Sustainable and Resilient Infrastructure

ISSN: 2378-9689 (Print) 2378-9697 (Online) Journal homepage: http://www.tandfonline.com/loi/tsri20

Defining resilience analytics for interdependent cyber-physical-social networks

Kash Barker, James H. Lambert, Christopher W. Zobel, Andrea H. Tapia, Jose E. Ramirez-Marquez, Laura Albert, Charles D. Nicholson & Cornelia Caragea

SUSTAINABLE AND RESILIENT INFRASTRUCTURE, 2017
VOL. 2, NO. 2, 59–67
http://dx.doi.org/10.1080/23789689.2017.1294859

Taylor & Francis
Taylor & Francis Group

SHORT COMMUNICATION

Defining resilience analytics for interdependent cyber-physical-social networks

Kash Barker[a], James H. Lambert[b], Christopher W. Zobel[c], Andrea H. Tapia[d], Jose E. Ramirez-Marquez[e], Laura Albert[f], Charles D. Nicholson[a] and Cornelia Caragea[g]

[a]School of Industrial and Systems Engineering, University of Oklahoma, Norman, OK, USA; [b]Department of Systems and Information Engineering, University of Virginia, Charlottesville, VA, USA; [c]Department of Business Information Technology, Virginia Polytechnic Institute and State University, Blacksburg, VA, USA; [d]College of Information Sciences and Technology, Pennsylvania State University, State College, PA, USA; [e]School of Systems and Enterprises, Stevens Institute of Technology, Hoboken, NJ, USA; [f]Department of Industrial and Systems Engineering, University of Wisconsin, Madison, WI, USA; [g]Computer Science and Engineering, University of North Texas, Denton, TX, USA

ABSTRACT
Theory, methodology, and applications of risk analysis contribute to the quantification and management of resilience. For risk analysis, numerous complementary frameworks, guidelines, case studies, etc., are available in the literature. For resilience, the documented applications are sparse relative to numerous untested definitions and concepts. This essay on resilience analytics motivates the methodology, tools, and processes that will achieve resilience of real systems. The paper describes how risk analysts will lead in the modeling, quantification, and management of resilience for a variety of systems subject to future conditions, including technologies, economics, environment, health, developing regions, regulations, etc. The paper identifies key gaps where methods innovations are needed, presenting resilience of interdependent infrastructure networks as an example. Descriptive, predictive, and prescriptive analytics are differentiated. A key outcome will be the recognition, adoption, and advancement of resilience analytics by scholars and practitioners of risk analysis.

ARTICLE HISTORY
Received 10 April 2016
Accepted 1 September 2016

KEYWORDS
Resilience; analytics; networks; disruptions; interdependencies

1. Introduction

Recent natural disasters have challenged our traditional approaches of planning for and responding to disruptive events. For example, Hurricane Sandy affected eight countries and U.S. states from Florida to Maine in October 2012 with property damage in the U.S. around $50 billion. At least 650,000 houses were damaged (Porter, 2013), and tens of thousands of people were left homeless in the wake of the storm (Barron, Lipton, & Rivera, 2012). Months after the storm, power had not been restored to many communities in the New York/New Jersey/Connecticut area (Manual, 2013).

Hurricane Sandy disabled the physical infrastructure networks that enable the heavily populated NY/NJ area to operate, including roads, public transit, electric power, and telecommunications. For example, one million cubic yards of debris was removed following Hurricane Sandy, much of which was impeding transportation networks (Lambert, Tsang, & Thekdi, 2013; Lipton, 2013). These physical infrastructures are aging and increasingly fragile (e.g. the ASCE report card rating U.S. infrastructure with a grade of 'D' American Society of Civil Engineers, 2013) and subject to breakdown, with massive consequences on the services, and ultimately the communities, that rely upon them. Combined with the potential that climate change will result in more frequent and severe storms, addressing resilience of interdependent infrastructure networks is all the more critical.

Hurricane Sandy demonstrated how disruptions to *infrastructure networks* can impact a variety of other networks, including, in particular, the *community networks* and *service networks* that interact with and depend on infrastructure networks to function properly. Figure 1 illustrates the interdependencies among these three network types. Infrastructure networks are defined as engineered cyber-physical systems that enable essential 'lifeline' services for society (e.g. transportation, electric power, communications). Service networks are defined as human systems that engage with these infrastructure systems during a disruption (e.g. emergency responders, humanitarian relief, debris removal) to enable the function of other networks. Community networks are defined

CONTACT Kash Barker kashbarker@ou.edu

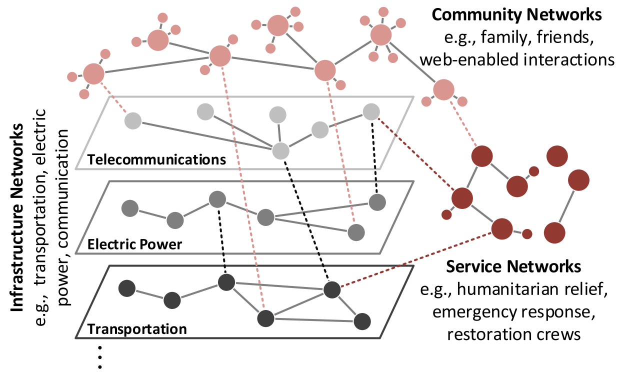

Figure 1. Interdependencies of infrastructure networks, with relationships of service and, ultimately, community networks.

as the interconnected society that the other networks support (e.g. relationships among people and communities). The resilience of one of these networks can affect the resilience of another, and the data generated for one network can potentially help us to understand the performance of another.

