Stochastic Measures of Network Resilience: Applications to Waterway Commodity Flows

Hiba Baroud,¹ Jose E. Ramirez-Marquez,² Kash Barker,¹,* and Claudio M. Rocco³

Given the ubiquitous nature of infrastructure networks in today’s society, there is a global need to understand, quantify, and plan for the resilience of these networks to disruptions. This work defines network resilience along dimensions of reliability, vulnerability, survivability, and recoverability, and quantifies network resilience as a function of component and network performance. The treatment of vulnerability and recoverability as random variables leads to stochastic measures of resilience, including time to total system restoration, time to full system service resilience, and time to a specific $\alpha$% resilience. Ultimately, a means to optimize network resilience strategies is discussed, primarily through an adaption of the Copeland Score for nonparametric stochastic ranking. The measures of resilience and optimization techniques are applied to inland waterway networks, an important mode in the larger multimodal transportation network upon which we rely for the flow of commodities. We provide a case study analyzing and planning for the resilience of commodity flows along the Mississippi River Navigation System to illustrate the usefulness of the proposed metrics.

KEY WORDS: Copeland Score; infrastructure; networks; resilience; stochastic ranking

NOMENCLATURE

$S_0$, original as-planned state of the network
$S_d$, disrupted state of the network
$S_r$, recovered state of the network
$e^j$, disruptive event befalling the network
$D$, set of $J$ possible disruptive events
$t_e$, time at which disruption $e^j$ occurs
$t_d$, time at which the network reaches its disrupted state
$t_s$, time at which recovery commences
$t_f$, time at which recovered state is reached

$\psi(t)$, system service function
$\eta(t_e | e^j)$, resilience of the network resulting from $e^j$ at time $t_e \in (t_s, t_f)$
$x_i(t)$, state variable describing the performance of the $i$th component at time $t$
$V^j_i(e^j) = V^j$, proportional reduction in the performance of the $i$th component due to $e^j$
$U^j_i(V^j_i(e^j)) = U^j_i(V^j_i)$, recovery time for the $i$th component due to $e^j$ and $V^j_i$
$s(e^j)$, recovery strategy following $e^j$
$\omega^j_i(e^j)$, order in which the recovery activity for the $i$th component is accomplished
$A_h$, set of strategies all occurring in parallel at order $h$
$W^j(e^j)$, overall recovery sets for all parallel sets
$T^j(e^j)$, stochastic time to total network restoration metric
$T_{e'i}(e'i)$, stochastic time to full network service resilience

$T_{e'i}(e'i)$, stochastic time to $\alpha \times 100\%$ resilience

1. INTRODUCTION AND MOTIVATION

Given the ubiquitous nature of infrastructure networks in today’s society, and given the need to plan for their sustained performance (e.g., the American Society for Civil Engineering report card describing U.S. infrastructure with grade of “D”), there is a global need to understand the ability of these networks to cope with disruptions (i.e., resilience). Such disruptions are considered to be inevitable, representing a shift in planning emphasis from prevention to resilience and recovery. While recent work has explored the topic of resilience and while the use of the term “resilience” has increased substantially in the literature in recent years, no standard definition or quantitative technique addressing this need has emerged. Generally, discussions of infrastructure resilience are largely qualitative and descriptive in nature with few usable metrics. Network-based models have been used to quantify resilience for interconnected transportation systems and communications systems, but they tend to be infrastructure- and network-specific. While different viewpoints about resilience exist within the research community, there is consensus that understanding, quantifying, and planning for system resilience is vital. Recent risk-based studies of infrastructure networks include those for electric power, highway transportation, and rail transportation.

Work presented here establishes a new paradigm for analyzing the resilience of networks, extending previous general system resilience modeling approaches of Whitson and Ramirez-Marquez; Henry and Ramirez; Barker et al.; Pant et al. This paradigm for describing resilience is introduced with the state transition described in Fig. 1, which suggests that a system can operate in three distinct states: (i) its original, as-planned state, $S_0$, (ii) its disrupted state, $S_d$, caused from a disruption to the system whether by attack, disaster, accident, or common failure, and (iii) its recovered state, $S_f$, that results from a recovery effort. The system operates in $S_0$ until a disruption occurs at $t_0$, and at time $t_d$, the system reaches its “maximum” disrupted state $S_d$. Recovery from the disruption commences at time $t_i$, and state $S_f$ is attained at time $t_i$ and is maintained thereafter, the recovered state is not necessarily the same as the initial state; more details on this are discussed below. Resilience is defined here in a general way as the time-dependent ratio of recovery over loss, or $\mathcal{R}(t) = \text{Recovery/Loss}(t_d)$. Note the notation for resilience, $\mathcal{R}$, as $R$ is historically reserved for quantifying reliability.

System performance is overlaid with the system state transition as in Fig. 2, providing a general approach for understanding and analyzing system resilience. In this figure, an increasing system service function $\psi(t)$ (i.e., where increasing values of $\psi(t)$ are preferable) describes the behavior or performance of the system at time $t$ (e.g., $\psi(t)$ could describe traffic flow for a highway network, throughout for a manufacturing facility). Resilience in the network at time $t$ is exhibited if and only if there is an external disruptive event, $e'$, that affects the original system state, $S_0$, in which state performance is measured as $\psi(t_0)$. Disrupted at time $t_d$, a period of degradation of length $(t_d - t_i)$ transitions the system to $S_d$ with corresponding performance $\psi(t_d)$. After a period of time of length $(t_e - t_d)$, the system restoration commences until the system reaches a stable recovered state, $S_f$, with corresponding performance described by $\psi(t_f)$. Note that state $S_f$ need not be the same as $S_0$, as the new state may reach an alternative ($\psi(t_f)$ may be lower, or perhaps higher) equilibrium level (e.g., the state of infrastructure following the 2010 earthquake in Haiti may be improved over predisruption levels, as part of the recovery activities aiming at helping the infrastructure system regain its functionality might at the same time be helping in improving the system). Set $D = \{e' | 1 \leq j \leq J\}$ describes the set of possible disruptive events that could befall the system at time $t_c$. Note that the actual identification of these events is outside of the scope of this study, but such events can be identified according to intelligence and expert evidence. Fig. 3 illustrates the resilience process but for decreasing service function (i.e., where decreasing values of $\psi(t)$ are preferable). Note that although the initial and final states are represented by rectangles of equal size in Figs. 1–3, this does not necessarily mean that they are the same state: the square merely symbolizes that the functionality state of the system is no longer disrupted.

