



HAL
open science

How Valid Are Social Vulnerability Models?

Samuel Rufat, Eric Tate, Christopher Emrich, Federico Antolini

► **To cite this version:**

Samuel Rufat, Eric Tate, Christopher Emrich, Federico Antolini. How Valid Are Social Vulnerability Models?. *Annals of the American Association of Geographers*, 2019, 109 (4), pp.1131-1153. 10.1080/24694452.2018.1535887 . halshs-02485376

HAL Id: halshs-02485376

<https://shs.hal.science/halshs-02485376>

Submitted on 28 Apr 2021

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

How Valid are Social Vulnerability Models?

Annals of the American Association of Geographers 109 (4):1131–53, doi:
[10.1080/24694452.2018.1535887](https://doi.org/10.1080/24694452.2018.1535887)

Samuel Rufat, Institut Universitaire de France, CY Cergy Paris University, France,
samuel.rufat@u-cergy.fr

Eric Tate, Geographical and Sustainability Sciences Department, University of Iowa, eric-tate@uiowa.edu

Christopher T. Emrich, College of Health and Public Affairs, University Of Central Florida, Christopher.Emrich@ucf.edu

Federico Antolini, Geographical and Sustainability Sciences Department, University of Iowa, federico-antolini@uiowa.edu

Abstract

Social vulnerability models are becoming increasingly important for hazard mitigation and recovery planning, but it remains unclear how well they explain disaster outcomes. Most studies using indicators and indices employ them to either describe vulnerability patterns or compare newly devised measures to existing ones. The focus of this article is construct validation, in which we investigate the empirical validity of a range of models of social vulnerability using outcomes from Hurricane Sandy. Using spatial regression, relative measures of assistance applicants, affected renters, housing damage, and property loss were regressed on four social vulnerability models and their constituent pillars while controlling for flood exposure. The indices best explained housing assistance applicants, while they poorly explained property loss. At the pillar level, themes related to access and functional needs, age, transportation, and housing were the most explanatory. Overall, social vulnerability models with weighted and profile configurations demonstrated higher construct validity than the prevailing social vulnerability indices. The findings highlight the need to expand the number and breadth of empirical validation studies to better understand relationships among social vulnerability models and disaster outcomes.

Keywords: social vulnerability, index, model, validation, spatial regression

INTRODUCTION

Disaster scholarship has long described processes that translate social, political, and economic marginalization into adverse human impacts (Blaikie et al. 1994; Fothergill, Maestas, and Darlington 1999; Laska and Morrow 2006). Geospatial modelers have quantified these relationships, using proxy demographic variables to construct indices of social vulnerability to natural hazards. Foundational studies (Clark et al. 1998; Cutter, Mitchell, and Scott 2000; Tapsell et al. 2002; Cutter, Boruff, and Shirley 2003) established methodologies that spawned a plethora of research that quantifies and analyzes spatial patterns of social vulnerability. The benefit of indices in reflecting multidimensionality, reducing complexity, and visualizing results is reflected in their growing promotion and use by government for hazard planning, setting priorities, and resource allocation (BRACE 2017; SCDRO 2017; WVDC 2017; FEMA 2018). Yet it remains unknown if pre-event descriptive patterns reflected in social vulnerability indices correspond with empirical post-disaster outcomes. Social vulnerability models are descriptive, but to what extent are they explanatory?

Descriptive modeling is advantageous for summarizing and visualizing social vulnerability, but determining their explanatory power requires observational data to test causal hypotheses (Shmueli 2010). Most index studies continue to emphasize describing social vulnerability patterns (Saldajeno et al. 2012; Hou et al. 2016; Colavito, Bjarnadottir, and Li 2017), reproducing index construction methods in varying geographic settings (Chen et al. 2013; Lawal and Arokoyu 2015; Roncancio and Nardocci 2016), and geospatial integration with physical hazard (Santos, del Rio, and Benavente 2013; Frigerio et al. 2016; Remo, Pinter, and Mahgoub 2016; Fischer and Frazier 2018). Despite the proliferation of social vulnerability indices, there remains limited statistical support for how effectively they actually measure social vulnerability.

Index validation requires modelers to make parameter choices regarding both the index configuration and the disaster outcome measure. The configuration of most social vulnerability indices is inductive, using factor analysis to identify latent statistical variables (Cutter, Boruff, and Shirley 2003; Rygel, O'Sullivan, and Yarnal 2006). Other configurations include hierarchical models that employ pillars to thematically organize indicators (Chakraborty, Tobin, and Montz 2005; Mustafa et al. 2011), and deductive models that apply a linear combination of indicators (Cutter, Mitchell, and Scott 2000; Wu, Yarnal, and Fisher 2002). The choice of configuration is important, because the statistical robustness has been found to vary with index structure (Tate 2012). For outcome measures, previous studies have evaluated physical damage, economic loss, mortality, migration, and rate of resident return. When selecting an outcome variable for a validation study, conceptually the analyst should be able to complete the statement: "*If you are more vulnerable, then [insert disaster outcome here].*" Unfortunately, hypothesized causal linkages between social vulnerability and selected outcome variables are often not explained. Collectively, both index configurations and outcome variables vary widely across previous validation studies, making it difficult to draw generalizable conclusions about the validity of social vulnerability indices.

The objective of this article is to assess the empirical validity of a range of social vulnerability models that are based on different approaches and assumptions. The first section describes index validation and findings from previous studies. Subsequent sections profile our analysis data: disaster outcome measures from Hurricane Sandy and social vulnerability models. The methods section describes our application of spatial multivariate regression. We report on the explanatory power of social vulnerability models and their constituent pillars in the results section, before concluding with a discussion of major findings and future research needs.

INDEX VALIDATION

Model validity is the degree to which a model adequately represents its underlying construct. For the construct of social vulnerability, to what extent do indices reflect its multidimensionality, interactivity, and causal processes? Previous validation studies for social vulnerability indices have been focusing on convergence, robustness, and construct.

Convergent validation assesses the level of agreement among alternative models measuring the same construct (O'Leary-Kelly and Vokurka 1998; Adcock 2001). Social vulnerability studies have applied statistical and geospatial methods to demonstrate similarity between a new and a widely-accepted model (Cutter et al. 2013; Holand and Lujala 2013; Hile and Cova 2015). The prevailing models in the US are the Social Vulnerability Index (SoVI[®]) created at the University of South Carolina (Cutter, Boruff, and Shirley 2003), and the identically named Social Vulnerability Index (SVI) developed at the US Centers for Disease Control (Flanagan et al. 2011). Although the use of an accepted index as a benchmark can demonstrate similarity in the operationalization of social vulnerability, convergent analysis alone is insufficient to demonstrate model validity. This is because two highly convergent models could also both poorly represent social vulnerability.

Robustness validation evaluates the degree to which a model is analytically and statistically sound. Index robustness is often assessed using uncertainty and sensitivity analyses.

Uncertainty analysis quantifies the variability in model outputs given changes in model inputs, and has been used to measure the stability of social vulnerability index scores and ranks (Jones and Andrey 2007; Tate 2013; Reckien 2018). Sensitivity analysis apportions total uncertainty among input parameters and has been used to identify the most influential stages of index construction (Schmidtlein et al. 2008; Tate 2012; Yoon 2012; Zhang and Huang 2013). Together, uncertainty and sensitivity analyses help measure the internal reliability of an index. However, reliability alone is an insufficient indicator of model validity, because a statistically robust index could also poorly represent social vulnerability.

Construct or empirical validation examines if hypothesized processes underlying a measure are borne out with empirical data. Previous empirical studies have collected post-disaster data and assessed index performance using statistical correlation analysis and ordinary least squares (OLS) regression. Decades of disaster case studies have documented how socially vulnerable populations disproportionately experience adverse impacts. Indicator selection for

social vulnerability indices is intended to reflect such outcomes, but to what degree do the index values correspond with empirical disaster data?

Is the model correct? In this case, the correctness is not so much a matter of its mathematical correctness but rather whether or not the algorithms represent what we intended to model in the real world. In other words, does the model provide a reasonable representation of the processes and spatial interactions of the real-world phenomena being examined? (DeMers 2002, p 177)

To assess the state of knowledge of the empirical validity of social vulnerability indices to natural hazards, we conducted a Web of Science search in August 2017. The search covered the years 2008 to 2017, using the following terms:

- ‘social vulnerability index’ [topic] OR ‘social vulnerability indicator’ [topic]
- AND ‘hazard’ OR ‘disaster’ [topic]
- NOT ‘climate change’ [topic]
- NOT ‘risk’ OR ‘resilience’ [title]

We excluded titles with the terms ‘risk’ and ‘resilience’ to focus on the construct of social vulnerability, and excluded the topic of ‘climate change’ to pinpoint papers based on specific hazard events and disasters as opposed to general climatic conditions. We read the abstracts of the 116 articles returned by the search. We removed articles that were purely descriptive, based on anthropogenic hazards (e.g., smoking), or conducted solely convergent or robustness validation. The remaining eleven empirical validation studies are summarized in Table 1. The body of work on empirical validation demonstrates that social vulnerability indicators have statistically explained a range of disaster outcomes, including physical damage, property loss, displacement, and migration. However, it is difficult to make generalizable conclusions due to the low number of studies and significant variation in methods. Moreover, the outcome variables were rarely described in terms of their conceptual or causal relationship with social vulnerability. Statistical analyses in only half of the studies controlled for hazard exposure or intensity, while only one study accounted for spatial dependence.

The knowledge gaps regarding the quantitative relationship between social vulnerability measures and disaster outcomes are suggestive of a research area that is still emerging. Given the increasing use of social vulnerability metrics in planning and decision making, there is a critical need to better characterize the ability of social vulnerability models to explain disaster outcomes. To do so, we constructed alternative models of pre-storm social vulnerability, and statistically compared them with post-disaster Federal Emergency Management Agency (FEMA) data from Hurricane Sandy.