While interconnectivities of cyber-physical-social networks are essential (e.g. the effectiveness and efficiency of service networks, such as emergency response, relies on physical networks, such as transportation and communications), interdependency makes them more vulnerable to disruptions and subject to cascading effects (Vespignani, 2010). Recognizing the inevitability of large-scale disruptions, emphasis has shifted from prevention to protection to resilience, or to 'the ability to adapt to changing conditions and withstand and rapidly recover from disruption.' This 'inevitability' has been demonstrated too often in the last decade. For example, in addition to Hurricane Sandy previously discussed, the August 2003 U.S. blackout caused transportation and economic network disruptions (Minkel, 2008), Hurricane Isabel devastated the transportation system of the Hampton Roads, VA, region in 2003 and overwhelmed emergency response (Smith & Graffeo, 2005), and the 2011 9.0 magnitude earthquake and Tsunami that struck Japan disrupted global supply chain networks (MacKenzie, Santos, & Barker, 2012), among many others.

Data describing the performance of such cyber-physical-social networks are particularly important before, during, and after a large disruption like Hurricane Sandy because of the central role these networks play in supporting the society's resilience as a whole. These data may come from sensors embedded in the physical infrastructure, or from cameras which monitor system performance, but they also may be generated at the service network level in such forms as data feeds from emergency services operations, or at the community network level in such forms as social media posts. Collecting, storing, and analyzing such data are increasingly recognized as important by community leaders across the nation. Data.gov, the central site for U.S. Government data, reports that there are 96 U.S. States, State related agencies, Cities, or Counties with 'open data' web sites (Open Government – Data.gov, n.d.), each of which contains many data-sets corresponding to a wide range of service, community, and infrastructure information. For example, in April 2013, the New York City Mayor's Office of Data Analytics was established to create a 'big data' warehouse to capture the data streams from 911/311 services, energy, and telecommunications statuses, and restoration crews, among others. Recently efforts like these have become part of 'Smart Cities' initiatives in which metropolitan areas capture data streams from a variety of sources, enabling monitoring of city environments. Although some of these efforts involve cities becoming more responsive to citizen demands and more efficient use of tax dollars, other uses can be seen to contribute toward building resilience. A variety of data sources, such as video from CCTV cameras, voice, social media, streaming data, sensor logs, traditional structured, and unstructured data can be fused together to support these efforts. This availability of data has the potential to inform decisions through the application of advanced analytical methods, or *analytics*, thereby improving resilience. As we define resilience as the ability to adapt to changing conditions and withstand and rapidly recover from disruption, it is essential for communities to gather baseline data about the systems and conditions within their environments, to receive alerts when normal thresholds are exceeded, and to visualize the progress toward recovery. Monitoring enables awareness, increased community resilience is made possible through increased awareness coupled with the ability to intervene in these systems at appropriate moments to reduce impact and duration of crises and increase system flexibility and durability.

A focus on resilience through data analytics has entered the national stage in the U.S. In 2012, the White House announced new initiatives in big data and announced more than $200 million in new commitments for big data research, thus making big data research a national priority (White House, 2012). We define big data to be extremely large data-sets that may be analyzed computationally to reveal patterns, trends, and associations, especially relating to human behavior and interactions. In 2012, Gartner (Gartner, 2012) updated its definition as follows: 'Big data is high volume, high velocity, and/or high variety information assets that require new forms of processing to enable enhanced decision-making, insight discovery, and process optimization. Big data represents the information assets characterized by such a high volume, velocity, and variety to require specific technology and analytical methods for its transformation into value.'

For the purposes of this work, big data can be characterized as coming from static sensors informing the power network, water utility, and transportation network, among others. It can also be sees as arising from dynamic sensors, such as mobile devices, vehicles, and air/water sensors. Lastly it can be found in digital traces via internet-based action, social media, and direct contributions from participants (De Mauro, Greco, & Grimaldi, 2016). In August of 2015 Siddhartha (Sid) Dalal, the Chief Data Scientist and Senior Vice President at AIG, stated in a talk to the American Statistical Association that new methodologies and technologies are enabling the collection and analysis of data to enhance real-time probabilistic risk analysis and global resilience (AMSTAT News, 2015). This emerging paradigm can enable agencies to more effectively manage risks associated with complex systems. Dalal stated that statistical models are playing an increasingly important role in risk analysis, resilience, and helping the United States and other countries around the globe mitigate the effects of natural and man-made disasters. Big data research continues to be a national research priority due to its potential to transform how our quality of life by better managing risks from natural disasters and other events (White House, 2014).

Due to the interconnectedness of the cyber-physical-social infrastructure networks, there is opportunity to examine their overall interdependent structure from a number of different perspectives, and to leverage the availability of different types of data sources in order to do so. For example, data from physical infrastructure networks can be used both to support the operations of the service networks and to directly provide the community networks with valuable information about damage to roads or bridges. Similarly, service network data can provide information to both the community networks and the infrastructure networks about ongoing recovery operations, such as debris removal. There is also opportunity for community networks to support both infrastructure and service network resilience.