To illustrate implementations of these two representations of resilience, Fig. 4 introduces two examples describing the resilience process as described...
by Figs. 2 and 3. Fig. 4(a) describes the recovery of employment for six different historic U.S. recessions, where the service function $\varphi(t)$ describes the percentage change in jobs held relative to the peak just prior to the recession (describing an increasing recovery as in Fig. 2). Fig. 4(b) illustrates the loss and restoration of electric power to households in southern New Jersey following Hurricane Sandy (illustrating recovery that is described by a decreasing service function $\varphi(t)$ as in Fig. 3). The recovery trend in those graphs is impacted by certain recovery activities that are helping in creating jobs in a case of recession or repairing the damaged power grid to restore the electric power in households.

Based on this description, system resilience is provided in Equation (1). Equation (1) shares similarities with the metric described by Ta et al., who, for freight transportation systems, stated that “the resilience of an intermodal component, $\alpha$, can be expressed as the postdisruption fraction of demand that can be satisfied by using specific resources while maintaining a prescribed level of service.” However, such metric is not time-dependent. The purpose of both the illustrative description of Fig. 2 and the corresponding mathematical formulation of Equation (1) is to convey a systems perspective to the general definition of resilience, or “the [time-dependent] capability of a strained body to recover its size and shape after deformation caused especially by compressive stress.” Furthermore, Fig. 2 highlights four areas of research that follow from the process described. Note that, generally, resilience
Fig. 3. Alternative depiction of state transitions describing a decreasing system service function, $\psi(t)$.

Fig. 4. Recent examples of the graphical depiction of resilience, including (a) percentage job changes (relative to the previous peak) during six historical U.S. recessions, and (b) the loss (and subsequent recovery) of power to households in southern New Jersey following Hurricane Sandy in 2012.

takes values between 0 and 1, with $\mathcal{J}_\psi(t_r|e') = 1$ implying a full resilient network. However, in cases where the final state is improved relative to the initial state, the resilience is greater than 1.

$$\mathcal{J}_\psi(t_r|e') = \frac{\psi(t_r|e') - \psi(t_d|e')}{\psi(t_0) - \psi(t_d|e')}, \quad t_r \in (t_s, t_f).$$

(1)

1.1. Reliability

In the absence of an external disruptive event, the period of time $(t_c - t_0)$ corresponds to the system time to failure, where at time $t_c$, a failure event occurs. Whenever these failure events are of a stochastic nature, research in reliability theory provides methods and techniques to model, analyze, and
optimize system behavior. Particularly for networks, reliability indices often include the probability of network connectivity and probability of source-sink connectivity. Network reliability computation has been found to be an NP-hard problem, with several assessment approaches proposed: factoring algorithms, minimal paths/cuts enumeration, and approximation techniques.

1.2. Vulnerability

Based on current vulnerability research, the interest in this area is to develop approaches that describe the interaction between a disruptive event and subsequent system performance to quantify the degradation of specific system performance/service functions, including human performance. By understanding such effects, one can then identify the system elements that when disrupted generate the highest damage. These elements are known as points of system vulnerability, examples of which include (i) identifying elements responsible for the maximum reduction of telecommunication performance for backbone Internet networks, (ii) understanding transportation network damage as a function of road link-loss, car accident, roadwork, and other jamming disruptions, and (iii) understanding the interaction of multiple points of vulnerability with multiobjective optimization. Further research in the area of network interdiction has extensively contributed to the analysis of vulnerability in a variety of network applications.

1.3. Survivability

To understand the effects of external events on the performance of the system, research in survivability focuses on developing approaches that can help the system become “robust to partially successful attack through general architecture features, through adaptability (ability to respond to unanticipated changes) and flexibility (ability to adapt to a range of adverse events without having to anticipate the particular response in advance).” Broadly, survivability focuses on techniques that maintain system service continuity so that potential disruptive events are no longer so. For example, Zolfaghari and Kaudel describe a framework for ensuring that networks maintain service despite failures, while Kalyoncu and Sankur estimate survivability as the ability of a telecommunications network to maintain service in the presence of failures. Unfortunately, in engineering-based applications, the concepts of resilience and survivability are often interchangeably used.

1.4. Recoverability

Alluded to previously, research in this area is related to understanding the ability of systems to recover after a disruptive damaging event. For example, Rose describes dynamic recoverability as related to “the speed at which an entity or system recovers from a severe shock to achieve a desired state.” While there are many studies related to this area (e.g., socioecological resilience, sociotechnical resilience), most work is oriented toward lessons learned and management results that are generally unquantifiable. Moreover, except for several analyses of intermodal freight systems, there is a void in research related to the stochastic behavior of recovery in general infrastructure networks.

