Modeling Uncertain and Dynamic Interdependencies of Infrastructure Systems Using Stochastic Block Models

Jin-Zhu Yu
Interdisciplinary Studies of Systems Engineering,
Vanderbilt University,
Nashville, TN 37235
e-mail: jinzhu.yu@vanderbilt.edu

Hiba Baroud[1]
Department of Civil and Environmental Engineering,
Vanderbilt University,
Nashville, TN 37235
e-mail: hiba.baroud@vanderbilt.edu

Modeling the resilience of interdependent critical infrastructure (ICI) requires a careful assessment of interdependencies as these systems are becoming increasingly interconnected. The interdependent connections across ICIs are often subject to uncertainty due to the lack of relevant data. Yet, this uncertainty has not been properly characterized. This paper develops an approach to model the resilience of ICIs founded in probabilistic graphical models. The uncertainty of interdependency links between ICIs is modeled using stochastic block models (SBMs). Specifically, the approach estimates the probability of links between individual systems considered as blocks in the SBM. The proposed model employs several attributes as predictors. Two recovery strategies based on static and dynamic component importance ranking are developed and compared. The proposed approach is illustrated with a case study of the interdependent water and power networks in Shelby County, TN. Results show that the probability of interdependency links varies depending on the predictors considered in the estimation. Accounting for the uncertainty in interdependency links allows for a dynamic recovery process. A recovery strategy based on dynamically updated component importance ranking accelerates recovery, thereby improving the resilience of ICIs. [DOI: 10.1115/1.4046472]

Keywords: resilience, interdependent networks, water and power networks, uncertain interdependencies, stochastic block model, dynamic importance ranking, social vulnerability index

1 Introduction

Modern society depends on critical infrastructure for security, safety, health, and overall well-being. Critical infrastructure systems include water distribution systems, transportation systems, and power grids, among other sectors. Recent changes in social and economic development as well as advances in technology have led to increasing interdependent connections between critical infrastructure systems [1]. On one hand, these interdependencies can enhance the overall efficiency and robustness of these systems under normal operations by providing a platform for information, services, and product sharing. On the other hand, highly connected critical infrastructure can be more vulnerable to hazards, whereby a disruptive event in one system can result in *cascading failures* across other connected infrastructure systems [2–4]. For example, a power outage cascades into the water supply system as pumping stations fail due to loss of power, and the lack of water for cooling impacts the generation of electricity. As such, interdependencies among infrastructures must be taken into account to understand the operational characteristics of infrastructure systems [5] and inform operations and future design of robust and resilient systems. One of the major challenges in evaluating the impact of disruptive events on interdependent critical infrastructure (ICIs) lies in understanding the influence of multiple interdependencies on the systemic performance after a disruption.

Rinaldi et al. [6] define interdependency as "the bidirectional relationship between two infrastructures through which the state of each infrastructure influences or is correlated to the state of the other." After formalizing the concept of interdependency, Rinaldi et al. [6] propose four principal classes of interdependencies: (1) physical interdependency, when energy or materials flow between two systems; (2) cyber interdependency, when information is transmitted between systems; (3) geographic interdependency, when the state of one system can be altered due to spatial proximity of another system; and (4) logical interdependency, when two systems influence each other via a mechanism that is not based on physical, cyber, or geographic connection. Although other classifications have also been proposed [7–10], this study employs the classification proposed by Rinaldi et al. [6] as the proposed approach can be directly extended to other types of interdependencies [11], which can be found in Ref. [12]. Modeling the resilience or recovery of ICIs has been extensively studied [9,13–19]. In most studies, ICIs are often modeled in a deterministic way by establishing interdependency links first and then evaluating the systemic performance. These links are established between randomly generated nodes [20], or based on either degree centrality [16] or spatial proximity (minimum Euclidean distance) [11,15,21–23]. One approach to assessing resilience under uncertainty is the use of Bayesian networks (BNs), such as Refs. [24–26]. The majority of studies that use BNs examine a single network instead of multiple interdependent networks. Furthermore, BNs are an acyclic graphical model and thus have great difficulty in handling the bidirectional relationship in interdependent networks.

The majority of existing studies do not consider the uncertainty associated with the interdependency between the networks and few studies can be found in the literature that addresses dynamic and uncertain interdependencies. Full knowledge of interdependencies across infrastructure networks, despite its key importance, is often not available due to lack of data [27]. Therefore, the topology of ICIs is subject to uncertainty. Neglect or improper characterization of this uncertainty can lead to underestimation or overestimation of system performance as the metrics cannot be assessed within an acceptable level of fidelity [28,29].

The objective of this paper is to model the resilience of ICIs by accounting for the uncertainty and dynamics of interdependencies

[1]Corresponding author.
Manuscript received May 31, 2019; final manuscript received February 14, 2020; published online March 27, 2020. Assoc. Editor: Geng Feng.

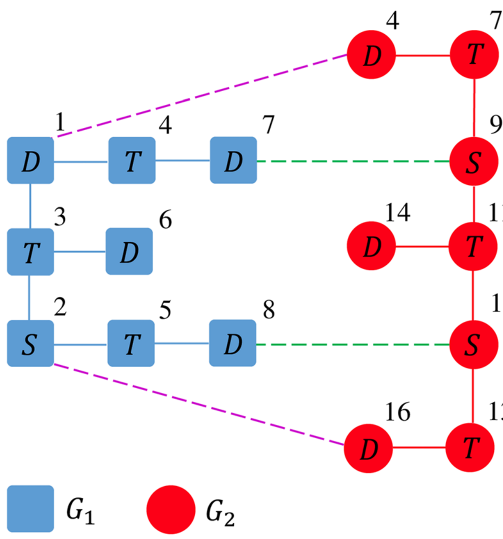

Fig. 1 A schematic of interdependent networks (dash lines represent interdependency links. Nodes labeled with *S*, *T*, *D* represent supply nodes, transmission nodes, and demand nodes, respectively).

after the disruption in order to guide restoration strategies. The specific contributions of this study are listed below:

(1) The application of statistical network models to evaluate uncertain network interdependencies. Specifically, the use of stochastic block models (SBMs) provides a probabilistic characterization of interdependency links between infrastructure networks.
(2) A formulation for estimating the likelihood that an interdependency link exists using multiple predictors to represent major factors influencing the presence of interdependency links. To avoid zero values in normalizing the variables that will be used in the denominator, a variant of min–max normalization called *truncated min–max normalization* is developed.
(3) A network recovery strategy based on dynamic ranking of component importance. Network restoration based on dynamic component ranking results in a faster recovery of ICIs when compared to static ranking.

The remainder of this paper is outlined as follows: In Sec. 2, we introduce the mathematical representation of interdependent networks, SBM, and network resilience. Section 3 outlines the approach to assessing the resilience of ICIs along with two restoration strategies based on static and dynamic component importance ranking. In Sec. 4, the proposed method is illustrated using interdependent water and power networks. Finally, Sec. 6 provides the conclusion along with a discussion for future work.

2 Background

2.1 Interdependent Networks. This study considers interdependent networks that are comprised of multiple individual networks connected by interdependency links. This work considers two individual networks; however, the approach can be extended to accommodate additional networks. Mathematically, $G = (G_1, G_2, E_{12}, E_{21})$ where $G_1 = (V_1, E_1)$ and $G_2 = (V_2, E_2)$ represent the individual networks and E_{12} and E_{21} represent the interdependency links from G_1 to G_2 and those from G_2 to G_1, respectively. The number of nodes and the number of links in G are denoted by $|V|$ and $|E|$, respectively. Each of the G_1 and G_2 is comprised of multiple supply nodes (S), transmission nodes (T), demand nodes (D), and the links that connect them (Fig. 1). The connectivity of a graph can be encoded using the *adjacency matrix* where rows and columns are labeled by nodes. Let $A_{|V|\times|V|}$ represent the adjacency matrix of the graph. If there is a link between node i and node j, $A_{ij} = 1$, and $A_{ij} = 0$ otherwise. Since self-edges are not possible, we assume $A_{ii} = 0$, which is a valid assumption for the network representing infrastructure systems. For the undirected graphs considered in this study, the adjacency matrix is symmetric. As an illustrative example, the adjacency matrix for the interdependent networks in Fig. 1 is shown in Fig. 2.