Guikema (2009) has argued that techniques of analytics can play an important role in infrastructure risk analysis but have limitations in predicting future behaviors based on past disruptive events. However, the growing availability of social media data and new techniques for analyzing these data introduce a number of interesting possibilities with respect to characterizing and managing resilience in this interconnected environment. For example, it may now be possible to gather data from static sensors about power outages, from service delivery responders as to power system repair, from social media as to the human experience of power outages, and from 911/311 calls concerning power outages causing secondary crises, all in near real-time. These data can be fused, geolocated, and visualized to provide improved situational awareness to city and emergency managers. Among other recent technologies, machine learning, data mining, and natural language processing have made leaps in extracting, processing, and classifying micro-blogged feeds, including detecting disruptions (Sakaki, Okazaki, & Matsuo, 2010), propagating rumors and misinformation (Mendoza, Poblete, & Castillo, 2010), assessing damage (Cresci, Tesconi, Cimino, & DellOrletta, 2015; Imran, Castillo, Lucas, Meier, & Vieweg, 2014), identifying needs (Caragea et al., 2011), analyzing sentiments (Caragea, Squicciarini, Stehle, Neppalli, & Tapia, 2014; Nagy & Stamberger, 2012), and identifying emotions (Schulz, Thanh, Paulheim, & Schweizer, 2013). These and other innovative methodologies will be essential for mining disaster data and help to characterize infrastructure and community networks, as well as to guide service networks, in the coming decade.

According to Meier (2013), disaster-affected communities are increasingly becoming the source of big (crisis) data during and following major disasters. During Hurricane Sandy over 20 million tweets were posted. Five thousand tweets were posted every second during the earthquake and subsequent Tsunami in Japan in 2011, resulting in 1.5 million tweets every five minutes. Meier (2013) finds that due to this surge in big social media data we now have the opportunity to better characterize, in real-time, the social, economic, and political processes that provide structure to our society. He described the rise of social media as a new nervous system for the planet, capturing the pulse of our social systems. However, the tools to leverage this massive amount of data to measure and support societal resilience have lagged behind the generation and widespread availability of social media data. We acknowledge that social media data have been leveraged to aid recovery and response efforts, however, these efforts have been ad hoc and piecemeal, largely driven by individuals (Gary, 2011).

2. Resilience analytics

With the above in mind, we define *resilience analytics* to be the systematic use of advanced data-driven methods to understand, visualize, design, and manage interdependent infrastructures to enhance their resilience and the resilience of the communities and services that rely upon them. The following discussion seeks specifically to characterize the role of social media data within the new, broader concept of resilience analytics, in order to argue for new frameworks for resilience of large-scale systems from a data-centric viewpoint. We begin by defining resilience in the context of cyber-physical-social infrastructure networks, and then consider the role of social media analytics in characterizing and enabling resilience in such interdependent infrastructure networks through the use of descriptive, predictive, and prescriptive techniques. Following a look at practical considerations, such as data collection and management, we expand our focus again to position social media-driven data analytics within the wider framework of the different types of data that our interdependent physical, service, and social networks will increasingly generate as they continue to co-evolve.

The ability to withstand, adapt to, and recover from a disruption is generally referred to as *resilience*, a definition with which many would largely agree (Aven, 2011; Ayyub, 2013; Haimes, 2009; Obama, 2011). Resilience is a concept that is increasingly gaining traction in government, industry, and academia (Hosseini, Barker, & Ramirez-Marquez, 2016; Park, Seager, Rao, Convertino, & Linkov, 2013). With respect to critical infrastructure, the Infrastructure Security Partnership (2011) noted that a resilient infrastructure sector would 'prepare for, prevent, protect against, respond or mitigate any anticipated or unexpected significant threat or event' and 'rapidly recover and reconstitute critical assets, operations, and services with minimum damage and disruption.'

Regarding the ability of society to cope with a disruption, the National Institute of Standards and Technology (2015) defined *community resilience* as 'the ability of a community to prepare for anticipated hazards, adapt to changing conditions, and withstand and recover rapidly from disruptions.' It is determined by community capacity for collective action as well as its ability for problem-solving and consensus building to negotiate coordinated response (Walker, Sayer, Andrew, & Campbell, 2010). Several works have explored community and social resilience from the perspective of social capital (the quantity and quality of social resources upon which people draw in pursuit of livelihoods) (Aldrich, 2012; Elliott, Haney, & Sams-Abiodun, 2010; Frankenberger, Mueller, Spangler, & Alexander, 2013; Magis, 2010; Wilson, Wiebe, & Hoffmann, 2005) and with dynamic and spatial dimensions (Béné, Wood, Newsham, & Davies, 2012; Norris, Stevens, Pfefferbaum, Wyche, & Pfefferbaum, 2008; Cutter et al., 2008). It is recognized that disruptive events tend to engage the community in social media activities, or 'a conversational, distributed mode of content generation, dissemination, and communication among communities' (Zeng, Hsinchun, Lusch, & Li, 2010). Meier (2013) strongly argues that social media can nurture social capital during disasters, in that by 'providing norms, information, and trust, denser social networks can implement a faster recovery' (Aldrich, 2012).

Situational awareness is a 'human mental process that can be enhanced using technology to access, analyze, and present information to have a greater understanding of existing conditions and how they will change over time.' (ESRI, 2008) Government agencies and response partners work to establish and maintain situational awareness to sustain general communications, gather intelligence from the field, execute logistical plans, track resources, send alerts and warnings, and perform general operations. If integrated with traditional data, social media can also help resilience planners achieve and maintain situational awareness in real-time.