To begin to address these current research gaps, the purpose, and main contribution, of this article is to extend the deterministic resilience metrics discussed in Ref. 19 into a stochastic context and to mathematically describe several new metrics in a network setting: (i) the time required for a network to return to a specific network state (with special interest in the original state, $S_0$) and (ii) the probability that resilience at time $t_r$ is greater than a target value $\alpha$ (with specific interest on $\alpha = 1$). After initial experimentation with these metrics, an inherent stochastic optimization problem related to resilience planning presents itself: identifying the optimal configuration for a known number of recovery activity sets that minimize the time to full network service resilience. This article examines a new heuristic approach to solve these optimization problems.

The remainder of the article is presented as follows: Section 2 shows the network description and background necessary to define resilience and cost-related metrics. Section 3 introduces two stochastic order optimization problems inherent to the metrics described in Section 2, as well as provides a heuristic for approximating the solution of such problems. Section 4 illustrates the metrics and optimization approach with an inland waterway network example, with Section 5 providing conclusions along with future areas of research.
2. MODELING NETWORK RESILIENCE

The rationale behind the description of the resilience process in Fig. 2 is to highlight that there is a need to provide metrics that allow for understanding and quantifying the effect of external disruptive events on critical infrastructure networks and their corresponding recovery activities from a stochastic perspective. This section introduces such metrics.

2.1. Defining the Network, Disruptive Events, and Recovery Activities

Let $G = (N, A)$ represent a network where $N$ is the set of nodes, and $A = \{i\} | 1 \leq i \leq m \}$ is the set of arcs. For link $i$ at time $t$, $x_i(t)$ is the state variable describing the performance of the link, possibly valuating different entities, including capacity, delay, or length, among others. The network state vector at time $t$, $\mathbf{x}(t) = (x_1(t), x_2(t), \ldots, x_m(t))$, denotes the state of all the links at time $t$. The entire system performance can be quantified with respect to an overall network performance measure. This service function, $\varphi(\mathbf{x}(t))$, which can be calculated for any possible realization of $\mathbf{x}(t)$, maps the network state vector into a network performance at time $t$. We also make an assumption that $\varphi(\mathbf{x}(t))$ is a monotonically increasing function of $\mathbf{x}(t)$ because it agrees with the practical understanding that the overall network performance is improved if link performance is high. As such, $\varphi(\mathbf{x}(t))$ can be defined in:

$$\varphi(t) = \{\varphi(\mathbf{x}(t)) \mid \varphi(\mathbf{x}(t)) > \varphi((\mathbf{y}(t)) \forall \mathbf{y}(t) > \mathbf{y}(t))\}. \tag{2}$$

Let $x_i(t_0)$ represent the as-planned state of the $i$th link prior to the onset of disruptive event $e^i$. Assume that the effect of $e^i$ at time $t_e$ is a proportional reduction in the performance of the $i$th component by $V_i^j(e^i) = V_i^j$ where $V_i^j \in [0, 1]$. As shown in Fig. 2, $V_i^j$ refers to a component’s vulnerability, or its ability to maintain performance after $e^i$. As such, the effect of $e^i$ on the state variable of link $i$ is provided in Equation (3). The decreasing network performance due to the disruptive event $e^i$ is seen in its response until time $t_1$ when the new level of the network service function is measured. Note that a complete reduction in the functionality of the link occurs when $V_i^j = 1$. The vector quantifying the disruptive effects of $e^i$ for all links is $\mathbf{V}^j = (V_1^j, \ldots, V_i^j, \ldots, V_m^j)$.

$$x_i(t_0) = (1 - V_i^j) x_i(t_0) \quad \tag{3}$$

Parameter $V_i^j$ is considered stochastic due to the uncertainty associated with the nature of event $e^i$ and the subsequent performance behavior of component $i$. Equation (4) governs the behavior of $V_i^j$ in the interval $[a,b] \in [0,1]$:

$$P\left(a < V_i^j \leq b\right) = \int_a^b f(v_i^j) dv_i^j. \tag{4}$$

As recoverability in Fig. 2 refers to the speed at which the network recovers, recoverability can manifest itself as the time required to recover the functionality of a link. Naturally, recovery time for the $i$th link would be a function of the initial effect, $V_i^j$, of the disruption on the component, or $U_i^j(V_i^j(e^i)) = U_i^j(V_i^j)$. Similar to the initial impact, recovery is also uncertain; therefore, $U_i^j(V_i^j)$ is a stochastic term. The probability that link $i$ recovers prior to time $t_r \in (t_i, t_f)$ is found in Equation (5). For this article, it is assumed that $x_i(t_r) = x_i(t_1)$ until the recovery time is met, suggesting a step function to repair. This assumption could be relaxed with a known trajectory (e.g., linear, convex, and concave) relationship describing $x_i(t), t \in (t_i, t_f)$.

$$P\left(t_r < U_i^j(V_i^j) \leq t_r\right) = \int_{t_i}^{t_f} f(u_i^j(V_i^j)) du_i^j \quad \tag{5}$$

We can devise a recoverability strategy, $s(e^i) = (s_1^i, s_2^i, \ldots, s_m^i)$, which is a vector of link recovery activities to restore the performance of the system following disruptive event $e^i$, where $s_i^i = s_i(e^i)$ is the recovery activity for link $i$ disrupted by event $e^i$. Each element of the recovery activity vector $s(e^i) = s^i$ is described by (i) the order in which recovery is performed for links, $s_i^i$, and (ii) the time required for recovery to occur, $U_i^j$. This is represented with a duplet, as shown in Equation (6):

$$s(e^i) = (a(e^i), \mathbf{U}(V^j(e^i)))$$

$$= \left((a_1^i, U_1^j), \ldots, (a_m^i, U_m^j)\right). \quad \tag{6}$$

We introduce $\tilde{V}_i^j$ in Equation (7) to count link $i$ among those links that are disrupted by $e^i$:

$$\tilde{V}_i^j = \begin{cases} 1 & \text{if } V_i^j > 0, \\ 0 & \text{otherwise}. \end{cases} \quad \tag{7}$$
As such, Equations (8) and (9) describe the \((o^i_j, U^i_j)\) duple in more detail, respectively.