2.2 Stochastic Block Models. The SBM is a probabilistic graphical model for describing and analyzing the structure of a network [30]. Pioneered by Holland et al. [31], SBM has been widely used in community detection in social networks [32–34]. In SBM, the nodes in a graph are divided into different blocks based on the class membership of the nodes. The probability of a link between two nodes depends on the blocks (class) to which the nodes belong. Let $\boldsymbol{z} = [z_1, z_2, \ldots, z_{|V|}]^T$ represent the class membership vector and $\boldsymbol{\theta}$ represent the *block probability matrix* [35], the matrix of probabilities of forming edges between blocks and within a block (Fig. 3).

Formally, SBM is defined as follows [31,35]:

Definition 1 (SBM). *$\boldsymbol{A}$ is generated according to an SBM with respect to $\boldsymbol{z}$ if and only if: 1. $\forall\, i \neq j$, A_{ij} are statistically independent. 2. $\forall\, i \neq j$ and $i' \neq j'$ with $z_i = z'_i$ and $z_j = z'_j$, A_{ij} and A_{ij} are identically distributed.*

Given Definition 1, the block probability matrix, and the adjacency matrix, the probability of a link connecting node i in block a and j in block b can be given by

$$P(A_{ij} = 1 | z_i = a, z_j = b) = \theta_{ab} \quad (1)$$

In Eq. (1), θ_{ab} represents the probability of forming a link between block a and block b. Equivalently, $A_{ij}|z_i = a, z_j = b \sim \text{Bernoulli}(\theta_{ab})$, meaning that entries of $\boldsymbol{A}$ can be modeled as statistically independent Bernoulli random variables [36]. Depending on the available dataset, SBM can be applied to a priori setting where the partition of nodes is predefined and a posteriori setting where the block partition is uncertain [35]. In this study, SBM is used in the a priori setting.

The SBM offers several computational and application advantages such as the ability to estimate missing links based on incomplete data [37], the integration of statistical and network properties, and the flexibility in the analysis of stochastic interdependent links. However, this model does not consider the heterogeneity of nodes besides their block membership. In order to account for such heterogeneity, SBM is modified to give a probabilistic estimate of the presence of interdependency links E_{12} and E_{21}. The estimation of the probability of these links can be considered as a regression problem in which node attributes are the predictors and the probability of A_{ij} is the response variable. For example, suppose the class membership and the distance between nodes are used as the predictors, the regression model is shown in the following equation:

$$P(A_{ij} = 1 | z_1, z_2) = f(z_1, z_2, d_{v_i v_j}) \quad (2)$$

In Eq. (2), $d_{v_i v_j}$ represents the distance between v_i and v_j. Once the data on other nodal attributes are available, Eq. (2) can be modified to include more predictors, and the model form can be identified by use of statistical methods for model selection.

2.3 Network Resilience. First introduced in ecology [38], the concept of resilience has triggered significant interest during the past decades in several other fields, ranging from psychology to engineering [39]. Within the engineering domain, multiple definitions of resilience have been proposed by different organizations, such as Refs. [40–42] as well as several scholars in the field of engineering resilience analysis, such as Refs. [39], [43], and [44].

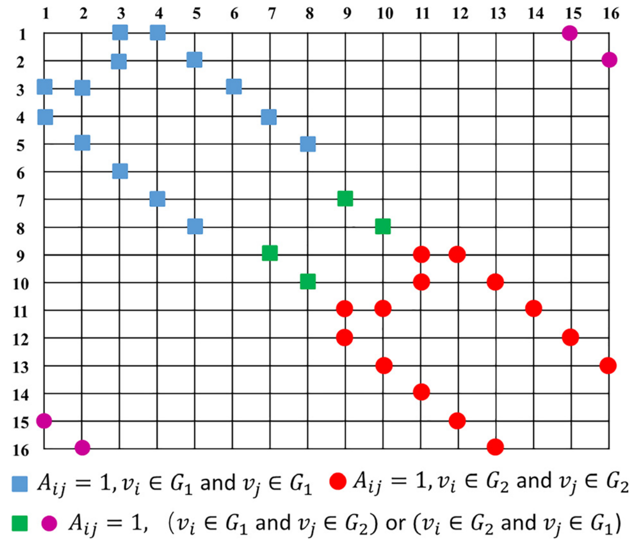

Fig. 2 Adjacency matrix for the interdependent networks in Fig. 1

In many of the definitions, resilience commonly refers to the performance of a system of networks before, during, and after a disruptive event [23,44].

Consider an engineering system where $\varphi(t)$ represents a service function describing system performance. Depending on the type of system, $\varphi(t)$ can describe network capacity, maximum flow, flowrate, or network connectivity [45]. The performance measure can also incorporate social and economic factors that indirectly impact the system performance and resilience. Figure 4 provides a general approach for understanding and analyzing system resilience where the y-axis is a measure of system performance. Initially, the system is at the original stable state with performance level $\varphi(t_0)$. Following the disruptive event e at time t_e, the system begins to degrade due to reduced or loss of functionality of components. The systemic performance decreases gradually due to cascading effects until the maximum loss is incurred at time t_d, after which the system enters the disrupted state. Starting at t_s, recovery activities begin to restore the network from its disrupted state, S_d. The recovery continues until time t_f when the system reaches a new stable state.

Systemic resilience can be calculated as the ratio of recovered performance to the maximum loss [44]. Let $R(t)$ denote the resilience at time t during the recovery process. $R(t)$ is computed as follows:

$$R(t) = \frac{\varphi(t|e) - \varphi(t_s|e)}{\varphi(t_0|e) - \varphi(t_s|e)} \tag{3}$$

In Eq. (3), $t \in (t_s, t_f)$; $R(t) \in [0, 1]$, with 1 indicating that the system has been fully recovered from the disruptive event to the original state.

3 Methodology

3.1 Estimation of Interdependency Between Networks. This study assumes that the interdependency links connect the demand node of one network to a supply node in another network. This assumption is realistic for many real-world ICIs. For example, in interdependent water and power network, the interdependency links are established as follows:

(1) The end-user node (demand node) in the water network and the power station (supply node) in the power network that requires water for generating steam.
(2) The end-user node (demand node) in the power network and the pumping station (supply node) in the water network that requires electric power [11].