While popular definitions for analytics vary, a common theme is that analytics inherently leverages data – and more often than not, large, unwieldy data. Laney (2001) introduced three dimensions of challenges for big data analytics. *Volume* denotes the vast size and scale of the data. The *velocity* dimension refers to the speed at which data is being created, which leads to the challenge of developing computing systems and algorithms that can cope with how fast new data is being created and somehow analyze it in near real-time. The third dimension of the big data challenge is *variety*, that is, data comes from many sources and in many forms. This includes sensor data, satellite imagery, video feeds, and social media updates: sources of data that one would expect to be collected before, during, and after a disruptive event. The ability to fuse these sources into commensurate data feeds (e.g. appropriate measurement scales, time frames, data granularity) is a non-trivial effort that adds complexity to the already difficult task of data preparation (e.g. outlier analysis, missing value imputation, data transformation, feature engineering) to support advanced analytics. And such issues may only increase in magnitude as infrastructure networks and their data collection structures become more autonomous.

In addition to increased and improved monitoring of social media, the recent five years have seen the development of numerous 'crowdsourcing' or 'public participation' technologies. Crowdsourcing is defined as a type of participative online activity in which an individual, an institution, a non-profit organization, or company proposes to a group of individuals of varying knowledge, heterogeneity, and number, via a flexible open call, the voluntary undertaking of a task (Estellés-Arolas & González-Ladrón-de-Guevara, 2012). Crowdsourcing takes advantage of Internet technologies and networks as well as the access, intelligence, knowledge, and time of participants. These new tools leverage networked digital technologies to enable scientists to bring research problems to non-scientists for participation and engagement, thus harnessing a distributed networked labor force. This distributed networked labor force may also come from new sectors of the population as the numbers of elders

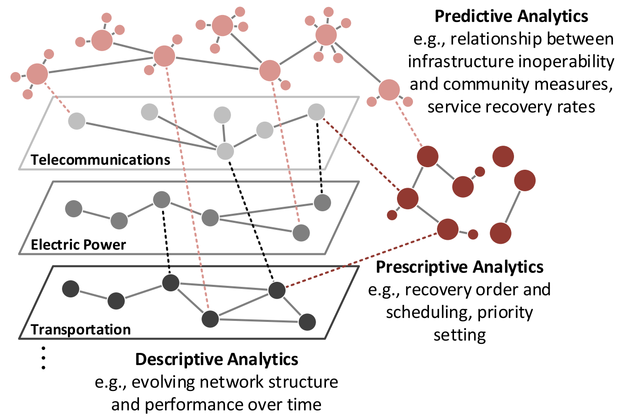

Figure 2. Descriptive, predictive, and prescriptive analytics to inform interdependent network behavior.

living long and well-passed retirement age increases. This population possesses the perfect trinity of socio-technical attributes leading to potentially excellent contributors to a crowdsourcing endeavor; i.e. they are highly educated, technical savvy, and have the luxury of free time (Beach & McKenzie, 2014; Kieboom, 2013). Crowdsourcing methods can be an effective proxy when physical sensor data are unavailable or misleading. Resilience analytics requires an understanding of how the emergent processes of crowd knowledge/labor and scientific discovery come together under the structures of networked computer platforms.

Social media data and crowdsourcing tools make it possible to facilitate descriptive, predictive, and prescriptive analytics, and thus enable researchers to better understand and enhance the behavior of interdependent networks after a disruptive event. Analytics is ultimately focused on improving decisions, and while data is required for facilitating analytics, data itself is not sufficient. The Institute for Operations Research and the Management Sciences (INFORMS) concisely defines analytics as 'the scientific process of transforming data into insight for making better decisions.' (INFORMS, 2015) Likewise, community resilience can be improved only if better decisions are being made before, during, and after disasters.

3. Research gaps

We offer several research directions for resilience analytics centered around the three perspectives commonly used to describe analytics: descriptive, predictive, and prescriptive analytics. Each of these perspectives plays a role in understanding the resilience of interdependent critical infrastructure networks, and some examples of the three perspectives of analytics are superimposed on the graphical depiction of interdependent networks in Figure 2.

3.1. Descriptive analytics

Descriptive analytics refers to techniques that effectively describe and possibly help to visualize the performance of the interdependent networks before, during, and after a disruptive event. In the case of the physical or operational characteristics of the infrastructure and service networks, baseline conditions can be easily characterized by gathering basic spatially explicit information, such as the location of nodes and links and their capacities in as-planned operating conditions, as well as infrastructure-related indicators of importance to particular regions (e.g. evacuation potential (in arterial miles/mi^2 National Research Council, 2006), and housing age (% built 1970–1994 Mileti, 1999). Information regarding the status of the infrastructure systems may be accessible real-time or near real-time during an event (e.g. neighborhoods with or without electric power service) and other data elements help to describe the longer-term infrastructure recovery (e.g. aerial imagery, number of new building permits issued, traffic data to help describe an improving transportation system). Both in the planning stages prior to a disruptive event and in the response phase after an event, the various service networks can be characterized by the numbers of crews available, amounts of resources (e.g. equipment, their dispatch locations, numbers of physicians Norris et al., 2008, shelter capacity Tierney, 2009, and medical capacity Auf der Heide & Scanlon, 2007). Baseline conditions for community networks can be described by spatially explicit populations and work locations, but also by indicators of resilient communities, such as: racial/ethnic inequality (difference in percentages) (Cutter et al., 2008; Norris et al., 2008), educational inequality (Morrow, 2008; Norris et al., 2008), previous disaster experience (Cutter et al., 2008), and the social vulnerability index (Cutter et al., 2008; Morrow, 2008; Tierney, 2009). Such metrics can be tracked during the post-event time frame to help quantify and qualify the recovery with respect to the baseline.