\[
o^i_j(e^i) = \left\{ o^i_j | o^i_j = h, h \in Z^+, \sum_i o^i_j \geq \sum_i V^i_j \right\}
\]

\[
U^i_j(e^i) = \left\{ U^i_j| P\left(t_i < U^i_j(V^i_j) \leq t_i\right) = \int_{t_i}^{t} f\left(u^i_j(V^i_j)\right) du^i_j \right\}
\]  

(8)

(9)

The order in which the recovery activity for link \(i\) is accomplished is represented by \(o^i_j\) and the time required to complete this activity is \(U^i_j\), and both are respective to the disruptive event, \(e^i\). Note that \(o^i_j\) represents the order of the starting time of the recovery activities. For example, if all disrupted links are repaired at the same time, \(o^i_j = 1, \forall i\). Also, \(U^i_j\) is a random variable described by its probability density function (pdf), \(f(u^i_j(V^i_j))\).

If the recovery orders are known and the probability distributions for the component recoveries are given, then we can devise the schedule for recovery. The set \(A_k = \{s^i_j | o^i_j = h, \forall i\}\) is the collection of all those components having the same order of recovery planning. The recovery planning activity schedule is thus given in Equation (10). Each element in set \(A_k\) shows those activities that are planned in parallel (i.e., all occurring at order \(h\)), while the different sets show the series planning of the overall recovery activities, \(W^p(e^i)\).

\[
W^p(e^i) = \left\{ A_1, A_2, \ldots, A_l | \sum_i V^i_j \right\},
\]

where \(p = 1, \ldots, P_L\).  

10

Special cases include scenarios where all the recovery activities are in series, \(l = V^i_j\), or when they are all in parallel, \(l = 1\). The number, \(P_L\), of possible recovery sets, \(W^p\), is governed by the different combinations of recovery activities the sets contain. If the number of element sets, \(A_k\), is fixed to \(L\), then the number of possible recovery sets can be represented in:

\[
P_L = \sum_{n_0=0}^{n-n_1-\ldots-n_L} \ldots \sum_{n_l=0}^{n-n_1-\ldots-n_l} \prod_{i=1}^{L} \left( n - n_1 - \ldots - n_l \right).
\]

(11)

Each combination in Equation (11) is counting the number of possible ways to recover a certain number of links for one particular order. The combinations are multiplied to cover all the orders, \(L\), and then summed over the possible values for the number of links in each set \(A_k\).

2.2. Network Resilience Metrics

In addition to the measure of resilience at time \(t_r\) after a disruption \(e^i\), \(A_{(t_r)}(e\) provided in Equation (1), there are three other temporal resilience metrics related to Equation (1), initially described in Pant et al.\(^{(21)}\) for the resilience of general systems.

First, the metric \(time\ to\ total\ network\ restoration, T_{(e^i)}\), records the total time spent from the point when recovery activities are started, at time \(t_r\), up to the time when all recovery activities are finalized. This is represented mathematically in Equation (12).

Based on \(T_{(e^i)}\), one can calculate the probability that total system restoration is finished before mission time \(t_s\) as \(P(t_s) = P(T_{(e^i)} \leq t)\).

\[
T_{(e^i)} = \sum_{i=1}^{L} U^i_j
\]

(12)

The second metric, \(time\ to\ full\ network\ service\ resilience, T_{p(\alpha)}(e^i)\), records the total time spent from the point when recovery activities are started, at time \(t_r\), up to the exact time, \(t_f\), when system service is completely restored (i.e., \(A_{p(\alpha)(e^i)} = 1\) and \(A_{p(\alpha)(\delta e^i)} < 1 \forall \delta > 0\)). From \(T_{p(\alpha)(e^i)}\), one can define the probability that system service restoration is finished before mission time \(t_f\) as \(P_{(t_f)} = P(T_{p(\alpha)(e^i)} \leq t_f)\). Note that \(T_{(e^i)} \geq T_{p(\alpha)(e^i)}\), or the time at which the network is fully restored is at least as lengthy as the time until a desired network resilience, \(A_{p(\alpha)(e^i)} = 1\), is achieved. For example, flows along a network can occur with full capacity despite not all arcs being restored: full capacity would suggest full network service resilience without the network being completely restored.

Finally, the metric \(Time\ to\ \alpha \times 100\%\ Resilience, T_{\alpha}(e^i)\), records the total time spent from the point when recovery activities commence, at time \(t_\alpha\), up to the exact time, \(t_\alpha\), when the system service is restored to \(\alpha \psi(x_{(0)}\) (i.e., \(A_{(t_\alpha)}(e^i) = \alpha\) and \(A_{(t_\alpha)}(\delta e^i) < \alpha \forall \delta > 0; \alpha \in (0, 1]\)). From \(T_{\alpha}\), one can define the probability that network service is restored by \(\alpha \times 100\%\), or \(\alpha \psi(x_{(0)})\), before mission time \(t_\alpha\) as \(P_{(t_\alpha)} = P(T_{(e^i)} \leq t_\alpha)\). This metric provides a means to compare different recovery strategies, determining which strategy achieves...
\( \alpha \times 100\% \) resilience the quickest holding resilience constant. Similarly, \( R_{\varphi}(t_r|e^i) \) can be found for different strategies, holding \( t_r \) constant.

