Risk Analysis, Vol. 39, No. 9, 2019 DOI: 10.1111/risa.13343

Quantifying Community Resilience Using Hierarchical Bayesian Kernel Methods: A Case Study on Recovery from Power Outages

Jin-Zhu Yu [1] **and Hiba Baroud** [1,2,*]

The ability to accurately measure recovery rate of infrastructure systems and communities impacted by disasters is vital to ensure effective response and resource allocation before, during, and after a disruption. However, a challenge in quantifying such measures resides in the lack of data as community recovery information is seldom recorded. To provide accurate community recovery measures, a hierarchical Bayesian kernel model (HBKM) is developed to predict the recovery rate of communities experiencing power outages during storms. The performance of the proposed method is evaluated using cross-validation and compared with two models, the hierarchical Bayesian regression model and the Poisson generalized linear model. A case study focusing on the recovery of communities in Shelby County, Tennessee after severe storms between 2007 and 2017 is presented to illustrate the proposed approach. The predictive accuracy of the models is evaluated using the log-likelihood and root mean squared error. The HBKM yields on average the highest out-of-sample predictive accuracy. This approach can help assess the recoverability of a community when data are scarce and inform decision making in the aftermath of a disaster. An illustrative example is presented demonstrating how accurate measures of community resilience can help reduce the cost of infrastructure restoration.

KEY WORDS: Community resilience; hierarchical Bayesian kernel model; power outage; predictive accuracy; stochastic dominance

1. INTRODUCTION

In the aftermath of a natural disaster, a loss in the performance of infrastructure systems can severely affect communities relying on these services. Evaluating how communities recover after the disruption of an infrastructure is critical to strategically inform restoration activities and resource allocation.

[1]Department of Civil and Environmental Engineering, Vanderbilt University, Nashville, TN, USA.
[2]Department of Earth and Environmental Sciences, Vanderbilt University, Nashville, TN, USA.
*Address correspondence to Hiba Baroud, Department of Civil and Environmental Engineering, Vanderbilt University, 274 Featheringill Hall, Nashville, TN 37212, USA; hiba.baroud@vanderbilt.edu.

A challenge in quantifying and predicting community resilience is the lack of recovery data from historical events. Traditional statistical learning tools rely on large data sets to provide accurate estimates (Yan & Haimes, 2010). As such, data scarcity on community recovery patterns results in challenges in measuring and predicting resilience metrics within an acceptable level of accuracy using existing statistical models. In this article, the rate of recovery in the aftermath of a disaster is used as the metric for community resilience. Such metric has been defined and used as a key indicator for measuring community recovery in prior studies on disaster management (Liu, Shi, Lu, & Wang, 2017). This work develops a new statistical learning method, the hierarchical Bayesian kernel model (HBKM), which integrates

the Bayesian property that updates predictions as data are dynamically obtained, the kernel function that makes nonlinear data more manageable, and the hierarchical property that allows strength and information borrowing across categories of data. The proposed model works well in cases where data are scarce and heterogeneous, which is common in disaster scenarios. This method is applied to quantify a community's recovery rate from power outages and is compared to other statistical methods.

The objective of this article is to evaluate the use of statistical methods to quantify and predict the recovery rate of communities when data are scarce. Cross-validation techniques are employed to assess the predictive accuracy of the models using the log-likelihood and root mean squared error (RMSE). A real-world case study is presented in which the recovery rate from power outages is modeled using data collected from the five most severe storms in Shelby County in Tennessee between 2007 and 2017.

The remainder of this article is structured as follows. Section 2 outlines prior work on community recovery modeling and Bayesian methods, the proposed model and comparative methods are introduced in Section 3, data collection and processing are presented in Section 4, Section 5 covers the results and discussion, and the conclusion is provided in Section 6.

2. BACKGROUND

2.1. Community Resilience

Community resilience is referred to as the capacity of a community to prepare for, respond to, and recover from a disruptive event. To evaluate a community's ability to respond and recover, it is important to define and identify metrics that measure community resilience.

Definitions of community recovery vary in the literature; however, they can be summarized into three categories: (1) returning to "predisaster" conditions, (2) attaining what would have occurred "without" the disaster, or (3) reaching a new stable state that may be different from the first two categories (Chang, 2010). In this article, full recovery—depending on the type of disaster, type of indicators, and characteristics of communities—is defined as either returning to predisaster conditions or reaching a new state that can be better or worse than the predisaster conditions.

Since community recovery in the aftermath of a disaster is a dynamic, multidimensional, and uncertain process (Chang, 2010; Yan & Haimes, 2010), measuring the recovery rates across all dimensions of a community is complex. In this article, the focus is on the effect of infrastructure restoration on the recovery of a community. Examples of infrastructure restoration include the number of public facilities and lifelines (e.g., bus, sewer and water systems, gas, and electricity) restored and buildings repaired. The resilience of these systems is used to calculate the recovery rate of a community. Certain infrastructure recovery indicators can be used as the basis of a community's resilience since they have significant and direct impacts on the social and economic recovery (Miles & Chang, 2003).

A wide range of community resilience indicators and metrics have been proposed in the literature and applied in practice. Although various metrics have been developed to measure the postdisaster recovery for different types of hazards (Horney, Dwyer, Aminto, Berke, & Smith, 2017), little research has been done to apply them to quantifying postdisaster community recovery. Liu et al. (2017) develop a qualitative approach to assess community recovery from earthquakes considering four dimensions (population, economic, building, and infrastructure) by extending the concepts of the resilience triangle; however, the method presented in their work is qualitative. Gillespie-Marthaler, Nelson, Baroud, and Abkowitz (2019) provide a comprehensive review of indicators and metrics with a classification scheme enabling the identification and selection of metrics for research and implementation. Although the majority of research and practice have implemented qualitative metrics, the use of quantitative metrics, although challenging, provides a number of advantages. Quantifying community resilience allows the use of data-driven methods to evaluate and predict resilience of communities, presenting future opportunities to integrate machine learning, artificial intelligence, and cyber-physical technologies in disaster preparedness and management. The work presented in this article constitutes a first step in the quantification of community resilience in which the recovery metrics are measured as a function of time or initial vulnerability.

2.2. Methodological Background

This section provides the background of hierarchical Bayesian (HB) models and Bayesian kernel methods.