Table 1. External Validation Studies for Social Vulnerability Indices

Study	Hazard	Social Vulnerability Predictor	Analysis Scale	Outcome Measure(s)	Method	Hazard Confounder	Statistically Significant Findings

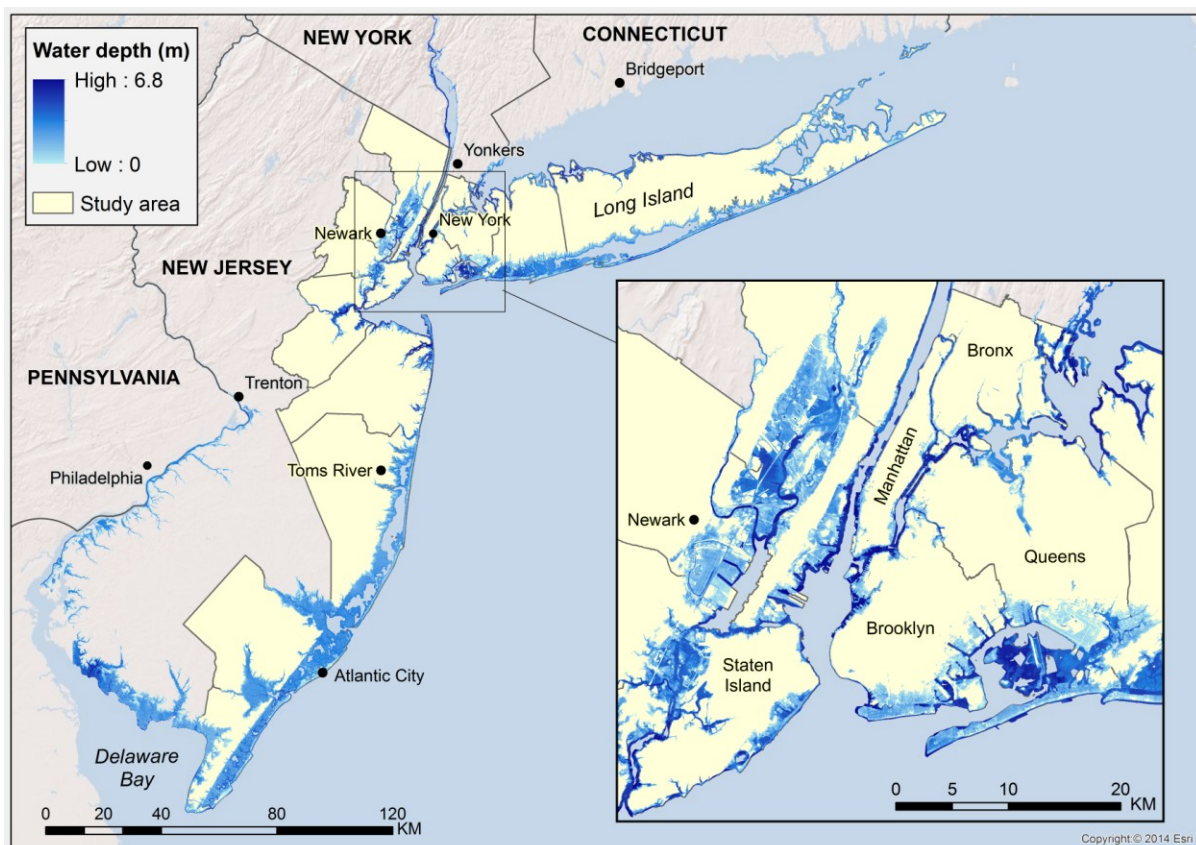
Abbas and Routray (2014)	River flood	Hierarchical index	Household	Health vulnerability	Correlation	None	Positive association† ($X^2=0.07$)
Bakkensen et al. (2017)	Severe weather	Deductive & inductive indices	County	Property loss, mortality, disaster declarations	OLS regression	Exposed elements, event magnitude, disaster frequency	Property loss: SVI**, So VI**, Declarations: So VI**
Burton (2010)	Hurricane wind & storm surge	Inductive factors	Census tract, block group	Damaged residential structures	OLS regression in catastrophic damage areas	Wind speed, % area inundated	Tract: (age & class)***, (Asian & agriculture workers)***, (Black & poor)**, rural**, Hispanic**; Block group: urban***, (Black & poor)***, (Asian & agriculture workers)***
Fekete (2009)	River flood	Inductive variables	Household	Displacement, shelter use	OLS regression	None	Displacement: urban***, home ownership**, number of rooms** Shelter: age**, home ownership**
Finch, Emrich, and Cutter (2010)	Hurricane storm surge	Inductive index	Census tract	Flood depth, return rate, homeowner assistance	Correlation, OLS regression	Flood depth	Rate of return: negative association*** ($r=-0.3$); So VI***
Khunwishit and McEntire (2012)	Hurricane	Deductive index	Household	Perceived disaster impact	OLS regression	None	Social vulnerability***
Liu and Li (2016)	River flood	Deductive index	Household	Mortality	Correlation	Only flood-affected homes surveyed	Positive association ($r=0.75$)*
Myers, Slack, and Singelmann (2008)	Hurricane	Inductive factors	County	Migration	OLS & spatial regression	Damage	Race and socioeconomic factor*
Schmidlein et al. (2011)	Earthquake	Inductive index	Census tract	Modeled economic loss and debris	OLS regression	Ground motion, distance	Loss: total***, income normalized***, area normalized***, income and area normalized***; Debris: total***, area normalized***
Tate et al. (2016)	River flood	Hierarchical index	Census block group	Property buyouts	Correlation	None	Positive association ($r=0.66$)***
Yoon (2012)	Multi-hazard	Hierarchical pillars, inductive factors	County	Property loss	OLS regression	None	Hierarchical: ascribed and achieved pillars**; inductive: poverty*

* $p < 0.05$; ** $p < 0.01$; *** $p < 0.001$; † p -value not specified

SANDY DISASTER OUTCOMES

"Superstorm Sandy" was the deadliest and most destructive hurricane of the 2012 season (Blake et al. 2013). At least 233 people were killed along the path of the storm, and at the estimated loss total of \$79 billion was surpassed only by Hurricane Katrina (FEMA 2013). In the US, Hurricane Sandy affected 24 states, with particularly severe damage in New York and New Jersey. On October 29, the New York City borough of Lower Manhattan flooded due to a storm surge of 14 feet (4.27m) (Figure 1). Approximately 100,000 homes on Long Island were severely damaged. Across the State of New York, fifty-three Sandy-related deaths were reported, with economic loss estimated at \$42 billion (*Hurricane Sandy's Rising Costs* 2012). Meanwhile, thirty-seven Sandy-related deaths were reported in New Jersey (Respaut et al. 2012), with an estimated loss total of \$37 billion.

Figure 1. Flood depths in the Hurricane Sandy study area (data: FEMA, 2015)



Impacted areas reaching pre-determined damage thresholds (Salkowe and Chakraborty 2009) can be declared Presidential Disaster Areas, enabling the flow of disaster recovery resources through FEMA's Individual Assistance (IA) Program.

Housing units represent nearly 70% of all built environment features across the United States (Comerio 1998; Phillips 2009). Damage to residential units and contents is therefore a suitable and frequently applied indicator of overall disaster impact. However, there may be substantial variation in the effect of that damage within an impact area due to population characteristics. Socially vulnerable areas are often associated with higher concentrations of mobile homes, more rental properties, lower educational attainment, and other factors rooted

in social processes, resulting in attenuated capacity to prepare for and respond to disaster events. Because homes in these places generally have lower property values, total economic loss in these places is likely to be lower compared to areas with more valuable homes. Accordingly, the use of relative impact measures may be more appropriate than absolute measures for social vulnerability analysis.

FEMA Outcome Data

The Presidential Disaster Declarations for Hurricane Sandy in New York and New Jersey (FEMA 2012a, 2012b) generated an array of flood hazard and impact data. The hazard data include 3-meter GIS grids of maximum water depth on October 31, 2012 (FEMA 2015). Within FEMA’s IA program, the Individuals and Housing program (IHP) provides financial assistance and repairs for residents with housing or contents damaged by presidentially declared disasters (Edgeley and Paveglio 2017). The IHP collects household-level information including number of occupants, age, income, access and functional needs (AFN), and damage to building and contents, among many other data points. There are two IHP sub-programs, Housing Assistance (HA), providing repair/replacement funding for real property (building) damage, and Other Needs Assistance (ONA), providing funds for personal property (contents) replacement, medical/dental expenses and other approved disaster related household expenses. IHP funds are available to residents who have been denied disaster loans due to affordability and meet other criteria set forth by the Stafford Act. Both HA and ONA target urgent unmet needs, in particular focusing on impacts not covered by insurance.

The FEMA data represent the most comprehensive and consistent data across the study area on impact, relief and recovery. As such they are the best available data to validate social vulnerability models in relation to this disaster. Outcome data from the IA program included the number of applicants, number of affected renters, damaged housing units, and building economic loss (Table 2) at the census tract level. We considered each outcome to be an indicator of short-term vulnerability as opposed to long-term recovery. We aggregated the water depth to the census tract level, and normalized the IA data using variables from the American Community Survey (ACS) for 2008-2012, a time frame representing pre-storm conditions. We normalized the IA variables to express them in relative as opposed to absolute terms. We expected each of the normalized Sandy outcomes to be positively related to social vulnerability.

Table 2. Sandy Outcome Measures (HUD 2015; FEMA 2015)

FEMA Variable (tract level)	Normalization	Outcome Variable
IA Applicants	Total population per tract	% Applicants
Renter Applicants with Personal Property Damages	Total renter households per tract	% Affected Renters
Applicants with Housing Flood Damage	Total non-seasonal housing units housing units per tract	% Damaged housing units
FEMA Verified Property Losses	Median house value	% Property loss
Maximum water depth	None	Maximum flood depth

The first two outcome measures in Table 2 describe impacts to people and their belongings. IA applicants can apply via phone, online, or in person at FEMA recovery centers opened across the affected area. We normalized the total number of applicants by the total population in each tract. This measure represents the initial “hand-raising” and assistance seeking activities undertaken by individuals who have been adversely impacted. For home repair, HA is available to homeowners but not renters. However, both homeowners and renters are eligible for ONA, through which they can be compensated for uninsured personal property (contents) loss. We utilized data for the number of rental units sustaining personal property loss, and normalized it by the number of renter households in each census tract.

The second two outcome measures describe relative impacts to residential housing. Property damage data are generated by FEMA when a household applies for IA and is deemed eligible based on ownership and evidence of storm-specific damage. We developed a relative measure of housing damage by normalizing total damaged houses by the total number of housing units in each census tract. FEMA inspection teams visit each eligible home and estimate economic loss to real property (building) and personal property (contents), producing two variables known as FEMA Verified Losses. We produced a relative measure of economic property losses by normalizing total real property (building) verified losses by the median house value in each census tract.

Analysis Scope and Scale

The spatially-varying nature of floods and the high spatial resolution of the FEMA depth grids (3m) suggested use of small spatial units. However, error in ACS demographic estimates generally increases as aggregation level decreases (Spielman, Folch, and Nagle 2014), suggesting the use of large spatial units. Meanwhile, the FEMA outcome data were provided at the census tract level. In the end we selected the tract scale for the analysis to balance tension among availability and error in the input data.