In order to model the uncertainty of interdependencies across ICIs, interdependency links are estimated probabilistically. The model shown in Eq. (4) is proposed to evaluate the probability an interdependency link exists based on a set of nodal attributes. These attributes can represent physical, economic, and social characteristics of the networks. The attributes considered in this paper account for the distance between the networks, the number of customers served by the nodes, and the vulnerability of the customers represented with the social vulnerability index (SoVI). Distance is used as a predictor as distance-based features are found to be significant in estimating the missing link [46]. The model includes the number of customers (modeled using the population) since disrupted components serving a large number of customers might be given higher priority during the restoration. The third predictor, SoVI, was originally developed to identify the characteristics of the population that render social communities vulnerable to external disturbances [47]. It is calculated based on a large number of factors, including socioeconomic status, age, house type, education level, race, among others. SoVI values range from 0 to 1 with higher values indicating a higher level of vulnerability. Social vulnerability can inform underlying and intangible interdependency links that would guide restoration activities to achieve community resilience by prioritizing vulnerable customers who might disproportionately suffer more damage from disruptions [48,49]

$$P(A_{ij} = 1) = \beta_0 + \beta_1 \times d_{v_i v_j}^{-1} + \beta_2 \times p_{v_{nd}} + \beta_3 \times s_{v_{nd}} \tag{4}$$

In Eq. (4), $A_{ij} = 1$ indicates the presence of an interdependency link between v_i and v_j (v_i and v_j belong to different individual networks). β_i $(i = 1, 2, 3)$ represents the regression coefficient, $p_{v_{nd}}$ represents the population served by the parent node of the

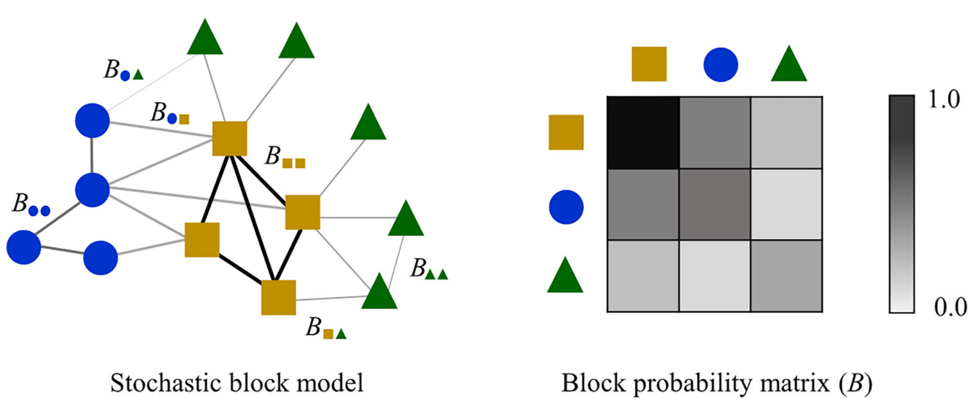

Fig. 3 An example of SBM with the corresponding block probability matrix

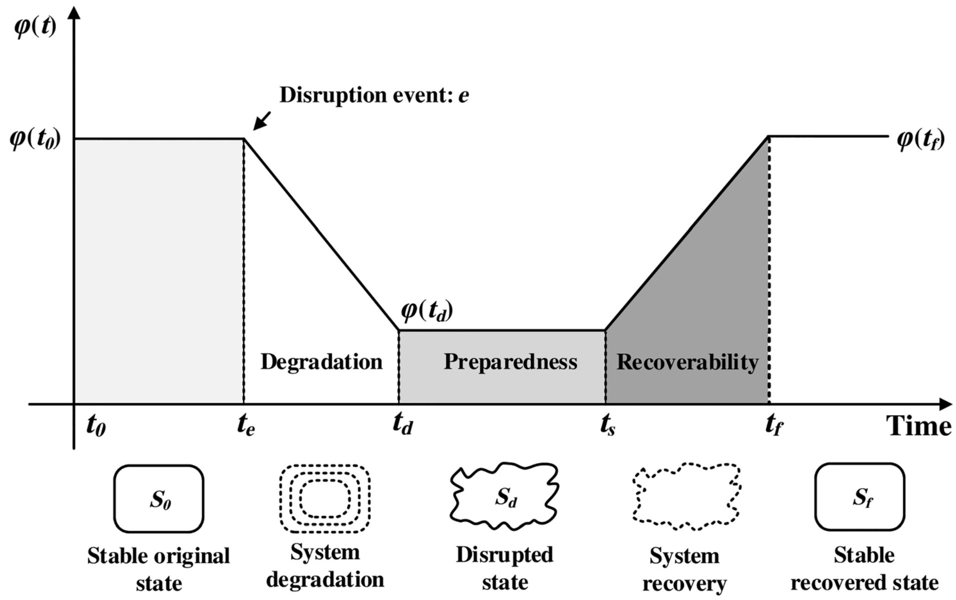

Fig. 4 System performance over time (adapted from Ref. [44])

interdependency link on which the dependent (child) node relies, and $s_{v_{nd}}$ represents the SoVI of the census tract where the parent node is located. Note that both $d_{v_i v_j}$ and $p_{v_{nd}}$ should be normalized before they are used in the model and the obtained probabilities should also be normalized such that they sum to one. The model can be conveniently refined when data about real-world interdependency links and nodal attributes are available.

Since the reciprocal of distance between nodes is used in Eq. (4), zero values must be avoided in the normalized distance. To this end, the min–max normalization is modified (Eq. (5)) to scale the distance data to the range $[\frac{\alpha}{1+\alpha}, 1]$. As $(\alpha/1+\alpha)$ should approach zero, a small value of α is preferred, i.e., $\alpha \ll 1$. Therefore, $(\alpha/1+\alpha) \approx \alpha$ indicates that the normalized data have an approximate lower bound equal to α

$$x' = \frac{x - x_{\min} + \alpha(x_{\max} - x_{\min})}{(1+\alpha)(x_{\max} - x_{\min})} \quad (5)$$

In Eq. (5), x represents the data to be normalized and x' represents data after normalization. This variant of min–max normalization can be referred to as truncated min–max normalization.

3.2 Component Importance Ranking and Restoration.

In this study, the sequence of infrastructure network restoration is determined according to the ranking of components. Resilience-based component importance ranking can help inform resource allocation and prioritization of repair activities when multiple components are damaged [50]. The importance is quantified by the relative resilience improvement of the interdependent networks after each component is restored individually. Two restoration sequences are proposed, the first one is based on *static ranking* and the second one is based on *dynamic ranking*. In the case of static ranking, the damaged components to be repaired are ranked only once before the start of recovery; therefore, the benchmark for resilience improvement is the resilience of the ICIs at the disrupted state $R(t_d)$. The importance of a component can be calculated using the following equation:

$$I^{c_i} = \frac{R^{c_i}(t_d) - R(t_d)}{R(t_d)} \quad (6)$$

In Eq. (6), I^{c_i} represents the importance of component i, $R^{c_i}(t_d)$ represents the resilience of the ICIs after component i is restored at time t_d before initiating recovery activities. The damaged components are then restored sequentially according to the ranking.

In the case of dynamic ranking, the damaged components are ranked at every time-step until all the components are restored. At the time t, the benchmark for resilience improvement is $R(t)$. Accordingly, the dynamic component importance can be computed using the following equation:

$$I^{c_i}(t) = \frac{R^{c_i}(t) - R(t)}{R(t)}, \; t \in (t_d, t_f) \quad (7)$$

In Eq. (7), $I^{c_i}(t)$ is the importance of component i at time t and $R^{c_i}(t)$ is the resilience of the ICIs after component i is restored at time t. When $t = t_d$, Eq. (7) becomes Eq. (6), meaning that static importance ranking is simply the initial dynamic importance ranking of all the damaged components. The steps for dynamic component importance ranking, coupled with component restoration, are summarized in Algorithm 1.