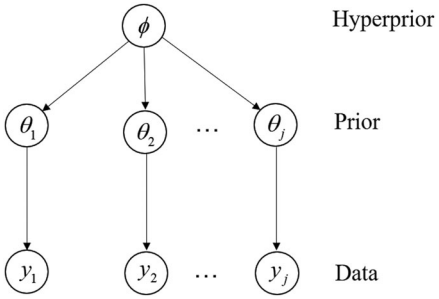


Fig. 1. Graphical representation of a hierarchical Bayesian model.

2.2.1. *Hierarchical Bayesian Methods*

Hierarchical methods are composed of multiple parameters that are representative of the structure of the problem or the data whereby a joint probability model for these parameters is used to reflect their dependence (Gelman et al., 2013). Specifically, to estimate the parameters of interest, HB methods use a multilevel model that relates the data to the probability distribution of the parameter using Bayes's theorem, shown in Equation (1).

$$p(\theta|y) = \frac{p(y|\theta)p(\theta)}{p(y)} \tag{1}$$

In Equation (1), θ represents the parameter of interest and y represents the data. Bayes's theorem is used to integrate the submodels under each level with the data to account for the uncertainty of the parameters of interest (Allenby, Rossi, & McCulloch, 2005). The result of this integration is the posterior distribution, which is a two-level HB model that is represented mathematically in Equation (2). Fig. 1 provides a graphical representation of the structure of a two-level hierarchical model.

$$p(\theta_j|\phi, y_j) = \frac{p(y_j|\phi, \theta_j)p(\theta_j, \phi)}{p(\phi, y_j)} \tag{2}$$

In hierarchical models, the data are categorized into groups, y_j represents one group of data that follows a distribution with parameter θ_j. Each θ_j in the prior is described by a hyperprior distribution with a set of hyperparameters, denoted as ϕ. Note that the distributions of different groups of data y_j depend on a common hyperprior, ϕ, only through θ_j. In the nonhierarchical model, ϕ is assumed to be deterministic. In the hierarchical model, however, the uncertainty of ϕ is accounted for using a probability distribution. This multilevel structure reduces the subjectivity of making assumptions about the priors by assigning a distribution to the prior parameter and updating it with observed data. The hyperpriors are generally specified by using expert elicitation, incorporating existing information, or using a proper noninformative distribution (Gelman, 2006).

HB methods provide advantages over nonhierarchical models: (i) they avoid overfitting, providing better predictive accuracy with fewer parameters (Gelman et al., 2013) and (ii) they are flexible in combining information across different groups of data. The latter property, which is unique to HB models, is referred to as "strength borrowing" and can be demonstrated using analytical derivations of the posterior calculations of the parameters (Yan & Haimes, 2010). Suppose a particular data group, j, lacks sufficient historical information to predict θ_j accurately. The HB model addresses this problem by borrowing strength from data collected in a similar subset of the hierarchical structure. All θ_js are governed by the same underlying distribution that will be updated with the entire data set y. As a result, the posterior distribution of each θ_j does not solely depend on the direct data point y_j; instead, it is dependent on the commonly updated hyperprior. The posterior distribution calculation is shown in Equation (3) (Yan & Haimes, 2010).

$$\begin{aligned} p(\theta_j|\phi, y) &= \frac{p(y|\phi, \theta_j)p(\theta_j, \phi)}{p(\phi, y)} \\ &\propto p(y|\phi, \theta_j)p(\theta_j, \phi) \\ &= p(y_j|\theta_j)p(\theta_j|\phi)p(\phi) \\ &\propto p(y_j|\theta_j)p(\theta_j|\phi) \end{aligned} \tag{3}$$

Although the HB method is powerful, the solution techniques are often computationally intensive, especially when the hierarchical models are too complicated to generate a closed form of posterior distributions. As such, Markov chain Monte Carlo (MCMC) methods are used by replacing the analytical solution with repetitive calculations that computers can perform over a large number of iterations (Allenby, Rossi, & McCulloch, 2005). Recent advancement in computational tools and efficiency of Monte Carlo algorithms resulted in HB models becoming more applicable and popular (Browne & Draper, 2006). In this article, rstan (Guo et al., 2016) is used to perform MCMC for hierarchical models.

2.2.2. *Bayesian Kernel Methods*

Bayesian kernel methods are a class of models in which kernel functions are integrated with Bayesian

modeling, producing probabilistic rather than deterministic estimates of the parameter of interest, and thus accounting for uncertainty in the underlying data (MacKenzie, Trafalis, & Barker, 2014). The approach was first developed in machine learning using the Gaussian kernel model (Schölkopf & Smola, 2001), which assumes a Gaussian prior for the response variable and uses the kernel matrix as the covariance matrix estimated from the explanatory variables.

Non-Gaussian extensions of the Bayesian kernel method result in models that are applicable to a wider range of problems. Montesano and Lopes (2009) develop the beta kernel model that uses a beta distribution as the conjugate prior distribution to predict the probability of a robot successfully grasping an object. MacKenzie et al. (2014) highlight two issues associated with Gaussian kernel models in binary classification: (i) imbalanced data sets and (ii) longitudinal data sets. Specifically, the complexity of Gaussian kernel models limits their ability to update parameters quickly as new data become available. They improve and generalize the beta Bayesian kernel model (BKM) to deal with heavily imbalanced data sets by adding weights to the kernel matrix in the computation of the posterior. The beta kernel model was found to outperform the relevance vector machine (RVM) and LASVM (an incremental learning version of support vector machine) if the data contain 50 or fewer data points. Also, the model frequently performs better than RVM and LASVM even if more data are available (Baroud, Francis, & Barker, 2016; MacKenzie et al., 2014). Other applications of the beta BKM include the prediction and ranking of resilience-based importance measures in infrastructure networks (Baroud & Barker, 2018) and the evaluation of the likelihood of global supply chain disruptions (Baroud et al., 2016).