Table 3. Descriptive Statistics for Sandy Outcome Measures (data: FEMA, 2015; ACS 2008-2012)

Indicator	Tract Count	Min	Max	Mean	Median	Moran's I
IA Applicants	1,084	1	2,851	219	41	0.64
Affected Renters	608	1	2,601	115.7	3.35	0.48
Damaged non-seasonal housing units	625	14	5,938	479	229	0.56
Real Property FEMA Verified Losses	1,006	\$209	\$34,939,476	\$1,614,341	\$95,283	0.43
Maximum flood depth	1,188	0.001	5.8	3.14	3.17	0.65

The ideal analysis scope would align with the geographical extent of the disaster, but how is this extent best defined? Plausible choices include an extent based on the physical hazard, disaster outcomes, or affected administrative units. Scope selection based on administrative grounds may align with decision making, but can muddle statistical analyses if there is a

spatial mismatch between decision-making units and the disaster effects. This occurred for Sandy flooding, as 73% of the census tracts in the disaster-declared counties had no flooding. Accordingly, we set analysis scope to encompass the set of tracts that both intersected the Sandy floodplain and had non-zero FEMA outcomes. Table 3 provides descriptive statistics for the outcomes based on this scope at the census tract scale. The skewness and spatial autocorrelation of impact data further confirms that the damages from Hurricane Sandy are spatially concentrated.

Table 4. Input Variables for the Social Vulnerability Models

Variable	Description	SoVI	Weighted	SVI	SVP
AGE	% Age dependent (under 5 + over 65)	✓			
AGE05	% Age 5 years and under		✓		✓
AGE17	% Age 17 years and under			✓	
AGE65	% Age 65 years and over		✓	✓	✓
ASIAN	% Asian population	✓	✓		
BLACK	% African-American population	✓	✓		
CROWDING	% Households with more people than rooms			✓	
DISAB	% Older than 5 years with a disability			✓	
EDU12LES	% Adult educational attainment less than Grade 12	✓	✓	✓	✓
ESLANG	% English as a second language	✓	✓	✓	✓
EXTRACT	% Extractive sector employment	✓	✓		
FAMMARR*	Married families	✓			
FEMALE	% Female population	✓	✓		
FEMPLBR	% Female employment	✓	✓		
FHHOLDS	% Female-headed households	✓	✓		✓
GQ	% Pop living in group quarters			✓	
HISP	% Hispanic population	✓	✓		
MDGRENT*	Median rent	✓	✓		✓
MEDAGE	Median Age	✓			
MHSEVAL*	Median home value	✓	✓		✓
MOHOME	% Mobile homes	✓	✓	✓	✓
MUTIUNIT	% Housing in multi-unit structures			✓	
NATAM	% Native American population	✓	✓		
NOAUTO	% Households with no vehicle	✓	✓	✓	✓
NURSRES	% Nursing home residents	✓	✓		✓
NWHITE	% Ethnic and race minorities			✓	✓
PERCAP*	Per-capita income	✓	✓	✓	
PERPUNIT	People per housing unit	✓	✓		✓
POPDENS	Population density				✓
POVTY	% Households in poverty	✓	✓	✓	✓
RENTERS	% Renters	✓	✓		✓
RICH200K*	% Annual Income >\$200K	✓	✓		
SERVICE	% Service sector employment	✓	✓		
SINGHOLDS	% Single parent households			✓	
SSBEN	% Social security income	✓	✓		✓
UNEMPLOY	% Unemployed	✓	✓	✓	✓
VACANT	% Vacant housing	✓	✓		✓
<i>Variable count</i>		27	26	15	18

SOCIAL VULNERABILITY MODELS

To explore the relationship between social vulnerability and Sandy outcomes, we applied four social vulnerability models: an inductive model based on factor analysis (SoVI), a hierarchical weighted model based on expert knowledge, a deductive model composed of thematic pillars (SVI), and a profile approach based on clusters (SVP). Collectively, these models employ the configurations used in the majority of social vulnerability models. All of our models use demographic data from the 2008-2012 ACS at the census tract scale. The input variables used for each model are provided in Table 4. Variables whose values increase as social vulnerability decreases were multiplied by -1 to reverse their directionality, and are denoted by an asterisk in Table 4. The models were constructed for all New York and New Jersey census tracts in the affected counties (n=3,947).

Social Vulnerability Index (SoVI)

Inductive index approaches apply factor analysis based on principal components analysis (PCA) to reduce an initial indicator set into a smaller number of latent factors. Inductive modeling for social vulnerability was popularized by the SoVI, originally constructed at the county scale for the US (Cutter, Boruff, and Shirley 2003). The SoVI algorithm has since been widely used for description of vulnerability patterns in different countries (Solangaarachchi, Griffin, and Doherty 2012; Siagian et al. 2014; Guillard-Goncalves et al. 2015; Roncancio and Nardocci 2016), and for decision making in disaster recovery (City of Cedar Rapids 2010; SCDRO 2017; WVDC 2017). At the census tract scale, SoVI uses twenty-seven demographic variables from the ACS (Table 4). The current SoVI method and variable lists can be found at www.sovius.org.

Table 5. SoVI Components

Component	% Variance Explained	High loading variables
(1) Socioeconomic status	27.8	Per capita income (0.89) Educational attainment (0.83) Poverty (0.67) Unemployment (0.64)
(2) Gender and race (Black)	11.2	Female employment (0.88) Females (0.76) African-Americans (0.65)
(3) Vehicle access and renters	8.0	No automobile (0.87) Renters (0.81)
(4) Age	6.8	Age dependency (0.89) Social security beneficiaries (0.83)
(5) Vacant housing	5.2	Vacant housing (0.69)
(6) Access and functional needs	4.0	Nursing home residents (0.90)
(7) Race & ethnicity	3.9	Native Americans (0.92)
(8) Rental housing cost	3.7	Median rent (0.91)

We standardized the input variables into z-scores and entered them into a PCA. A Varimax rotation was applied to increase the interpretability of the component loadings. Using the Kaiser criterion (eigenvalues greater than 1.0), eight principal components were extracted, which accounted for 71% of variance in the original variable set. Based on the high loading variables, we interpreted the components to describe the social vulnerability factor represented by each component. The index was then computed by summing the factor scores

in each census tract. Table 5 profiles the components, their explained variance, and correlations with their highest loading variables (greater than 0.6).

Weighted Model

Hierarchical configurations aggregate indicators into pillars that share an underlying dimension of social vulnerability (e.g. socioeconomic status, health). The pillars are then aggregated to create the index. Such indices require a greater level of theoretical organization than more data-driven inductive models. We constructed a hierarchical model consisting of twenty-six variables organized into five pillars (Table 6). Vulnerability mapping and index design often rely on participatory approaches or expert knowledge for the selection and weighting of variables (Bankoff, Frerks, and Hilhorst 2004). We applied findings from a study that created weights for SoVI variables using subject matter experts in a modified Delphi Method (Emrich 2005). Using these weights, we computed weighted indicators as the product of the variable weights and normalized indicator values (min-max scaling). The weighted indicators were then summed within each pillar to create pillar scores, and the pillar scores averaged to create the index.

Table 6. Weighted Hierarchical Index Configuration

Pillar	Variable	Description	Weight %*
Socioeconomic status	PERCAP	Per capita income	15.9
	POVTY	% Poverty	14.9
	MHSEVAL	Median value of owner-occupied housing	8.8
	UNEMPLOY	% Unemployed	5.5
	RICH200K	% Families earning > US \$200,000	3.6
	SERVICE	% Service sector employment	1.7
	EXTRACT	% Extractive sector employment	1.5
	EDU12LES	% Adult educational attainment less than Grade 12	1.4
	MDGRENT	Median gross rent	1.9
Population structure	AGE65	% Age under over 65 years	7.9
	AGE05	% Age under 5 years	5.9
	PERPUNIT	People per housing unit	2.4
	FHHOLDS	% Female-headed households	1.9
	FEMLBR	% Female employment	0.8
	FEMALE	% Female population	0.8
Race & Ethnicity	BLACK	% African-American population	5.2
	HISP	% Hispanic population	2.9
	NATAM	% Native American population	1.5
	QASIAN	% Asian	0.9
Access & functional needs	NURSRES	% Nursing home residents	3.9
	SSBEN	% Social security income	3.7
	ESLANG	% English as a second language	1.9
	NOAUTO	% No automobile	0.9
Housing structure	RENTERS	% Renter occupied units	3.2
	VACANT	% Vacant housing	0.5
	MOHOME	% Mobile homes	0.5

*Weights derived from Emrich 2005.

Social Vulnerability Index (SVI)

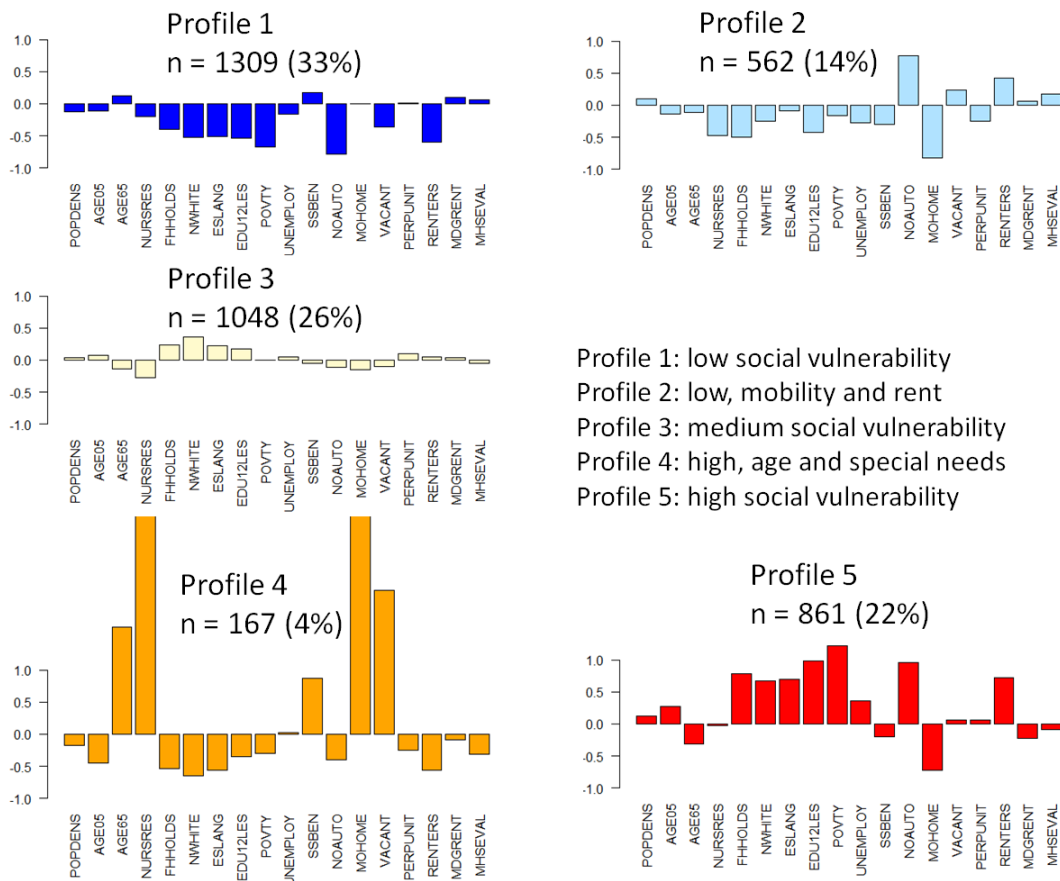
The Social Vulnerability Index (SVI), designed at the US Centers for Disease Control, is comprised of fifteen indicators (Table 4), which are normalized and summed to create the

