Sustainable Cities and Society 36 (2018) 272–285

Contents lists available at ScienceDirect

Sustainable Cities and Society

journal homepage: www.elsevier.com/locate/scs

The Resilience to Emergencies and Disasters Index: Applying big data to benchmark and validate neighborhood resilience capacity

Constantine E. Kontokosta*, Awais Malik

Dept. of Civil & Urban Engineering and Center for Urban Science & Progress, New York University, 1 MetroTech Center, 19th Floor, Brooklyn, NY 11201, United States

ARTICLE INFO

Keywords:
Urban resilience
Emergency management
Disaster recovery
Community resilience
Big data
Anomaly detection

ABSTRACT

Resilience planning and emergency management require policymakers and agency leaders to make difficult decisions regarding which at-risk populations should be given priority in the allocation of limited resources. Our work focuses on benchmarking neighborhood resilience by developing a unified, multi-factor index of local and regional resilience capacity: the Resilience to Emergencies and Disasters Index (REDI). The strength of the REDI methodology is the integration of measures of physical, natural, and social systems – operationalized through the collection and analysis of large-scale, heterogeneous, and high resolution urban data – to classify and rank the relative resilience capacity embedded in localized urban systems. Feature selection methodologies are discussed to justify the selection of included indicator variables. Hurricane Sandy is used to validate the REDI scores by measuring the recovery periods for neighborhoods directly impacted by the storm. Using over 12,000,000 records for New York City's 311 service request system, we develop a proxy for neighborhood activity, both pre- and post-event. Hurricane Sandy had a significant and immediate impact on neighborhoods classified as least resilient based on the calculated REDI scores, while the most resilient neighborhoods were shown to better withstand disruption to normal activity patterns and more quickly recover to pre-event functional capacity.

1. Introduction

Hurricane Sandy's devastation led to 147 deaths, over 650,000 homes destroyed, and left 8.5 million people along the U.S. East Coast without power (Sullivan, 2012). In New York City alone, Sandy caused 43 deaths and over $19 billion in damage (Bloomberg, 2013). This historic storm particularly exposed New York City's vulnerability to coastal flooding, and left 6800 evacuees assigned to shelters, including 1800 patients from chronic care facilities (Gibbs & Holloway, 2013). The aftermath of this extreme event resulted in a comprehensive set of proposals aimed at achieving resilience through 'protection' and 'accommodation' in the event of a similar disaster (McArdle, 2014). Hurricane Sandy brought resilience improvement measures to the forefront of urban policy and planning, making the process of identifying vulnerable communities and quantifying their resilience capacity critical to effective emergency management and long-term resilience investments.

Resilience planning and emergency management require policymakers and agency leaders to make difficult decisions regarding which at-risk populations should be given priority in the allocation of limited resources. By extension, policymakers need to understand how community resilience capacity changes over time, how different communities compare, and how to quantify the impact of resilience-related investments and mitigation strategies. Our work presents a methodology to use diverse, large-scale urban data to identify, quantify, and benchmark neighborhood resilience through a unified, multi-factor index of local and regional resilience capacity: the Resilience to Emergencies and Disasters Index (REDI).

The REDI score is a benchmark of relative neighborhood resilience capacity within and between municipalities, and can be used to prioritize investment and funding needs across multiple dimensions of physical, social, economic and environmental conditions. The index is also intended to measure progress over time in increasing local resilience capacity, and to provide a performance measure to estimate the return on investment of resilience capacity-building measures. The strength of the proposed REDI score is that it combines measures of physical, natural, and social infrastructure systems to classify and rank the relative resilience capacity embedded in localized urban systems at high spatial resolution. This approach recognizes the importance of critical local community attributes that impact the ability to respond to and recover from emergencies and disasters.

The neighborhood REDI score is normalized on a scale of 1–100 to measure the deviation of a given neighborhood from a reference region mean. The reference region can be shifted to coincide with the political

* Corresponding author.
E-mail addresses: ckontokosta@nyu.edu (C.E. Kontokosta), awais.malik@nyu.edu (A. Malik).

http://dx.doi.org/10.1016/j.scs.2017.10.025
Received 23 June 2017; Received in revised form 13 October 2017; Accepted 21 October 2017
2210-6707/

and governance boundary of a particular agency's jurisdiction. A score of 100 represents the highest relative resilience capacity. ArcGIS and Python (Pandas) serve as the spatial data integration and visualization platforms, while temporal data are visualized using Tableau, and ARIMA time-series analysis is performed using Minitab. In this exploratory study of New York City, we collect and analyze a range of data sources provided by city, state and federal agencies, including the Department of City Planning, Department of Transportation, Metropolitan Transit Authority, Office of Emergency Management, Federal Emergency Management Agency, Department of Information Technology & Telecommunications, Department of Finance, and the United States Census Bureau, among others. The REDI score methodology is designed to be scalable to all municipalities across the U.S. and potentially global cities, provided the requisite underlying data are available at the appropriate spatio-temporal resolution. The methodology also allows for customization of the REDI algorithm to account for local priorities through indicator weighting.

This paper begins with a literature review of previous work relevant to community resilience metrics and measurement, and highlights the challenges involved in creating a unified resilience capacity index. The subsequent section describes the REDI methodology: the collection and integration of publicly-available data, the selection and extraction of relevant indicator variables, the formulation of REDI scores for each neighborhood, and the data cleaning process required to remove outliers and areas with incomplete information. Following the methodology description, a discussion of the major findings of the analysis and its implications for urban resilience policy, planning, and decision support is presented. The paper continues with a validation model of the REDI methodology using "big data" in the form of 12,000,000 "311" service request records before, during, and after Hurricane Sandy in New York City. Using service request data as a proxy for neighborhood activity patterns, the recovery period from this event is measured and compared for the most resilient and least resilient communities identified by their REDI scores. The paper concludes with a discussion of limitations and recommendations for future research.