Baroud, Barker, and Lurvey (2013) further extend non-Gaussian BKMs to account for count data by developing the Poisson BKM that integrates prior information from experts' knowledge with historical data and attribute information to calculate the probability distribution of the rate of occurrence of a particular event. The model assumes that the rate of occurrence, λ, which has a Poisson likelihood function, follows a gamma (α, β) prior distribution. Since gamma distribution is a conjugate prior for the Poisson distribution, the posterior distribution of λ is also a gamma distribution. The posterior of λ is estimated based on the kernel matrix that measures the similarities of the explanatory variables between the test data (new data) and the training data (historical data). The kernel function is represented by K_j, which is the $n_j^{\text{test}} \times n_j^{\text{train}}$ kernel matrix for the jth group of the data, $\boldsymbol{Y}_j$ is a $n_j^{\text{train}} \times 1$ vector containing the output data associated with n_j^{train} observations of $\boldsymbol{X}$, and $\boldsymbol{V}_j$ is an $n_j^{\text{train}} \times 1$ vector of ones. The posterior distribution of λ is then gamma (α'_j, β'_j) with the parameters of the posterior computed using Equations (4) and (5).

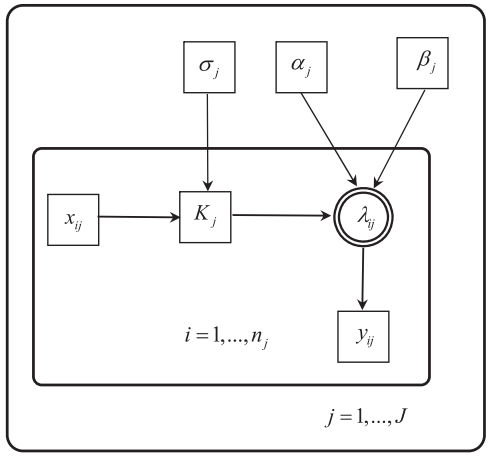


Fig. 2. Graphical representation of the Bayesian kernel model.

$$\alpha'_j = K_j Y_j + \alpha_j \tag{4}$$

$$\beta'_j = K_j V_j + \beta_j \tag{5}$$

The parameters $\boldsymbol{\alpha}_j$ and $\boldsymbol{\beta}_j$ in the prior distribution of λ are deterministic and are determined based on expert elicitation or empirically using the method of moments. A graphical representation of the Poisson BKM is shown in Fig. 2. In this graphical representation (and all graphical representations shown thereafter in this article), the directed arrows represent direct influence between the nodes, the square refers to deterministic values, and the circle refers to probabilistic variables. For this particular Poisson BKM displayed in Fig. 2, the data y are divided into J groups with n_j data points in each group. Within the jth group, y_{ij} is governed by λ_{ij} following a posterior distribution determined by parameters α_j and β_j and the kernel matrix K. The kernel matrix is calculated using the value of predictors and the tuned kernel parameter σ_j of each group.

Later, Baroud (2015) conducted a comparative analysis of the Poisson BKM and the Poisson and negative binomial generalized linear models (GLMs) using several goodness-of-fit and predictive accuracy

measures. The results of the analysis show higher predictive accuracy for the Poisson BKM for small data with a small number of predictors.

2.3. Applications of HB Models

HB models have been widely used in risk analysis due to their ability to account for the uncertainty in priors and their flexibility in combining information from multiple sources. This class of models provides accurate estimations of the parameters of interest along with their posterior distributions (Polettini & Stander, 2004; Qin, Ivan, Ravishanker, & Liu, 2005) and synthesizes information across several data sources (Coull, Mezzetti, & Ryan, 2003; MacNab, 2003) into a unified framework. Various studies have used HB models in the field of risk analysis, such as dose administration, data protection, and transportation safety (Andrade & Teixeira, 2015; Faes, Geys, Aerts, & Molenberghs, 2006; Qin, Ivan, Ravishanker, & Liu, 2005). In these studies, HB models provide a full assessment of risk by accounting for the hierarchical and clustered nature of the data. However, few studies can be found that explicitly make use of HB models to handle problems with scarce data. Migon and Moura (2005) use HB models to improve the estimation of model parameters for the number of insurance claims by borrowing strength using age as a predictor when information on a patient's medical record is not available or very difficult to obtain. Other studies adopt HB methods to deal with data sparseness problems. Yan and Haimes (2010) apply HB models as well as HB methods with cross-classified random effects (CHB) to address the data sparseness problem in systems risk and safety analysis for extreme and rare events. They demonstrate that the HB and CHB models provide flexible and theoretically sound frameworks for integrating various sources of information and complementing the sparse data in risk and safety analysis. Another study is focused on modeling species communities distributions (Ovaskainen & Soininen, 2011) where the hierarchical model combines species-specific models to overcome challenges of statistical power in the case of rare species.

Although HB models have been applied to sparse data situations, the use of HB methods in handling data scarcity is rarely considered in the literature. This article extends HB models to accommodate situations where historical data are not large enough to make accurate predictions using classical statistics models.

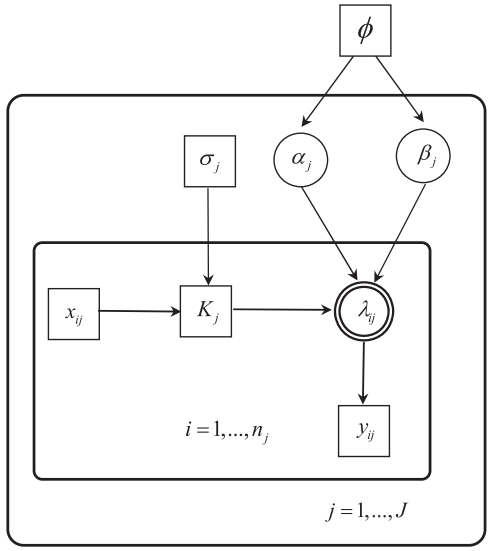


Fig. 3. Graphical representation of the hierarchical Bayesian kernel model.

3. MODELING APPROACH

In this section, HBKM is developed to analyze the rate of recovery of communities after a disaster. Also, two models, the HB regression model (HBRM) and the Poisson GLM, are presented that are used to assess and compare the performance of the proposed approach. The parameter of interest is the recovery rate that can be determined depending on the case study, data availability, and indicators of community resilience. In this article, the recovery rate is calculated as a function of the number of people without power.