2.3. Illustrative Example 1

To illustrate the notation and metrics discussed in Sections 2.1 and 2.2, consider the network shown in Fig. 5 (extended from Refs. 19 and 54) with 12 arcs and 7 nodes. The arc labels represent the arc index number and arc capacity, respectively. The service function for this network is the maximum network capacity between nodes S and T.

A single disruptive event, \( e^1 \), is considered. This event renders arcs 1 through 4 completely inoperable \( (V_i^e = 1, i \leq 4) \) with the remaining arcs unaffected \( (V_i^e = 0, i > 4) \). A single recovery activity set is assumed, \( W^i(e^1) = \{s_1^1, s_2^1, s_3^1, s_4^1\} \). In the deterministic recovery time resilience analysis of Henry and Ramirez-Marquez, \((^{19}\) recovery activities had a fixed time duration of 10 time units. Fig. 6(a) illustrates the trajectory of resilience over time for this network, disruptive event, and recovery activity set when a deterministic time duration is assumed. Note that full network resilience is achieved after the first three recovery activities (the recovery of arcs 1, 2, and 3), while full network restoration does not occur until all four disrupted arcs are recovered after all four recovery activities. As such, in the deterministic recovery case, \( T_{\varphi}(e^1) > T_{\varphi(x(t_0))}(e^1) \).

To implement the metrics of Section 2.2, consider a uniform distribution for arc recovery time, \( U_i^j \sim \text{UNI}(8,12), i = 1, \ldots, 5 \). The parameters of the uniform distribution are in time units. Further specification on the unit of recovery time can be elicited by experts in the area of application. For instance, a transportation network link disruption might require days to be recovered, whereas one disrupted link or node in a power grid can recover in a few hours. The trajectory of resilience under the stochastic recovery time assumption is provided in Fig. 6(b), where interval representations replace point estimates following a discrete event simulation of \( W^i \) with 1,000 iterations chosen for illustration purposes (while sufficiently reaching stability).

Approximate cumulative distribution function (cdf) representations of \( T_{\varphi}(e^1) \) and \( T_{\varphi(x(t_0))}(e^1) \) are provided in Fig. 7 for \( W^i \). Note that there is a difference in the scale of the axes: \( T_{\varphi}(e^1) \) is approximately bounded by the interval \((32, 46)\), while \( T_{\varphi(x(t_0))}(e^1) \) is approximately bounded by the interval \((24, 35)\). Note that \( P(T_{\varphi(x(t_0))}(e^1) > t) \leq P(T_{\varphi}(e^1) > t), \forall t \) such that \( T_{\varphi(x(t_0))}(e^1) \) is less than \( T_{\varphi}(e^1) \) in terms of stochastic dominance.\((^{55}\) An obvious conclusion from such a result is that the network has built in redundancy to satisfy flow demand (i.e., the network does not have to be completely restored for full network resilience to be achieved). Further, if \( \alpha = 0.75 \), the time to achieve 75% resilience lies between 16 and 24 time units, according to Fig. 7(c).

3. OPTIMIZING NETWORK RESILIENCE

The metrics described previously underscore the importance of understanding network resilience as a function of (i) the order in which recovery activities are performed and (ii) the time required to complete each recovery activity. That is, modifications to the resilience strategy, via changes to the configuration of the recovery activity sets or changes to the recovery activity time, impact metrics of network resilience, time to network restoration, time to full network resilience, and time to a desired level of network resilience.

3.1. Optimization Problem

The optimization problem considers identifying the optimal configuration of the recovery activity sets, \( W^p \), to minimize the time to full network service resilience. A condition for the optimality of recovery activity set \( W^p \) is shown in:

\[
T_{\varphi(x(t_0))}(e^1) \leq T_{\varphi(x(t_0))}(e^1) \quad (W^p) \quad \left\{ W^1, W^2, \ldots, W^p \right\}.
\]  \( (13) \)
Such an optimization problem is of a stochastic order in nature.\textsuperscript{56,57} The problem is to identify the optimal configuration of a known number, $L$, of recovery activity sets, $W^l$, according to:

$$P \left( T_{\psi(x(t_0))} \left( e^l \right) < t \right) \geq P \left( T_{\psi(x(t_0))} \left( e^l \right) < t \right),$$

$$\forall t \in (t_d, t_f).$$

(14)

Possible recovery activity sequences (dictated by the recovery sets) can be characterized in terms of the “goodness” by the corresponding cdf for $T_{\psi(x(t_0))}(e^l)$.

To choose the optimal $T_{\psi(x(t_0))}(e^l)$ such that Equation (14) holds, the Copeland Score (CS) method is introduced. The CS is a technique used to rank objects characterized by a set of attributes.\textsuperscript{58} The technique assumes that the ranking of the objects could be defined without considering decision-maker preferences and it is considered a nonparametric approach. The CS is computed based on pairwise comparisons between objects in a set and is defined as the difference between the number of times an object $a$ is better (with respect to attribute $q_k$) than the other objects and the number of times that object $a$ is worse (with respect to the same attribute $q_k$) to the other objects. $C_k(a,b)$ provides a value based on a comparison between link $a$ and link $b$ for attribute $q_k$, $k = 1, \ldots, \Omega$, performed according to the rule in Equation (15). Note that a minimum to the objective is desired. Before the first attribute, $q_1$, $C_0(a,b)$ is initialized at zero, and Equation (15) iterates through all $\Omega$ attributes:

$$C_k(a, b) = \begin{cases} C_{k-1} (a, b) + 1 & q_k (a) < q_k (b) \\ C_{k-1} (a, b) - 1 & q_k (a) > q_k (b) \\ C_{k-1} (a, b) & q_k (a) = q_k (b). \end{cases}$$

(15)