index. The SVI has elements of both deductive and hierarchical design. In appearance the index is hierarchical because the indicators are conceptually organized into the four themes of socioeconomic status, household composition & disability, minority status & language, and housing & transportation. However, the SVI is functionally a deductive index because the four themes are mathematically ignored in the aggregation of individual indicators to create the index. The SVI is intended for use by emergency planners and public health officials to identify places and populations susceptible to environmental hazards (Flanagan et al. 2011; ATSDR 2017; BRACE 2017).

Social Vulnerability Profiles (SVP)

An alternative approach to indices that quantify social vulnerability magnitude is to produce spatially-varying typologies of social vulnerability, or Social Vulnerability Profiles (SVP). Following the methodology described in Rufat (2013), we constructed a set of vulnerability profiles using the same data set (Table 4). The correlation analysis was used to reduce the number of variables. This process limited collinearity, prevented implicit weighting, strengthened statistical power, and preserved a balance between the different dimensions of vulnerability to around five demographic, five socioeconomic and five cultural or institutional variables (Rufat 2013). For example, the share of female-headed households was used to represent all the gender variables due to the strong correlation among them, and the share of non-white, as a surrogate for all population identifying with ethnic and race minorities, resulting in the selection of only eighteen variables (Table 4).

Figure 2. Social Vulnerability Profiles (SVP)



Profile 1: low social vulnerability
 Profile 2: low, mobility and rent
 Profile 3: medium social vulnerability
 Profile 4: high, age and special needs
 Profile 5: high social vulnerability

The SVP combines factor analysis and clustering to produce spatially-compact vulnerability profiles instead of a single aggregated value like an index. The input variables were entered into a PCA, and five components (70% of the total variance) were extracted based on the Kaiser criterion. The factor scores were then used as a distance matrix for Hierarchical Clustering using Ward's method. The larger threshold in Ward's level index pointed to five as the ideal number of clusters. For each cluster, the eighteen variables were standardized using mean and SD. Each profile is interpreted according to the underrepresentation or overrepresentation of each variable, their association or mutual exclusion, and the resulting impacts on the strengthening (or reduction) of vulnerability (Figure 2). For example, while profile 4 is associating the elderly and nursing home residents, vacant housing, mobile homes and to a lesser extent people dependent on social security, and was interpreted as the "high social vulnerability, age dependency and special needs", profile 5 was over representing all other vulnerability indicators and was labeled "high social vulnerability" based on those important interactions across all vulnerability dimensions.

Model Comparison

Figure 3 displays the spatial distribution of the four models. For the SoVI (Figure 3A), weighted (Figure 3B), and SVI (Figure 3C) models, we classified high values of social vulnerability as the top 20% of index scores, and low social vulnerability as the bottom 20%. The five Social Vulnerability Profiles (SVP) are mapped in Figure 3D.

The four models are consistent in the places they identify as the most and least socially vulnerable. On the one hand, the Bronx, Brooklyn, and Newark have high modeled social vulnerability across all models. These places are identified by SVP Profile 5, and are associated with higher rates of poverty, renters, and female-headed households, and lower rates of educational attainment, English proficiency, and vehicle access. On the other hand, the center of Long Island falls within the low vulnerability category for all of the models, while Manhattan and the New Jersey coast are also classified as low by most models. The spatial distribution of social vulnerability is the most concentrated for the weighted and SVP models, and more diffuse for SoVI. Overall, visual inspection of the maps suggests a convergence in the social vulnerability models.

Figure 4 illustrates a Pearson correlation matrix of the social vulnerability indices and Sandy outcomes, with statistically significant correlations ($p < 0.01$) highlighted in blue for positive correlations and in orange for negative correlations. The three indices are positively related, with coefficients ranging from 0.35 to 0.75. The weighted model and SVI are the most convergent, in agreement with pattern similarity between Figures 3B and 3C. However, the indices have no significant correlation with the Sandy outcomes, except for an unexpected negative relationship between SVI and the share of affected renters. Overall, despite differences in model structure, input indicators, weighting, and aggregation schemes, correlation analysis indicates that the three indices exhibit convergent validity.

Figure 3. Social vulnerability indices based on (A) SoVI, (B) weighted model, (C) SVI, and (D) SVP.

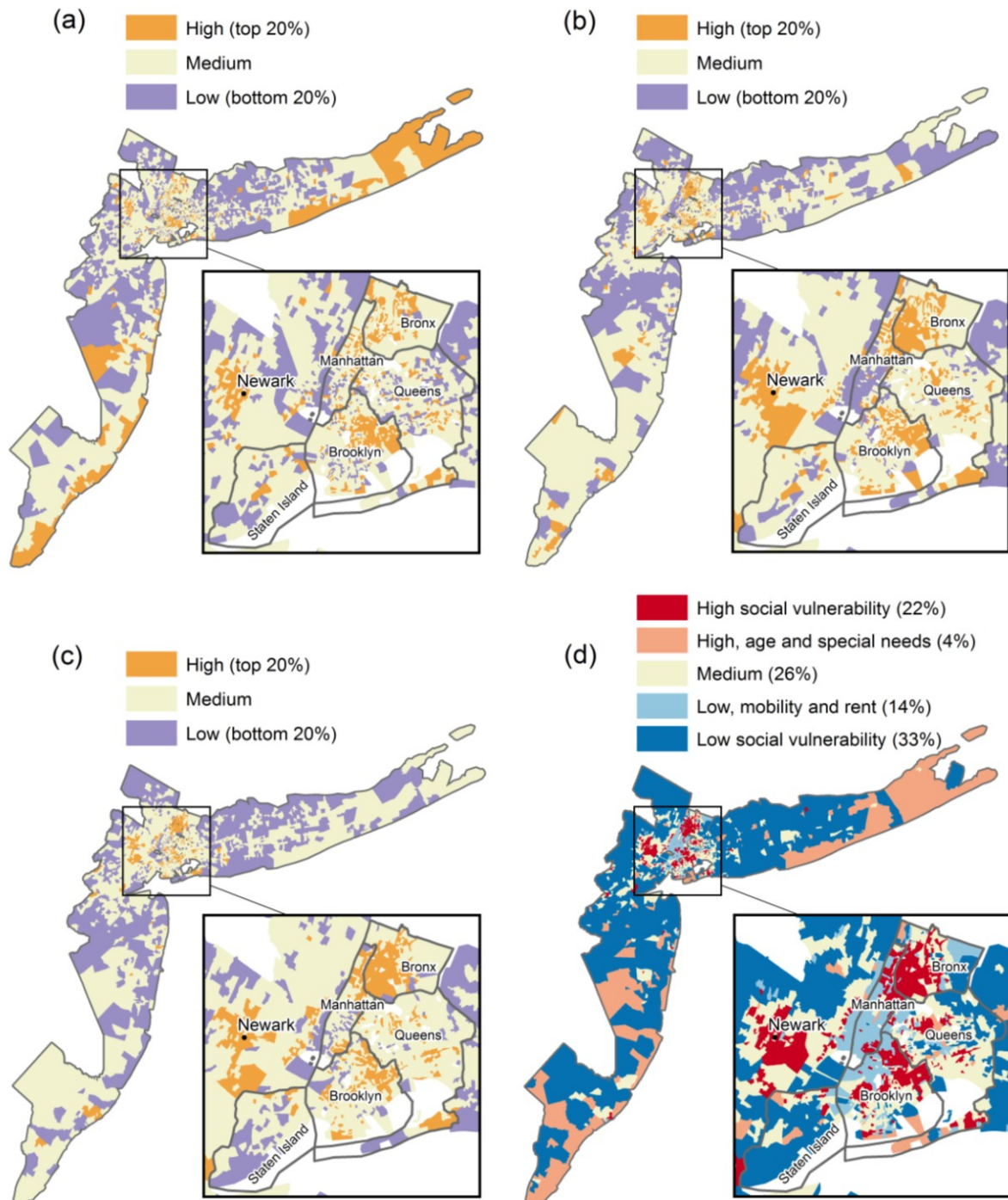
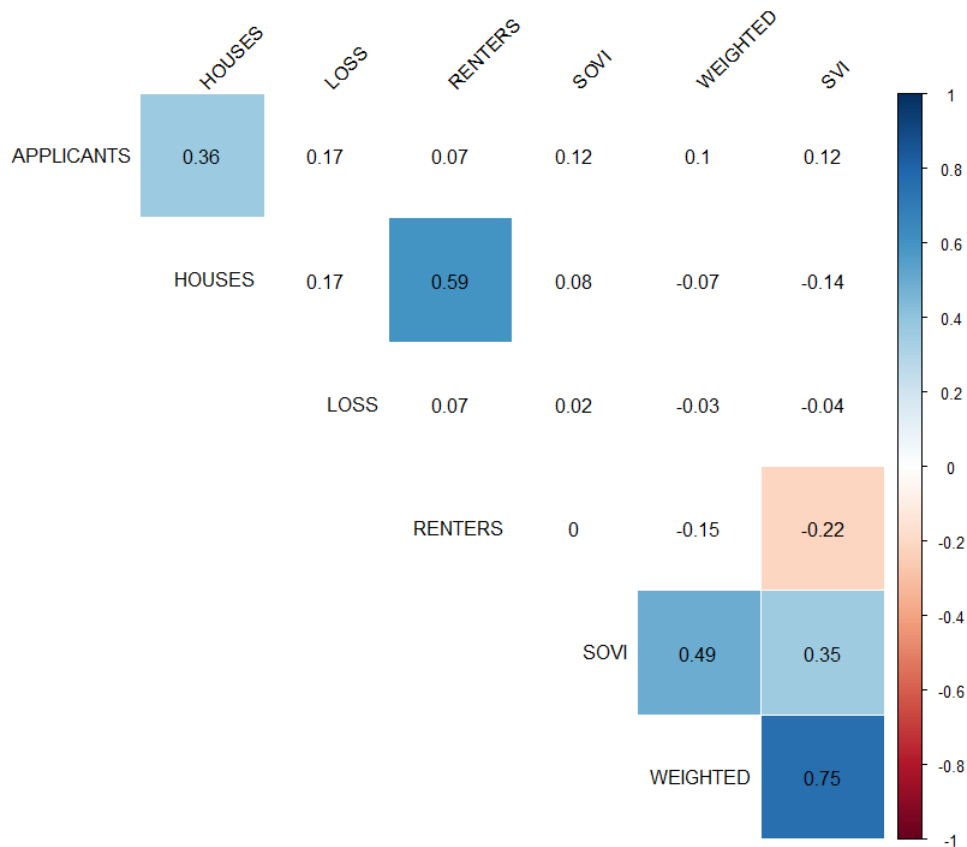