3.1. Hierarchical Bayesian Kernel Model

HBKM integrates the HB model with the kernel function to allow for strength borrowing and to incorporate attribute information. Compared to the BKM, the HBKM improves the accuracy of the estimation of α_j and β_j in the prior distribution of λ using an extra layer of Bayesian updating. As a result, the values of α_j and β_j are probabilistic and are estimated using the full data set. The graphical representation of HBKM is shown in Fig. 3 where α_j and β_j are shown in a circle to indicate that they are random parameters described by the same hyperprior ϕ, whereas in Fig. 2, α_j and β_j are shown in a square to indicate that they are scalars. Note that the data set does not necessarily need to be grouped in the BKM because α_j and β_j are scalars.

The HB model used to obtain the updated hyperparameters in HBKM is given in Equation (6), where ϕ represents the common hyperprior following a half Cauchy distribution.

$$\begin{aligned} &y_{ij} \sim \text{Poisson}\left(\lambda_{ij}\right), \quad i = 1, 2, \ldots, n_j, \quad j = 1, 2, \ldots, J \\ &\lambda_{ij} \sim \text{Gamma}\left(\alpha_j, \beta_j\right) \\ &\alpha_j \sim \text{Cauchy}\left(0, 5\right) \quad (\alpha_j > 0), \quad \beta_j \sim \text{Cauchy}\left(0, 5\right) \quad (\beta_j > 0) \end{aligned} \tag{6}$$

The radial basis function is chosen to be the kernel function shown in Equation (7), where x_{ij}^{train} and x_{ij}^{test} represent the ith data points in the training and test sets of the jth group, respectively. This function is widely used as it generates a kernel matrix with (i) all elements between zero and one, facilitating the interpretation of a similarity measure between the covariates of data points, (ii) full rank (Schölkopf & Smola, 2001), and (iii) only one free parameter to be tuned to the optimal value, hence reducing computation time (Baroud, 2015).

$$k\left(x_{ij}^{\text{test}},\ x_{ij}^{\text{train}}\right) = e^{-\frac{\left\|x_{ij}^{\text{test}} - x_{ij}^{\text{train}}\right\|^2}{2\sigma_j^2}} \tag{7}$$

The parameter σ_j in the kernel function is tuned by maximizing the log-likelihood function shown in Equation (8), where the final term can be ignored in the tuning process since it does not depend on the parameter of interest.

$$L(\hat{\lambda}_j) = \sum_{i=1}^{n_j} \left[y_i \ln(\hat{\lambda}_j) - \hat{\lambda}_j - \ln\left(y_i!\right)\right] \tag{8}$$

In the hierarchical version of the Poisson BKM, Equations (4) and (5) become Equations (9) and (10), respectively, where α_j^* and β_j^* now represent the mean values of the updated distributions of α_j and β_j, respectively.

$$\alpha_j' = K_j Y_j + \alpha_j^* \tag{9}$$

$$\beta_j' = K_j V_j + \beta_j^* \tag{10}$$

3.2. Comparative Analysis

To investigate the performance of HBKM, a comparative analysis is conducted where the proposed model is compared with two other statistical models, GLM and HBRM. Cross-validation is used where a subset of the data is held out while the models are fitted to the rest of the data and then tested using the held out test set. Both models are presented in this subsection.

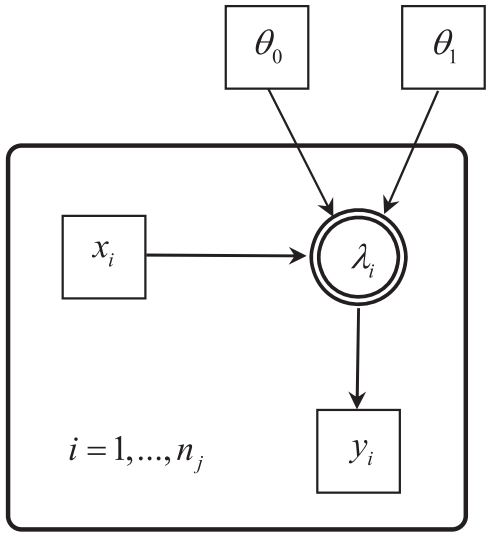


Fig. 4. Graphical representation of the generalized linear model.

The Poisson GLM is a classical approach used for count data. Each data point of y is fit with a Poisson distribution and the value of y given new attribute information is estimated using the mean rate of occurrence λ. The mean rate λ is thus the parameter of interest and the log of λ is expressed as a linear function of predictors, $\boldsymbol{x}$. In the case of one predictor, the coefficients θ_0 and θ_1 in the model are assumed to be deterministic and are calculated using the least square method. The model is shown in Equation (11) and Fig. 4.

$$\begin{aligned} &y_i \sim \text{Poisson}\left(\lambda_i\right), \quad i = 1, 2, \ldots, n_j \\ &\log\left(\lambda_i\right) = \theta_0 + \theta_1 x_i \end{aligned} \tag{11}$$

The other model used in the comparison is HBRM. This model, shown in Equation (12) and Fig. 5, where n_j is the number of data points in the jth group of the data set, also assumes that the log of the mean rate is a linear function of the predictors, $\boldsymbol{x}$. However, the coefficients in the log-linear function are assumed to be probabilistic and follow a noninformative normal distribution. Half-Cauchy (0, 5), a weakly informative distribution recommended by Gelman et al. (2013), is used to model the hyperparameters. The coefficients θ_j represent the rate of recovery in different groups of the data. The coefficients are indirectly correlated through the common hyperprior and the intercept θ_0 remains unchanged

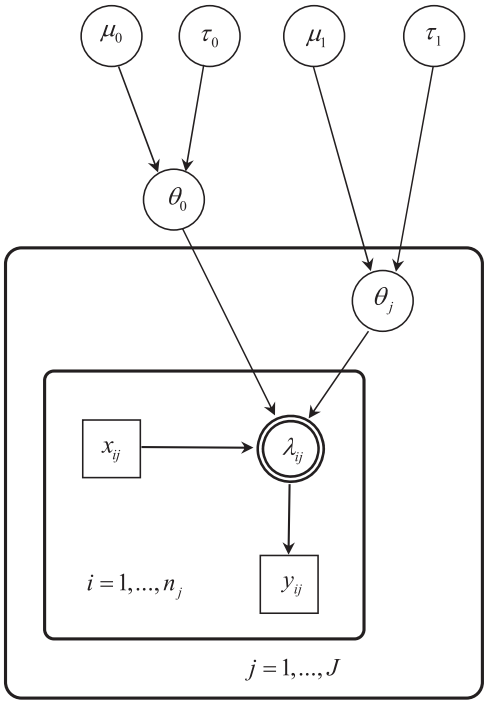


Fig. 5. Graphical representation of the hierarchical Bayesian regression model.