Algorithm 1 Dynamic component importance ranking for resilience assessment

Input: Adjacency matrix, node type, node coordinates, component failure probability, etc.
Output: The resilience over the restoration process $R(t)$, $t \in (t_d, t_f)$.
Compute the initial resilience $R(t_d)$.
for $t = 1$ to T **do** ▷ T: time needed to restore all the damaged components
  for $i = 1$ to N_t^r **do** ▷ N_t^r: # of remaining damaged components at time t
    Rank the remaining components based on the importance calculated by Eq. (7).
    Return the resilience after restoring the component with the highest importance, c_t.
  end for
  Remove the component c_t from the list of remaining components.
  Calculate the resilience according to Eq. (8) and record it as $R(t)$.
end for

3.3 Resilience Assessment of Interdependent Networks.

Network performance is determined by the ratio of the number of functional demand nodes to the total number of demand nodes of each network. The resilience of the interdependent networks is calculated as the weighted average of the resilience of individual networks

$$R(t) = \sum_{k=1}^{K} w_k \frac{n_t^k - n_d^k}{n_0^k - n_d^k},\ t \in (t_d, t_f) \tag{8}$$

In Eq. (8), K is the number of infrastructure networks. w_k is the weight for individual network k with $\sum_{k=1}^{K} = 1$. n_t^k, n_d^k, and n_0^k represent the number of functional demand nodes of network k at time t, t_d, and t_0, respectively.

As a network component in real-world ICIs can lose functionality due to (i) direct physical damage caused by the disruptive event or (ii) loss of necessary supply from the components of other networks, the following definitions are proposed to differentiate between the two cases.

Definition 2. *A network component is operable if the physical entity represented by the component is not damaged by the disruptive event.*

Definition 3. *A network component is functional if the component is operable and can receive supply from other components to maintain functionality.*

In a single network, we assume that an operable demand node is functional as long as it is connected to a functional supply node according to the two definitions, so the functionality of a demand node can be evaluated using shortest path algorithms, such as Dijkstra's algorithm [51], and Floyd-Warshall algorithm [52]. The Floyd-Warshall algorithm is used in this study because it can find the shortest path between all pairs of nodes simultaneously. However, in interdependent networks, the functional supply node in an individual network, for example, G_1, must also be connected to a demand node in the other network G_2, as shown in Fig. 1. The following theorem shows the conditions that must be satisfied for a node to be functional in interdependent networks (the proof is provided in the Appendix).

Theorem 1. *In undirected interdependent networks represented by the underlying graph G, a demand node v is functional if v is contained in an operable cycle that consists of at least one interdependency link or if v is connected to such a cycle.*

Based on this *sufficient* condition, a key step to determine the functionality of a node is to detect functional cycles in the interdependent networks. Building on the *depth first search* (DFS) [53] algorithm for detecting cycles in graphs and the Floyd-Warshall algorithm for checking connectivity, the steps for detecting the functional nodes in interdependent networks are provided in Algorithm 2.

Algorithm 2 Identification of functional nodes

Input: Adjacency matrix A of the interdependent networks
Output: An array that contains the identity (ID) of functional nodes
1 Apply DFS to detect cycles.
2: Select cycles that contain at least one interdependency link (The start node and end node belong to different individual networks).
3: Apply Floyd-Warshall algorithm to find the distance from all nodes to the nodes contained in the selected cycles.
4: If the distance value is greater than zero, then the corresponding node is functional; otherwise, the node is not functional.

Once the functional nodes are detected, the resilience of the interdependent networks at time t is obtained using Eq. (8). Considering the uncertainty of interdependency links and the failure of network components, the process for assessing the resilience of interdependent networks is proposed as follows:

(1) Import data on the interdependent networks, including the adjacency matrix, distance between nodes (normalized), population of the census tract where each node is located (normalized), and SoVI, among others.
(2) Estimate interdependency links based on Eq. (4).
(3) Draw samples of interdependency links according to their respective probabilities.
(4) Define disruption scenario and calculate the disruption intensity at the site of each network component.
(5) Calculate failure probability of each network component (including the estimated interdependency links) using empirical equations from HAZUS, a standardized methodology provided by the Federal Emergency Management Agency (FEMA) to estimate potential losses from multiple types of hazards [54]. For simplicity, components are assumed to be inoperable once they are damaged, i.e., partial functionality is not considered.
(6) Generate a sufficient number of network configurations, i.e., the possible network topology after the damage of a subset of components by the disruptive event e according to different probabilities. The randomness of the network configurations can be modeled by first comparing a random vector $\boldsymbol{u}$ ($\boldsymbol{u} \sim U(0, 1)$) to the vector containing the failure probability of each component, and then, if the failure probability of a certain component is greater than the random number drawn from $U(0, 1)$, the component is assumed to be inoperable.
(7) Generate new interdependency links if the links or nodes on the interdependency links are damaged. New interdependency links are considered to account for the interdependency that emerges in the aftermath of the disruption.
(8) Rank components based on the static ranking or dynamic ranking (Algorithm 1).
(9) Restore one component and record resilience at each time step t.

4 Case Study

4.1 Data. The system of two interdependent power and water networks shown in Fig. 5 is used to illustrate the proposed model in this study. The water distribution network includes six elevated storage tanks, nine pumping stations, 34 intermediate delivery nodes, and 71 water pipes while the power network (modified from Ref. [55]) consists of 14 gate stations, {23} 23-kV substations, and {22} 12-kV substations, respectively. Gate stations and pumping stations are considered to be the supply facilities, 23-kV substations, and storage tanks as transmission facilities, and 12-kV substations and intermediate delivery nodes as demand facilities. It should be noted that intermediate delivery nodes are the intersection points of water pipes and they are assumed to be undamaged after earthquakes since there do not exist large-scale facilities at the site of these nodes. In this study, physical interdependency is considered since pumping stations rely on the 12-kV substations for power supply [56] while the gate stations require clean water to generate high-pressure steam to drive the turbines. Note that the power stations may rely on the river nearby instead of the water distribution network for water to be used for cooling purposes [22].

Data on the population and SoVI at the census tract level where each node of the interdependent networks is located are publicly available through the Census Bureau [57] and Centers for Disease Control and Prevention [58], respectively.

4.2 Disruption Intensity and Component Fragility. The disruptive event considered in this case study is a hypothetical earthquake. An earthquake centered at $(35.3\,N,\ 90.3\,W)$ (the maximum probable earthquake [59]) is used to calculate the seismic intensity at the site of each component. In calculating the failure probability of each component given the earthquake scenario, the fragility curve of power and water network facilities under earthquakes is adapted from HAZUS. In HAZUS, five damage states are defined: none (ds_1), minor (ds_2), moderate (ds_3), extensive (ds_4), and complete (ds_5). Each damage state corresponds to one component fragility curve. This case study adopts the damage