Figure 4. Correlation heatmap of social vulnerability and Sandy outcomes



Previous convergence validation studies have employed visual map inspection to compare results of a new and established social vulnerability index, as well as methods including difference mapping, statistical correlation, PCA diagnostics, and radar charts (Cutter et al. 2013; Holand and Lujala 2013; Hile and Cova 2015). Although useful, visual assessment and statistical association are inadequate measures of the alignment of social vulnerability indices models with disaster outcomes. Figure 4 suggests that using correlations might be a misleading approach to validation, as convergence is stronger among the indices than with the Sandy outcomes. As such, we focused on multivariate regression to evaluate the empirical validity of social vulnerability indices.

CONSTRUCT VALIDATION

For empirical validation, we created a set of ordinary least squares (OLS) regression models, using the normalized Sandy outcomes (Table 2) as dependent variables and social vulnerability measures as independent variables. We controlled for hazard severity using the maximum water depth during Sandy, and applied a natural log transform to all outcomes variables to reduce skewness. Regression analyses were then performed at the index and the sub-index levels.

At the index level, we constructed twelve multivariate OLS regression models to test the explanatory power of the three indices with each of the four Sandy outcomes. We constructed another sixteen regression models at the pillar level: four social vulnerability models by four

outcome variables. The pillar models used the SoVI components, weighted pillars, SVI themes, and SVP profiles as explanatory variables. The SVP is a more disaggregated measure, to compare it to the three indices at the index level, a linear parametric one-way ANOVA and a Kruskal-Wallis non-parametric test were used on the residuals of a regression between each Sandy outcome and maximum water depth to measure if the mean and median of each outcome significantly differ from one vulnerability profile to another while controlling for hazard severity.

As regression diagnostics, we applied the Multicollinearity Condition Number to evaluate multicollinearity, and the Breusch-Pagan test to assess heteroscedasticity in the residuals. If the Moran's I statistic on the OLS residuals indicated spatial dependence, spatial regression was conducted using a weights matrix based on queen contiguity. The Lagrange Multipliers, robust and non-robust accordingly, were used to test for the form of spatial dependence, guiding a decision to perform a spatial lag or spatial error regression. To assess model fit, adjusted R^2 (OLS) and pseudo R^2 (spatial regression) statistics were reported. However, because these two measures are not directly comparable, the Akaike Information Criterion (AIC) was computed. A reduction in AIC from OLS to spatial regression indicates improvement in model fit.

Among the social vulnerability models and outcome measures, we performed over one hundred independent statistical tests. As the number of statistical tests increases, so does the likelihood of finding a significant relationship just by chance. Such inflation of Type I error in multiple testing can be countered using a Bonferroni correction, in which the significance level for hypothesis testing is divided by the total number statistical tests performed (Bland and Altman 1995). Bonferroni correction in this study would entail dividing significance levels by 100, i.e., considering p-values of 0.0001 to be the threshold to maintain a significance level of 0.01 across all tests. Bonferroni critics have argued that the approach inflates Type II error (Morgan 2007), lacks guidelines for what constitutes a family of statistical tests (Cabin and Mitchell 2000), should only be applied to the universal hypothesis of a study (Armstrong 2014), and may not be required for exploratory analyses (Bender and Lange 2001). As a result, we report variable significance across a wide range of values to enable readers to choose whether to apply a Bonferroni correction (i.e. discarding 0.01 and/or 0.001 p-values) according to preferred balance between Type I and Type II errors.

RESULTS

Model-Level Validation

Table 7 summarizes the results at the index level, based on regression models using the indices and water depth as independent variables, and the Sandy outcomes as dependent variables. We considered an index to explain a disaster outcome if it is both statistically significant and has a positive coefficient. For all twelve regression models, the OLS residuals were significantly spatially auto-correlated, so we applied spatial regression. The AIC statistic decreased in each instance, indicating an improvement in model fit over OLS. Moving from OLS to a spatial regression model tended to reduce the magnitude of the index

beta coefficients and/or reduce the level of statistical significance. The opposite occurred for the water depth variable.

Table 7. Social Vulnerability Indices and Sandy Outcomes

Index	Statistic	Ln %Applicants		Ln %Damaged Homes		Ln %Property Loss		Ln %Affected Renters	
		OLS	Spatial	OLS	Spatial	OLS	Spatial	OLS	Spatial
SoVI	Depth	-0.007	0.02	0.03*	0.04***	0.12**	0.11***	-0.008	0.03*
	Index	0.26**	0.14** (error)	0.18*	0.11* (lag)	0.26	0.14 (lag)	0.12	0.04 (lag)
	Adjusted/ Pseudo R ²	0.02	0.69	0.03	0.44	0.03	0.32	0.005	0.43
	AIC	4760	3747	1963	1704	5064	3796	2484	2230
Hierarchical	Depth	-0.007	0.01	0.02*	0.05****	0.11***	0.14*** *	-0.03	0.02
	Index	0.21*	0.04** (lag)	0.01*	0.03* (error)	-0.05	0.02 (error)	0.56***	0.39*** (lag)
	Adjusted/ Pseudo R ²	0.008	0.59	0.003	0.47	0.02	0.33	0.06	0.46
	AIC	4770	3764	2038	1749	5072	3497	2451	2200
SVI	Depth	-0.009	0.01	0.02*	0.06****	0.11****	0.13*** *	-0.01	0.03*
	Index	0.11**	0.03 (lag)	-0.06*	-0.04** (error)	-0.16***	-0.14** (error)	-0.33***	-0.22*** (lag)
	Adjusted/ Pseudo R ²	0.01	0.59	0.02	0.47	0.04	0.33	0.14	0.48
	AIC	4764	3762	2031	1746	5058	3788	2394	2160
	N	1084	1084	625	625	1006	1006	608	608

* 0.01 ** 0.001 *** 0.0001 **** 0.00001

The Table 7 results indicate substantial variation in the explanatory power of the social vulnerability indices. This contrasts with the model convergence indicated by Figures 3 and 4, revealing the insufficiency of correlation and convergence studies for validating indices. Applicants had significant and positive relationships with the SoVI and weighted model. Based on the pseudo R² values, applicants also had the best model fit for each index. Housing damage was positively and significantly associated with the SoVI and weighted models, suggesting that irrespective of flood hazard, socially vulnerable populations incur a greater proportion of physical housing damage. However, the relationship between housing damage and the SVI was significant but negative. Property loss has the weakest relationship with modeled social vulnerability, as a negative association with the SVI was the only significant relationship.

The weighted index was positively and significantly related to all outcomes except property loss, demonstrating flexibility in explaining a variety of Sandy outcomes. SoVI was positively and significantly related to applicants and housing unit damage. The SVI had negative relationships with housing damage, property loss, and renters, contrary to our hypotheses. The SVI had the weakest explanatory power of the social vulnerability

assessments. At the other end, the SVP is the only model with significant positive relationships with all the Sandy outcomes, even in the case of a Bonferroni correction. The ANOVA and the Kruskal-Wallis tests reveal that each one of the outcomes vary significantly between the 5 vulnerability profiles when controlling for exposure with p-values below 0.0001 in all cases. If applying a Bonferroni correction, the weighted index has more significant relationships with the Sandy outcomes than the SoVI model and percent affected renters is the outcome best explained by the indices.

Pillar-Level Validation

Empirical validation at the pillar level enables examination of which sub-dimensions of social vulnerability models align with disaster outcomes. We use the term ‘pillar’ to broadly refer to the SoVI components, SVI themes, weighted model pillars, and SVP profiles, all which are indicator aggregations below the index level. As with the index-level regressions, we used the four Sandy outcomes as dependent variables and controlled for hazard severity using water depth. The Moran’s I statistic for OLS residuals was significant for all OLS models, so spatial regression models were created. The AIC statistic decreased in each case, indicating improved model fit. Similar to the index-level findings, accounting for spatial dependence in the regression tended to reduce beta coefficient magnitude and/or statistical significance for the pillars, while increasing them for water depth.

Table 8. SoVI Components and Sandy Outcomes

	Ln %Applicants		Ln %Damaged Homes		Ln %Property Loss		Ln %Affected Renters	
	OLS	Lag	OLS	Lag	OLS	Lag	OLS	Lag
Water depth	-0.009	0.01	0.04**	0.04***	0.12***	0.10****	0.02	0.05*
(1) <i>Socio-economic status</i>	0.11	0.06	0.02	-0.02	-0.32**	-0.22*	-0.47***	-0.33***
(2) <i>Gender, race</i>	0.52***	0.19**	0.10	0.05	0.35*	0.16	0.10	0.06
(3) <i>No auto, renters</i>	0.42***	0.15***	-0.17**	-0.11*	-0.28**	-0.18	-0.81***	-0.52***
(4) <i>Age</i>	0.14*	0.04*	0.21**	0.10**	0.28*	0.15	0.49***	0.28***
(5) <i>Vacant housing</i>	-0.08	0.07	0.06	0.06	0.43**	0.29*	0.16**	0.12*
(6) <i>Access and functional needs</i>	0.13	-0.02	0.15*	0.08	-0.007	-0.06	0.16	0.06
(7) <i>Native American</i>	0.13	0.06	0.04	0.07	0.17	0.12	-0.01	0.04
(8) <i>Median rent</i>	0.13	0.009	-0.01	-0.01	0.09	0.04	0.06	0.09
Adjusted/Pseudo R ²	0.10	0.70	0.07	0.46	0.09	0.34	0.26	0.51
AIC	4678	3740	1956	1701	5016	4771	2311	2128
Moran’s I	0.67****		0.50****		0.40****		0.45****	
Breusch-Pagan	16.5	17.6	6.1	13.9	18.9	19.4	4.5	5.7
Condition number	5.6	5.6	7.2	7.2	5.8	5.8	7.2	7.2
N	1084	1084	625	625	1006	1006	608	608

* 0.01 ** 0.001 *** 0.0001 **** 0.00001

Table 8 presents the multivariate regression results for the SoVI components. Overall, flood depth was highly significantly related to the housing outcomes variables but had much weaker relationships with human-centric outcomes. Applicants and renters had significant

relationships with some of the SoVI components, and the models for these outcomes also had the highest pseudo R². The model fit for the applicants variable (pseudo R² = 0.7) was much higher than for the other Sandy outcomes. The applicants were disproportionately female and black, renters and lacked car access, and were children and/or elderly. Affected renters were explained by four of the SoVI components, but the relationship was negative with two of them (renters & vehicle and socioeconomic status).