across all groups. MCMC is used to compute the posterior distribution of the parameters.

$$
\begin{aligned}
&y_{ij} \sim \text{Poisson}\left(\lambda_{ij}\right), \quad i = 1, 2, \ldots, n_j, \quad j = 1, 2, \ldots, J \\
&\log(\lambda_{ij}) = \theta_0 + \theta_j x_{ij} \\
&\theta_0 \sim N(\mu_0, \tau_0), \quad \theta_j \sim N(\mu_1, \tau_1) \\
&\mu_0 \sim N(0, 5), \quad \mu_1 \sim N(0, 5) \\
&\tau_0 \sim \text{Cauchy}\ (0, 5), \quad \tau_1 \sim \text{Cauchy}\ (0, 5) \quad (\tau_0 > 0, \tau_1 > 0)
\end{aligned} \tag{12}
$$

3.3. Performance Assessment

Cross-validation is used to evaluate the out-of-sample predictive error. Specifically, 35% of the data are used as the training set, 35% as the tuning set to optimize the parameter in the kernel function, and 30% as the test set. The samples are randomly selected over 100 iterations. In performing MCMC for the HB models, four chains are deployed with 3,000 iterations in each chain where 30% of the iterations are for warm-up and 70% are for sampling. All the models are also compared to having no model, which simply uses the mean of the training data set, that is, the training and tuning sets in each group, as the prediction for each group. Four model performance metrics are computed: (i) log-likelihood, (ii) RMSE, (iii) relative frequency of the highest log-likelihood, and (iv) relative frequency of the lowest RMSE across all iterations. The mean, probability density function (PDF), and cumulative density function (CDF) of the RMSE and log-likelihood are used to compare the performance of each model. In evaluating the models based on the CDF of the log-likelihood and RMSE, second-order stochastic dominance (SSD) is used to compare the distribution of the values of these metrics. SSD provides a comprehensive technique to compare CDFs. For two random variables X and Y with corresponding CDF F_X and F_Y, if $F_X(s) \leq F_Y(s)$ holds for any real number $s \in (-\infty, +\infty)$, then X has first-order stochastic dominance over Y, denoted as $Y \preceq X$. However, if Equation (13) holds, then X has SSD over Y, denoted as $Y \preceq^2 X$.

$$
\int_{-\infty}^{t} F_X(s)\, ds \leq \int_{-\infty}^{t} F_Y(s)\, ds \tag{13}
$$

In the domain of risk-based decision making, a rational risk-averse decisionmaker who prefers the expectation of a random return $E(X)$ to the random return X itself (Friedman & Savage, 1968) will choose the alternative that leads to X if a higher value of X is preferred and $Y \preceq X$ or $Y \preceq^2 X$. More details on how the first-order stochastic dominance and SSD imply risk aversion are provided in Nie, Wu, and Homem-de-Mello (2012). In evaluating the predictive accuracy of the models considered in the case study, the model for which the RMSE CDF is dominated by the RMSE CDFs of other models is considered to be the best since a smaller RMSE is preferred. Similarly, a higher value of the log-likelihood is preferred.

4. CASE STUDY

The proposed model is applied to a case study in which community resilience is related to the number of customers without power. The goal is to model the stochastic recovery of a community as a function of time after a storm and the intensity of the storm. The models are applied to assess the recovery from power outages in Shelby County, Tennessee. The population of Shelby County, which is approximately 934,600 (U.S. Census Bureau, 2016), is served by Memphis Light, Gas, and Water (MLGW), the largest three-service municipal utility in the United States. Shelby County, which includes Memphis, the second largest city in Tennessee, is vulnerable to hazards including earthquakes, flooding, and storms. Assessing the community recoverability can assist

Table I. Initial Number of People Without Power and Size of Data Set for Each Storm

Storm	Initial Number of People Without Power	Number of Data Points
S1 (May 2017)	188,000	14
S2 (June 2009)	133,000	9
S3 (April 2011)	63,500	6
S4 (April 2011)	64,000	6
S5 (August 2015)	67,842	6

local agencies and the utility in protecting customers against adverse impacts of disasters by improving preparedness for future events. In this case study, the recoverability of power after storms is explored to provide assistance to the utility company in strategic restoration.

Power outage data are collected from MLGW's news releases that are posted on the company's website and updated intermittently during and after a storm. The data set includes the number of people without power given the time after each of the 15 documented storms that hit Shelby County between 2007 and 2017. Data from the five most intense storms, referred to as S1, S2, S3, S4, and S5, are included in this case study.

Fig. 6 illustrates the community recovery rate in Shelby County after each of the five storms considered in this case study. As noted earlier, the proposed method is applicable to any definition of recovery rate. In this specific example, the original data of the number of people without power at a given time are translated into the number of people without power per 1,000 customers. The model yields the predicted number of people without power per 1,000 at the desired time, and the recovery rate is calculated as $1 - (\frac{C}{1,000})$, where C represents the number of customers without power obtained from the model. Note that for S1 and S2, the two strongest storms, the community took almost double the time to achieve full recovery.

Table I provides a summary of the initial number of people without power after each storm as well as the number of recorded data points during the recovery process.

The models presented in Section 3 are used to model the recovery rate from power outages. The number of groups corresponds to the variable J in Equations (6), (11), and (12), which is equal to five, that is, the number of storms considered. The number of data points available for each storm is represented by n_j. To obtain the posterior distribution efficiently, the MCMC technique is implemented. rstan is used to perform the MCMC simulation for HBKM and HBRM, which do not support missing values in the data matrix. To maintain consistency given the differences in data structures and sizes and address missing data, a nonlinear interpolation is used to impute the missing data points and ensure that all data subsets have the same structure and size. The nonlinear interpolation is based on the exponential or logarithmic model fitted for each subset of data representing one storm. Gaussian noise is added to each interpolated data point to minimize the sensitivity of the model performance to the interpolation approach and to have realistic imputed data assuming that observations usually fluctuate around a true model.