Using social media analytics to assess changes after the onset of a disruptive event, however, involves gathering and analyzing more dynamic information than this from online discussions about community water, wastewater, roads, gas, transportation, electric power, and communication services. Social media data are an indicator of human perception of physical infrastructure systems, not indicators of the systems themselves. It is this difference that enables researches to compare the data from physical, social, and service indicators to see key differences. Changes in frequency and intensity of discussion after a disruption can be used to inform a new description of the interdependent networks on both the infrastructure network level and service network level, and thus to provide a more dynamic characterization of the networks than that offered by the baseline indicators. The crisis and recovery information volunteered through social network platforms can also enable the formulation of a more detailed view of the interaction with the interdependent set of networks on the community network level, and it will provide a more complete picture of the community's resilience over time than that given by the standard static indicators described above. This community-level information also allows for relating geographic regions to communities exhibiting 'less-resilient' characteristics through dynamic social media expressions of help-seeking behavior (e.g. emergency services, policing), threat, and response.

Given the nature of this problem domain, descriptive analytics entails effectively quantifying and communicating the various aspects of dynamic systems, including their performance metrics and the behavior of the individuals (e.g. residents, business owners, decision makers) affected by events, as well as the numerous uncertainties involved.

3.2. Predictive analytics

Predictive analytics involves models that help to determine complex patterns and relationships among variables to quantify the likelihood of future events and thus reduce the associated uncertainty. Predictive analytics is enabled by the same type of data discussed in relation to descriptive analytics. In traditional studies of engineering-based resilience, indicators typically measure the loss (e.g. random forest models of hurricane-induced power outages Nateghi, Guikema, & Quiring, 2014) and subsequent recovery of resources (e.g. proportional hazards models of electric power recovery Barker & Baroud, 2014). There exists a need to combine these traditional measures with more modern, dynamic measures of community behavior that measure an increase in activity during times of crisis against a non-crisis benchmark. Essentially, there is a need to relate dynamic infrastructure behavior to what is being witnessed in the community, which thus provides a mean to predict community performance from infrastructure performance. For example, community sentiment is a function of how the individuals in the community are challenged by transportation or communication (or other) disruptions. Communities with different levels of racial diversity, income and education levels, and other demographics may be affected differently by similar infrastructure disruptions. Likewise, recovery and adaption may be unique given the socioeconomic characteristics of a community. Predictive analytics is an approach to transform what we know about our dynamic environment into various models of how changes impact other components and behaviors and thereby uncover fundamental relationships within the interdependent system.

Furthermore, the ability to predict how the performance of one network impacts the performance of another enables the ability to predict how the resilience (or lack thereof) of one network may impact resilience in another. For example, this includes the scheduling of network recovery operations based on dependencies within the service network (e.g. certain equipment is needed by two crews at once) and interdependencies imposed by real-time situations within the whole network (e.g. debris must be removed from transportation links before telecommunication restoration can commence). This involves trading off the repair of different cyber-physical-social networks depending on the importance of particular networks and the importance of particular regions.

Further, a related area of techniques in diagnostic analytics can be used for causal analysis after a disruption to gain insight into why resilient behavior was not observed.

3.3. Prescriptive analytics

While descriptive analytics relate to the current or historic states of system, and predictive analytics attempt to quantify future states, *prescriptive analytics* provides guidance on how to achieve desirable outcomes. That is, given a set of potential interventions or strategies, prescriptive analytics are mathematical tools that provide a quantifiable way of identifying and evaluating a feasible course of actions to best achieve specified objectives. The models involved in prescriptive analytics include optimization, stochastic optimization, simulation, and various hybrids of such approaches. Prescriptive analytics can guide pre-disaster resource allocation to reduce the potential effects of disruptive events as well as to aid post-disaster recovery efforts and priorities within the interdependent networks. For example, a growing prescriptive analytics problem lies in the recovery of infrastructure networks (e.g. the post-disaster scheduling of restoration crews, and the removal of debris from transportation networks) (Aksu & Ozdamar, 2014; Celik, Ergun, & Keskinocak, 2015; Gonzalez, Duenas-Osorio, Sanchez-Silva, & Medaglia, 2016; Matisziw, Murray, & Grubesic, 2010; Nurre, Cavdaroglu, Mitchell, Sharkey, & Wallace, 2012). There are numerous complexities involved in even such a straightforward process, however, since limited resources can cause the diversion of funds from one area to another and result in significant trade-offs associated with engineering, political, and societal impacts. The ability to quantify these potentially conflicting impacts may be enhanced by the descriptive analytics capability. For example, the social vulnerability of certain neighborhoods within the broader community could be used to inform the priorities in multiple objective optimization. The large uncertainties regarding (real-time) scheduling of restoration crews can potentially be reduced through an intelligent use of predictive analytics. Predictive models might be used to establish probabilities associated with completion times or crew availability as inputs to a stochastic optimization model. A compelling case, therefore, can be made for incorporating descriptive and predictive models within a unified decision-making paradigm.

