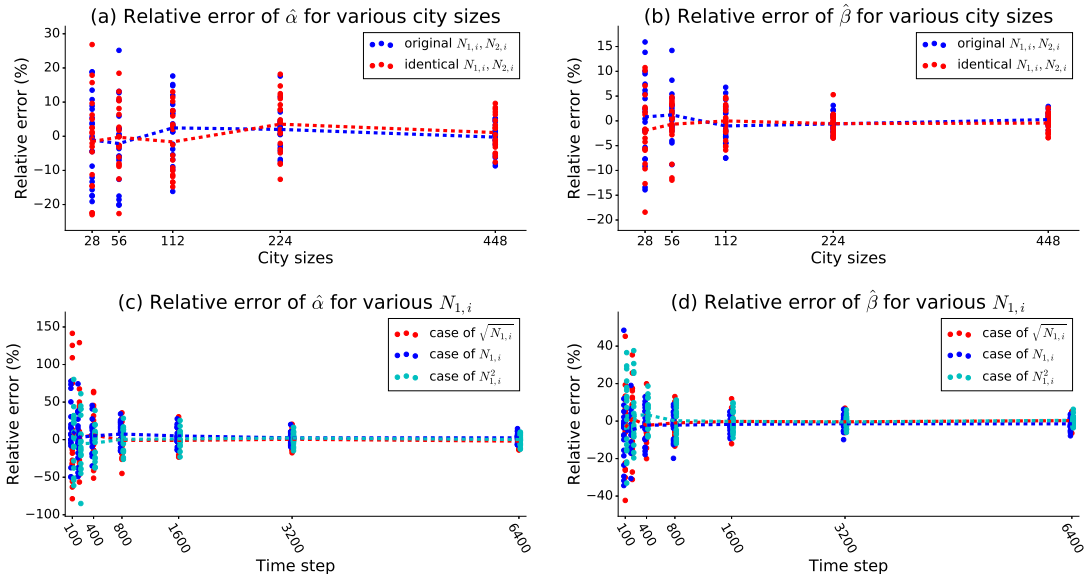
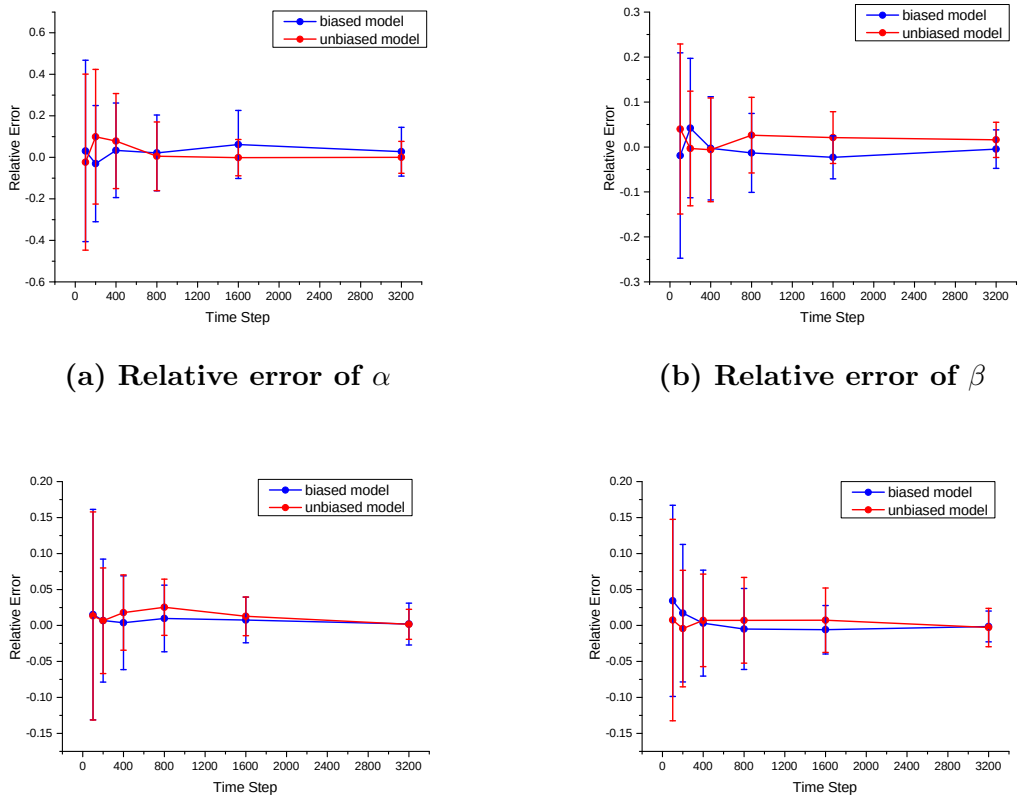


Figure 4.11: Parameter recovery in various scenarios.

where ϵ_i is the error of parameter uniformly distributed within the range: $[-0.1, 0.1]$. In this case, the parameter of each house has different values due to the random error, while in previous case, all houses have identical parameter values. We repeat the same experiment in this biased model. The variance of the biased model is larger than the unbiased model which is intuitive since the error of the input indeed effects the output of MLE. Both models have the same trend for the size of training data. This trend is consistent with the asymptotic normality of MLE discussed above. The normalized estimates follow a normal distribution and the variance becomes smaller as the sample size increases. However, actually the size of training data can not be infinity, so the variance of the estimates always exists. This is a significant limit for our predictive model. We have to consider the trade off between running performance and prediction accuracy.

In order to demonstrate more information, we compare the mean value and standard variance for the unbiased and biased model. In Figure 4.12, the x-axis is for five discrete time steps: $100, 200, 400, 800$, 1600 and 3200. The y-axis is the relative error for the trained parameters. The curve shows the mean values of relative error for both models along with the standard deviation. The standard deviation of both models is very large when data size is small. As the size increases, the mean value



(a) Relative error of α

(b) Relative error of β

(c) Relative error of γ

(d) Relative error of δ

Figure 4.12: Parameter recovery sensitivity.

of relative error becomes closer to zero and the variance also decreases. In general, the unbiased model performs slightly better than the biased model. However, when the time step is 3200, the difference between them is not obvious. It is intuitive that when the biased error of the input is larger, the error for the biased model becomes clearer. The result is sensitive for the input bias. Hence, it is necessary to make sure the accuracy of training data. No matter the training data is from simulation or from historical observation, we can not avoid error but only mitigate it. This is another predicting limit for our model. We have to make sure the accuracy of input data.

4.3.2 Parameter recovery precision in global risk network

Here we show how to apply the presented approach to a disparate, real-life dataset of the global risk network, which, as with fires in cities, exhibits spreading

risk activation [46]. Using CARP, we estimate hidden parameters of global risks previously modeled using an Alternating Renewal Process [46]. Experts from the World Economic Forum 2013 Global Risks Report [47] define the properties of 50 global risks grouped into five categories: economic, environmental, geopolitical, societal, and technological. These assessments include the likelihood, impact of materialization, and connections of each risk. We take the advantage of this crowd-sourcing assessment to build an interconnected network to simulate risk propagation through the system.

In the global risk network, each risk has binary states (normal and active), and the state transitions also follow Poisson processes. The difference is that in the global risk network, there are three state transitions instead of four in the fire-propagation model: 1, internally triggering process from normal to active state; 2, externally triggering process from normal to active state; 3, recovery process from active to normal state. Hence, we have three control parameters (α, β, γ) and recover the optimal parameter values from historical events. Based on the historical occurrences of each risk in a 156-month period and using maximum likelihood estimation, we recover the following parameter values: $\alpha = 0.003038$, $\beta = 0.00117$ and $\gamma = 3.5561$. They control internal risk materialization (α), external risk materialization (β) and recovery (γ) processes, and their detailed definitions can be found in Ref. [46]. Once we obtain the parameter values and set the initial situation, we can simulate the stochastic process of model evolution. As in case of the fire-propagation model presented here, we used these recovered parameters as the ground truth parameters for establishing the recovery process precision.

We apply the same method used in Figure 4.10 to this global risk network. In the first step of precision evaluation, we produce alternative historical ground truth datasets by using parameter values recovered from real historical data to simulate the model for a selected period of time repeatedly with different random generator seeds. Then, we recover parameter values using MLE on the simulated time series and compute their deviation from the values used in simulations. In this experiment, we use 120, 240, 480 and 960 monthly time steps of data (representing 10, 20, 40 and 80 years) for parameter recovery. We create 50 independent historical datasets.

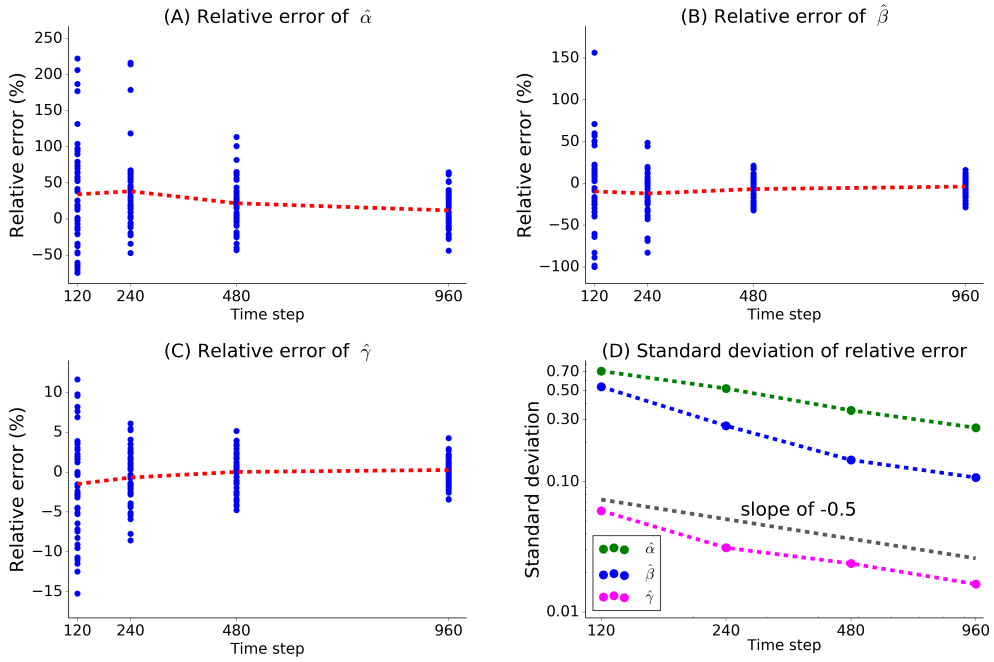


Figure 4.13: Parameter recovery error in the global risk network.

Figure 4.13 shows the similar trend of relative error of parameter recovery as for the fire propagation model shown in Figure 4.7. In particular, the variance of relative error for $\hat{\alpha}$ and $\hat{\beta}$ is obviously larger than γ indicating that it is hard to split the combined effects of two interwoven triggering processes with limited information. Here, the first process is the internal risk materialization controlled by parameter α. The other process is external risk materialization controlled by parameter β. We can only observe the state transition from normal to materialized state without knowing the actual reason for it. In parameter recovery, if α is over estimated, β will be under estimated and vice versa. In contrast, the state transition from materialized to normal state depends only on the control parameter γ. So the variance of γ is much smaller than the other two parameters. When we increase the amount of training data, the mean value and variance of relative errors decrease quickly (approaching zero).