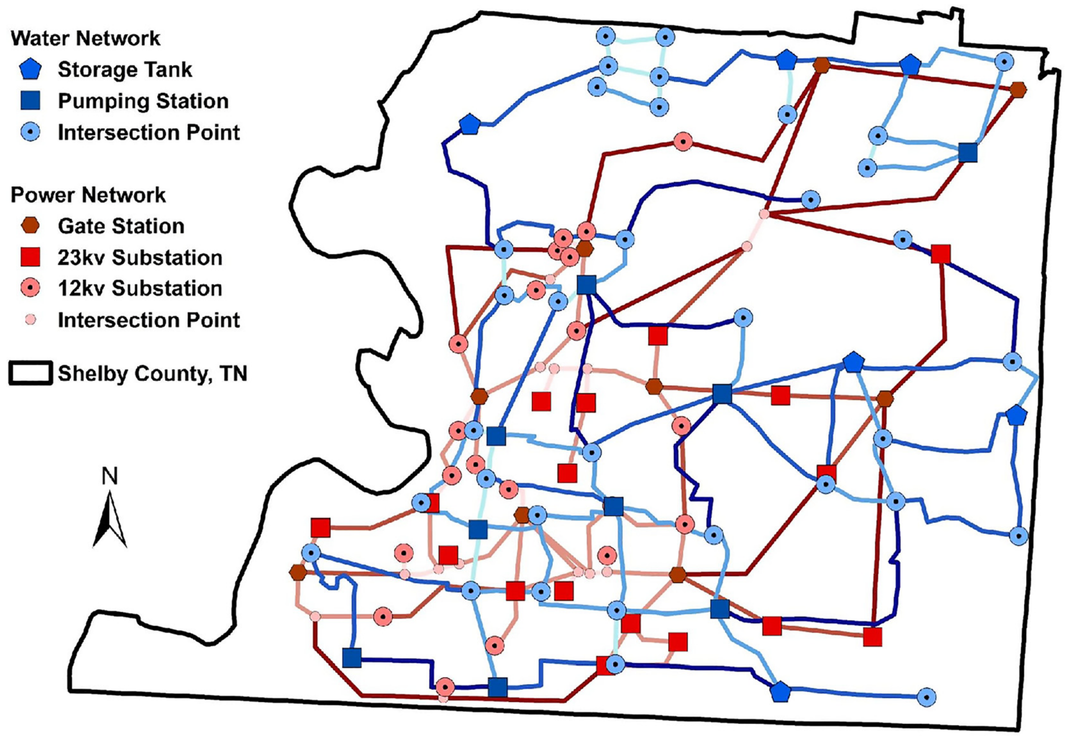

Fig. 5 Interdependent water and power networks of Shelby County, TN

state ds_5. More details on the determination of seismic intensities at the site of network components and the failure probability of components can be found in the Earthquake Model Technical Manual [60]. Components of the power network are assumed to respond in a similar way as water network to seismic events. This assumption can be relaxed when data about the fragility of power network components become available.

4.3 Interdependency Links. The estimated parameters of the probabilistic model for interdependency links from Eq. (4) are $\beta_0 = 0$, $\beta_1 = 0.5$, $\beta_2 = \beta_3 = 0.25$, $\alpha = 0.01$. For one particular node in each network, the probabilities of interdependency links between this node and components from the other network are shown in Figs. 6 and 7. Specifically, the likelihood of an interdependency link is described by the probability using the estimation model from Eq. (4) under two scenarios. The first scenario only considers the physical interdependence by including the distance as the sole predictor, while the second scenario incorporates social aspects such as population and SoVI besides the distance. It is noted that the interdependency links with the highest probability of occurrence are different under the two scenarios, especially in Fig. 7 where the two rankings of power nodes are entirely different. In addition, the distribution for these probabilities of all possible links becomes a bit flatter when social aspects are considered. The reason is that the ranking of nodes based on the population or SoVI is strikingly different from that based on the distance. The change in the probability of interdependency links after including the social attributes indicates that the social aspects do not differentiate between possible interdependency links like the geographical distance does. Collecting additional information allows for a more comprehensive model with additional predictors to identify the contributing factors to the existence of an interdependency link.

4.4 Resilience Assessment. Without loss of generality, this study assumes that the water network and power network weigh equally in calculating the resilience of the interdependent networks, thus $w^k = (1/2)$ with $K = 2$. Due to the probabilistic failure of network components, each of the possible network structures generated by each estimation of interdependency links can have a myriad of new network structures at time t_d after the disruption. To characterize the randomness in the network structure at time t_d, 100 simulation runs are used to obtain the mean value and the range of resilience for each potential scenario of interdependency links between the networks. During the recovery process, it is assumed that one component can be restored at each time-step, which can be modified to represent other possible restoration strategies. The resilience curves from the disrupted stage to the new stable stage under different seismic intensities (Fig. 8) show the response of different network structures to the same disruptive event.

Throughout the recovery process, the mean value, lower bound, and upper bound of resilience based on dynamic ranking are greater than or equal to those based on static ranking. Dynamic ranking yields a more rapid recovery process and improved resilience in the early stage of restoration. Further, the overall range under the static ranking approach is much larger than the dynamic ranking approach. This outcome suggests that accounting for the dynamic nature of interdependencies by updating the ranking of components at each time-step given the new configuration of the network is critical to improving the resilience of these systems.

5 Discussion

Interdependencies between infrastructure networks not only lead to cascading failures but also impact the recovery process. Failure to fully capture interdependencies can lead to inaccurate estimation of the resilience, which would mislead utility managers in making suboptimal restoration plans that would result in increased disruption duration and subsequently higher repair costs [29,61]. These interdependency links are uncertain and can change over time in response to disruption, restoration activities, and reconfiguration of infrastructure networks. As such, modeling the resilience of ICIs should account for the inherent uncertainty and dynamic behavior of interdependency links after a disruption. The outcome of this work is a data-driven stochastic method that estimates the likelihood that an interdependency link exists based

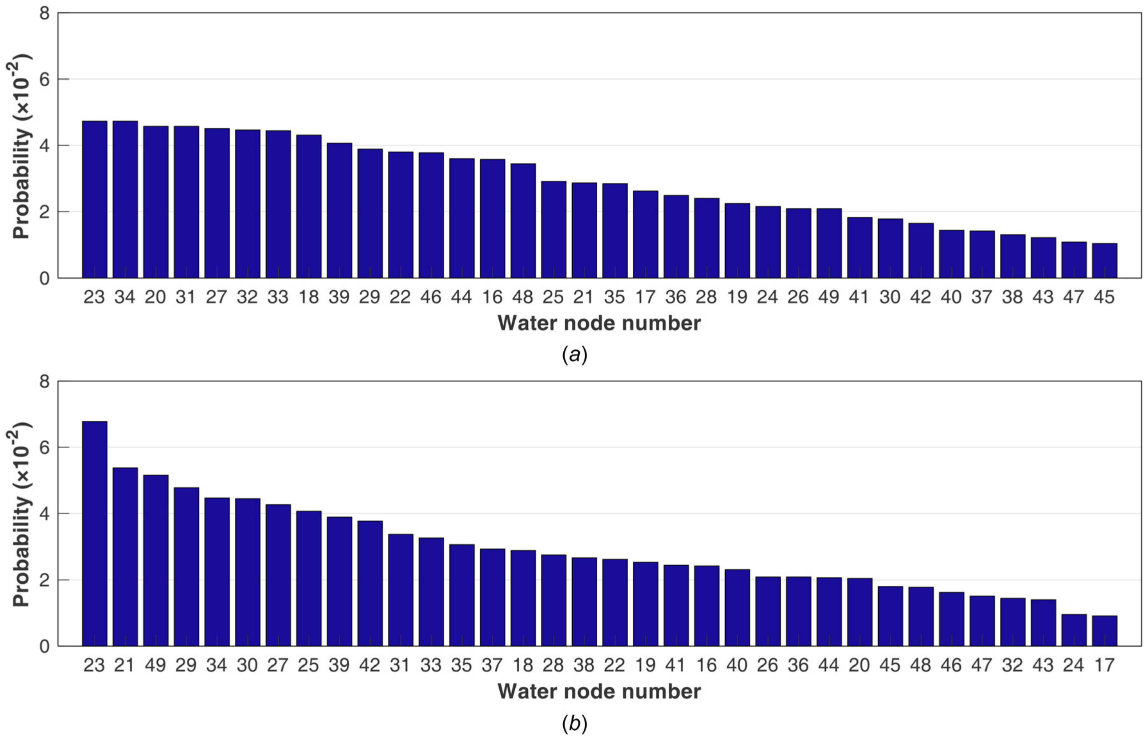

Fig. 6 Probability of interdependency links between water nodes and gate stations: (*a*) with population and SoVI and (*b*) without population and SoVI

on the behavior of the networks after a disruption. The calculation of the probability of an interdependency link incorporates multiple predictors describing geographic proximity, physical connection, and social impact. It is therefore possible to predict how interdependency links emerge and disappear during and after a disruption. Given this additional information on infrastructure network behavior, restoration strategies can be adapted in real time following a dynamic ranking process for the most important components. In addition, this approach provides the ability to identify factors that impact the existence and strength of an interdependency link. This work expands on existing studies of ICIs where only geographic and physical interdependence are considered [21–23]. For example, the incorporation of SoVI in the estimation of interdependency links and consequent dynamic ranking of components significantly improves the resilience of ICIs by decreasing the restoration time after a disruption. The restoration