2. Literature review

The notion of "resilience" has been the subject of several contested definitions (Aldunce, Beilin, Handmer, & Howden, 2014), resulting in divergent views on what it should encompass (Cutter, 2016a; Linkov & Florin, 2016; Linkov & Palma-Oliveira, 2017), and how it should be measured (Cutter, Ash, & Emrich, 2014; Winderl, 2014). Holling (1973) describes resilience as a "measure of the persistence of systems and ability to absorb change and disturbance and still maintain the same relationships between populations and state variables." The National Academies of Sciences defines the resilience of a system as "its ability to plan and prepare for, absorb, respond to, and recover from disasters and adapt to new conditions" (Ganin et al., 2016). The National Institute of Standards and Technology (NIST) characterizes community resilience as "the ability of a community to prepare for anticipated hazards, adapt to changing conditions, and withstand and recover rapidly from disruptions" (López-Cuevas, Ramírez-Márquez, Sanchez-Ante, & Barker, 2017). To operationalize these definitions, Bruneau et al. (2003) propose four key dimensions of community resilience: robustness, redundancy, resourcefulness and rapidity. The "resilience capacity" of a community, therefore, is the inherent set of features that enable that community to effectively respond to and recover from extreme events (Foster, 2012).

To add to the epistemological complexity around the term resilience, *vulnerability* – the propensity of systems to incur adverse shocks – is often considered the opposite side of the same coin (Bates, Angeon, & Ainouche, 2014; Gallopin, 2006; Weichselgartner, 2001), as regions of greater vulnerability tend to be the least resilient (Bergstrand, Mayer, Brumback, & Zhang, 2015). In urban environments, both the susceptibility of a system to potential harm (vulnerability) and its ability to quickly and effectively 'bounce back' from any damage (resilience capacity) must be accounted for in efforts to create the "sustainable networks of physical systems and human communities" (Godschalk, 2003) that define resilient cities (Meerow, Newell, & Stults, 2016). The focus of this study is on measuring resilience capacity, independent of neighborhood vulnerability, although it is recognized that many indicators of resilience capacity may also influence an area's risk exposure.

Urban communities can be viewed as the complex, dynamic interactions of physical, social, economic, and environmental systems (Cavallaro, Asprone, Latora, Manfredi, & Nicosia, 2014; McPhearson et al., 2016; Norris, Stevens, Pfefferbaum, Wyche, & Pfefferbaum, 2008). Community boundaries are considered porous, with most residents moving frequently across and between neighborhoods for work and leisure activities (Berkes & Ross, 2013). Consequently, a neighborhood's resilience capacity cannot be considered completely independent of the resilience capacity of its surrounding neighborhoods. Since most spatial geographies are defined by political or U.S. Census boundaries, it is important to pair any quantitative neighborhood resilience measures with local knowledge of the region through partnerships with the municipal government and the public that resides in those neighborhoods.

Several indices, frameworks, and conceptual models have been developed to quantify resilience (see Appendix A). Most of these models either lack data at the spatial granularity needed to adequately represent urban neighborhoods – resulting in several studies that instead use counties or other large administrative divisions (Bergstrand et al., 2015; Cimellaro, Solari, & Bruneau, 2014; Cutter, Boruff, & Shirley, 2003; Lam, Reams, Li, Li, & Mata, 2015; Miles & Chang, 2011; Sherrieb, Norris, & Galea, 2010) – or the methods do not define the community's spatial boundaries at all, thereby implying a spatially-scalable property of the proposed methodologies (Cutter et al., 2008). Other approaches are specifically applied to certain communities, such as coastal regions, making them less generalizable to other geographic areas (Fox-Lent, Bates, & Linkov, 2015; Islam, Swapan, & Haque, 2013; Orencio & Fujii, 2013; Razafindrabe, Parvin, Surjan, & Shaw, 2009). Feature selection methods used to justify the indicators included in previous models range from literature reviews and applied domain knowledge to statistical correlation analysis. One significant limitation of previous work is the omission of any empirical validation method to test the efficacy of the proposed methods (Bakkensen, Fox-Lent, Read, & Linkov, 2017; Cai, Lam, Zou, Qiang, & Li, 2016; Tate, 2012).

Existing approaches to measuring neighborhood resilience can be categorized into one or more of four domains: (1) social infrastructure and community connectivity, (2) physical infrastructure, (3) economic strength, and (4) environmental conditions (Jordan & Javernick-Will, 2013). The social infrastructure and community connectivity domain consists of demographic indicators and social services that either signify social vulnerability or highlight community cohesiveness. Suggested social resilience enhancement strategies include increased civic engagement to enable effective and efficient emergency management operations (Aldrich & Meyer, 2015; Burnside-Lawry & Carvalho, 2015; Godschalk, 2003; Magis, 2010), and greater policy engagement leading to effective land use zoning changes and infrastructure investments in vulnerable regions (Dale, Ling, & Newman, 2010; Ernstson et al., 2010; Wagner, Chhetri, & Sturm, 2014). A number of social vulnerability measurement techniques have been proposed (Cutter et al., 2003; Madrigano, Ito, Johnson, Kinney, & Matte, 2015; Sherrieb et al., 2010; Van Zandt et al., 2012). Tate (2012) provides an assessment of the major social vulnerability index configurations using global sensitivity analysis, and highlights the significant challenge of attempting to validate these indices using external data. López-Cuevas et al. (2017) proposed a method to benchmark the "mood steady state" of a community using online social networks (OSNs), and measured community resilience by gauging the changes in these steady states when a disaster occurred.