In addition to previous the precision evaluation, we obtain multiple recovered-parameters from different lengths of alternative historical data to predict the number of risk materialization for a selected period of time repeatedly with different random

generator seeds. First, based on the ground truth parameters recovered from the real historical data (α, β, γ), we generate 120, 240, 480 and 960 monthly time steps of alternative historical data (representing 10, 20, 40 and 80 years) for parameter recovery. Then, 125 cases of parameter recovery $(\alpha_i, \beta_i, \gamma_i, i = 1...125)$ are finished for each period of time. Next, using the recovered parameters from both real and alternative historical data, we complete 20 realizations for a prediction of 4 periods: 120, 240, 480 and 960. In each period, we use the average value of risk materialization among all 20 realizations to measure the performance of each simulation case. Thereafter, we compute the relative error of the average risk materialization between the case with alternative recovered-parameters $(\alpha_i, \beta_i, \gamma_i, i = 1...125)$ and the ground truth parameters (α, β, γ). In the end, we remove the 39 simulation cases with the worst performance (with the largest absolute relative error). The remaining 86 cases (68.8% of 125 cases) of results determine the $\pm\sigma$ boundary for the performance of estimated parameters.

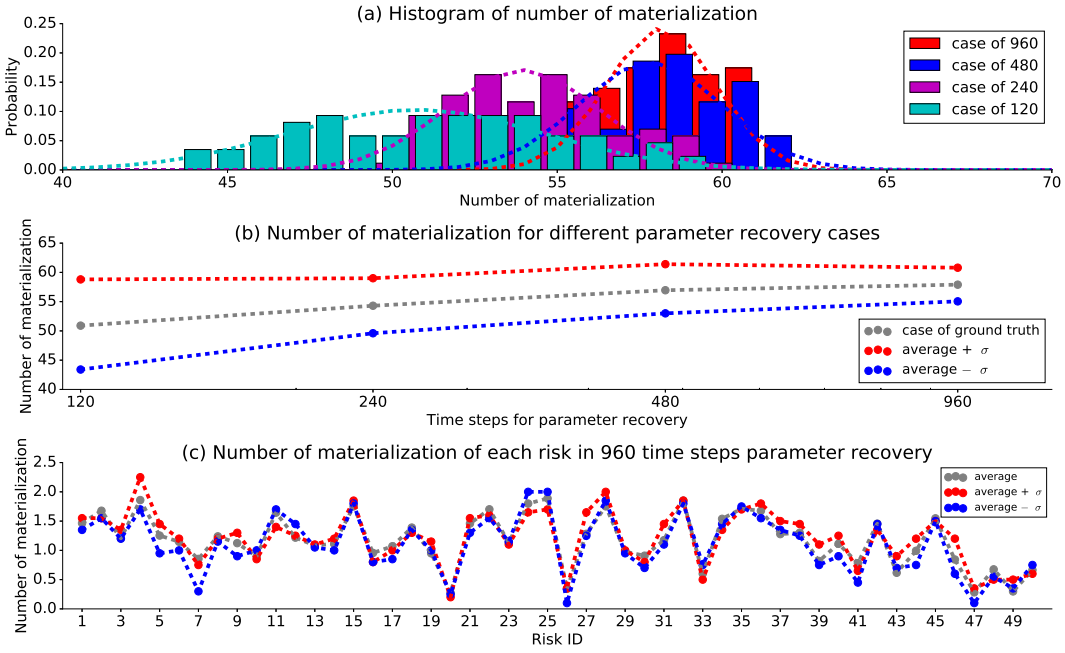


Figure 4.14: Performance of recovered parameters in the global risk network.

Figure 4.14a shows a histogram of number of materialization in each case. Dash curve represents a Gaussian fitting over the histogram. As the length of

simulation period increases, the distribution of number of materialization gradually approaches a normal distribution. Meanwhile, the distribution shifts to the right and gets close to a steady level. In the case with longer time steps for parameter recovery and prediction simulation, the predictability is more consistent and variance shrinks generally. Figure 4.14b shows the boundary of $\pm\sigma$ performance for number of materialization in each period. It is obvious that the distance between the boundary is decreasing as we increase the length of period, which implies a stronger confidence of prediction. Figure 4.14c shows the number of materialization for each risk in the case of 960 time steps. The difference between three cases is very small indicating a consistent prediction for estimated parameters.

4.4 Conclusion

The CARP model is used to simulate and then recover parameters of heterogeneous stochastic processes. First, we created a model of fires in the cities that we use to illustrate our approach. Using assumed parameters values, we generated several historical datasets and used them to measure parameter recovery precision. The results confirmed that the accuracy of our method increases as the amount of data increases even in the presence of parameters hidden from direct observations. The limits of the parameter recovery precision are caused by the stochastic nature of the modeled processes, so the variance of recovery always exists regardless of the size of historical data.

Applying our approach to real-life cases, we started with recovery of the model parameter values based on unique and limited real-life ground truth data. Then, using these values as ground truth, we finished simulations to create many alternative historical datasets. Then using these historical datasets, the parameters are recovered by applying MLE method. Next, we compared the results to the assumed ground truth values to measure the accuracy of recovery. The standard deviation of relative error of parameter recovery exhibits a power-law decay with an exponent value of -0.5 as the training data size increases. The resulting statistics enable us to verify the reliability of predictions based on originally recovered parameters. We did so by comparing original predictions to predictions based on parameters differing

from their average values by the desired multiple of their standard deviations as was demonstrated for the city model. We recorded the duration of each state (normal, on-fire, and recovery) and the number of emerging fires within the simulated time period. The largest relative error of these variables is just below 2%.

In conclusion, we showed that the CARP model is a novel approach to predict and simulate risks. It is particularly useful for modeling cascading catastrophic events and thus has potential applications for analyzing local and global risks. Local risks were demonstrated using simulated fires in cities. However, the CARP model was also successfully used to model global risks in earlier work [46]. A better understanding of global networked risks is critical to predicting and mitigating them [20]. Most of the world's critical infrastructure forms a complex, interconnected network prone to cascading failures with potentially devastating consequences to global stability [85]. Quantifying the limits of risk prediction, which are bounded by the amount of data, may inform earlier planning and thus potential mitigation of spreading risks spread and their adverse consequences.

CHAPTER 5
PREDICTABILITY OF CASCADES IN GEOMETRICALLY CONCENTRATED POWER GRID NETWORK

Recently, cascading failures in infrastructure systems have been studied intensively because of their ability to inflict tremendous damage to our society. Among basic infrastructure systems, the electric power system plays an important role, and as such, has drawn substantial attention in the past decades [33], [34], [36], [86]–[88]. The industries of gas, oil, water, agriculture, banking and finance, transportation, telecommunication and so on all depend heavily on electrical power supply [14]. A small breakdown in power grid may trigger huge cascading failures and propagate quickly through the entire system. Once the system breaks down extensively, it may cause the social unrest and extensive financial losses. The 2003 North America blackout and 2011 Southwest blackout are two good examples. A tiny and localized disruption may be exaggerated by the topological structures of a system and cause tremendous financial losses and potentially life-threatening situations. Moreover, the cascading failures travel at the speed of electricity in the system, which makes it hard to react to them in time for arresting their spread.

Due to the huge impacts that a blackout would have on our society, there is a strong need to study the cascading failures in a real world power grid system. A great deal of research has focused on theoretical networks. However, a real world power grid is much more complicated and operates under severe constraints, unlike standard models. Geometrical constraint is a very significant feature of the real world system since the transmitting lines have a limited length and are therefore mostly connected with adjacent power stations located within a practical distance. It is too expensive to transmit the current flow directly between two distant power stations. Practically, the connections have to follow physical restrictions and the current flow has to be delivered through multiple stations along the paths between the original location and destination. Similarly, the cascading failure inside a geo-

metrically constrained network also emerges locally at first and then propagates to distant parts of the system. Our study in this chapter models the resistance in a spatially embedded power grid and demonstrates how cascading failure behaves in different scenarios.

5.1 Introduction

In addition to previous studies about cascades in the scale-free networks [4], [9], [11], [16], [31], recent research has considered geometrical constraints in real-world power grids, which could influence the propagation of cascades.

Zhou, et al created a flow-base model in the European electrical power grid system, and simulated how the flow travels across countries [89], [90]. The authors showed a high correlation (more than 90%) between the simulation results and the public data of cross-border flow. The methodology has been applied to the European power grid [89]–[91]. After collecting the information of transmission stations (such as location, capacity, types, and connectivity), they computed the value of flow crossing the boundary of countries. To verify the results, they modified the model by varying the value of generation and capacity for each power plant. The simulated results show a good approximation to the public data in various scenarios.

Asztalos et al. studied the cascades in the spatially embedded random networks [91], [92]. Some previous studies on cascades consider the case that flow only travels along the shortest path from an original station to the destination. A more realistic scenario is that the flow can travel through all possible paths between any two stations. Moreover, in a spatial network, one node can only be connected by a direct power line with neighbors within certain geometric distances. A distant connection is prohibitively expensive in a real world power grid. Unlike Erdős-Rényi and Scale-free networks, a non-monotonic behavior between the relative size of the surviving giant component G and the tolerance parameter α can be observed in a spatially constraint network such as the random geometric graph and an empirical European power system. The damage does not decrease monotonically as the capacity increases. The authors also proposed effective mitigation strategies: removing preemptive nodes and introducing altruist nodes, which they tested in an empirical

spatially embedded power grid.

5.2 The UCTE network definition

Inspired by the previous studies, we focus on exploring the cascades in a spatial network triggered by a regional attack. It is difficult to predict the cascading failures directly in a real world system because of its high complexity. Our purpose is to study the limits of cascading prediction in an actual system. Mitigations against a single node failure can be effectively achieved. However, as the complexity (such as the number of initiators) increases, we find it is harder to predict the cascading damage. In some cases, the mitigation of cascade cannot be enhanced by only increasing the tolerance parameter. We concentrate our study on the European power transmission system, and the data is from the Union for the Co-ordination of Transmission of Electricity (UCTE) [89]–[91].

In this real word network, there are $N = 1254$ transmission stations. The average degree is $\langle k \rangle = 2.889$, and the number of edges is 1812. When such a system is in the functional state, all transmission stations and lines conduct flow below their capacity. We consider the resistance of each node and use the direct current (DC) power flow approximation in this resistor network to simplify the calculations. When an initial attack triggers the cascade, the overloaded parts will be removed, and the flow distribution will be recalculated. The new distribution may cause additional failures in other functional areas of the system. The cascade stops when there are no further failures, and the remaining system becomes stable again. Once the cascade ends, the relative size of the surviving giant component $G = N'/N$ is used to measure the damage of cascades, where N' is the number of surviving nodes and N is the number of original nodes [91]. A small value of G indicates a severe cascading damage.

5.3 Model dynamics

Asztalos et al. analyzed the evolution dynamics in this spatially embedded power grid with distributed flow [91], [92]. In this electricity system, current flow is transmitting along all possible paths and among all nodes. The total flow going

through a node is denoted as the load. As stated in [9], the capacity is defined as:

$$C_i = (1 + \alpha)l_i, \quad i = 1, 2, ...N, \tag{5.1}$$

where $\alpha \geq 0$ is the tolerance parameter, N is the number of nodes and l_i is the load of node i. When the load exceeds the capacity, this node becomes overloaded and flow redistribution triggers successive failures. In this resistor model, *Asztalos et al.* used the direct current (DC) power flow approximation to calculate load distribution. Because of the huge cost, the capacity of nodes and edges cannot be increased unrestrictedly so that the tolerance parameter α should stay in a reasonable range. Generally speaking, the removal of nodes with the small load will not have too much influence on the remaining power grid. Hence, the following overloads are unlikely to occur. In contrast, the failure of large-load nodes could cause severe damages in the end.