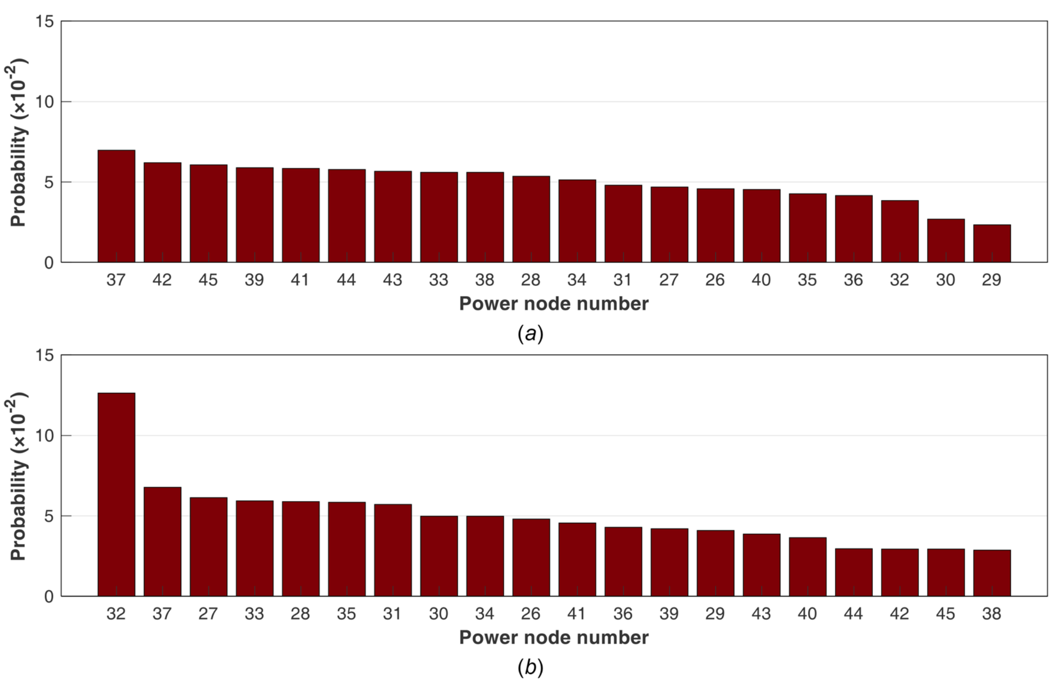

Fig. 7 Probability of interdependency links between power nodes and pumping stations: (*a*) with population and SoVI and (*b*) without population and SoVI

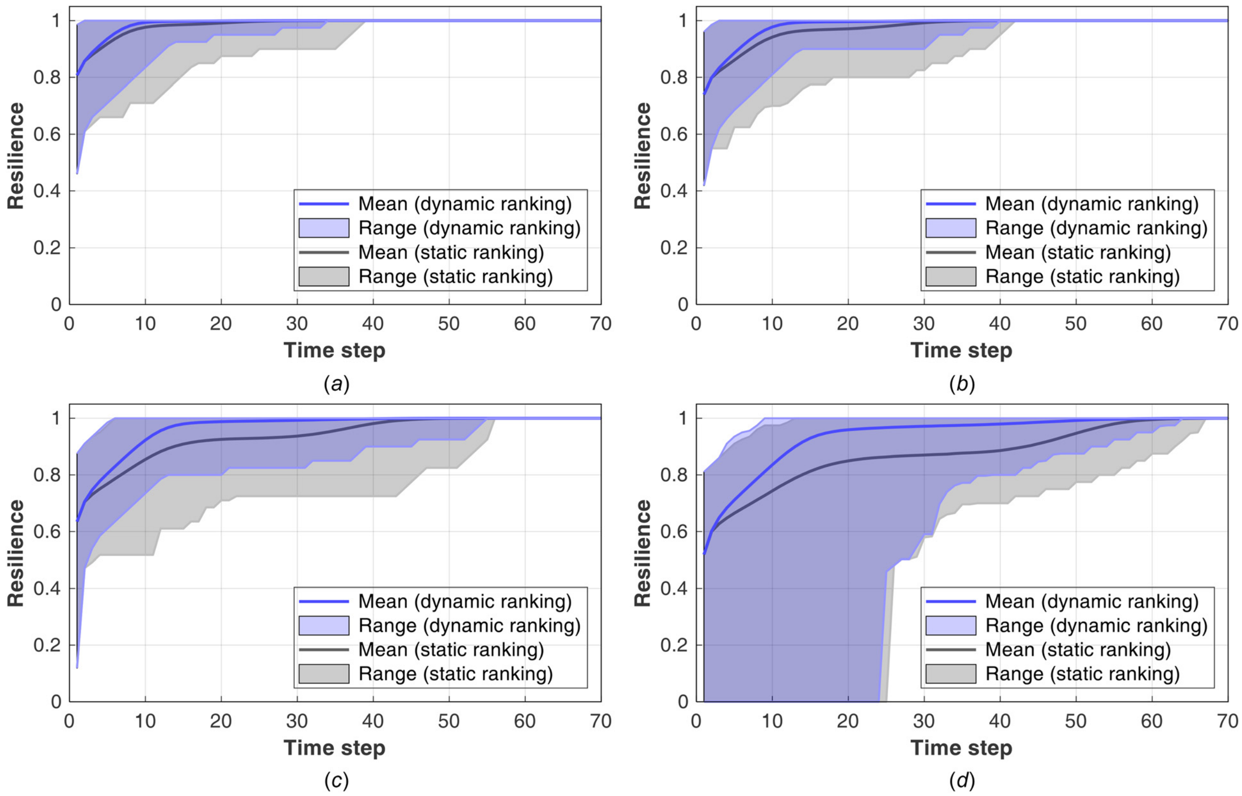

Fig. 8 Resilience curves with lower/upper bounds under different seismic intensities: (*a*) magnitude = 5.0, (*b*) magnitude = 6.0, (*c*) magnitude = 7.0, and (*d*) magnitude = 8.0

activities in this case are informed by the updated prioritization of components based on the needs of the communities measured here by SoVI.

The approach presented in this paper is flexible that it can be generalized to other types of ICIs and the probability of interdependency links can incorporate additional and different factors and variables. For example, in order to capture other sources of uncertainty such as the success of the restoration or secondary failures, the probability of an interdependency link or the reconfiguration of the networks can be dependent on stochastic variables describing these phenomena.

A limitation of this approach is that topology-based metrics for resilience provide a basic understanding of the dynamics of ICIs after disruptive events. Future work will consider building a full network flow model for a more accurate representation of the performance and behavior of ICI after a disruption under uncertainty. Additionally, incorporating more nodal attributes in the probabilistic model for estimating interdependency links will help with the interpretation of the model.

6 Conclusion

This paper presents a new approach founded in stochastic block models to capture the uncertainty associated with interdependency links of ICIs. A case study using real-world data evaluates the interdependent water and power networks of Shelby County, TN. The results show that estimation of the presence of interdependency links based on distance between network components can be significantly different from the estimation based on distance, population, and SoVI. This work demonstrates the importance of evaluating uncertain interdependency links by comparing two restoration strategies, one based on dynamic ranking given updates on the dynamic behavior of interdependencies and the other is static. The recovery based on dynamic component importance ranking results in faster restoration and improved resilience.

Acknowledgment

The authors thank Professor Songling Shan from the Department of Mathematics at Illinois State University for her help with improving the proof of Theorem 1.