Integrating perspectives of vulnerability and recoverability into the larger context of infrastructure network decision-making before, during, and after a disruptive event will help to support prescriptive decisions that lead to more resilient networks. Furthermore, accounting for community resilience will help to balance the dimensions of infrastructure network resilience with dimensions of social capital in the prescriptive models that are used to optimize investment decisions. This requires research into new data-driven interdependent network formulations, which are amenable to mathematical optimization and can benefit from both descriptive and predictive analytics to parameterize the model and to characterize uncertainty in the inputs.

4. Concluding remarks

Resilience analytics has the potential to do much good for vulnerable and fragile communities, including how agencies prepare for and recover from disasters. Data cannot improve resilience on its own. Data needs to be transformed into information through modeling that ultimately is used to support decisions. Above, we have described 'resilience analytics' as the data-driven process for supporting resilience through the application of descriptive, predictive, and prescriptive modeling.

Data may be derived from sensors that monitor cyber-physical systems, or they could come from a different, but growing, source: social media. According to Meier (2013), improving ways for communities to communicate internally and externally is an important part of building more resilient societies. This explains why social media and big data are central to growing more resilient societies, and understanding how resilient societies are strengthened by resilient physical infrastructure.

In summary, methodology should be central to society and community resilience. Analytics is ultimately focused on improved decision-making, and while data are a necessary ingredient, data are not the end goal. To address resilience it is imperative that we understand how communities are affected by infrastructure and how infrastructure is affected by communities. Community and societal needs are central to resilience, and for this reason social aspects of recoverability need to be supported. This can be achieved through the use of advanced analytical methods, which provides the opportunity to understand, forecast, and improve the performance of service networks using data from novel data types, such as social media and crowdsourcing.

Disclosure statement

No potential conflict of interest was reported by the authors.

Funding

This work was supported by Division of Civil, Mechanical and Manufacturing Innovation [grant number 1541165], [grant number 1541155].

Notes on contributors

Kash Barker is an associate professor and Anadarko Petroleum Corporation Presidential Professor in the School of Industrial and Systems Engineering at the University of Oklahoma. His research interests deal with the reliability, resilience, and economic impacts of infrastructure networks.

James H. Lambert is a research professor in the Systems and Information Engineering Department and the Department of Engineering and Society at the University of Virginia. He is Diplomate of Water Resources Engineering, a licensed Professional Engineer, and Fellow of the Society for Risk Analysis (SRA), American Society of Civil Engineers, and IEEE. He served as President of the SRA in 2015-2016 and is Chair of the Fifth SRA World Congress on Risk to convene in Capetown in 2019.

Christopher W. Zobel is RB Pamplin Professor of Business Information Technology in the Pamplin College of Business at Virginia Tech. His primary research interests are in the areas of disaster operations management, particularly supply chain resilience, and environmental decision making.

Andrea H. Tapia is Associate Professor of Information Sciences and Technology at the Pennsylvania State University. Her research deals with disaster or crisis informatics, where she studies the information discovery, needs, use, and sharing in disaster settings. Her recent work focuses on making data generated by individuals or social networks via mobile information technologies useful to decision-makers within large institutions.

Jose E. Ramirez-Marquez is Director of the Enterprise Science and Engineering Division and is an associate professor in the School of Systems & Enterprises at Stevens Institute of Technology. His research efforts focus on the analysis and computation of system operational effectiveness, reliability and vulnerability analysis as the basis for designing system resilience, and evolutionary computation for the optimization of complex problems associated with system performance and design.

Laura Albert is Assistant Dean of Graduate Studies at the University of Wisconsin-Madison. Her research interests are in the field of operations research, with a particular focus on discrete optimization with application to homeland security and emergency response.

Charles D. Nicholson is an assistant professor in the School of Industrial and Systems Engineering at the University of Oklahoma. His research interests are in complex network optimization, machine learning, and in their intersection, especially as relates to resilient systems.

Cornelia Caragea is an assistant professor in the Department of Computer Science and Engineering at the University of North Texas, where she directs the Machine Learning Research Laboratory. Her research interests are in machine learning, information retrieval, and natural language processing to classify and aggregate tweets and text messages from disaster events.

ORCID

Charles D. Nicholson http://orcid.org/0000-0002-7023-8802

References

Aksu, D. T., & Ozdamar, L. (2014). A mathematical model for post-disaster road restoration: Enabling accessibility and evacuation. *Transportation Research Part E: Logistics and Transportation, 61*, 56–67.

Aldrich, D. P. (2012). *Building resilience: Social capital in post-disaster recovery*. Chicago, IL: University of Chicago Press.

American Society of Civil Engineers (2013). *Report Card for America's Infrastructure 2013*. Retreived from http://www.infrastructurereportcard.org/

AMSTAT News. (2015). Analytics for resilience: Improving global security through real-time analysis of complex risks. Retrieved April 7, 2016, from https://www.amstat.org/newsroom/pressreleases/JSM2015-SPAIG-Dalal.pdf

Auf der Heide, E., & Scanlon, J. (2007). Health and medical preparedness and response. In W. L. Waugh & K. Tierney (Eds.), *Emergency management: Principles and practice for local government* (pp. 183–206). Washington, DC: International City Managers Association.