Asztalos et al. assumed that the flow can be transmitted in both directions and denote the edge flow between node i and its neighbor j as I_{ij} [91]. The sign of l_{ij} indicates the direction of flow propagation. A positive value means the flow travels from node i to node j and vice versa. Hence the actual value of flow going through node i should be the absolute value of all out-going or in-going edge flows of node i. The formula calculates the load of node i is shown in [91]:

$$l_i = \frac{1}{2} \sum_j \left| l_{ij} \right| . \tag{5.2}$$

When I_i exceeds C_i, node i becomes overloaded and is removed from the system together with its edges. To calculate the load distribution, the authors used the Kirchhoff's law and the Ohm's law. However, since the information of generators (source) and consumers (target) inside the system is unknown, each node can be a source node, and the target node will be randomly selected from remaining $N - 1$ nodes. One unit flow travels through each source-target pair. After averaging all $N - 1$ targets for one particular source and summing up all N possible sources,

Asztalos et al. computed the load value of node i as stated in [91], [92]:

$$l_{ij} = \frac{1}{N-1}\sum_{s,t=1}^{N}\left|I_{ij}^{st}\right|, \quad l_i = \frac{1}{N-1}\sum_{s,t=1}^{N}\left|I_i^{st}\right|, \tag{5.3}$$

where l_{ij} is the edge load of between node i and j, l_i is the load of node i. When the total amount of initial flow is I, the I_{ij}^{st} is defined in [91]:

$$I_{ij}^{st} = A_{ij} I \left(G_{is} - G_{it} - G_{js} + G_{jt}\right). \tag{5.4}$$

where A is the adjacency matrix of the power grid, G is the inverse of the *Laplacian* matrix. For any source node i and target node j, the expression to compute the load of node i is as stated in [91]:

$$I_i^{st} = \frac{1}{2}\sum_{j}\left|I_{ij}^{st}\right|, \tag{5.5}$$

Once we get the value of I_{ij}^{st} and I_i^{st}, we sum up all source-target pairs and get the value of l_{ij} and l_i. In this thesis, we use this methodology to focus on geometrically regional attacks, in which multiple adjacent nodes fail initially to trigger the cascades. In a more realistic case, if we can identify multiple source and target nodes, we can rewrite the Equation 5.3:

$$l_{ij} = A_{ij}\left(\sum_{k}^{N} G_{ik} I_k - \sum_{k}^{N} G_{jk} I_k\right) = A_{ij}\sum_{k}^{N}\left(G_{ik} - G_{jk}\right) I_k, \tag{5.6}$$

where I_k is the initial current flow of node k. If node k is a generator, I_k is positive which means that node k generates flow into the network; if node k is a consumer, I_k is negative; for other transmitting nodes, the value of I_k is zero. In this thesis, we use the above formulas introduced in [91], [92] to calculate the flow distribution and expand the analysis to more complicated scenarios.

5.4 Cascades triggered by single node failure

Cascading processes are triggered by initial failures of the system. The simplest case of initial failure is removing single node. In this experiment, we detect the relationship between the severity of cascading damage and the tolerance parameter α. The relative size of the giant component denoted as G is used to measure the loss of the cascade [91]. If G is close to zero, it means a huge damage in the system. If G is close to 1, the system encounters a little destruction. It is interesting to study how the value of α determines the value of G. Figure 5.1 shows the sizes of cascading failure as α increases [91]. To trigger the cascading process, we remove the node with the highest load initially. The value of α grows from 0 to 1.0 with an increment of 0.05. There is an obvious non-monotonic behavior between the cascading damage and the capacity allocation. When α changes from 0.4 to 0.45, the reduction of G indicates that the cascading damage in a higher capacity system is more severe than that of a smaller capacity system. When α is between 0.45 and 0.65, the cascading failures cause more damages than other cases. This phenomenon is counter-intuitive since in general, more capacity should protect more parts of the system. However, the realistic scenario behaves in an opposite manner because of the complex topology of the real world system.

In reality, the system capacity cannot be increased unrestrictedly. We have to consider the balance between the cost and benefit. In this thesis, instead of the uniform capacity, we propose a different method to allocate a stochastic random capacity to the power grid. In this method, the total capacity is identical to the uniform allocation. The additional capacity $\triangle C_i$ is drawn randomly from a uniform distribution with a mean value αl_i and a width σ. We assign the value of σ proportional to the mean value and guarantee the total additional capacity is fixed: $\sum_j \triangle C_i$. In Figure 5.2, we show the cascading damage in different cases of stochastic capacity allocations. In this experiment, we compare three different values of σ and finish 100 independent realizations to detect how the value of G changes as a function of σ. Three different stochastic capacity allocations are compared by varying the values of width σ: $\sigma = 0.25$ (green), $\sigma = 0.5$ (blue), and $\sigma = 1.0$ (red). In each case, we finish 100 realizations. The black dots show the results from the

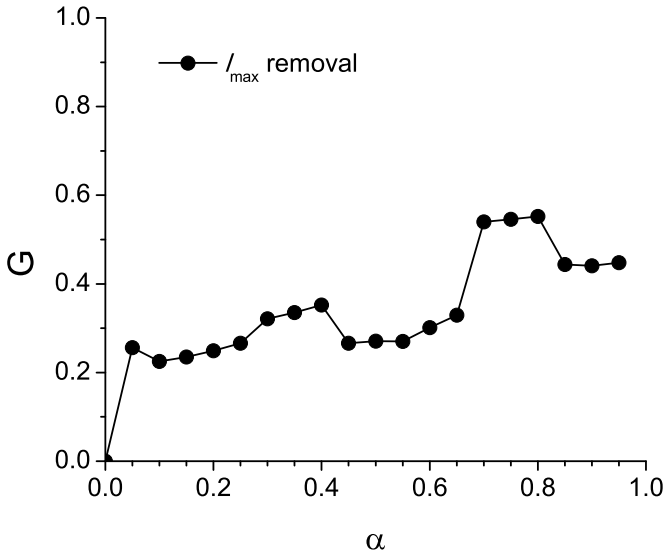


Figure 5.1: Cascades in the UCTE network triggered by a single-node removal.

uniform capacity case as shown in Figure 5.1. The node with the highest load is removed from power grid to trigger the cascade. The value of tolerance parameter α varies from 0 to 1.0 with a incremental step of 0.05. In general, a higher value of σ leads to a smaller damage without additional cost since the total capacity is preserved. For a baseline comparison, we also plot the initial case when the same amount of relative additional capacity is assigned to each node (black filled symbols). The results clearly demonstrate that the fixed-cost stochastic distribution of resources (capacities) allows for identifying particular realizations which provide superior protection against cascading failures in the power grid.

Figure 5.3 demonstrates the survived power grid caused by the removal of the highest load node in case of the uniform capacity allocation when the tolerance is $\alpha = 0.45$, and the width of the stochastic search space is $\sigma = 1/2$. The best-case scenario reflects the highest protection obtained from the stochastic capacity allocation, and the worst-case scenario shows the lowest protection obtained from stochastic capacity allocation. It is evident that the best case scenario has many more surviving nodes than other two cases. Since the total capacity remains unchanged in these cases, the best performance scenario protects the whole system

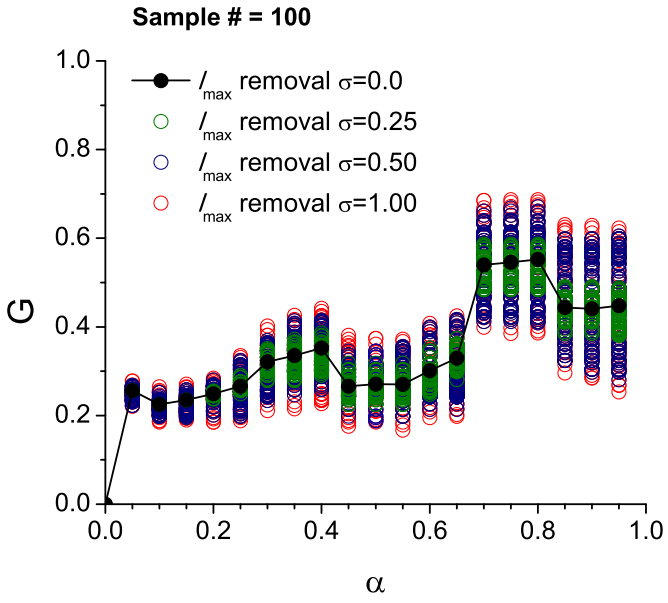


Figure 5.2: Cascades with different stochastic capacity allocations.

most based on the value of G. This result gives us a better understanding of the stochastic capacity allocation, which plays a critical role in determining the severity of cascading damage. Even for the fixed total capacity and the same value of the tolerance parameter, more randomness in the allocation makes the system obtain a higher potential ability to protect more parts of the system. In some scenarios, "fuse" nodes receive a smaller capacity and become more vulnerable to overload. The failure of "fuse" nodes blocks the propagation of cascading failures. We will discuss this issue more in this chapter.

5.5 The cascades triggered by spatially-localized regional attacks

In addition to previous results of cascades triggered by a single node removal [91], we consider the cascades triggered by regional attacks. In a real world case, the initial attack in a power grid could be the failures of multiple adjacent nodes. Taking into account the potential huge loss of cascades, we detect the impact of a geometrically constrained triggering in the UCTE network.

We define the radius r of a region to be the largest distance between the

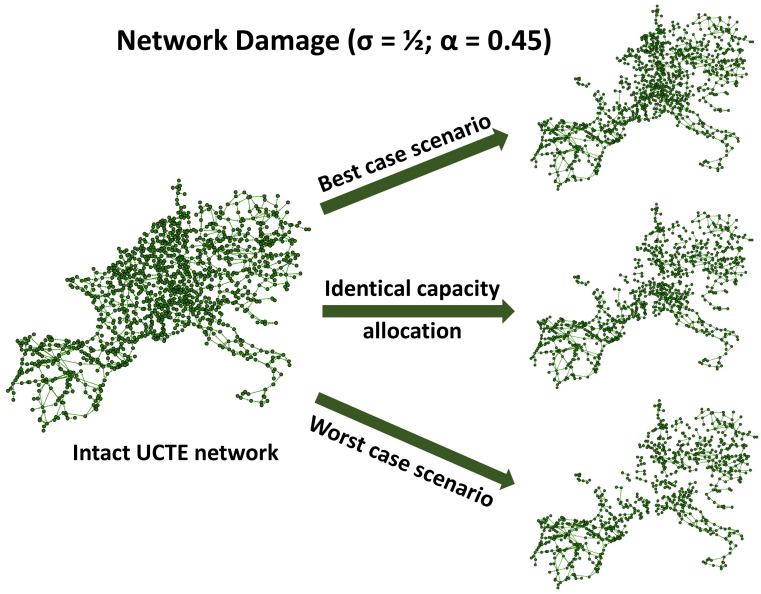


Figure 5.3: Cascading damage comparison in the UCTE network.

geometrical center and any node inside this region. In this experiment, we select a region with a radius of r at the border between France and Spain which consists of 9 nodes. If we expand the radius of the region to be $2r$, $3r$, $5r$ and $10r$, the number of nodes inside this region becomes 12, 20, 37 and 111. Consequently, we simulate the cascading failures triggered by a regional attack in the European electricity system and study how the cascading damage varies with different sizes of attacking region. The attacking regions with different sizes are shown in Figure 5.4. The topological features of the system determine the impact of adjacent initiators. In a larger region such as the $10r$ region, the area becomes larger, but the initiators are less spatially concentrated. This change increases the complexity and randomness of initiators, which makes the cascading failures more intricate.