The method by Al-Sharrah\textsuperscript{58} dictates that the CS of object $a$ is obtained by adding $C_j(a,b)$ over all $b$, each representing the other objects, as shown in Equation (16). The object with the largest CS value is assumed to stochastically dominate all other objects with respect to the set of attributes. Note that CS assumes that all attributes have the same importance. If this assumption is not valid for a decision-maker, then other approaches could be considered (e.g., the ordered weighted averaging\textsuperscript{59} weighted distributions\textsuperscript{60} and a weighted version of the CS method\textsuperscript{61}).

$$CS(a) = \sum_{b \neq a} C_\Omega(a, b)$$

(16)

Each activity sequence, dictated by the recovery sets, produces an approximate cdf for $T_{\psi(x(t_0))}(e^l)$. To make comparisons of these stochastic terms, selected percentiles of $T_{\psi(x(t_0))}(e^l)$ are treated as attributes, naturally lending to the application of the CS method. As such, Equation (15) is applied across selected percentiles, where attribute $q_k$ refers to percentile $k$ of the cdf. Noting that in this case, for a specific percentile $k$, if $T_{\psi(x(t_0))}(e^l)$ of sequence $a$ is less than $T_{\psi(x(t_0))}(e^l)$ of sequence $b$, then sequence $a$ is better than (preferred to) sequence $b$ at percentile $k$. After the comparison of all sequences and across all selected percentiles, the sequence ranked highest according to Equation (16) is the optimal sequence.
Further, a multiobjective formulation that minimizes the time to full network resilience with time required to achieve $\alpha \times 100\%$ resilience could also result from this analysis, balancing the time to full recovery with the time required to, for example, recover some basic level of network functionality.

3.2. Illustrative Example 2

The optimization formulation, as well as the use of the CS to address the issue of stochastic ordering, is applied to the network example in Fig. 5, where, again, the service function describes the maximum network flow between nodes $S$ and $T$. A single disruptive event, $e^1$, results in arcs 1–4 being completely inoperable ($V_i^1 = 1, i \leq 4$) with the remaining arcs unaffected ($V_i^1 = 0, i > 4$). Four different recovery activity sets are considered:

i. $W^1 = \{A_1, A_2, A_3, A_4\}$ such that $A_1 = \{s_1^1\}$, $A_2 = \{s_2^2\}$, $A_3 = \{s_3^3\}$, $A_4 = \{s_4^4\}$, this recovery set performs recovery activities with the sequence 1-2-3-4,

ii. $W^2 = \{A_1, A_2, A_3, A_4\}$ such that $A_1 = \{s_1^1\}$, $A_2 = \{s_2^3\}$, $A_3 = \{s_3^4\}$, $A_4 = \{s_4^2\}$, this recovery set performs recovery activities with the sequence 1-3-4-2,

iii. $W^3 = \{A_1, A_2, A_3, A_4\}$ such that $A_1 = \{s_1^2\}$, $A_2 = \{s_2^1\}$, $A_3 = \{s_3^3\}$, $A_4 = \{s_4^4\}$, this recovery set performs recovery activities with the sequence 2-1-3-4, and

iv. $W^4 = \{A_1, A_2, A_3, A_4\}$ such that $A_1 = \{s_1^1, s_2^3\}$, $A_2 = \{s_3^4\}$, $A_3 = \{s_3^1\}$, $A_4 = \{\emptyset\}$, this recovery set performs the sequence 1-3-4 in series while in parallel with 2.

Stochastic recovery activity time for each activity is assumed to be uniformly distributed, $U_i^j \sim UNI(8,12), i = \{1,2,3,4\}$. Network resilience over time for each recovery activity set is plotted in Fig. 8.

The approximate cdfs of the time to full network restoration and time to full network resilience for
recovery activity sets $W^1$ through $W^4$ are presented in Fig. 9. Sets $W^1$, $W^2$, and $W^3$, which feature recovery activities that are all performed in series, are very similar with respect to full network restoration time in Fig. 9(a), with $W^4$ being significantly less due to its parallel recovery activities. The behavior of time to full network resilience produces a slightly different result, as shown in Fig. 9(b): $W^4$ produces a significantly less recovery time to network resilience, $W^2$ significantly longer, with $W^1$ and $W^3$ in between and very similar to each other.

Stochastic ranking with the CS can more effectively distinguish between a set of very similar recovery time cdfs (e.g., sets $W^1$, $W^2$, and $W^3$ in Fig. 9a). To make a selection, we implement the CS approach\(^{(58)}\) as a heuristic for selecting competing recovery activity sequences for the cases in which there is no clear “better” option in terms of stochastic order. To illustrate, Table I enumerates the 24 possible recovery activity sequences when considering only activities in series. Recovery sets $W^1$, $W^2$, and $W^3$ refer to sequences 21, 20, and 18, respectively.

<table>
<thead>
<tr>
<th>$l$</th>
<th>Sequence</th>
<th>$l$</th>
<th>Sequence</th>
<th>$l$</th>
<th>Sequence</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>4-3-2-1</td>
<td>9</td>
<td>3-2-4-1</td>
<td>17</td>
<td>2-1-4-3</td>
</tr>
<tr>
<td>2</td>
<td>4-3-1-2</td>
<td>10</td>
<td>3-2-1-4</td>
<td>18</td>
<td>2-1-3-4</td>
</tr>
<tr>
<td>3</td>
<td>4-2-3-1</td>
<td>11</td>
<td>3-1-2-4</td>
<td>19</td>
<td>1-3-2-4</td>
</tr>
<tr>
<td>4</td>
<td>4-2-1-3</td>
<td>12</td>
<td>3-1-4-2</td>
<td>20</td>
<td>1-3-4-2</td>
</tr>
<tr>
<td>5</td>
<td>4-1-2-3</td>
<td>13</td>
<td>2-3-4-1</td>
<td>21</td>
<td>1-2-3-4</td>
</tr>
<tr>
<td>6</td>
<td>4-1-3-2</td>
<td>14</td>
<td>2-3-1-4</td>
<td>22</td>
<td>1-2-4-3</td>
</tr>
<tr>
<td>7</td>
<td>3-4-2-1</td>
<td>15</td>
<td>2-4-3-1</td>
<td>23</td>
<td>1-4-2-3</td>
</tr>
<tr>
<td>8</td>
<td>3-4-1-2</td>
<td>16</td>
<td>2-4-1-3</td>
<td>24</td>
<td>1-4-3-2</td>
</tr>
</tbody>
</table>