Table II. Example of a Data Set for One Storm (S4)

Time After the Storm (Hours)	Number of People Without Power After the Storm/10^3	
	Original Data	Data After Imputation
2.75	844	844
6.25	–	728
9.75	688	688
13.75	–	515
17.75	438	438
23.00	–	403
28.00	–	299
33.75	328	328
41.75	–	158
49.75	–	98
57.75	109	109
65.75	–	38
73.75	–	34
81.75	13	13

The mean of the Gaussian noise is 0 and the standard deviation is 10% of the fitted value. Table II provides an example of the original data along with the data after imputation. After imputing the missing data, each group in the data set has 14 data points, and the numbers of data points in the training set, tuning set, and test set are 5, 5, and 4, respectively.

5. RESULTS AND DISCUSSION

5.1. Model Accuracy

The proposed model in this research can ultimately help disaster responders and decisionmakers

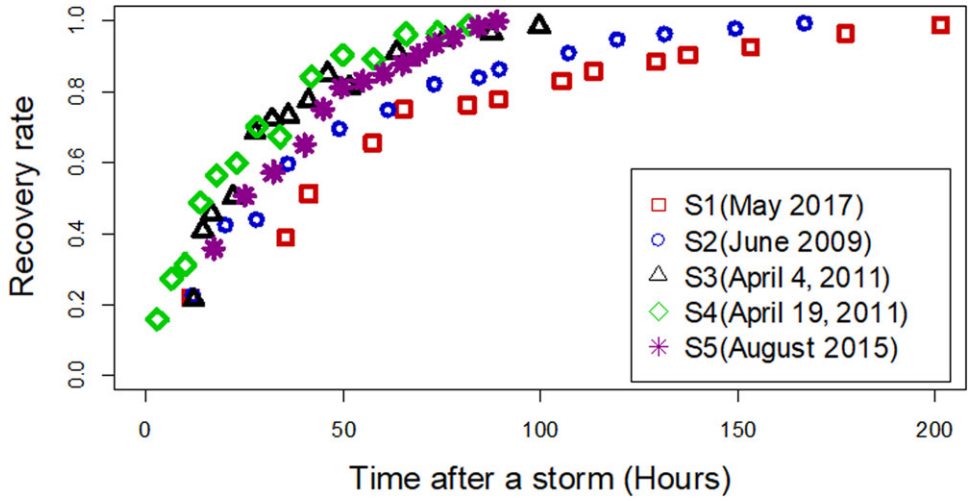


Fig. 6. Data of the community recovery rate for the five strongest storms.

Table III. Comparison on Mean Prediction Accuracy

Metric	HBKM	HBRM	GLM	No Model
Log-likelihood	25,120.0	24,559.3	24,900.5	23,059.9
RMSE	82.2	92.1	99.3	240.8

to understand how communities respond to and recover from disasters. As such, it is important that the model provides a high predictive accuracy whereby it is possible to forecast the recovery trajectory of a community given information on the storm intensity, depicted in this work by its length, as well as other community characteristics. This section investigates the predictive accuracy and computation time of the models using the power outage recovery data of Shelby County. Two evaluation metrics, the log-likelihood and RMSE, are computed using the cross-validation approach outlined earlier. One iteration of the cross-validation uses 8,400 samples to build the posterior distribution and calculate one value for the RMSE and log-likelihood using the mean of the posterior distribution as a point estimate. This process is repeated for 100 iterations and the mean values of the log-likelihood and RMSE for each model are shown in Table III, while the PDFs and CDFs of the two metrics are provided in Figs. 7 and 8, respectively.

Results indicate that, for the Shelby County power recovery data, all three models exhibit much higher predictive accuracy than having "no model," with HBKM providing the best predictive accuracy. Specifically, HBKM has an average RMSE of 82.2, whereas HBRM and GLM have an average RMSE of 92.1 and 99.3, respectively. Having no model yields the largest mean RMSE.

Although mean values of the evaluation metrics provide a general idea of the performance of the models, a more comprehensive way to assess these metrics across methods is to evaluate their distributions. From the CDFs of the log-likelihood shown in Fig. 7, it can be seen that having any model is better than not having a model since higher values of the CDF are preferred. The CDF for the "no model" is consistently above the CDFS of the other three models. However, differentiating between the three models' overall performance is not trivial and cannot be done visually. Fig. 8 displays the plots of the PDF and CDF for the RMSE of each model. GLM produces the lowest variance in the values of RMSE while HBRM has the largest variance. Similar to the log-likelihood distributions, the CDFs of RMSE show that having any model is better than not having a model because low values of RMSE are preferred. The distribution corresponding to "no model" has first-order stochastic dominance over the three models (i.e., the CDF of "no model" is consistently below the CDFs of other models). However, among the three models, the dominance changes as a function of the value of RMSE whereby GLM dominates for RMSE values less than 100 while HBRM dominates otherwise.

To compare the overall performance of the models across all possible values of RMSE and log-likelihood, it is necessary to examine the integral of their CDF to determine the SSD.