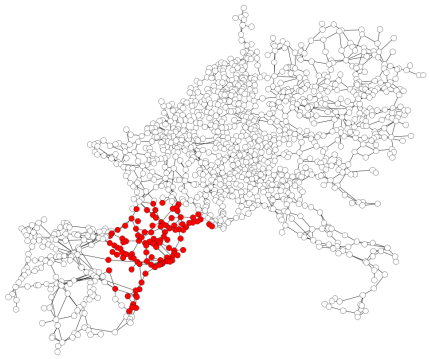
To find the correlation between the cascading damage and the node degree together with the initial load, we preserve identical numbers of initiators for all cases within different regional areas. In other words, 9 nodes are randomly selected from the attack region as initiators to trigger the cascade.

(a) Attack region of r

(b) Attack region of $2r$

(c) Attack region of $3r$

(d) Attack region of $5r$



(e) Attack region of $10r$

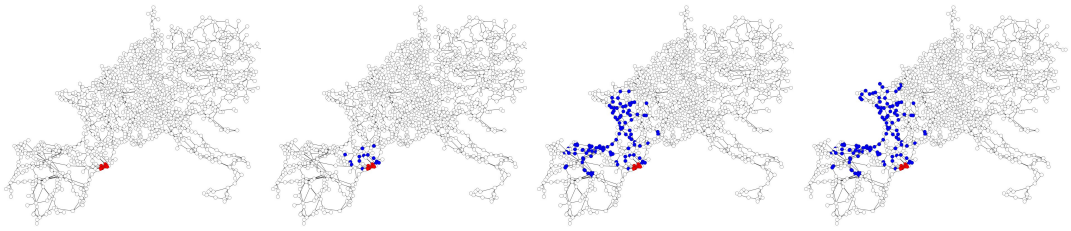
Figure 5.4: Different attack regions in the UCTE network.

5.5.1 Cascade with different tolerance parameters

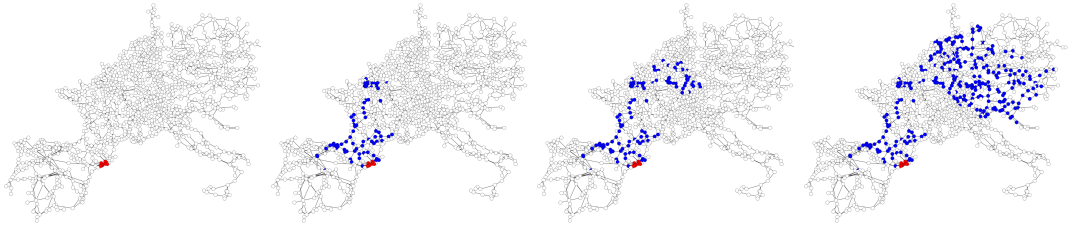
We start with the case of a r region and assign the tolerance parameter α to be 0.1. Figure 5.5 shows the propagation of the cascading failure across the system triggered by a 9-node regional attack. The initiators are located at the border between France and Spain, which are marked in red. We set the value of the tolerance parameter to be 0.1 and 0.15.

In the first case, as shown in Figure 5.5a, after initiators are failed, the flow of the initiators is redistributed in the system. In step 2, according to the new flow allocation, several nodes in normal operating conditions exceed their capacities. Therefore, these nodes are overloaded in this step and marked in blue. These overloaded nodes are located very close to initiators, mostly at the southwest of France. Then the flow distribution is redistributed again, and more overloaded nodes appear. In step 3, many more nodes become overloaded. The north of Spain and west of France suffer huge damages. The propagation of cascading failures is evidently traveling far from the original attack region to Eastern Europe. In the final step, only a few nodes are overloaded. The remaining system becomes stable again, and the cascading process stops. In the end, the number of overloaded nodes is 117 and the size of the giant component is $N' = 949$. Compare with the original size of 1254, the relative size of the giant component is $G = 0.7568$. In this case, the cascade caused by the regional attack only affects parts of Western Europe and does not influence distant parts of the system. We can find a different cascading path in the second case.

It is interesting that cascade causes much more severe damages in the UCTE network with a higher capacity. In the case of $\alpha = 0.15$, the cascade travels to Eastern Europe and has a profound impact even at that distance. However, in the case of $\alpha = 0.1$, only the nodes in northern Spain and western France are overloaded, which is more localized than the first scenario. It seems counter-intuitive because a power grid with a higher capacity should be more resistant to cascading failures. However, in reality, a higher capacity system suffers from a more severe cascading damage. The reason might be the complex topological features in a practical system. When we increase the capacity of nodes, several critical nodes can survive from the



(a) Tolerance parameter is 0.1



(b) Tolerance parameter is 0.15

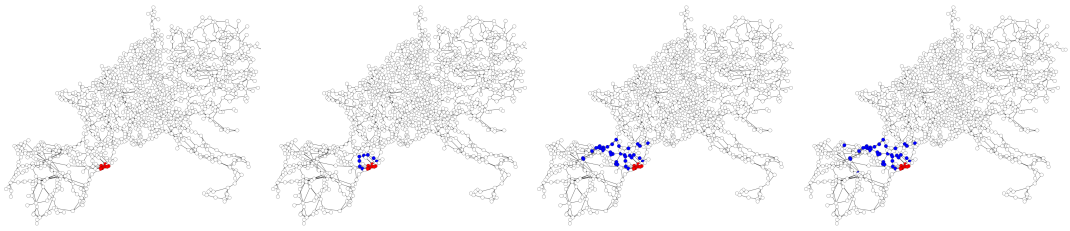
Figure 5.5: Cascade failures of the r region attack in the UCTE network.

cascading process and let the flow propagate to its neighbors. The excessive flow may cause a catastrophe in the remaining power grid. If the failure of these critical nodes can block the spreading of cascade and save more nodes eventually, we denote these nodes as "fuse" nodes, which play a significant role in determining the severity of cascade. In the case of smaller capacity, more "fuse" nodes are overloaded so as to protect more functional nodes in Eastern Europe effectively.

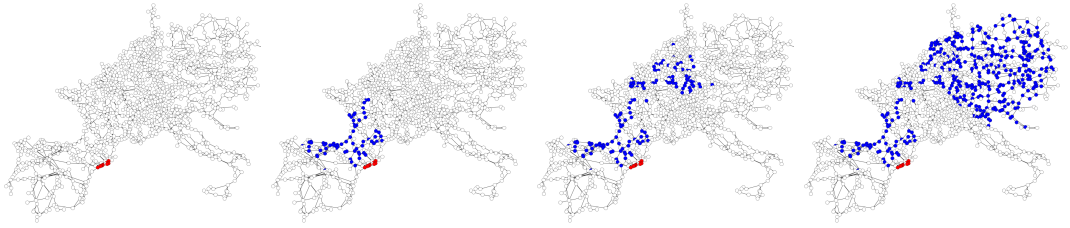
5.5.2 Cascades in different attack regions

Besides the original r region, we also study the cascading failures triggered in larger regions. Since larger regions have more than 9 nodes, there are many permutations to select the 9 initiators. Different permutations of initiators lead to different cascading processes. We find that in some scenarios, the cascade causes huge damages, but some other scenarios behave oppositely.

Figure 5.6 shows two cascading scenarios within the $2r$ region. Because the 9 initiators they have are different, these two cascading paths are totally different. In the low damage scenario, few nodes are overloaded in the end, and the cascade does not cause large damages. Most overloaded nodes are located at the border of France



(a) Small damage scenario



(b) Large damage scenario

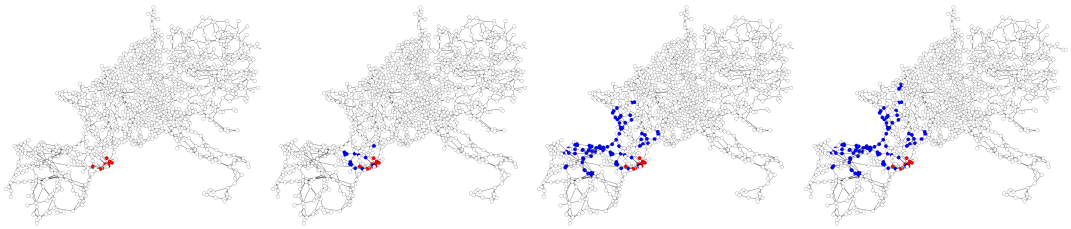
Figure 5.6: Cascade failures of the $2r$ region attack in the UCTE network.

and Spain. However, in the high damage scenario, the cascade spreads very quickly to Eastern Europe. Finally, almost all nodes in Eastern Europe are out of control. The results of different cascades differ so dramatically that even their initial failures are very close in a small region. This phenomenon shows the non-deterministic and intricate features of the network topology, which make it difficult to predict a cascading propagation in a real world system. This type of behavior also exists in a larger region such as $3r, 5r$, and $10r$.

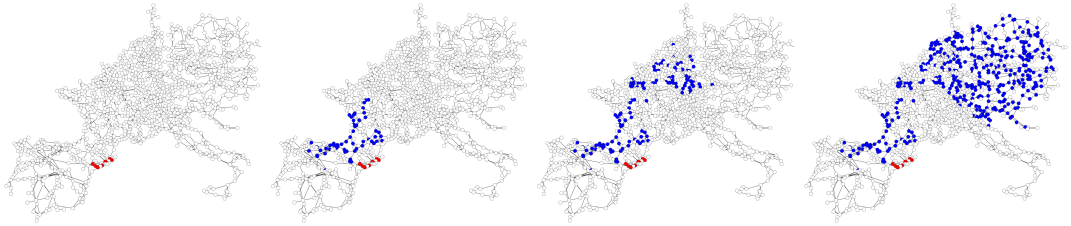
Figures 5.7 to 5.9 show a small and large cascading damage in larger initial regions: $3r, 5r$, and $10r$. Every cascade has 9 initiators randomly selected from the original region. The cascading propagation path is unique for each set of initiators.

5.6 Correlation between initiators and cascade failures

As shown in the previous subsections, the cascading process is determined mainly by the choice of initiators, and the variance of cascading damage is significant among different cases. To predict the cascading damage, we find the underlying connections between the features of initiators and cascading failures. We define the sum of degree and load of initiators as two important features. In this experiment,

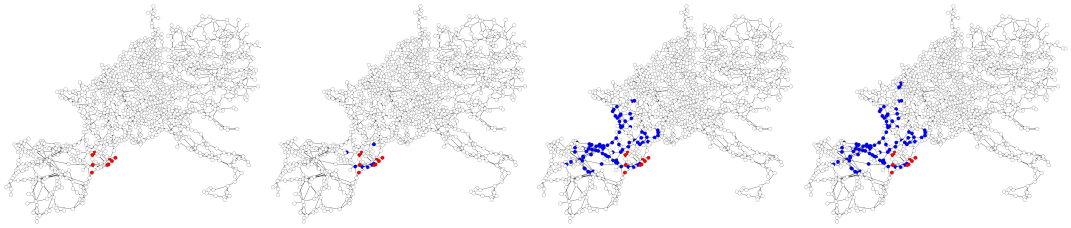


(a) Small damage scenario

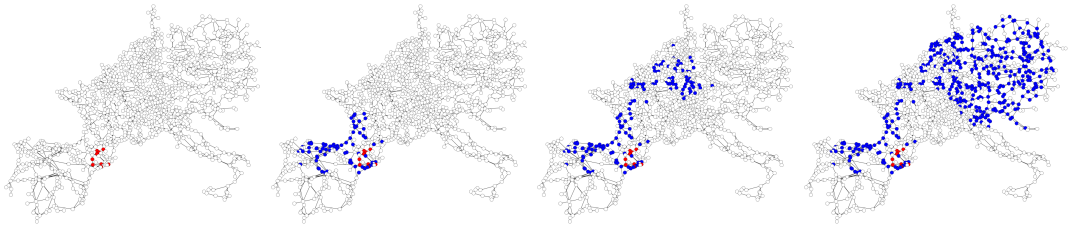


(b) Large damage scenario

Figure 5.7: Cascade failures of the $3r$ region attack in the UCTE network.