Funding Data

- National Science Foundation (Grant No. 1944559).

Appendix: Proof of the Theorem

Definition 4 (Functional path and functional cycle). *A path or cycle is operable if all the nodes and edges they contain are operable.*

Proof of Theorem 1. For the sake of contradiction, assume that a demand node v_x of G_1 can be functional even if (a) no operable cycles that contain interdependency links exist in G, or (b) the

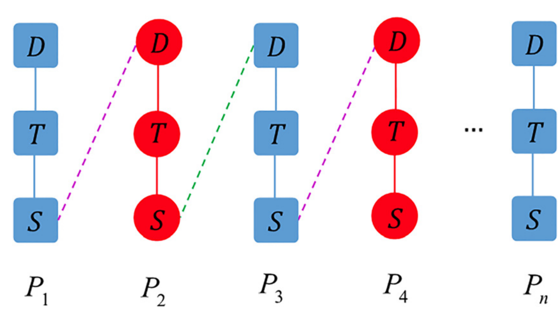

Fig. 9 Assumed functional longest path in interdependent networks without cycles

demand node is neither a part of nor connected to such a cycle. In case (a), since v_x is functional, there must exist an operable path from v_x to a supply node in G_1. Let P be the longest functional path that starts from v_x. Since v_x is functional, the supply node closest to v_x in G_1 must be adjacent to a functional demand node in G_2. Following this logic, P must be made up of D–T–S paths that connect a demand node, transmission node, and supply node by sequence, such as $P_1, P_2, \ldots, P_n$ in Fig. 9. Note that by definition, repeated nodes are not allowed in P otherwise a cycle will be complete. Let v_y be the endpoint of the last D–T–S path in P. Because v_y is the endpoint, it does not receive supply from the demand node of another network, so v_y is not functional, contradicting that P is functional. In case (b), suppose a functional demand v_x is connected to some supply node v_z in G_1 and v_z is not part of any cycle that consists of an interdependency link. Using the same logic of case (a), v_z will not be functional, leading to a contradiction. ■

References

[1] Pederson, P., Dudenhoeffer, D., Hartley, S., and Permann, M., 2006, "Critical Infrastructure Interdependency Modeling: A Survey of U.S. and International Research," Idaho National Laboratory, Idaho Falls, ID, Report No. INL/EXT-06-11464, p. 27.

[2] Vespignani, A., 2010, "Complex Networks: The Fragility of Interdependency," Nature, **464**(7291), p. 984.

[3] Bashan, A., Berezin, Y., Buldyrev, S. V., and Havlin, S., 2013, "The Extreme Vulnerability of Interdependent Spatially Embedded Networks," Nat. Phys., **9**(10), pp. 667–672.

[4] Danziger, M. M., Shekhtman, L. M., Bashan, A., Berezin, Y., and Havlin, S., 2016, "Vulnerability of Interdependent Networks and Networks of Networks," *Interconnected Networks*, Springer, Cham, Switzerland, pp. 79–99.

[5] Rinaldi, S. M., 2004, "Modeling and Simulating Critical Infrastructures and Their Interdependencies," Proceedings of the 37th Annual Hawaii International Conference on System Sciences, Big Island, HI, Jan. 5–8, p. 8.

[6] Rinaldi, S. M., Peerenboom, J. P., and Kelly, T. K., 2001, "Identifying, Understanding, and Analyzing Critical Infrastructure Interdependencies," IEEE Control Syst., **21**(6), pp. 11–25.

[7] Zimmerman, R., 2001, "Social Implications of Infrastructure Network Interactions," J. Urban Technol., **8**(3), pp. 97–119.

[8] Dudenhoeffer, D. D., Permann, M. R., and Manic, M., 2006, "CIMS: A Framework for Infrastructure Interdependency Modeling and Analysis," Proceedings of the 38th Conference on Winter Simulation, Winter Simulation Conference, Monterey, CA, Dec. 3–6, pp. 478–485.

[9] Lee Ii, E. E., Mitchell, J. E., and Wallace, W. A., 2007, "Restoration of Services in Interdependent Infrastructure Systems: A Network Flows Approach," IEEE Trans. Syst., Man, Cybern., Part C, **37**(6), pp. 1303–1317.

[10] Goldbeck, N., Angeloudis, P., and Ochieng, W. Y., 2019, "Resilience Assessment for Interdependent Urban Infrastructure Systems Using Dynamic Network Flow Models," Reliab. Eng. Syst. Saf., **188**, pp. 62–79.

[11] González, A. D., Dueñas-Osorio, L., Sánchez-Silva, M., and Medaglia, A. L., 2016, "The Interdependent Network Design Problem for Optimal Infrastructure System Restoration," Comput.-Aided Civ. Infrastruct. Eng., **31**(5), pp. 334–350.

[12] Ouyang, M., 2014, "Review on Modeling and Simulation of Interdependent Critical Infrastructure Systems," Reliab. Eng. Syst. Saf., **121**, pp. 43–60.

[13] Dueñas-Osorio, L., Craig, J. I., Goodno, B. J., and Bostrom, A., 2007, "Interdependent Response of Networked Systems," J. Infrastruct. Syst., **13**(3), pp. 185–194.

[14] Hernandez-Fajardo, I., and Dueñas-Osorio, L., 2013, "Probabilistic Study of Cascading Failures in Complex Interdependent Lifeline Systems," Reliab. Eng. Syst. Saf., **111**, pp. 260–272.

[15] Ouyang, M., and Wang, Z., 2015, "Resilience Assessment of Interdependent Infrastructure Systems: With a Focus on Joint Restoration Modeling and Analysis," Reliab. Eng. Syst. Saf., **141**, pp. 74–82.

[16] Chai, W. K., Kyritsis, V., Katsaros, K. V., and Pavlou, G., 2016, "Resilience of Interdependent Communication and Power Distribution Networks Against Cascading Failures," IFIP Networking Conference (IFIP Networking) and Workshops, Vienna, Austria, May 17–19, pp. 37–45.

[17] Guidotti, R., Chmielewski, H., Unnikrishnan, V., Gardoni, P., McAllister, T., and van de Lindt, J., 2016, "Modeling the Resilience of Critical Infrastructure: The Role of Network Dependencies," Sustainable Resilient Infrastruct., **1**(3–4), pp. 153–168.

[18] Liu, X., Ferrario, E., and Zio, E., 2017, "Resilience Analysis Framework for Interconnected Critical Infrastructures," ASCE-ASME J. Risk Uncertainty Eng. Syst., Part B, **3**(2), p. 021001.

[19] Sharma, N., Tabandeh, A., and Gardoni, P., 2019, "Recovery Optimization of Interdependent Infrastructure: A Multi-Scale Approach," 13th International Conference on Applications of Statistics and Probability in Civil Engineering, Seoul, South Korea, May 26–30.

[20] Di Muro, M., La Rocca, C., Stanley, H., Havlin, S., and Braunstein, L., 2016, "Recovery of Interdependent Networks," Sci. Rep., **6**(1), p. 22834.

[21] Ouyang, M., Hong, L., Mao, Z.-J., Yu, M.-H., and Qi, F., 2009, "A Methodological Approach to Analyze Vulnerability of Interdependent Infrastructures," Simul. Modell. Pract. Theory, **17**(5), pp. 817–828.

[22] Zhang, Y., Yang, N., and Lall, U., 2016, "Modeling and Simulation of the Vulnerability of Interdependent Power-Water Infrastructure Networks to Cascading Failures," J. Syst. Sci. Syst. Eng., **25**(1), pp. 102–118.

[23] Almoghathawi, Y., Barker, K., and Albert, L. A., 2019, "Resilience-Driven Restoration Model for Interdependent Infrastructure Networks," Reliab. Eng. Syst. Saf., **185**, pp. 12–23.