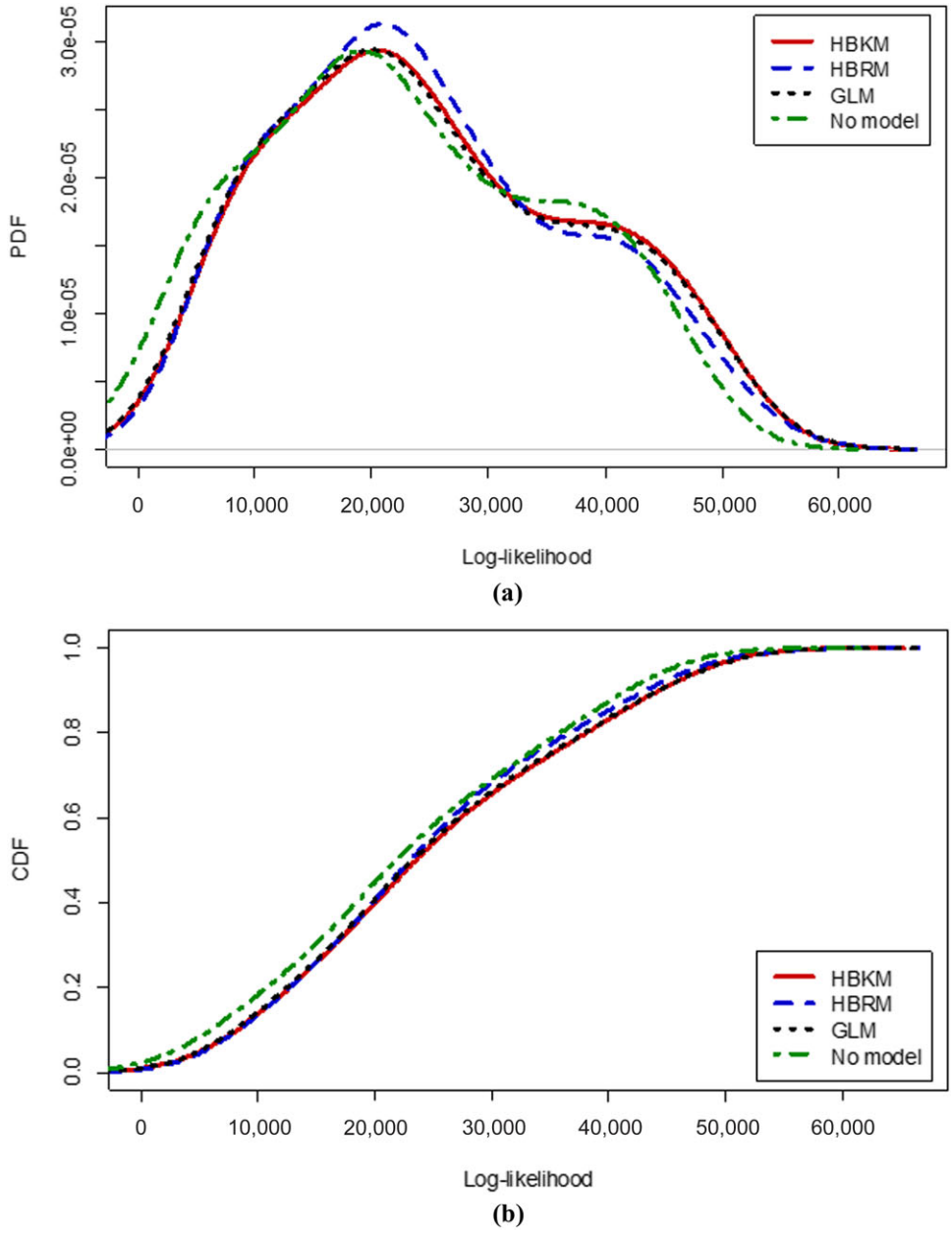


Fig. 7. (a) PDF and (b) CDF of the log-likelihood for each model.

Fig. 9 displays the integral of the CDF of log-likelihood and RMSE. The integral of the CDF of the log-likelihood of HBKM is consistently smaller than that of HBRM and GLM, as a result, HBKM has second-order dominance over the other models and is, hence, preferred. The integral of the CDF of RMSE for HBKM is larger than that of HBRM and GLM for large RMSE values (higher than 200). As such, HBRM and GLM have SSD over HBKM for large values of RMSE, indicating that while HBRM is more accurate for small values of error, HBKM does a better job at keeping the error within an acceptable range.

In addition to the predictive accuracy, the computation time of the different models is examined. Specifically, the mean computation time for one iteration of each model averaged over 50 realizations is provided in Table IV. It is clear that GLM