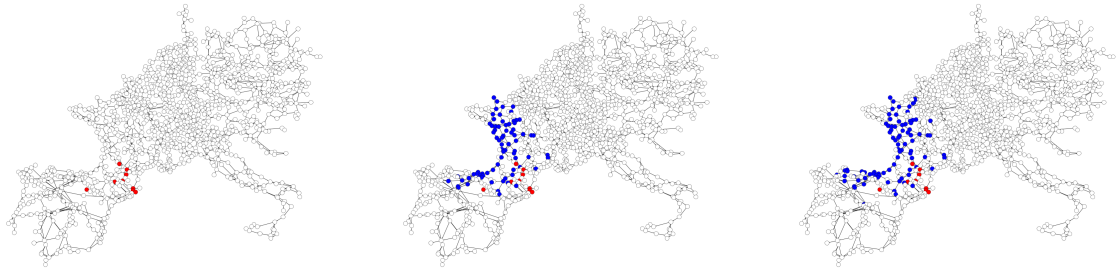


(a) Small damage scenario

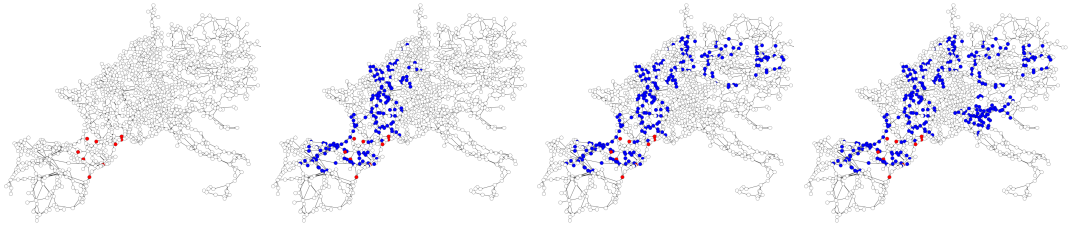


(b) Large damage scenario

Figure 5.8: Cascade failures of the $5r$ region attack in the UCTE network.



(a) Small damage scenario

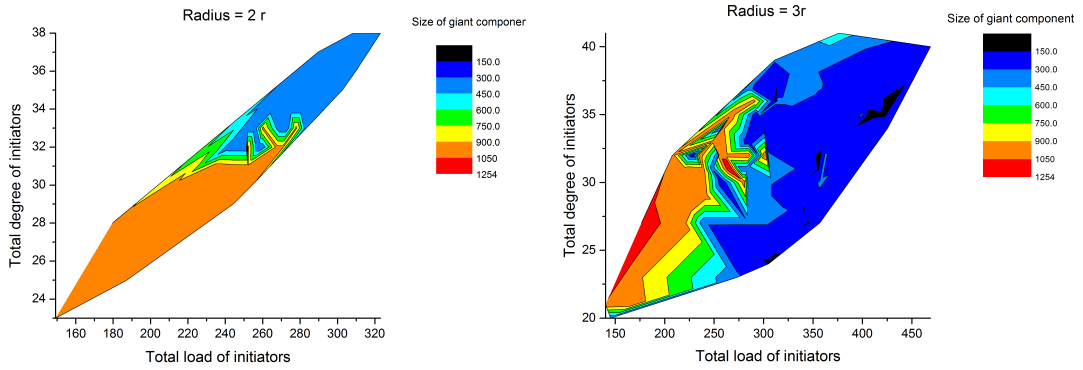


(b) Large damage scenario

Figure 5.9: Cascade failures of the $10r$ region attack in the UCTE network.

we focus on the $2r$ and $3r$ regions, and finish 100 independent realizations in each case. 9 initiators are randomly selected and the value of tolerance parameter α is 0.1. We record these features of initiators together with the relative size of the surviving giant component to analyze their correlation. Figure 5.10 shows the relationship between G and two features: total degree and load of initiators. The contour plot shows a bimodal distribution of G. The cascade leads to either a severe or a limited damage.

In Figure 5.10a, when the total load is smaller than 260 and the total degree is smaller than 30, the value of G is larger than 900 so that $G > 0.718$. However, large loads and degrees lead to a large damage $G < 0.4$. This behavior is consistent with two extreme cascading processes shown in the previous subsection, which indicates a positive correlation with the damage of cascade. In Figure 5.10b, the boundary of two phases is not as smooth as that in the case of a $2r$ region shown in Figure 5.10a. In the case of a larger attacking region, it is more likely for a cascading process



(a) case of $2r$ **(b) case of $3r$**

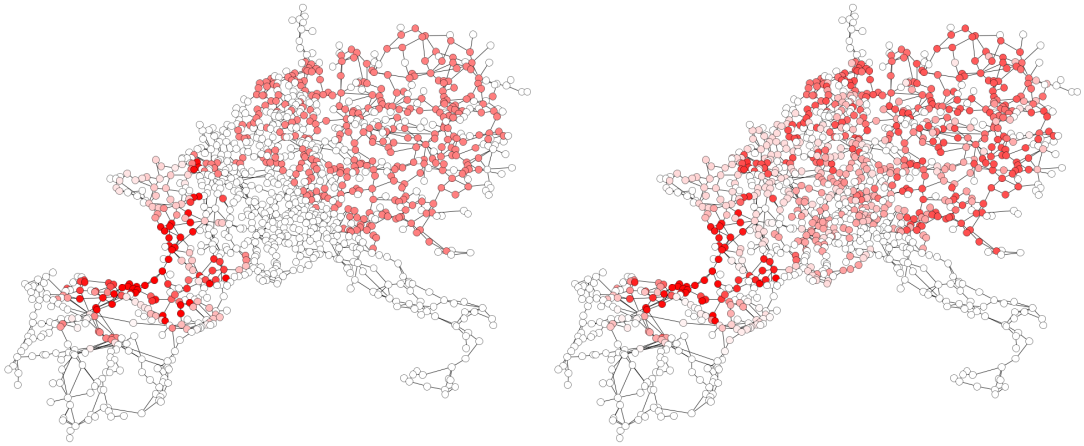
Figure 5.10: Relationship between cascades and initiators in $2r$ and $3r$ cases.

to cause a more severe damage since the area of dark color in the contour plot is greater than that of light color. As the region expands, it is more difficult for us to predict because of a more complex topological structure.

In addition to the contour plots, we count the frequency of node overload among 100 realizations for both cases as shown in Figure 5.11. Darker color implies a node becomes overloaded more frequently than lighter nodes. In both cases, the nodes at the border of France and Spain, together with most nodes in Eastern Europe all suffer a high frequency of overload. In the case of a $3r$ region, nodes in the center of Europe also fail frequently, which shows cascading failures are more wide-spreading than the case of a $2r$ region. It is interesting that many nodes are rarely overloaded because of the specific topological structure in a real world power grid.

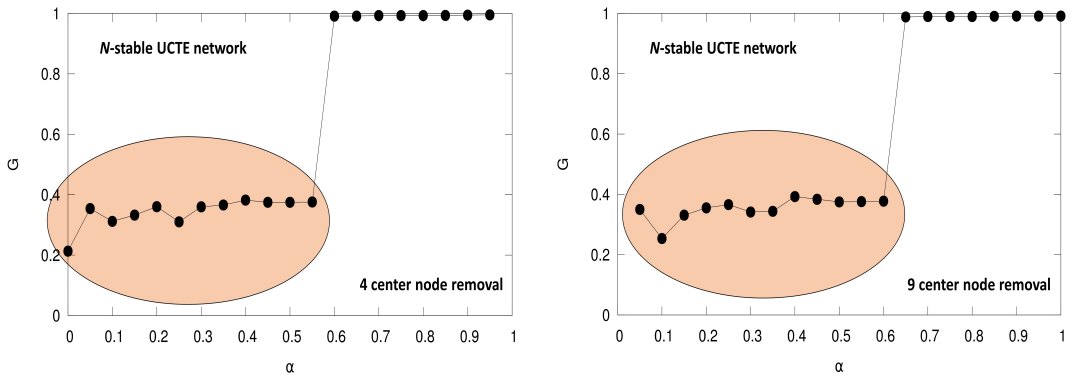
5.7 Phase transition in cascading failures with multiple initiators

As shown in Figure 5.1, the cascading damage measured by G in a spatially embedded network exhibits a non-monotonic behavior with parameter α in the case of single node removal [91]. We observe a similar performance of cascading damages when focusing on the cases of multiple initiators. Interestingly, unlike the single failure triggered cascades, the distribution of the cascading damage reflects a clear



(a) case of $2r$ **(b) case of $3r$**

Figure 5.11: Overload frequency in the $2r$ and $3r$ cases.

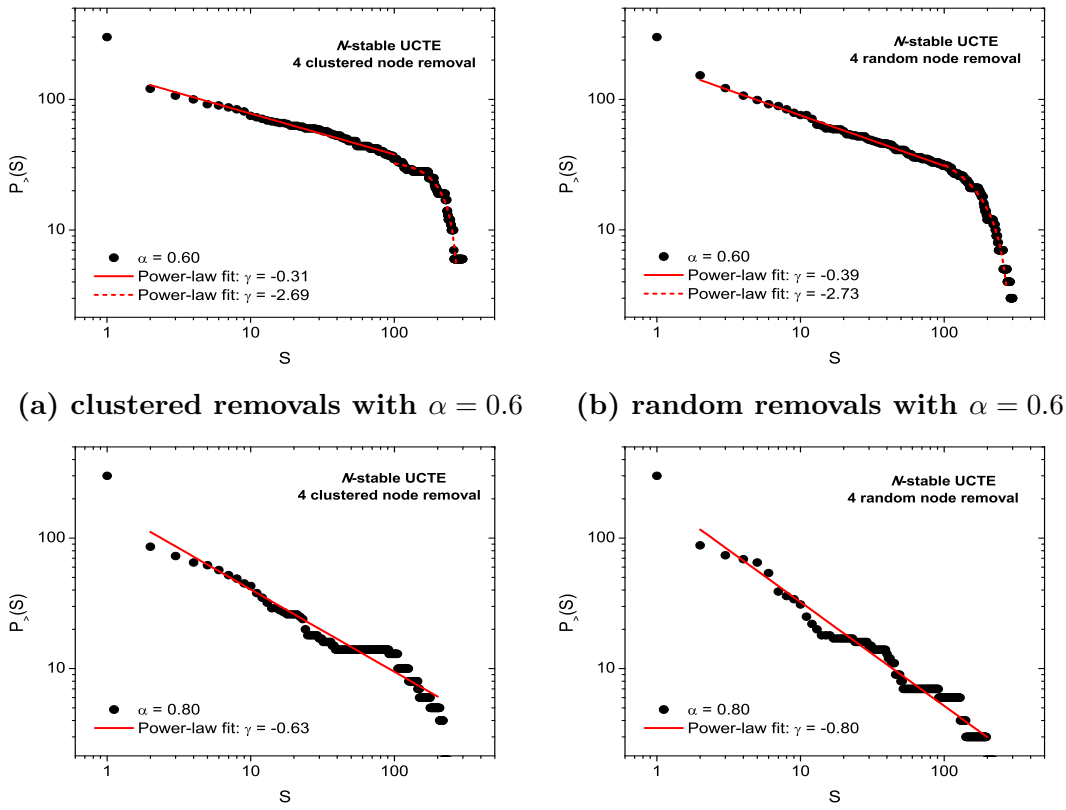


(a) 4 center node removal **(b) 9 center node removal**

Figure 5.12: Phase transition with increasing protections in the UCTE network.

phase transition in the cases of multiple initiators. When we increase α beyond the critical value, there is a dramatic jump for the value of G, which reflects a phase transition from vulnerable state to resistant state. This phenomenon exists in both cases of the regional or the random multiple node removals.