Aven, T. (2011). On some recent definitions and analysis frameworks for risk, vulnerability, and resilience. *Risk Analysis, 31*, 515–522.

Ayyub, B. (2013). Systems resilience for multihazard environments: Definition, metrics, and valuation for decision making. *Risk Analysis, 34*, 340–355.

Barker, K., & Baroud, H. (2014). Proportional hazards models of infrastructure system recovery. *Reliability Engineering and System Safety, 124*, 201–206.

Barron, J., Lipton, E., & Rivera, R. (2012). With $200 million in US housing aid, officials begin relocating the displaced. *New York Times*. Retrieved from http://www.nytimes.com/2012/11/06/nyregion/housing-relocation-begins-after-hurricane-sandy.html?_r=0

Beach, B., & McKenzie, D. (2014). Population ageing and the voluntary sector: Key figures and projected trends. *Commission on the Voluntary Sector & Ageing*, London, UK.

Béné, C., Wood, R. G., Newsham, A., & Davies, M. (2012). *Resilience: New Utopia or New Tyranny? Reflection about the potentials and limits of the concept of resilience in relation to vulnerability reduction programmes*. Brighton, UK: Institute for Development Studies.

Caragea, C., McNeese, N., Jaiswal, A., Traylor, G., Kim, H.-W., Mitra, P., … Yen, J. (2011). *Classifying text messages for the Haiti earthquake*. 8th International Conference on Information Systems for Crisis Response and Management, Lisbon.

Caragea, C., Squicciarini, A., Stehle, S., Neppalli, K., & Tapia, A. (2014). *Mapping moods: Geo-mapped sentiment analysis during hurricane sandy*. 11th International Conference on Information Systems for Crisis Response and Management, University Park, PA.

Celik, M., Ergun, O., & Keskinocak, P. (2015). The post-disaster debris clearance problem under incomplete information. *Operations Research, 63*, 65–85.

Cresci, S., Tesconi, A., Cimino, A., & DellOrletta, F. (2015). *A ligustically-driven approach to cross-event damage assessment of natural disasters from social media messages*. International World Wide Web Conference Committee (IW3C2), Florence.

Cutter, S. L., Barnes, L., Berry, M., Burton, C., Evans, E., & Tate, E. (2008). A place-based model for understanding community resilience to natural disasters. *Global Environmental Change, 18*, 598–606.

De Mauro, A., Greco, M., & Grimaldi, M. (2016). A formal definition of big data based on its essential features. *Library Review, 65*, 122–135.

Elliott, J. R., Haney, T., & Sams-Abiodun, P. (2010). Limits to social capital: Comparing network assistance in two new orleans neighborhoods devastated by Hurricane Katrina. *Sociological Quarterly, 51*, 624–648.

ESRI. (2008). Public safety and homeland security situational awareness. Retrieved from http://docplayer.net/2005281-An-esri-white-paper-February-2008-public-safety-and-homeland-security-situational-awareness.html

Estellés-Arolas, E., & González-Ladrón-de-Guevara, F. (2012). Towards an integrated crowdsourcing definition. *Journal of Information Science, 38*, 189–200.

Frankenberger, T., Mueller, M., Spangler, T., & Alexander, S. (2013). *Community resilience: Conceptual framework and measurement feed the future learning agenda*. Rockville, MD: Westat.

Gartner, Inc (2012). Big data. Retrieved from http://www.gartner.com/it-glossary/big-data/

Gary, J. (2011). Alabama tornados: Twitter, facebook, other social media make a big mark in disaster response, relief. *The Birmingham News*.

Gonzalez, A. D., Duenas-Osorio, L., Sanchez-Silva, M., & Medaglia, A. L. (2016). The interdependent network design problem for optimal infrastructure system restoration. *Computer-Aided Civil and Infrastructure Engineering, 31*, 334–350.

Guikema, S. D. (2009). Natural disaster risk analysis for critical infrastructure systems: An approach based on statistical learning theory. *Reliability Engineering and System Safety, 94*, 855–860.

Haimes, Y. Y. (2009). On the definition of resilience in systems. *Risk Analysis, 29*, 498–501.

Hosseini, S., Barker, K., & Ramirez-Marquez, J. E. (2016). A review of definitions and measures of system resilience. *Reliability Engineering and System Safety, 145*, 47–61.

Imran, M., Castillo, C., Lucas, J., Meier, P., & Vieweg, S. (2014). AIDR: Artificial intelligence for disaster response. In *Proceedings of the 23rd International Conference on World Wide Web, Seoul, South Korea*.

INFORMS (2015). What is analytics? Retrieved from https://www.informs.org/About-INFORMS/What-is-Analytics

Kieboom, M. (2013). *From problem to opportunity: Crowdsourcing the wise elders. The Wicked Notes #3*. Retrieved from https://www.kl.nl/wp-content/uploads/2014/04/wickednotes3.pdf

Lambert, J. H., Tsang, J., & Thekdi, S. (2013). Risk-informed investment for tropical cyclone preparedness of highway signs, signals, and lights. *Journal of Infrastructure Systems, 19*, 384–394.

Laney, D. (2001). *3-D data management: Controlling data volume, velocity and variety*. Stamford, CT: META Group.

Lipton, E. (2013, April 24). Cost of storm-debris removal in city is at least twice the US average. *The New York Times*.