When the renters/vehicle factor is mapped, it becomes clear it is pinpointing Manhattan. Here, high numbers of renters and people without cars is an indicator of a densely developed and expensive downtown area with high reliance on public transportation as opposed to suggestive of high social vulnerability. This result highlights the importance of context-specificity in social vulnerability drivers and disaster phase (Rufat et al. 2015) when interpreting social vulnerability indices and selecting validation outcome measures.

The factors of gender/race, age, and vacant housing had positive relationships with each Sandy outcome they significantly explained. Socioeconomic status had negative relationships with the property loss and affected renters. As previously described, all input indicators were pre-processed prior to index construction to ensure directionality with social vulnerability. As a result, the negative beta coefficients for socioeconomic status mean that as socioeconomic status increases, so did the percentage of property loss and the percentage of affected renters. The 8 SoVI factors have seven significant positive relationships with the outcomes (only two after a Bonferroni correction) and four significant negative associations (two accordingly). However, only the first five components of SOVI have explanatory power, demonstrating a decline in factor significance with the variance explained by the PCA.

Table 9. Weighted Model and Sandy Outcomes

	Ln %Applicants		Ln %Damaged Homes		Ln %Property Loss		Ln %Affected Renters	
	OLS	Lag	OLS	Error	OLS	Error	OLS	Error
Water depth	-0.0005	0.02**	0.04**	0.08****	0.15****	0.18****	0.04*	0.08****
Socioeconomic status	-0.001***	-0.0004**	-0.0001	0.0001	0.002****	0.003****	-0.0002	-0.0002
Population structure	-0.0001	0.002	0.002	0.001	0.01***	0.007**	0.005**	0.002
Race & ethnicity	0.002**	0.001	0.0005	0.0002	-0.004****	-0.004****	-0.0004	-0.0006
Access & functional needs	0.01***	0.001***	0.001	0.001**	-0.01**	-0.01***	0.002**	0.002**
Housing	0.002	0.001	0.001***	-0.002**	-0.007**	-0.008****	-0.01****	-0.01****
Adjusted/Pseudo R ²	0.07	0.57	0.07	0.29	0.14	0.32	0.32	0.46
AIC	4412	3767	1995	1870	5120	4928	2315	2205
Moran's I	0.54****		0.31****		0.29****		0.29****	
Breusch-Pagan	8.3	9.0	7.5	8.7	9.6	12.3	7.9	10.5
Condition number	14.3	14.3	17.2	17.2	18.1	18.1	14.7	14.7
N	1084	1084	625	625	1006	1006	608	608

* 0.01 ** 0.001 *** 0.0001 **** 0.00001

Table 9 reports the multivariate regression results for the pillars of the weighted model. Flood depth was a highly significant predictor of all Sandy outcomes. Property loss had the most significant relationships across the pillars, but only the relationships with socioeconomic status and population structure were positive. This contrasts with the index-level finding that property loss was the only non-significant model for the weighted index. The low model fit (pseudo $R^2=0.32$), indicates other factors have a strong effect on relative property loss. By contrast, applicants had the best model fit (pseudo $R^2 = 0.57$) of the outcome measures. As the percentage of applicants increased, the only positively and significantly related pillar was access and functional needs. Meanwhile, the housing damage and affected renter outcomes were not predicted by any of the pillars in the hypothesized direction.

The pillars of the weighted model have five significant positive relationships with the Sandy outcomes (only two after a Bonferroni correction) and six significant negative associations (four accordingly). Across the pillars, access and functional needs had a significant relationship with all four outcomes, and in a positive direction with all but property loss. The housing and race/ethnicity pillars were the weakest predictors, as all relationships were either non-significant or negative.

Table 10. SVI Themes and Sandy Outcome Measures

	Ln %Applicants		Ln %Damaged Homes		Ln %Property Loss		Ln %Affected Renters	
	OLS	Lag	OLS	Error	OLS	Lag	OLS	Lag
Water depth	-0.02	0.01	0.03*	0.06****	0.13****	0.11****	0.02	0.05*
Socioeconomic status	0.36	0.16	-0.18	-0.44	2.03**	1.21	-0.23	-0.28
Household composition & disability	0.29	0.24	0.30	0.01	0.39	0.22	0.85*	0.39
Minority & language	-0.51	-0.30	-0.43	-0.41	-3.86****	-2.19****	-1.97****	-0.88**
Housing & transportation	1.05*	0.25	-0.62**	-0.58**	-0.02	-0.24	-1.58****	-1.33****
Adjusted/Pseudo R ²	0.02	0.70	0.04	0.48	0.09	0.34	0.24	0.51
AIC	4761	3763	2020	1740	5003	4768	2323	2128
Moran's I	0.72****		0.53****		0.38****		0.45****	
Breusch-Pagan	2.9	9.3	2.5	11.8	6.2	14.1	6.3	6.4
Condition number	9.5	9.5	10.8	10.8	10.4	10.4	9.3	9.3
N	1084	1084	625	625	1006	1006	608	608

* 0.01 ** 0.001 *** 0.0001 **** 0.00001

Table 10 summarizes the regression results for the SVI themes. Overall, the SVI themes have poor explanatory power, as no outcome measure had a positive and significant relationship with any theme. This finding aligns with the results at the index level. The Housing and Transportation theme includes *Housing in structures with ten or more units* and *Households with no vehicle*. Similar to the results for SoVI factor of renters/vehicle access, the SVI Housing and Transportation theme is identifying Manhattan. However, because the theme also includes the variables of *Households with more people than rooms* and *Persons in*

institutionalized group quarters, this SVI theme is also identifying other places, complicating interpretation of the negative beta coefficient.

Table 11 presents the regression results at the profile level of the SVP, with four of the five vulnerability profiles entered into the regression as dummy variables. We discarded the medium profile (profile 3) to avoid multicollinearity issues, resulting in evaluation of the two low and two high vulnerability profiles. Similar to Table 9, the Sandy housing-related outcomes of damage and property loss were highly significantly related to flood depth. Applicants had the best model fit (pseudo $R^2 = 0.71$), and overall the SVP has a better model fit with the outcomes (slightly higher pseudo R^2 and lower AIC) than the pillars of the SoVI, SVI, and weighted indices.

Table 11. SVP Profiles and Sandy Outcome Measures

	Ln %Applicants		Ln %Damaged Homes		Ln %Property Loss		Ln %Affected Renters	
	OLS	Lag	OLS	Error	OLS	Error	OLS	Lag
Water depth	0.02	0.01	0.02*	0.06*****	0.11*****	0.14*****	0.004	0.04
Profile 1 <i>low vulnerability</i>	-0.71***	-0.21***	-0.13*	-0.05	-0.30*	-0.006	-1.29*****	-0.73*****
Profile 2 <i>low, mobility & rent</i>	-0.17**	-0.03*	-0.41**	-0.19**	-0.08	-0.41	-0.35*	-0.14*
Profile 4 <i>high, age & special needs</i>	0.81***	0.48**	0.40**	0.44****	1.55*****	0.85****	1.96*****	1.37*****
Profile 5 <i>high vulnerability</i>	0.08*	0.05	0.16	0.10	0.59*	0.69*	0.42*	0.21
Adjusted/ Pseudo R ²	0.32	0.71	0.05	0.49	0.15	0.43	0.21	0.48
AIC	3748	2763	1923	1661	4448	4189	2256	2057
Moran's I	0.71*****		0.53*****		0.42*****		0.46*****	
Breusch-Pagan test	6.9	10.8	8.5	14.6	9.7	13.9	5.1	5.9
Multicollinearity condition number	6.4	6.4	7.3	7.3	6.4	6.4	7.2	7.2
N	1084	1084	625	625	1006	1006	608	608

* 0.01 ** 0.001 *** 0.0001 **** 0.00001

Contrary to the previous models, the SVP consistently performs as expected. The sign of the profile coefficients consistently aligns with the direction of hypothesized relationships between social vulnerability and disaster outcomes. The high vulnerability profiles (4 & 5) have only positive significant relationships with the Sandy outcomes and the low vulnerability profiles (1 & 2) have only negative significant associations. Yet, the only significant relationship with the “high vulnerability” profile (Profile 5) is with property value, a finding that may not be deemed significant after a Bonferroni correction. However, the “high vulnerability, age and special needs” profile (Profile 4) has robust positive links with all the outcomes, indicating a high level of construct validity.

DISCUSSION

The study objective was to assess the empirical validity of a wide range of social vulnerability models using FEMA impact data from Hurricane Sandy. Our comprehensive study design included multiple configurations of social vulnerability, multiple normalized outcome measures, statistical control for hazard severity, and justification for hypothesized relationships between vulnerability and outcomes. Still, attributing the cause of nonsignificant and/or negative statistical relationships is challenging, because they can occur due to at least three reasons:

- 1) Social vulnerability models or pillars that are weak proxies for social vulnerability processes
- 2) Outcome measures that inadequately represent social impacts
- 3) Flawed conceptual relationships between social vulnerability and disaster outcomes

We found that explanatory power varied substantially across social vulnerability models for a given disaster outcome, and across outcomes for a given model. The variation across models demonstrates that the configuration of a social vulnerability index has a strong influence on its empirical validity. The variation across outcome variables demonstrates that disaster outcome measures differ in their efficacy as indicators of human impact. Given the multidimensional nature of models and the current weak state of theorized relationships between social vulnerability and specific disaster outcomes, a search for the ideal unidimensional validation measure may prove elusive. There is a need for greater theoretical understanding to aid the construction of social vulnerability models and the selection of validation measures.