Figure 5.12 illustrates the phase transitions in cascades caused by multiple center node removals. There are two cases: 4 center node removal in the UCTE network and 9 center node removal in the UCTE network. The pink highlighted



(a) clustered removals with $\alpha = 0.6$ **(b) random removals with $\alpha = 0.6$**

(c) clustered removals with $\alpha = 0.8$ **(d) random removals with $\alpha = 0.8$**

Figure 5.13: Cascade size distributions in the phase transition regime.

areas depict the region before the transition, after which the network becomes fully protected against failures. The encircled highlighted regions depict the tolerance values, where the system suffers from large-scale cascading damages. Increasing the protection beyond these values, we can observe the system undergoes a phase transition, after which the entire network almost survives from the cascades. Compared with the behaviors of G in a single node removal case, large capacity allocation protects the network better with respect to multiple node removals.

Next, we analyze the distribution of cascading size in the phase transition regimes, where, according to prior research, such distributions are characterized by power-law tails [93]. The number of overloaded nodes S is used to measure the cascading size. In Figure 5.13, we plot the cumulative frequency distribution of S. Although S is not equal to $N - G$, there is a linear correlation between these two quantities, and plotting data as a function of G or S provides similar

results. For each particular value of S, we count the frequency of cascades with a larger number of overloaded nodes than current S. The cumulative cascade size distributions exhibit a power-law tail $P_{>}(S) \sim S^{-\gamma}$ with the exponent of γ, and the cascading size distributions can be found as $p(S) \sim S^{-(\gamma+1)}$. We focus on cascading failures in the UCTE network triggered by spatially clustered sets of 4 nodes or random sets of 4 nodes distributed throughout the network with the value of α fixed within the phase transition region. Figure 5.13 illustrates a power-law tail distributions and indicates that increasing the tolerance parameter α from $\alpha = 0.60$ to $\alpha = 0.80$ makes the distribution of the cascading size more abrupt, producing a higher power-law exponent. This observation can be explained intuitively; by increasing the tolerance parameter, we increase the protection in the system, and reduce the number of large cascades, thus increasing the number of small failures. As seen in these figures, for $\alpha = 0.60$ both cascade size distributions triggered by clustered sets of 4 nodes and random sets of 4 nodes exhibit a dual plateau. Thus, in Figure 5.13a and Figure 5.13b, we analyze these as two separate event regimes: the small event regime, where cascading damage is small-scale ($1 \leq S \leq 200$), and the large event regime, where the resulted failures are large ($201 \leq S$). We can see that both triggering methods produce similar cascade size distributions, with power-law exponents of $\gamma \approx 0.3$ in the small event regime, and $\gamma \approx 2.7$ in the large event regime. Our results are in agreement with previous work [93], where the probability density of conductance changes also follow a power-law, with two different regimes, reported both for synthetic networks and the Norway power grid system. Moreover, the values of the reported power-law exponents in the two distinct regimes are close to the values observed in our work in the UCTE network.

5.8 Conclusion

In this chapter, we study the cascades in a geometrically concentrated load-based power grid. We start with studying the cascade failures due to a single node removal (highest load). In some cases, cascades in the network with a higher capacity could cause a larger damage. However, in general, a higher capacity can indeed protect more nodes from overloading during the cascading process. To mitigate the

damage of cascades, we introduce a stochastic capacity allocation in addition to a uniform allocation. In the new method, the value of capacity is drawn from a uniform distribution. By varying the width of normal distribution, we control the randomness of the capacity allocation. Moreover, the total capacity of the system is preserved the same as the uniform capacity case. In some stochastic allocations, the damage of cascading failures becomes smaller than the uniform capacity case which means the new allocation method with fixed total value indeed protects the entire system better.

Also, we consider a regional attack to trigger the cascades instead of removing a single node such as a node with the highest degree or load. Unlike standard models, the complicated topology in a real world scenario makes it harder to predict the damage of a cascade. We demonstrate the non-monotonic influence of the tolerance parameter. A system with a higher capacity, unfortunately, causes much more overloads than a lower capacity system. The hidden "fuse" nodes play an important role in blocking the cascade spreading and protecting the whole system. It is not easy to find these "fuse" nodes since they are not the same in different scenarios and many factors determine the role of the "fuse" nodes, such as locations, tolerance parameters, connectivity, and the initial load of the system. Given more information, we can find more clues about the "fuse" nodes and apply a greedy algorithm to find these nodes iteratively. In addition, we also study the relationship between the surviving giant component and the initiators. We find the relationship between the cascading damage G and the total degree or load of all initiators. The results show a bimodal distribution: the damage is either very small or large. To make the power grid stable, it is better to keep the dangerous region in a low level of degree and load so that to reduce the vulnerability of it. In a smaller attack region, the boundary of two states is more evident than that in a larger region, which makes it easier for us to protect this area. Once the size of dangerous region increases, it becomes more difficult to identify the boundary of two phases. The area of a high damage is much larger than that of a small damage. To prevent the wide spreading of the cascades, we need to keep the initiators localized within a spatial region. Otherwise, the cascading damage will be amplified by the complex topology of the power grid

and the cascade would propagate far away very quickly. Similarly, there is also a non-monotonic behavior and a phase transition for the cascading damage as a function of the capacity in the cases of multiple node removals. When the tolerance parameter α is relatively small, the cascading failure leads to a large-scale damage. Once α exceeds certain critical value, the entire system can almost survive from the cascading failures. The distribution of overloaded nodes S can be approximated using a power law distribution. We can observe this behavior both in the cases of clustered node removal and random node removal. This phenomenon provides us a better understanding of cascading failure due to multiple initiators. In most scenarios, cascading failure just results in a low damage. Hence, we should pay more attention to the large-scale cascading damage.

As we all know the importance of the power grid system, it is necessary to find an effective strategy to protect it. Given more information about our real world system (such as the location of the region under potential attack, the identification of power generators and consumers, the actual tolerance parameters for each transmitting station), we can make a more accurate prediction of a future cascade to mitigate the damage.

CHAPTER 6
CONCLUSION

In this thesis, we studied the cascading failures in the abstract and real-world networks. Our motivation is to detect the properties of cascades and predict their future behaviors. Significantly, the cascading diffusion in a world-wide financial crisis has different patterns than that in an electrical blackout. Since the behaviors are not the same in every system, we apply distinct methodologies to different networks. At first, we introduced related studies of the cascading failures in various scenarios, such as simulating the cascading process in a standard or real-world network, revealing the robustness of networks, evaluating mitigation strategies in particular networks, and so on. Then, we utilized the Alternating Renewal Process to propose a quantitative method for the global risk network. Our study provides an analytical guidance to manage the potential risks in a coupled abstract system. By evaluating the precision of parameter recovery, we established a limit for the predictability of our methodology. To expand our analysis, we focused on a real-world power grid system with spatial constraints. We concluded that several key factors, such as the capacity allocation, initiators, and spatial correlation of a real-world system, significantly determine the severity of cascading damage. The full contributions are described as follows.

6.1 Contributions

In Chapter 3, we formulated a novel method to quantitatively analyze the cascading failures in the global risk network, which is a binary-state stochastic model for the propagation of risk materialization. Each risk switches between binary states (normal and materialized). The model dynamics is governed by three state transitions: internal triggering, external triggering, and recovery, which are assumed to be independent Poisson processes. Each state transition has a specific control parameter. To build the network, we took advantage of crowd-sourcing assessments of global risks, including the likelihood, impact of risk materialization, and connec-

tions of risks. Professional evaluations act as a good starting point for our study to differentiate the specific risks that are critical to the whole system. If we delete the connection of risks or ignore the distinctive likelihoods, the prediction of network activity becomes less accurate than that of the full model. We collected the historical events of each risk as the training dataset. Since the model dynamics is Markovian, each node may alternate between binary states at each discrete time step. We implemented the maximum likelihood estimation to obtain the optimal values for control parameters. The convex shape of the likelihood function ensures the estimated parameters are globally optimal. After that, we analyzed the robustness, persistence, and contagion potential of global risks according to the simulation results. Based on these measurements, we identified the most persistent risks, which stay materialized longer than others, such as the "Severe income disparity" and "Chronic fiscal imbalances," contribute the most to the activity level (the number of active risks) of the whole network. Moreover, we demonstrated that reducing the likelihood of risks better mitigates the performance than cutting the connections does. In addition, when we considered a fix-expense mitigation strategy, the best mitigation performance comes from the case of controlling half of the network. Therefore, our study produces a quantitative analysis for the stakeholders to deeply understand the interacted network and efficiently manage the potential large-scale risks.

In Chapter 4, we investigate the limit of predictability of the Cascading Alternating Renewal Process (CARP) model, which simulates the stochastic processes. To reveal the prediction accuracy of the methodology described in Chapter 2, we applied a similar method to model the fire propagation in an artificial city. In the city structure, there are three types of houses: large, medium and small. Each type of house reacts differently for state transitions. Unlike the global risk network, there are four state transitions in the fire-propagation model: internal triggering, external triggering, fire extinguishing, and recovery. Similarly to the global risk network, each state transition requires a control parameter. Since the model dynamics is Markovian, we intentionally assigned the ground-truth values for the parameters and generated the time series of model evolution as the historical dataset of the model. Then, we completed parameter recovery from various historical datasets

and compared the recovered parameters with the ground-truth parameters. Even for the hidden parameters, our method splits their combined effects and makes an excellent approximation of the ground-truth values, which is difficult since we can only observe their combined effects explicitly, but do not know the actual reason for this direct observation. After investigating various scenarios of the model, we concluded that as the historical dataset increases, the variance of the relative error of parameter recovery declines in a power-law decay. However, the variance cannot be eliminated because the stochastic properties of the model dynamics determine the limit of parameter estimation. In this research, we verified the reliability of predictions of the CARP model and demonstrated the ability of the CARP model to simulate the small-scale (the fire propagation in a city) and worldwide risks (global risks network).