[24] Garvey, M. D., Carnovale, S., and Yeniyurt, S., 2015, "An Analytical Framework for Supply Network Risk Propagation: A Bayesian Network Approach," Eur. J. Oper. Res., **243**(2), pp. 618–627.

[25] Yodo, N., and Wang, P., 2016, "Resilience Modeling and Quantification for Engineered Systems Using Bayesian Networks," ASME J. Mech. Des., **138**(3), p. 031404.

[26] Hosseini, S., and Barker, K., 2016, "Modeling Infrastructure Resilience Using Bayesian Networks: A Case Study of Inland Waterway Ports," Comput. Ind. Eng., **93**, pp. 252–266.

[27] Chen, C., Tong, H., Xie, L., Ying, L., and He, Q., 2016, "Fascinate: Fast Cross-Layer Dependency Inference on Multi-Layered Networks," Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, San Francisco, CA, pp. 765–774.

[28] Der Kiureghian, A., and Ditlevsen, O., 2009, "Aleatory or Epistemic? Does It Matter?," Struct. Saf., **31**(2), pp. 105–112.

[29] Yu, J.-Z., and Baroud, H., 2019, "Quantifying Community Resilience Using Hierarchical Bayesian Kernel Methods: A Case Study on Recovery From Power Outages," Risk Anal., **39**(9), pp. 1930–1948.

[30] Anderson, C. J., Wasserman, S., and Faust, K., 1992, "Building Stochastic Blockmodels," Soc. Networks, **14**(1–2), pp. 137–161.

[31] Holland, P. W., Laskey, K. B., and Leinhardt, S., 1983, "Stochastic Blockmodels: First Steps," Soc. Networks, **5**(2), pp. 109–137.

[32] Žiberna, A., 2014, "Blockmodeling of Multilevel Networks," Soc. Networks, **39**, pp. 46–61.

[33] Barbillon, P., Donnet, S., Lazega, E., and Bar-Hen, A., 2017, "Stochastic Block Models for Multiplex Networks: An Application to a Multilevel Network of Researchers," J. R. Stat. Soc.: Ser. A, **180**(1), pp. 295–314.

[34] Abbe, E., 2018, "Community Detection and Stochastic Block Models: Recent Developments," J. Mach. Learn. Res., **18**(177), pp. 1–86.

[35] Xu, K., 2015, "Stochastic Block Transition Models for Dynamic Networks," 18th International Conference on Artificial Intelligence and Statistics, San Diego, CA, May 9–12, pp. 1079–1087.

[36] Kolaczyk, E. D., 2017, *Topics at the Frontier of Statistics and Network Analysis:(Re) Visiting the Foundations*, Cambridge University Press, Cambridge, UK.

[37] Mahoney, M. W., 2016, "Lecture Notes in Spectral Graph Methods," preprint arXiv:1608.04845.

[38] Holling, C. S., 1973, "Resilience and Stability of Ecological Systems," Annu. Rev. Ecol. Syst., **4**(1), pp. 1–23.

[39] Ayyub, B. M., 2015, "Practical Resilience Metrics for Planning, Design, and Decision Making," ASCE-ASME J. Risk Uncertainty Eng. Syst., Part A, **1**(3), p. 04015008.

[40] National Infrastructure Advisory Council, 2009, "Critical Infrastructure Resilience: Final Report and Recommendations," National Infrastructure Advisory Council, U. S. Department of Homeland Security, Washington, DC, Report.

[41] National Research Council, 2012, *Disaster Resilience: A National Imperative*, National Academies Press, Washington, DC.

[42] Presidential Policy Directive, 2013, *Critical Infrastructure Security and Resilience*, White House, Washington, DC.

[43] Haimes, Y. Y., 2009, "On the Definition of Resilience in Systems," Risk Anal.: An Int. J., **29**(4), pp. 498–501.

[44] Henry, D., and Ramirez-Marquez, J. E., 2012, "Generic Metrics and Quantitative Approaches for System Resilience as a Function of Time," Reliab. Eng. Syst. Saf., **99**, pp. 114–122.

[45] Zhang, X., Mahadevan, S., Sankararaman, S., and Goebel, K., 2018, "Resilience-Based Network Design Under Uncertainty," Reliab. Eng. Syst. Saf., **169**, pp. 364–379.

[46] Hasan, A., Zaki, M., and J, M., 2011, "A Survey of Link Prediction in Social Networks," *Social Network Data Analytics*, Springer, Boston, MA, pp. 243–275.

[47] Cutter, S. L., Boruff, B. J., and Shirley, W. L., 2003, "Social Vulnerability to Environmental Hazards," Soc. Sci. Q., **84**(2), pp. 242–261.

[48] Highfield, W. E., Peacock, W. G., and Van Zandt, S., 2014, "Mitigation Planning: Why Hazard Exposure, Structural Vulnerability, and Social Vulnerability Matter," J. Plann. Educ. Res., **34**(3), pp. 287–300.

[49] Peacock, W. G., Van Zandt, S., Zhang, Y., and Highfield, W. E., 2014, "Inequities in Long-Term Housing Recovery After Disasters," J. Am. Plann. Assoc., **80**(4), pp. 356–371.

[50] Baroud, H., and Barker, K., 2018, "A Bayesian Kernel Approach to Modeling Resilience-Based Network Component Importance," Reliab. Eng. Syst. Saf., **170**, pp. 10–19.

[51] Dijkstra, E. W., 1959, "A Note on Two Problems in Connexion With Graphs," Numer. Math., **1**(1), pp. 269–271.

[52] Floyd, R. W., 1962, "Algorithm 97: Shortest Path," Commun. ACM, **5**(6), p. 345.

[53] Tarjan, R., 1972, "Depth-First Search and Linear Graph Algorithms," SIAM J. Comput., **1**(2), pp. 146–160.

[54] FEMA, 2013, *Multi-Hazard Loss Estimation Methodology: Earthquake Model*, Federal Emergency Management Agency, Washington, DC.

[55] Shinozuka, M., Rose, A., and Eguchi, R., 1998, "Engineering and Socioeconomic Impact of Earthquakes: An Analysis of Electricity Lifeline Disruptions in the New Madrid Area Monograph 2," Multidisciplinary Center for Earthquake Engineering Research, Red Jacket Quadrangle, State University of New York at Buffalo, Buffalo, NY.
[56] González, A. D., Chapman, A., Dueñas-Osorio, L., Mesbahi, M., and D'Souza, R. M., 2017, "Efficient Infrastructure Restoration Strategies Using the Recovery Operator," Comput.-Aided Civ. Infrastruct. Eng., **32**(12), pp. 991–1006.
[57] Census Bureau, 2018, "Census Data," Department of Commerce, Census Bureau, Washington, DC, accessed Mar. 2, 2020, https://www.census.gov/quickfacts/fact/table/US/PST045219
[58] CDC, 2016, "Social Vulnerability Index (SVI)," Department of Health and Human Services, Washington, DC, accessed Mar. 2, 2019, https://svi.cdc.gov/data-and-tools-download.html
[59] Harmsen, S., Frankel, A., and Petersen, M., 2003, "Deaggregation of US Seismic Hazard Sources: The 2002 Update," U.S. Geological Survey, Reston, VA, Report No. 2003-03-440.
[60] FEMA, 2013, "HAZUS-MH 2.1 User Manual: Earthquake Model Technical Manual," Federal Emergency Management Agency, Washington, DC.
[61] Zhang, D-M., Du, F., Huang, H., Zhang, F., Ayyub, B. M., and Beer, M., 2018, "Resiliency Assessment of Urban Rail Transit Networks: Shanghai Metro as an Example," Saf. Sci., **106**, pp. 230–243.