MacKenzie, C. A., Santos, J. R., & Barker, K. (2012). Measuring international production losses from a disruption: Case study of the Japanese earthquake and tsunami. *International Journal of Production Economics, 138*, 293–302.

Magis, K. (2010). Community resilience: An indicator of social sustainability. *Society and Natural Resources, 23*, 401–416.

Manual, J. (2013). The long road to recovery: Environmental health impacts of hurricane sandy. *Environmental Health Perspectives, 121*, A152–A159.

Matisziw, T. C., Murray, A. T., & Grubesic, T. H. (2010). Strategic network restoration. *Networks and Spatial Economics, 10*, 345–361.

Meier, P. (2013). What is big (crisis) data? Retrieved November 13, 2015, from http://irevolution.net/2013/06/27/what-is-big-crisis-data/

Mendoza, M., Poblete, B., & Castillo, C. (2010). Twitter under crisis: Can we trust what we RT? In *Proceedings of the First Workshop on Social Media Analytics, Washington, DC*.

Mileti, D. S. (1999). *Disasters by design: A reassessment of natural hazards in the United States*. Washington, DC: Joseph Henry.

Minkel, J. R. (2008, August 18). The 2003 Northeast blackout-five years later. *Scientific American*, www.scientificamerican.com

Morrow, B. (2008). *Community resilience: A social justice perspective* (CARRI Research Report 4). Oak Ridge: Community and Regional Resilience Institute.

Nagy, A., & Stamberger, J. (2012). *Crowd sentiment detection during disasters and crises*. 9th International Conference on Information Systems for Crisis Response and Management, Vancouver.

Nateghi, R., Guikema, S. D., & Quiring, S. M. (2014). Power outage estimation for tropical cyclones: Improved accuracy with simpler models. *Risk Analysis, 34*, 1069–1078.

National Institute of Standards and Technology. (2015). Community resilience planning guide for buildings and infrastructure systems. NIST Special Publication 1190.

National Research Council. (2006). *Facing hazards and disasters: Understanding human dimensions*. Washington, DC: National Academies Press.

Norris, F. H., Stevens, S. P., Pfefferbaum, B., Wyche, K. F., & Pfefferbaum, R. L. (2008). Community resilience as a metaphor, theory, set of capacities, and strategy for disaster readiness. *American Journal of Community Psychology, 41*, 127–150.

Nurre, S. G., Cavdaroglu, B., Mitchell, J. E., Sharkey, T. C., & Wallace, W. A. (2012). Restoring infrastructure systems: An integrated network design and scheduling problem. *European Journal of Operational Research, 223*, 794–806.

Open Government – Data.gov. (n.d.). Retrieved October 11, 2015, from https://www.data.gov/open-gov/

Park, J., Seager, T. P., Rao, P. S. C., Convertino, M., & Linkov, I. (2013). Integrating risk and resilience approaches to catastrophe management in engineering systems. *Risk Analysis, 33*, 356–367.

Porter, D. (2013). Hurricane sandy was second-costliest in US history, report shows. Retrieved March 8, 2013, from http://www.huffingtonpost.com/2013/02/12/hurricane-sandy-second-costliest_n_2669686.html

Sakaki, T., Okazaki, M., & Matsuo, Y. (2010). Earthquake shakes twitter users: Real-time event detection by social sensors. In *Proceedings of the 19th International Conference on World Wide Web*, (pp. 851–860). Raleigh, NC.

Schulz, A., Thanh, T. D., Paulheim, H., & Schweizer, I. (2013). A fine-grained sentiment analysis approach for detecting crisis related microposts. *10th International Conference on Information Systems for Crisis Response and Management*, Baden-Baden, Germany.

Smith, C. M., & Graffeo, C. S. (2005). Regional impact of Hurricane Isabel on emergency Departments in Coastal Southeastern Virginia. *Academic Emergency Medicine, 12*, 1201–1205.

The Infrastructure Security Partnership. (2011). *Regional disaster resilience guide for developing an action plan*. Reston, VA: Technical Report, American Society of Civil Engineers.

Tierney, K. (2009). *Disaster response: Research findings and their implications for resilience measures* (CARRI Research Report 6). Oak Ridge: Community and Regional Resilience Institute.

Vespignani, A. (2010). Complex networks: The fragility of interdependency. *Nature, 464*, 984–985.

Walker, B., Sayer, J., Andrew, N. L., & Campbell, B. (2010). Should enhanced resilience be an objective of natural resource management research for developing countries? *Crop Science, 50*, S-10–S-19.

White House (2011). Presidential policy directive/PPD-8: National preparedness, Washington, DC: Department of Homeland Security.

White House. (2012). Obama administration unveils "Big data" initiative: Announces $200 million in new R&D investments, Press release. *Office of Science and Technology Policy, Executive Office of the President, Washington, DC*.

White House. (2014). *Big data: Seizing opportunities, preserving values*. Washington, DC: Executive Office of the President.

Wilson, T., Wiebe, J., & Hoffmann, P. (2005). Recognizing contextual polarity in phrase-level sentiment analysis. In *Proceedings of the Conference on Human Language Technology and Empirical Methods in Natural Language Processing*, (pp. 347–354). Vancouver, BC.

Zeng, D., Hsinchun, C., Lusch, R., & Li, S. (2010). Social media analytics and intelligence. *IEEE Intelligent Systems, 25*, 13–16.