In Chapter 5, we study the cascading failures in a resistor power grid with spatial constraints. In a real-world system, the locations of power stations and the length of transmitting lines have to obey spatial limitations based on a cost-benefit consideration, so the properties of cascades in a real-world system are different from those in the case of standard structures. The conserved flow in the European power grid is governed by Kirchhoff's and Ohm's law. We applied a direct current (DC) approximation to calculate the resistance without considering the phase of current flow. At first, we detected the influence of a single-node (with the highest load) removal on the cascading damage. The relative size of surviving giant component exhibits a non-monotonic behavior as a function of the tolerance parameter because the "fuse" nodes play critical roles in determining the propagation of failures. Despite the non-monotonic behavior, a significantly higher capacity level preserves more parts of the system. Instead of a uniform tolerance parameter for the whole system, we introduced a stochastic capacity allocation to mitigate the cascading damage. In this method, the excessive capacity of each node is randomly selected from a designated normal distribution, and the total capacity is the same as in the uniform tolerance parameter case. In addition, we also analyzed the cascading failures triggered by regional attacks, where initiators are located in a geometric region. Similarly, in certain cases, a larger capacity may not protect the system better

than a smaller capacity case. Moreover, we detected a bimodal distribution of the cascading size as a function of the total degree and load of the initiators. As the spatial concentration of the initiators weakens (initiators are far from each other), the features of the initiators worsen the predictability for the cascading damage. In reality, multiple power stations and transmitting lines may fail simultaneously due to internal reasons or external attacks. Our study provides a better understanding on controlling the potential risks and mitigating the damages in a spatially embedded real-world power grid. Decision makers need to focus on the high-density regions with a high spatial correlation, and reduce the connectivity and load level in these regions to design efficient strategies.

6.2 Future work

There are several ways to extend our current research. In the global risk network, we can update the network structure and risk evaluations by using the latest expert assessments. An updated crowd-sourcing assessments could reduce the error of parameter recovery. The experts from many fields and organizations, such as the World Economic Forum, evaluate the properties of global risks evaluations in real time. Moreover, to improve the quality of historical occurrence of each risk, we can utilize machine learning algorithms, such as natural language processing, to detect the keywords of references on the Internet to collect the historical data for parameter recovery instead of collecting data manually. This improvement could be significant since the quality of training data critically determines the quality of parameter estimation, which will enhance our future analysis on the resilience and persistence of global risks. In addition, it is interesting to find how global risks behave regionally, such as in Asia, Europe, North America, and so on. We can analyze the coupled risks in a smaller scale to increase the prediction accuracy. For the small-scale risk analysis (fire propagation in a city), we need to collect the real-world data to build the city structure and use the real-word fire data to estimate the control parameters. In this thesis, we arbitrarily assigned ground-truth values for these parameters, which may be unrealistic to some extent. Given more valuable data related to the fire propagation, we can make our model to simulate the fire

propagation better. Furthermore, we can also implement the regression analysis to detect the underlying patterns between the house properties and fire occurrences. For the real-world power grid, we can extend current analysis by considering the case of multiple generators and consumers. Hence, we need to revise the expressions to calculate the load distribution and design a balanced strategy for the power demand and supply. Although the analysis is more complicated in the case of multiple generators and consumers, we can benefit from the additional efforts to understand a real-world system more comprehensively.

REFERENCES

[1] S. H. Strogatz, "Exploring complex networks," *Nature*, vol. 410, no. 6825, pp. 268-276, Mar. 2001.

[2] R. Albert and A.-L. Barabási, "Statistical mechanics of complex networks," *Rev. Mod. Phys.*, vol. 74, no. 1, pp. 47-97, Jan. 2002.

[3] M. E. J. Newman, "The structure and function of complex networks," *SIAM Rev.*, vol. 45, no. 2, pp. 167-256, Jan. 2003.

[4] D. J. Watts, "A simple model of global cascades on random networks," *Proc. Natl. Acad. Sci*, vol. 99, no. 9, pp. 5766-5771, Apr. 2002.

[5] D. Centola, V. M. Eguíluz, and M. W. Macy, "Cascade dynamics of complex propagation," *Physica A*, vol. 374, no. 1, pp. 449-456, Jan. 2007. Accessed on: Apr. 21, 2017. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S0378437106007679.

[6] R. Albert, H. Jeong, and A.-L. Barabási, "Error and attack tolerance of complex networks," *Nature*, vol. 406, no. 6794, pp. 378-382, Jul. 2000.

[7] T. Roukny, H. Bersini, H. Pirotte, G. Caldarelli, and S. Battiston, "Default cascades in complex networks: topology and systemic risk," *Sci. Rep.*, vol. 3, no. 2759, Sep. 2013. Accessed on: Apr. 21, 2017. [Online]. Available: https://www.nature.com/articles/srep02759.

[8] X. Huang, I. Vodenska, S. Havlin, and H. E. Stanley, "Cascading failures in bi-partite graphs: model for systemic risk propagation," *Sci. Rep.*, vol. 3, no. 1219, Feb. 2013. Accessed on: Apr. 21, 2017. [Online]. Available: https://www.nature.com/articles/srep01219.

[9] A. E. Motter and Y. C. Lai. "Cascade-based attacks on complex networks," *Phys. Rev. E.*, vol. 66, pp. 065102-065105, Dec. 2002.

[10] P. Crucitti, V. Latora, and M. Marchiori, "Model for cascading failures in complex networks," *Phys. Rev. E*, vol. 69, no. 4, pp. 045104-045107, Apr. 2004.

[11] J. Wu, Z. Gao, and H. Sun, "Cascade and breakdown in scale-free networks with community structure," *Phys. Rev. E*, vol. 74, no. 6, pp. 066111-066115, Dec. 2006.

[12] R. Cohen, K. Erez, D. ben-Avraham, and S. Havlin, "Resilience of the Internet to Random Breakdowns," *Phys. Rev. Lett.*, vol. 85, no. 21, pp. 4626-4628, Nov. 2000.

[13] R. Cohen, K. Erez, D. ben-Avraham, and S. Havlin, "Breakdown of the Internet under Intentional Attack," *Phys. Rev. Lett.*, vol. 86, no. 16, pp. 3682-3685, Apr. 2001.

[14] S. M. Rinaldi, J. P. Peerenboom, and T. K. Kelly, "Identifying, understanding, and analyzing critical infrastructure interdependencies," *IEEE Control. Syst. Mag.*, vol. 21, no. 6, pp. 11-25, Dec. 2001.

[15] L. Daqing, J. Yinan, K. Rui, and S. Havlin, "Spatial correlation analysis of cascading failures: Congestions and Blackouts," *Sci. Rep.*, vol. 4, no. 5381, Jun. 2014. Accessed on: Apr. 21, 2017. [Online]. Available: https://www.nature.com/articles/srep05381.

[16] S. V. Buldyrev, R. Parshani, G. Paul, H. E. Stanley, and S. Havlin, "Catastrophic cascade of failures in interdependent networks," *Nature*, vol. 464, no. 7291, pp. 1025-1028, Apr. 2010.

[17] S. Pinnaka, R. Yarlagadda, and E. K. Çetinkaya, "Modelling robustness of critical infrastructure networks," in *Proc. 11th IEEE Int. Conf. Des. Rel. Commun. Networks*, Kansas City, MO, 2015, pp. 95-98.

[18] M. Parandehgheibi and E. Modiano, "Robustness of interdependent networks: The case of communication networks and the power grid," in *Proc. IEEE Global Commun. Conf.*, Atlanta, GA, 2013, pp. 2164-2169.

[19] M. Parandehgheibi, E. Modiano, and D. Hay, "Mitigating cascading failures in interdependent power grids and communication networks," in *Proc. IEEE Int. Conf. Smart Grid Commun.*, Venice, 2014, pp. 242-247.

[20] D. Helbing, "Globally networked risks and how to respond," *Nature*, vol. 497, no. 7447, pp. 51-59, May 2013.

[21] A. Vespignani, "Complex networks: The fragility of interdependency," *Nature*, vol. 464, no. 7291, pp. 984-985, Apr. 2010.

[22] H. Cramér, *Mathematical Methods of Statistics*. Princeton, NJ, USA: Princeton Univ. Press, 1946.

[23] M. H. DeGroot, *Probability and Statistics*, 4th ed. Boston, MA, USA: Addison-Wesley, 2012.

[24] J. A. Rice, *Mathematical Statistics and Data Analysis*, 3rd ed. Belmont, CA, USA: Duxbury Press, 2006.

[25] A. P. Dempster, N. M. Laird, and D. B. Rubin, "Maximum likelihood from incomplete data via the EM algorithm," *J. Roy. Stat. Soc.*, vol. 39, no. 1, pp. 1-38, Jan. 1977.

[26] A.-L. Barabási and R. Albert, "Emergence of scaling in random networks," *Sci.*, vol. 286, no. 5439, pp. 509-512, Oct. 1999.

[27] D. J. Watts and S. H. Strogatz, "Collective dynamics of 'small-world networks," *Nature*, vol. 393, no. 6684, pp. 440-442, Jun. 1998.

[28] P. Erdős and A. Rényi, "On random graphs," *Publ. Math. Debrecen*, vol. 6, no. 1, pp. 290-297, Nov. 1959.

[29] P. Erdős and A. Rényi, "On the evolution of random graphs," *Publ. Math. Inst. Hung. Acad. Sci.*, vol. 5, pp. 17-61, Oct. 1960.

[30] A.-L. Barabási, R. Albert, and H. Jeong, "Mean-field theory for scale-free random networks," *Physica A*, vol. 272, no. 1-2, pp. 173-187, Oct. 1999. Accessed on: Apr. 21, 2017. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S0378437199002915.

[31] V. Latora and M. Marchiori, "Efficient behavior of small-world networks," *Phys. Rev. Lett.*, vol. 87, no. 19, pp. 198701-198704, Oct. 2001.

[32] M. E. J. Newman and M. Girvan, "Finding and evaluating community structure in networks," *Phys. Rev. E*, vol. 69, no. 2, pp. 026113-026127, Feb. 2004.

[33] P. Hines, K. Balasubramaniam, and E. C. Sanchez, "Cascading failures in power grids," *IEEE Potentials*, vol. 28, no. 5, pp. 24-30, Sep. 2009.

[34] P. Hines, J. Apt, and S. Talukdar, "Large blackouts in North America: historical trends and policy implications," *Energy Policy*, vol. 37, no. 12, pp. 5249-5259, Dec. 2009.

[35] I. Dobson, B. A. Carreras, and D. E. Newman, "A loading-dependent model of probabilistic cascading failure," *Probab. Eng. Inf. Sci.*, vol. 19, no. 1, pp. 15-32, Jan. 2005.

[36] I. Dobson, B. A. Carreras, V. E. Lynch, and D. E. Newman, "Complex systems analysis of series of blackouts: cascading failure, critical points, and self-organization," *Chaos*, vol. 17, no. 2, p. 026103-026115, Jun. 2007.

[37] E. K. Çetinkaya, A. M. Peck, and J. P. G. Sterbenz, "Flow robustness of multilevel networks," in *Proc. 9th IEEE Int. Conf. Design Rel. Commun. Networks*, Budapest, 2013, pp. 274-281.

[38] Y. Shunkun, Z. Jiaquan, and L. Dan, "Prediction of cascading failures in spatial networks," *PLoS One*, vol. 11, no. 4, p. e153904, Apr. 2016. Accessed on: Apr. 21, 2017. [Online]. Available: http://journals.plos.org/plosone/article?id=10.1371/journal.pone.0153904.

[39] J. Zhao, D. Li, H. Sanhedrai, R. Cohen, and S. Havlin, "Spatio-temporal propagation of cascading overload failures in spatially embedded networks," *Nature Commun.*, vol. 7, no. 10094, Jan. 2016. Accessed on: Apr. 21, 2017. [Online]. Available: https://www.nature.com/articles/ncomms10094.

[40] Y. Koç, M. Warnier, R. E. Kooij, and F. M. T. Brazier, "A robustness metric for cascading failures by targeted attacks in power networks," in *Proc. 10th IEEE Int. Conf. Networking, Sensing and Control*, Evry, 2013, pp. 48-53.

[41] R. Albert, I. Albert, and G. L. Nakarado, "Structural vulnerability of the North American power grid," *Phys. Rev. E*, vol. 69, no. 2, pp. 025103-025106, Feb. 2004.