These 24 sequences can then be ranked according to their CSs obtained by considering approximated percentiles of $T_{r(x|t|e)}(e^j)$ at 21 different percentile attributes $q_i, i = [0.00, 0.05, \ldots, 1.00]$ (percentile increments of 0.05, starting at 0 and ending at 1). As desired to minimize the time to recovery, the highest CS is preferred.

We first consider the time to full network restoration, as noted earlier in Fig. 9(a). Recovery
sets that are in series behave similarly; therefore, we use the CS method to distinguish between the cumulative distributions. The CS for the cumulative time to full network restoration is computed for all 24 sequences and sequence 22 (1-2-4-3) provides the best overall time to full network restoration with a CS equal to 233, whereas sequence 8 (3-4-1-2) exhibits the longest time with the smallest CS of $-383$. The cdfs of time to full network restoration for sequences 8 and 22 are plotted in Fig. 10.

With the time to full network resilience, the difference between the best and the worst recovery sets is larger due to the nature of the network being redundant. Hence, for some sets, the time to full network resilience is smaller than the time to full network restoration. That is, full resilience is achieved before all the links are recovered. In that case, the best overall recovery set is sequence 10 (3-2-1-4) with a CS equal to 435, while the worst is sequence 8 (3-4-1-2) with a CS of $-427$. The cdfs of time to full network resilience for sequences 8 and 10 are plotted in Fig. 10.

The different metrics of time to full network resilience and time to full network restoration could result in entirely different recovery set rankings depending on the structure of the network (e.g., a network with redundancy may achieve full network resilience much more quickly than it achieves full network restoration). As such, the two metrics may be competing objectives. To compare the behavior of
the two metrics, the difference between the ranks for each of the 24 sequences is plotted in Fig. 12 along with a network redundancy indicator illustrating the number of links recovered by the time the network achieved full resilience for each sequence. Note that the rank of the recovery set increases (i.e., the difference of the ranks is negative) whenever resilience is reached before full network restoration. A higher rank, indicated by a larger CS (i.e., faster time of recovery/resilience), occurs for situations where redundancy is present or where a smaller number of links recovered by the time the network regains its resilience. This is the case for sets 10, 11, 14, 18, 19, and 21. Note that recovery set 8 has the same rank for both the time to total network restoration and the time to network resilience, and in fact, set 8 performs the worst for both metrics. However, the CS is not equal, since it is compared with a different set of recovery activities for each metric but the overall cumulative distribution is nearly the same.

4. CASE STUDY: INLAND WATERWAY NETWORK

The methodologies discussed in this research can be applied to a number of critical infrastructure
networks dotting today’s global landscape. A particular infrastructure network of interest here is a transportation network, which can have a significant impact to pedestrian traffic and/or especially commodity flows when disrupted. Despite the prominence of inland waterways in the economies of North America, Europe, and Asia, little work has been done in describing disruptions in this particular transportation network. Most risk-based studies address coastal waterways. And the authors have found no work published in quantifying the resilience of inland waterway networks with respect to commodity flows. The importance of the 25,000 miles of commercially navigable U.S. waterways for transporting goods is likely to grow in the future as barge transportation represents a cheaper and environmentally friendlier alternative to already highly congested truck and train transportation. An expansion of inland waterways to deal with larger shipments (i.e., containers) has been proposed, leading to an increased need in addressing container security and the malevolent man-made attacks that could go along with unsecured containers. Container security will become even more important when the planned Panama Canal expansion project opens in 2014, enabling bigger ships and more containers from Asian to Atlantic and Gulf coastal ports and their associated inland waterway networks.

The case study presented in this article focuses on the Mississippi River Navigation System. The nation’s economy depends strongly on this waterway network with 47 billion tons of annual commodity flow circulating through the network. The National Waterway Network (NWN) is composed of a large number of links and nodes. A link represents either a shipping lane or simply a path in open water, and a node could be a facility such as a port, lock, dam, or perhaps another intermodal terminal. The case study analyzes the resilience of a known number of links that might go completely inoperable due to a disruptive event. A U.S. Army Corps of Engineers database was used to construct the Mississippi River Navigation System network shown in Fig. 13. A subset of the nation’s entire waterway network of 6,906 links, the Mississippi River Navigation System, includes 3,046 links and 1,545 nodes.

The Mississippi River is prone to different types of disruptive events, including periods of drought that have regularly occurred throughout history causing the closure of sections in the river and leaving stranded many barges and other vessels. For example, the most recent drought incident occurred in August 2012 causing nearly 100 barges and other vessels to wait and forcing shipment weight controls. Another type of common disruptive event that could render parts of the network is...
hurricanes that can impact facilities such as dams due to extreme flooding (e.g., Hurricane Isaac in 2012\cite{70}). Closing sections of the river impacts the nation’s economy by incurring losses to a large number of industries relying on the shipments that are being delayed, with macro-level, interdependent losses becoming quite significant\cite{53,71}. As such, resilience planning for inland waterway networks is of high importance.