Model Validity

Previous empirical validation studies found SoVI to be a significant predictor of property loss, resident rate of return, and earthquake debris at the tract scale (Finch, Emrich, and Cutter 2010; Schmidlein et al. 2011), and for property loss at the county scale (Bakkensen et al. 2017). We found SoVI to be significantly related to the applicants and housing damage, but not property loss or affected renters. At the component level, previous studies found socioeconomic status to have significant relationships in the expected direction with property loss (Yoon 2012) and migration (Myers, Slack, and Singelmann 2008) at the county scale, and with housing damage at the tract scale (Burton 2010). In our case, we found that socioeconomic status failed to explain housing damage; it did for property loss but not in the expected direction. In short, there is little consistency across studies in social vulnerability dimensions and disaster outcome types (human, housing) found to be empirically valid.

We featured SoVI and SVI in the study design because they are the most prominent social vulnerability configurations, in both recognition and application. Analysis of Sandy data raises questions about the construct validity of each approach. For SoVI, there were notable differences in the explanatory power of the index and its constituent factors. The three components that explain the least amount of variance weren't significantly related to any of

the Sandy outcomes. One reason could be the reliance on the Kaiser criterion in the SoVI algorithm for determining the number of PCA components to retain. A leading critique of the Kaiser criterion is that it leads to over-extraction of components (Costello and Osborne 2005; Ledesma and Valero-Mora 2007). Although a higher factor count may be useful for descriptive purposes, including non-influential factors in the SoVI aggregation and/or assigning them equal weights may decrease both parsimony and explanatory power. More research is needed on this topic.

The explanatory power of the SVI was decidedly poor. The only significant relationships between the SVI and the Sandy outcomes were negative. This also occurred at the sub-index level, as most of the SVI themes had either non-significant or negative relationships with the disaster outcomes. The SVI is being promoted for use by public health officials and planners to identify socially vulnerable areas and populations (ATSDR 2018). But based on the analysis of Sandy outcomes, the construct validity of the SVI is weak.

Our study also evaluated alternatives to SoVI and SVI. Compared to these more established models, the weighted index based on expert knowledge had higher validity, explaining both human and housing related impacts. However, previous work found the rankings of hierarchical indices to be highly sensitive to the choice of weighting scheme (Tate 2012).

Although the SVP is not an aggregated measure, it had the highest explanatory power, significantly explaining all the Sandy outcomes in the expected direction. These findings indicate a very strong alignment of construct and measure. SVP provides both quantification and qualification of social vulnerability, not only indicating if each profile has a high or low vulnerability but also explaining why, thus unraveling the spatial distribution of the dominant vulnerability drivers. It directs attention to the locally convergent characteristics giving rise to social vulnerability, the intersectionality of social vulnerability drivers, and clusters of different vulnerability profiles, addressing why some places might be more vulnerable than others (Rufat 2013).

Statistical Model Specification

Our review of previous validation studies (Table 1) found that only half controlled for hazard severity, and only one accounted for spatial dependence. Our study analyzed twenty-eight multivariate regression models that included both. Water depth was statistically significant in most of these models, while spatial dependence was detected in every single one. If hazard influence and spatial dependence existed in the previous studies, but were not examined, the results may overinflate the strength of the relationships between social vulnerability and disaster outcomes.

Given the expanding use of social vulnerability measures, there is a broad need for further empirical validation, exploring other input variables, disaster outcomes, disaster stage and hazard types.

CONCLUSION

There is a mismatch between the rising application of social vulnerability models and understanding of their empirical validity. The current focus on descriptive and convergent studies is poorly suited to advance knowledge on this front. SoVI and the SVI are the most prominent social vulnerability configurations and are being promoted for use by public health officials and planners to identify socially vulnerable places and populations. However, our empirical analysis from Hurricane Sandy raises questions about their construct validity. The weighted model based on expert knowledge performed slightly better. The SVP had the highest empirical validity; we thus recommend the use of a profiles approach. Overall, there is a need for additional studies focused on the construct validity of social vulnerability models and measures. The number of such studies is surprisingly low, particularly given the rising profile of social vulnerability indices in all phases of hazard planning and decision making. Results prove that acute caution when using social vulnerability indices for decision making is still highly advised until a more complete assessment of validity is completed. Specifically troublesome is the lack of validation surrounding the freely available CDC-SVI model when compared to the more robust profiles approach.

Models can achieve validity both through empiricism and acceptance. The latter is currently the dominant mode. The failure to thoroughly validate the leading models against empirical disaster outcomes introduces a potentially serious problem: internal validity and convergent validity based on an accepted index are illusory if the model is empirically weak. Ideally model development should proceed by first producing internally robust and empirically valid models, and only then proceed to use the best performers as the baseline for convergence studies. The main research needs are thus to identify which social vulnerability models consistently explain disaster outcomes across studies. This will require empirical studies in different places, with variation among models and indicators, outcome measures, hazards, and temporal and analysis scale.

ACKNOWLEDGEMENTS

This research was partially supported by funds from the U.S. National Science Foundation (1333190 and 1707947), under the Humans, Disasters, and the Built Environment program.

REFERENCES

- Abbas, H. B., and J. K. Routray. 2014. Vulnerability to flood-induced public health risks in Sudan. *Disaster Prevention and Management* 23 (4):395-419.
- Adcock, R. 2001. Measurement validity: A shared standard for qualitative and quantitative research. *American political science review* 95 (3):529-546.
- Armstrong, R. A. 2014. When to use the Bonferroni correction. *Ophthalmic and Physiological Optics* 34 (5):502-508.
- ATSDR. *The Social Vulnerability Index (SVI)*. Agency for Toxic Substances and Disease Registry 2017 [cited August 8, 2017. Available from <http://svi.cdc.gov>.
- . *Fact Sheet: The Social Vulnerability Index (SVI)*. Agency for Toxic Substances and Disease Registry 2018 [cited March 8, 2018. Available from <https://svi.cdc.gov/Documents/FactSheet/SVIFactSheet.pdf>.

- Bakkensen, L. A., C. Fox-Lent, L. K. Read, and I. Linkov. 2017. Validating resilience and vulnerability indices in the context of natural disasters. *Risk Analysis* 37 (5):982-1004.
- Bankoff, G., G. Frerks, and D. Hilhorst. 2004. *Mapping vulnerability: disasters, development, and people*: Routledge.
- Bender, R., and S. Lange. 2001. Adjusting for multiple testing—when and how? *Journal of clinical epidemiology* 54 (4):343-349.
- Blaikie, P., T. Cannon, I. Davis, and B. Wisner. 1994. *At risk: natural hazards, people's vulnerability and disasters*. London: Routledge.
- Blake, E. S., T. B. Kimberlain, R. J. Berg, J. P. Cangialosi, and J. L. B. II. 2013. Tropical Cyclone Report: Hurricane Sandy (AL182012) 22 – 29 October 2012 National Hurricane Center
- Bland, J. M., and D. G. Altman. 1995. Multiple significance tests: the Bonferroni method. *Bmj* 310 (6973):170.
- BRACE. *Social Vulnerability Index*. University of Illinois at Chicago, School of Public Health 2017 [cited August 18, 2017. Available from <https://braceillinois.uic.edu/climate-change-health/people-places-at-risk/svi/>].
- Burton, C. G. 2010. Social Vulnerability and Hurricane Impact Modeling. *Natural Hazards Review* 11 (2):58-68.
- Cabin, R. J., and R. J. Mitchell. 2000. To Bonferroni or not to Bonferroni: when and how are the questions. *Bulletin of the Ecological Society of America* 81 (3):246-248.
- Chakraborty, J., G. A. Tobin, and B. E. Montz. 2005. Population Evacuation: Assessing Spatial Variability in Geophysical Risk and Social Vulnerability to Natural Hazards. *Natural Hazards Review* 6 (1):23-33.
- Chen, W. F., S. L. Cutter, C. T. Emrich, and P. J. Shi. 2013. Measuring Social Vulnerability to Natural Hazards in the Yangtze River Delta Region, China. *International Journal of Disaster Risk Science* 4 (4):169-181.
- City of Cedar Rapids. 2010. OTHER SOCIAL EFFECTS REPORT: City of Cedar Rapids, Iowa - Flood of 2008.
- Clark, G., S. Moser, S. Ratick, K. Dow, W. Meyer, S. Emani, W. Jin, J. Kasperson, R. Kasperson, and H. Schwarz. 1998. Assessing the Vulnerability of Coastal Communities to Extreme Storms: The Case of Revere, MA., USA. *Mitigation and Adaptation Strategies for Global Change* 3 (1):59-82.
- Colavito, K. H., S. Bjarnadottir, and Y. Li. 2017. Social Vulnerability Mapping Considering Hurricane Hazards in a Changing Climate. *Structures Congress 2017: Business, Professional Practice, Education, Research, and Disaster Management*:110-121.
- Comerio, M. C. 1998. *Disaster hits home: New policy for urban housing recovery*. Berkeley: University of California Press.
- Costello, A. B., and J. W. Osborne. 2005. Best practices in exploratory factor analysis: Four recommendations for getting the most from your analysis. *Practical assessment, research & evaluation* 10 (7):1-9.
- Cutter, S. L., B. J. Boruff, and W. L. Shirley. 2003. Social Vulnerability to Environmental Hazards. *Social Science Quarterly* 84 (1):242-261.
- Cutter, S. L., C. T. Emrich, D. P. Morath, and C. M. Dunning. 2013. Integrating social vulnerability into federal flood risk management planning. *Journal of Flood Risk Management* 6 (4):332-344.
- Cutter, S. L., J. T. Mitchell, and M. S. Scott. 2000. Revealing the Vulnerability of People and Places: A Case Study of Georgetown County, South Carolina. *Annals of the Association of American Geographers* 90 (4):713-737.
- DeMers, M. N. 2002. *GIS Modeling in Raster*. New York: Wiley.
- Edgeley, C. M., and T. B. Paveglio. 2017. Community recovery and assistance following large wildfires: The case of the Carlton Complex Fire. *International Journal of Disaster Risk Reduction* 25:137-146.
- Emrich, C. T. 2005. Social Vulnerability in US Metropolitan Areas: Improvements in Hazard Vulnerability Assessment., Department of Geography, University of South Carolina, Columbia, SC.