[42] M. E. J. Newman, "Scientific collaboration networks. II. shortest paths, weighted networks, and centrality," *Phys. Rev. E*, vol. 64, no. 1, pp. 016132-016138, Jun. 2001.

[43] A. E. Motter, "Cascade control and defense in complex networks," *Phys. Rev. Lett.*, vol. 93, no. 9, pp. 098701-098704, Aug. 2004.

[44] C. M. Schneider, A. A. Moreira, J. S. Andrade, S. Havlin, and H. J. Herrmann, "Mitigation of malicious attacks on networks," *Proc. Natl. Acad. Sci.*, vol. 108, no. 10, pp. 3838-3841, Mar. 2011.

[45] E. Atalay, A. Hortaçsu, J. Roberts, and C. Syverson, "Network structure of production," *Proc. Natl. Acad. Sci*, vol. 108, no. 13, pp. 5199-5202, Mar. 2011.

[46] B. K. Szymanski, X. Lin, A. Asztalos, and S. Sreenivasan, "Failure dynamics of the global risk network," *Sci. Rep.*, vol. 5, no. 10998, Jun. 2015. Accessed on: Apr. 21, 2017. [Online]. Available: https://www.nature.com/articles/srep10998.

[47] L. Howell *et al.*, "Global Risks 2013," *World Econ. Forum*, Jan. 2013. Accessed on: Apr. 21, 2017. [Online]. Available: http://reports.weforum.org/global-risks-2013.

[48] D. C. Brabham, "Crowdsourcing as a model for problem solving: an introduction and cases," *Convergence*, vol. 14, no. 1, pp. 75-90, Feb. 2008.

[49] E. Estellés-Arolas and F. González-Ladrón-de-Guevara, "Towards an integrated crowdsourcing definition," *J. Inform. Sci.*, vol. 38, no. 2, pp. 189-200, Apr. 2012.

[50] M. K. Poetz and M. Schreier, "The value of vrowdsourcing: can users really compete with professionals in generating new product ideas?," *J. Prod. Innovation Manage.*, vol. 29, no. 2, pp. 245-256, Mar. 2012.

[51] J. Prpić, A. Taeihagh, and J. Melton, "The fundamentals of policy crowdsourcing," *Policy & Internet*, vol. 7, no. 3, pp. 340-361, Sep. 2015. Accessed on: Apr. 21, 2017. [Online]. Available: http://onlinelibrary.wiley.com/doi/10.1002/poi3.102/full.

[52] J. Prpić and P. Shukla, "Crowd science: measurements, models, and methods," in *Proc. 49th IEEE Hawaii Int. Conf. Syst. Sci.*, Koloa, HI, 2016, pp. 4365-4374.

[53] P. W. Holland, K. B. Laskey, and S. Leinhardt, "Stochastic blockmodels: First steps," *Social Networks*, vol. 5, no. 2, pp. 109-137, Jun. 1983. Accessed on: Apr. 21, 2017. [Online]. Available: http://www.sciencedirect.com/science/article/pii/0378873383900217.

[54] D. R. Cox and H. D. Miller, *The Theory of Stochastic Processes*. London, United Kingdom: Methuen, 1965.

[55] F. Beichelt, *Stochastic Processes in Science, Engineering and Finance*. Boca Raton, FL, USA: CRC Press, 2006.

[56] Y. Pawitan, *In All Likelihood: Statistical Modelling And Inference Using Likelihood*. Oxford, United Kingdom: Oxford Univ. Press, 2001.

[57] D. R. Mandel and A. Barnes, "Accuracy of forecasts in strategic intelligence," *Proc. Natl. Acad. Sci.*, vol. 111, no. 30, pp. 10984-10989, Jul. 2014.

[58] E. W. Steyerberg *et al.*, "Assessing the performance of prediction models: a framework for some traditional and novel measures," *Epidemiology*, vol. 21, no. 1, pp. 128-138, Jan. 2010. Accessed on: Apr. 21, 2017. [Online]. Available: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3575184.

[59] G. W. Brier, "Verification of forecasts expressed in terms of probability," *Mon. Wea. Rev.*, vol. 78, no. 1, pp. 1-3, Jan. 1950.

[60] X. Lin, A. Moussawi, G. Korniss, J. Bakdash and B. K. Szymanski, "Limits of risk predictability in a cascading alternating renewal process model," submitted for publication.

[61] E. Banks, *Catastrophic Risk: Analysis and Management*. New York, NY, USA: John Wiley & Sons, 2005.

[62] N. N. Taleb, *The Black Swan: The Impact of the Highly Improbable*. New York, NY, USA: Random House Publishing Group, 2007.

[63] E. Paté-Cornell, "On 'black swans and 'perfect storms: risk analysis and management when statistics are not enough," *Risk Anal.*, vol. 32, no. 11, pp. 1823-1833, Nov. 2012.

[64] A. A. Ganin *et al.*, "Operational resilience: concepts, design and analysis," *Sci. Rep.*, vol. 6, no. 19540, Jan. 2016. Accessed on: Apr. 21, 2017. [Online]. Available: https://www.nature.com/articles/srep19540.

[65] D. Linders and W. Schoutens, "A framework for robust measurement of implied correlation," *J. Comput. Appl. Math.*, vol. 271, pp. 39-52, Dec. 2014. Accessed on: Apr. 21, 2017. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S0377042714001757.

[66] J. Reason, *Human Error*. Cambridge, United Kingdom: Cambridge Univ. Press, 1990.

[67] P. E. Auerswald, L. M. Branscomb, T. M. L. Porte, and E. O. Michel-Kerjan, Eds. *Seeds of Disaster, Roots of Response: How Private Action Can Reduce Public Vulnerability*. Cambridge, United Kingdom: Cambridge Univ. Press, 2006.

[68] E. Çinlar, "Markov renewal theory," *Advances Appl. Probab.*, vol. 1, no. 2, pp. 123-187, Oct. 1969.

[69] D. R. Cox and H. D. Miller, *The Theory of Stochastic Processes*. Boca Raton, FL, USA: CRC Press, 1977.

[70] J. M. Dickey, "The renewal function for an alternating renewal process, which has a Weibull failure distribution and a constant repair time," *Rel. Eng. & Syst. Safety*, vol. 31, no. 3, pp. 321-343, Jan. 1991. Accessed on: Apr. 21, 2017. [Online]. Available: http://www.sciencedirect.com/science/article/pii/095183209190075I.

[71] J. A. M. van der Weide and M. D. Pandey, "A stochastic alternating renewal process model for unavailability analysis of standby safety equipment," *Rel. Eng. & Syst. Safety*, vol. 139, pp. 97-104, Jul. 2015. Accessed on: Apr. 21, 2017. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S0951832015000630.

[72] F. Beichelt, *Stochastic Processes in Science, Engineering and Finance*. Boca Raton, FL, USA: CRC Press, 2006.

[73] J. Roldán and D. A. Woolhiser, "Stochastic daily precipitation models: 1. a comparison of occurrence processes," *Water Resour. Res.*, vol. 18, no. 5, pp. 1451-1459, Oct. 1982.

[74] D. A. Woolhiser and G. G. S. Pegram, "Maximum likelihood estimation of Fourier coefficients to describe seasonal variations of parameters in stochastic daily precipitation models," *J. Appl. Meteor.*, vol. 18, no. 1, pp. 34-42, Jan. 1979.

[75] M. R. Sampford, "The truncated negative binomial distribution," *Biometrika*, vol. 42, no. 12, pp. 58-69, Jun. 1955.

[76] B. Freireich and J. Li, "A renewal theory approach to understanding interparticle coating variability," *Powder Technol.*, vol. 249, pp. 330-338, Nov. 2013.

[77] U. Mann, "Analysis of spouted-bed coating and granulation. 1. batch operation," *Ind. Eng. Chem. Proc. Des. Dev.*, vol. 22, no. 2, pp. 288-292, Apr. 1983.

[78] S. M. Samuels, "A characterization of the Poisson process," *J. Appl. Probab.*, vol. 11, no. 1, pp. 72-85, 1974.

[79] J. A. Ferreira, "Pairs of renewal processes whose superposition is a renewal process," *Stochastic Process. and Their Applicat.*, vol. 86, no. 2, pp. 217-230, Apr. 2000.

[80] D. Silvestrov and A. Martin-Löf, *Modern Problems in Insurance Mathematics*. New York, NY, USA: Springer, 2014.

[81] X. Yang and A. P. Petropulu, "The extended alternating fractal renewal process for modeling traffic in high-speed communication networks," *IEEE Trans. Signal Process.*, vol. 49, no. 7, pp. 1349-1363, Jul. 2001.

[82] S. A. Steele, D. Tranchina, and J. Rinzel, "An alternating renewal process describes the buildup of perceptual segregation," *Front. Comput. Neurosci.*, vol. 8, pp. 166-178, Jan. 2015. Accessed on: Apr. 21, 2017. [Online]. Available: http://journal.frontiersin.org/article/10.3389/fncom.2014.00166/full.

[83] H. W. Lilliefors, "On the Kolmogorov-Smirnov test for normality with mean and variance unknown," *J. Amer. Stat. Assoc.*, vol. 62, no. 318, pp. 399-402, Jun. 1967.

[84] F. J. Massey, "The Kolmogorov-Smirnov test for goodness of fit," *J. Amer. Stat. Assoc.*, vol. 46, no. 253, pp. 68-78, Mar. 1951.

[85] C. Perrow, *Normal Accidents: Living with High Risk Technologies*. Princeton, NJ, USA: Princeton Univ. Press, 1999.

[86] S. Soltan, D. Mazauric, and G. Zussman, "Cascading failures in power grids: analysis and algorithms," in *Proc. 5th Int. Conf. Future Energy Syst.*, Cambridge, 2014, pp. 195-206.

[87] T. Verma, W. Ellens, and R. E. Kooij, "Context-independent centrality measures underestimate the vulnerability of power grids," *Int. J. Critical Infrastructures*, vol. 11, no. 1, pp. 62-81, Jan. 2015.

[88] M. Rahnamay-Naeini, Z. Wang, N. Ghani, A. Mammoli, and M. M. Hayat, "Stochastic analysis of cascading-failure dynamics in power grids," *IEEE Trans. Power Syst.*, vol. 29, no. 4, pp. 1767-1779, Jul. 2014.

[89] Q. Zhou and J. W. Bialek, "Approximate model of European interconnected system as a benchmark system to study effects of cross-border trades," *IEEE Trans. Power Syst.*, vol. 20, no. 2, pp. 782-788, May 2005.

[90] N. Hutcheon and J. W. Bialek, "Updated and validated power flow model of the main continental European transmission network," in *Proc. IEEE Grenoble Conf.*, Grenoble, 2013, pp. 1-5.

[91] A. Asztalos, S. Sreenivasan, B. K. Szymanski, and G. Korniss, "Cascading failures in spatially-embedded random networks," *PLoS One*, vol. 9, no. 1, p. e84563, Jan. 2014. Accessed on: Apr. 21, 2017. [Online]. Available: http://journals.plos.org/plosone/article?id=10.1371/journal.pone.0084563.

[92] A. Asztalos, S. Sreenivasan, B. K. Szymanski, and G. Korniss, "Distributed flow optimization and cascading effects in weighted complex networks," *Eur. Phys. J. B*, vol. 85, no. 8, pp. 288-297, Aug. 2012.

[93] J. Ø. H. Bakke, A. Hansen, and J. Kertész, "Failures and avalanches in complex networks," *Europhys. Lett.*, vol. 76, no. 4, pp. 717-723, Oct. 2006.