The best recovery activity set that would minimize the time to regain full network resilience, $T_{\varphi(x(t_0))}(e^i)$, is sought. We consider a simple case where no recovery activities are performed in parallel, as considering parallel activities would require an account of recovery costs (thus a second objective). The time to full network resilience would be the minimum of the maximum time that could be spent to recover from the disruptive event.

The U.S. Army Corps of Engineers provides extensive data sets on the waterway network, including aggregate upstream and downstream commodity flows, as well as separate specific commodity categories, including coal, petroleum products, and farm products, among others, for each link in the waterway network. Five links are randomly picked from one region (circled in Fig. 13) in the Mississippi River Navigation System from the same region, with the following IDs: 200210, 200310, 200500, 200708, and 200800. Assume that one event occurred, $e^i$, in a specific place where those links are located. The links are considered completely inoperable due that event, meaning that they were closed and shipments cannot pass through those paths, leads to Equation (17), suggesting no randomness in vulnerability parameter $V_i$:

$$U_i = \begin{cases} 1 & i = 200210, 200310, 200500, 200708, 200800, \\ 0 & \text{otherwise.} \end{cases}$$

(17)

The performance of the network, measured by the aggregate commodity flow, is decreased by 395,770 tons each day those links are inoperable. The recovery time (in days) for each link is assumed to follow a triangular distribution, $U_i \sim \text{TRI}(a_i, c_i, b_i)$ where $a_i$ is the minimum value, $b_i$ is the maximum, and $c_i$ is the mode or most
likely value. Index $i$ denotes the link ID, with $i = 200210, 200310, 200500, 200708, 200800$. The parameters of the triangular distribution are ideally based on expert elicitation, with several methods proposed to encode key values (minimum, maximum, or most likely) or probabilities.\(^{(72)}\)

There are $5! = 120$ possible recovery activity sets for those five links, with very similar probability distribution and cdfs. In addition, we assume that the time to full network resilience and the time to full network restoration coincide. The reason for this is the structure of the network in which the flow follows one particular path going from port A to port B. Since the approximations of both the pdf and the cdf are similar to each other for all recovery sets, the CS is used to differentiate between the 120 cumulative distributions. The cdfs are constructed using 1,000 simulation runs, which, based on trial and error, is a good balance between run time and stability of results. The recovery set with the highest CS is the recovery set 93, $W^{93} = \{S_{200310}(e^1), S_{200210}(e^1), S_{200800}(e^1), S_{200500}(e^1), S_{200708}(e^1)\}$, with $CS(93) = 1,911$. The recovery set with the lowest CS rank is set 53, $W^{53} = \{S_{200500}(e^1), S_{200708}(e^1), S_{200210}(e^1), S_{200310}(e^1), S_{200800}(e^1)\}$, with $CS(53) = -1,735$.

The approximate pdfs for both activity sets are represented in Fig. 14; on average, the minimum time to recovery is 33.23 and the maximum time is 48.05 time units. The cumulative distributions of the time to full recovery for the best and worst recovery sets are shown in Fig. 15. The distributions perform similarly and are very closely related; however, the CS method was helpful in determining which sets

---

**Fig. 14.** Time to full network resilience, $T_{\phi(x(t_0))}(e^1)$, results in histogram (pdf approximation) form for 1,000 simulations of recovery sets (a) $W^{93}$ and (b) $W^{53}$.

**Fig. 15.** Approximate cumulative distribution of time to full network resilience, $T_{\phi(x(t_0))}(e^1)$, comparing recovery sets $W^{93}$ (solid line) and $W^{53}$ (dotted line).
perform best in terms of minimizing the time to full recovery. Such distinction is essential for decision-makers even though the alternatives have similar characteristics and might lead to similar outcomes in terms of optimizing the time to full recovery.

5. CONCLUDING REMARKS

This article introduces stochastic metrics of network resilience that allow a quantitative analysis under uncertainty of the time needed for a disrupted network to regain full operation after a disruptive event. An optimization problem to determine the most effective link recovery sequence is presented, aiming to minimize the time to full network resilience. The problem was heuristically solved using the CS method, which is used here as a technique to determine stochastic dominance.

Such metrics can be applied to any kind of network. As a case study, an inland waterway transportation network of the Mississippi River Navigation System is considered. The analysis shows that the methodologies presented here can be of great use to the decision-making authorities and risk managers overseeing the reliability and resilience of critical infrastructures to disruptive events. Minimizing the time to full recovery can decrease the economic losses incurred by the closure of a few sections in the waterway network. It is also a means to compare different approaches for implementing policy safeguards.

The current interest of the government in safeguarding critical infrastructures renders it imperative to understand the tradeoffs between system hardening and service restoration (i.e., resilience process) as a function of time and cost. Different strategies, such as parallel recovery sets, can improve resilience by further decreasing the time to full recovery. However, such strategies come at an increased cost that was kept to its minimum in this case study. Therefore, further research will develop a multiobjective optimization problem in which the decisionmaker is interested in improving the resilience time by considering parallel recovery activities but without incurring large costs.

Risk management of critical infrastructures incorporates the analysis of interdependent systems. That is, the smallest alteration in a recovery strategy could incur more significant changes in multiregional, multi-industry economic losses as these account for the losses in the disrupted system as well as the industries relying on that system. In the inland waterway case study, the disrupted network incurs a loss of 395,770 tons of commodity flow for one day. Considering the wider spread impacts of all the industries relying on such a commodity would significantly increase that number. The analysis of such wider spread losses is a topic for future work. And the CS method will be used to distinguish the slight stochastic differences among several alternatives that would decrease economic losses.

ACKNOWLEDGMENTS

The authors acknowledge the comments and suggestions of the associate editor and reviewers that greatly improved this article.

REFERENCES