- Fekete, A. 2009. Validation of a social vulnerability index in context to river-floods in Germany. *Natural Hazards and Earth System Sciences* 9 (2):393-403.
- FEMA. *New Jersey Hurricane Sandy (DR-4086): Incident Period: October 26, 2012 - November 08, 2012, Major Disaster Declaration declared on October 30, 2012* 2012a [cited January 18, 2018. Available from <https://www.fema.gov/disaster/4086>.
- . *New York Hurricane Sandy (DR-4085): Incident Period: October 27, 2012 - November 09, 2012, Major Disaster Declaration declared on October 30, 2012* 2012b [cited January 18, 2018. Available from <https://www.fema.gov/disaster/4085>.
- . 2013. Hurricane Sandy FEMA After-Action Report, ed. FEMA.
- . *FEMA MOTF Hurricane Sandy Impact Analysis* 2015 [cited November 2, 2017. Available from <https://www.arcgis.com/home/item.html?id=307dd522499d4a44a33d7296a5da5ea0>.
- . *National Risk Index* 2018 [cited March 5, 2018. Available from <http://riskindex.atkinsatg.com/Home/Index>.
- Finch, C., C. T. Emrich, and S. L. Cutter. 2010. Disaster disparities and differential recovery in New Orleans. *Population and Environment* 31 (4):179-202.
- Fischer, A. P., and T. G. Frazier. 2018. Social Vulnerability to Climate Change in Temperate Forest Areas: New Measures of Exposure, Sensitivity, and Adaptive Capacity. *Annals of the American Association of Geographers* 108 (3): 658-678.
- Flanagan, B. E., E. W. Gregory, E. J. Hallisey, J. L. Heitgerd, and B. Lewis. 2011. A Social Vulnerability Index for Disaster Management. *Journal of Homeland Security and Emergency Management* 8 (1):23.
- Fothergill, A., E. G. Maestas, and J. D. Darlington. 1999. Race, ethnicity and disasters in the United States: A review of the literature. *Disasters* 23 (2):156-173.
- Frigerio, I., S. Ventura, D. Strigaro, M. Mattavelli, M. De Amicis, S. Mugnano, and M. Boffi. 2016. A GIS-based approach to identify the spatial variability of social vulnerability to seismic hazard in Italy. *Applied Geography* 74:12-22.
- Guillard-Goncalves, C., S. L. Cutter, C. T. Emrich, and J. L. Zezere. 2015. Application of Social Vulnerability Index (SoVI) and delineation of natural risk zones in Greater Lisbon, Portugal. *Journal of Risk Research* 18 (5):651-674.
- Hile, R., and T. J. Cova. 2015. Exploratory Testing of an Artificial Neural Network Classification for Enhancement of the Social Vulnerability Index. *Isprs International Journal of Geo-Information* 4 (4):1774-1790.
- Holand, I. S., and P. Lujala. 2013. Replicating and Adapting an Index of Social Vulnerability to a New Context: A Comparison Study for Norway. *Professional Geographer* 65 (2):312-328.
- Hou, J. D., J. Lv, X. Chen, and S. W. Yu. 2016. China's regional social vulnerability to geological disasters: evaluation and spatial characteristics analysis. *Natural Hazards* 84:S97-S111.
- HUD. 2016. *Sandy Damage Estimates Based on FEMA IA Registrant Inspection Data*. US Department of Housing and Urban Development 2015 [cited October 28, 2016 2016]. Available from <https://catalog.data.gov/dataset/sandy-damage-estimates-based-on-fema-ia-registrant-inspection-data>.
- Hurricane Sandy's Rising Costs*. [Editorial]. New York Times, November 27, 2012 2012 [cited. Available from <https://www.reuters.com/article/us-storm-sandy-deaths/factbox-storm-sandy-blamed-for-at-least-132-deaths-in-u-s-canada-idUSBRE8AF0ZX20121116>.
- Jones, B., and J. Andrey. 2007. Vulnerability Index Construction: Methodological Choices and Their Influence on Identifying Vulnerable Neighborhoods. *International Journal of Emergency Management* 4 (2):269-295.
- Khunwishit, S., and D. A. McEntire. 2012. Testing Social Vulnerability Theory: A Quantitative Study of Hurricane Katrina's Perceived Impact on Residents living in FEMA Designated Disaster Areas. *Journal of Homeland Security and Emergency Management* 9 (1):17.
- Laska, S., and B. H. Morrow. 2006. Social vulnerabilities and Hurricane Katrina: an unnatural disaster in New Orleans. *Marine technology society journal* 40 (4):16-26.
- Lawal, O., and S. B. Arokoyu. 2015. Modelling social vulnerability in sub-Saharan West Africa using a geographical information system. *Jamba-Journal of Disaster Risk Studies* 7 (1):11.

- Ledesma, R. D., and P. Valero-Mora. 2007. Determining the number of factors to retain in EFA: An easy-to-use computer program for carrying out parallel analysis. *Practical assessment, research & evaluation* 12 (2):1-11.
- Liu, D. L., and Y. Li. 2016. Social vulnerability of rural households to flood hazards in western mountainous regions of Henan province, China. *Natural Hazards and Earth System Sciences* 16 (5):1123-1134.
- Morgan, J. F. 2007. p Value fetishism and use of the Bonferroni adjustment. *Evidence-based mental health* 10 (2):34.
- Mustafa, D., S. Ahmed, E. Saroch, and H. Bell. 2011. Pinning down vulnerability: from narratives to numbers. *Disasters* 35 (1):62-86.
- Myers, C. A., T. Slack, and J. Singelmann. 2008. Social vulnerability and migration in the wake of disaster: the case of Hurricanes Katrina and Rita. *Population and Environment* 29 (6):271-291.
- O'Leary-Kelly, S. W., and R. J. Vokurka. 1998. The empirical assessment of construct validity. *Journal of Operations Management* 16 (4):387-405.
- Phillips, B. D. 2009. *Disaster recovery*. Boca Raton: CRC press.
- Reckien, D. 2018. What is in an index? Construction method, data metric, and weighting scheme determine the outcome of composite social vulnerability indices in New York City. *Regional Environmental Change* 18 (5):1439-1451.
- Remo, J. W. F., N. Pinter, and M. Mahgoub. 2016. Assessing Illinois's flood vulnerability using Hazus-MH. *Natural Hazards* 81 (1):265-287.
- Respaut, R., D. Skariachan, A. Longstreth, and B. Goldberg. 2012. Factbox: Storm Sandy blamed for at least 132 deaths in U.S., Canada. *Reuters*, November 12, 2012.
- Roncancio, D. J., and A. C. Nardocci. 2016. Social vulnerability to natural hazards in Sao Paulo, Brazil. *Natural Hazards* 84 (2):1367-1383.
- Rufat, S. 2013. Spectroscopy of Urban Vulnerability. *Annals of the Association of American Geographers* 103 (3):505-525.
- Rufat, S., E. Tate, C. G. Burton, and A. S. Maroof. 2015. Social vulnerability to floods: Review of case studies and implications for measurement. *International Journal of Disaster Risk Reduction* 14:470-486.
- Rygel, L., D. O'Sullivan, and B. Yarnal. 2006. A Method for Constructing a Social Vulnerability Index: An Application to Hurricane Storm Surges in a Developed Country. *Mitigation and Adaptation Strategies for Global Change* 11 (3):741-764.
- Saldajeno, P. B., L. M. Florece, R. D. Lasco, and M. T. H. Velasco. 2012. Vulnerability Assessment of Upland Communities in Sibalom Natural Park, Antique, Using Capital-based Approach. *Journal of Environmental Science and Management* 15 (2):1-12.
- Salkowe, R. S., and J. Chakraborty. 2009. Federal disaster relief in the US: The role of political partisanship and preference in presidential disaster declarations and turndowns. *Journal of Homeland Security and Emergency Management* 6 (1).
- Santos, M., L. del Rio, and J. Benavente. 2013. GIS-based approach to the assessment of coastal vulnerability to storms. Case study in the Bay of Cadiz (Andalusia, Spain). *Journal of Coastal Research*:826-831.
- SCDRO. 2017. South Carolina Action Plan for Disaster Recovery, ed. South Carolina Disaster Recovery Office: South Carolina Department of Commerce.
- Schmidlein, M. C., R. C. Deutsch, W. W. Piegorsch, and S. L. Cutter. 2008. A sensitivity analysis of the Social Vulnerability Index. *Risk Analysis* 28 (4):1099-1114.
- Schmidlein, M. C., J. M. Shafer, M. Berry, and S. L. Cutter. 2011. Modeled earthquake losses and social vulnerability in Charleston, South Carolina. *Applied Geography* 31 (1):269-281.
- Shmueli, G. 2010. To explain or to predict? *Statistical science* 25 (3):289-310.
- Siagian, T. H., P. Puhadi, S. Suhartono, and H. Ritonga. 2014. Social vulnerability to natural hazards in Indonesia: driving factors and policy implications. *Natural Hazards* 70 (2):1603-1617.
- Solangaarachchi, D., A. L. Griffin, and M. D. Doherty. 2012. Social vulnerability in the context of bushfire risk at the urban-bush interface in Sydney: a case study of the Blue Mountains and Ku-ring-gai local council areas. *Natural Hazards* 64 (2):1873-1898.

- Spielman, S. E., D. Folch, and N. Nagle. 2014. Patterns and causes of uncertainty in the American Community Survey. *Applied Geography* 46 (0):147-157.
- Tapsell, S. M., E. C. Penning-Rowsell, S. M. Tunstall, and T. Wilson. 2002. Vulnerability to flooding: health and social dimensions. *Philosophical Transactions of the Royal Society of London A: Mathematical, Physical and Engineering Sciences* 360 (1796):1511-1525.
- Tate, E. 2012. Social vulnerability indices: a comparative assessment using uncertainty and sensitivity analysis. *Natural Hazards* 63 (2):325-347.
- . 2013. Uncertainty Analysis for a Social Vulnerability Index. *Annals of the Association of American Geographers* 103 (3):526-543.
- Tate, E., A. Strong, T. Kraus, and H. Y. Xiong. 2016. Flood recovery and property acquisition in Cedar Rapids, Iowa. *Natural Hazards* 80 (3):2055-2079.
- Wu, S. Y., B. Yarnal, and A. Fisher. 2002. Vulnerability of coastal communities to sea level rise: a case study of Cape May County, New Jersey, USA. *Climate Research* 22:255-270.
- WVDC. 2017. West Virginia Community Development Block Grant Disaster Recovery Action Plan, ed. West Virginia Department of Commerce. Charleston, WV.
- Yoon, D. K. 2012. Assessment of social vulnerability to natural disasters: a comparative study. *Natural Hazards* 63 (2):823-843.
- Zhang, N., and H. Huang. 2013. Social vulnerability for public safety: A case study of Beijing, China. *Chinese Science Bulletin* 58 (19):2387-2394.