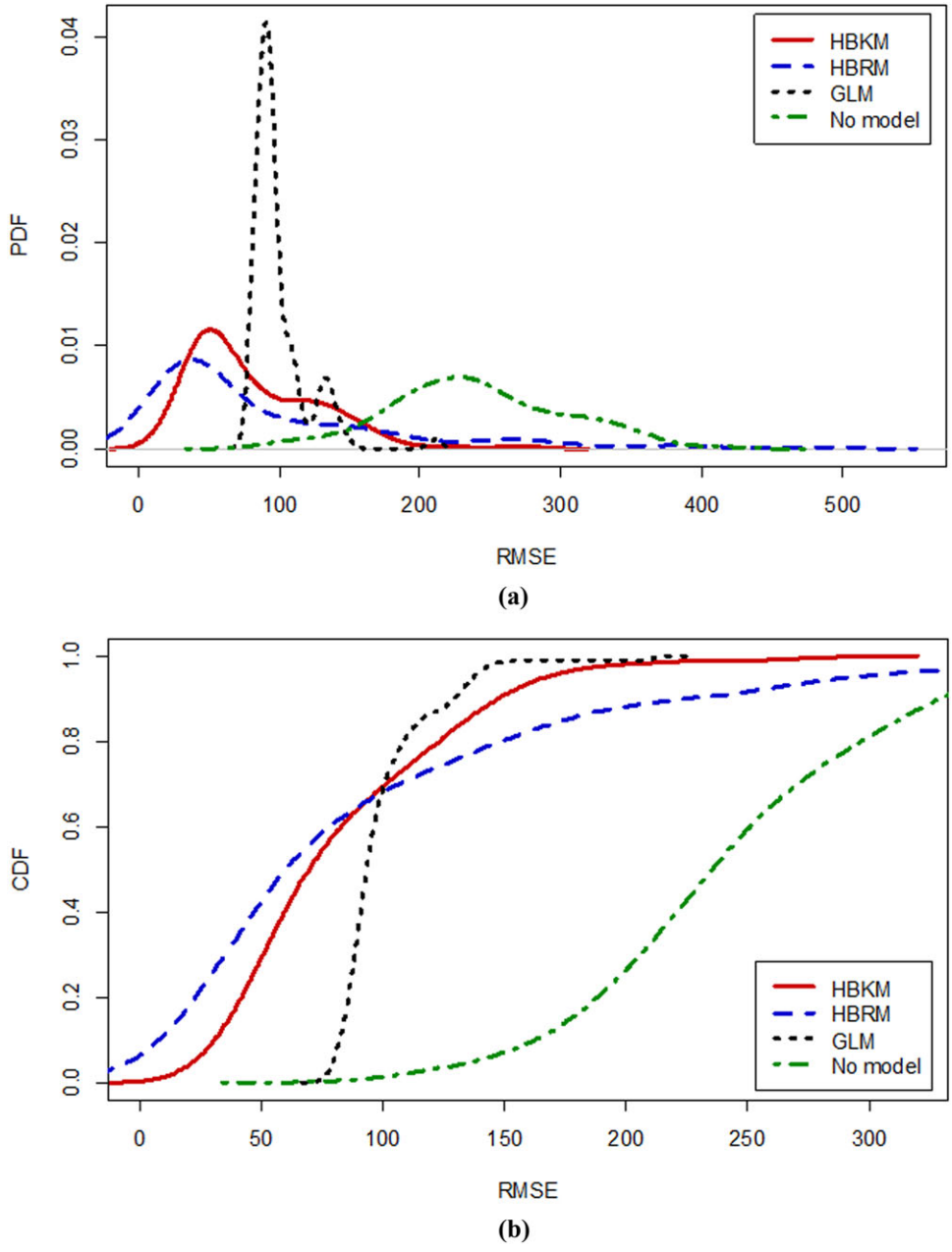
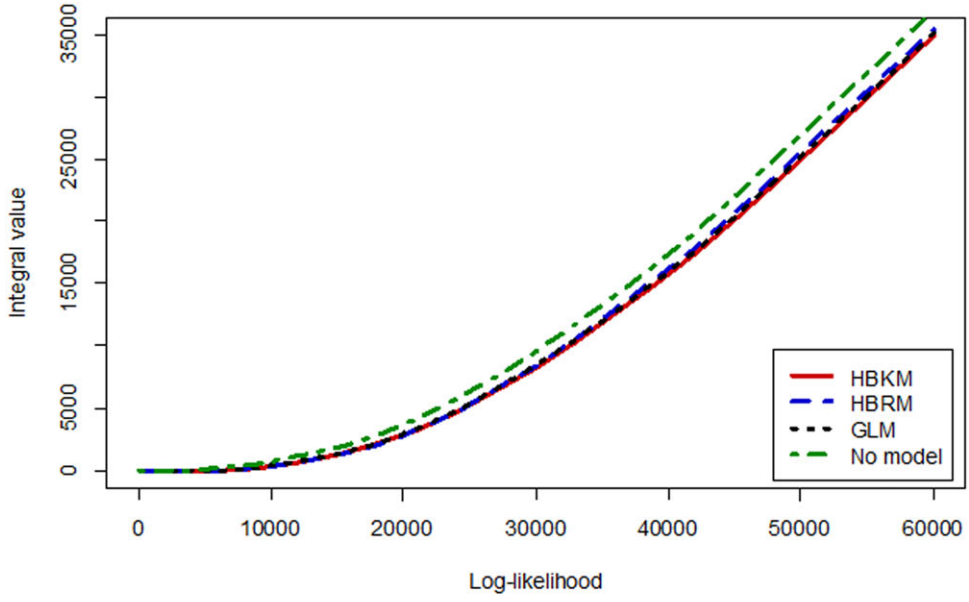


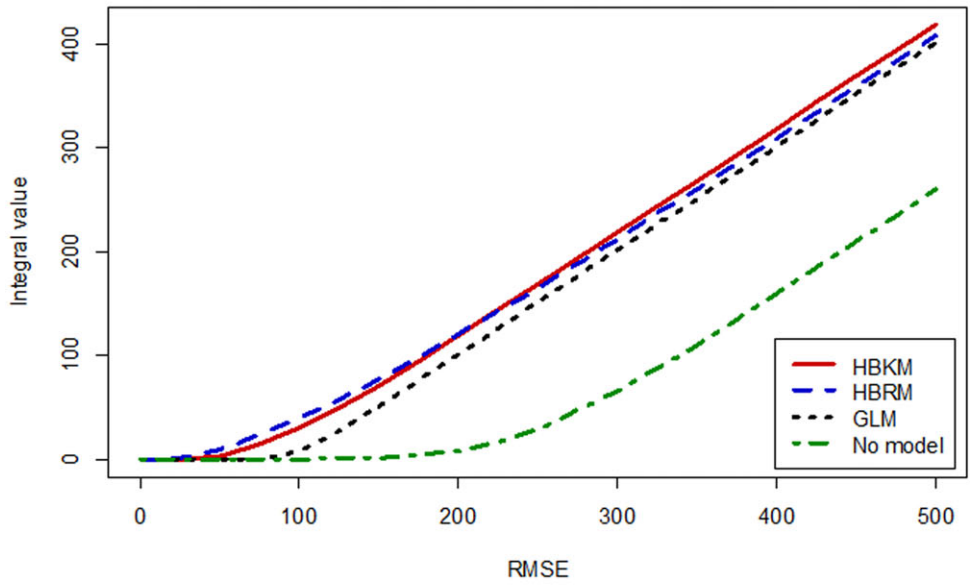
Fig. 8. (a) PDF and (b) CDF of the RMSE for each model.

and having "no model" provide the fastest computation results. The MCMC simulation for the hierarchical models contributes to a longer computation time. HBRM is the most computationally intensive method, requiring slightly less than one minute on average for each iteration, while HBKM provides similar if not better predictive results for a significantly smaller amount of computation time. When such models are scaled to larger systems and interconnected networks, the computation time combined with the predictive accuracy will make a significant difference in the choice of the model.

Finally, the frequencies of achieving the best prediction out of the 100 iterations in the hold-out cross-validation are plotted in Fig. 10. According to this measure, HBRM provides the highest frequency of the best predictive accuracy metrics, that is,



(a)



(b)

Fig. 9. Integral of the CDF of the (a) log-likelihood and (b) RMSE.

smallest RMSE and largest log-likelihood, followed by HBKM, both outperforming GLM.

Overall, disaster responders' choice of model for carrying out the predictive analytics for community resilience will depend on a number of factors, such as the model performance, the decisionmaker's risk attitude, and the implementation of the model as a predictive and decision support tool. A decisionmaker's risk attitude might impact the weight placed on prediction error as opposed to computation time or interpretability. Although HBKM can be outperformed in a few instances by HBRM, HBKM requires less computational effort with very little difference in the accuracy metrics.

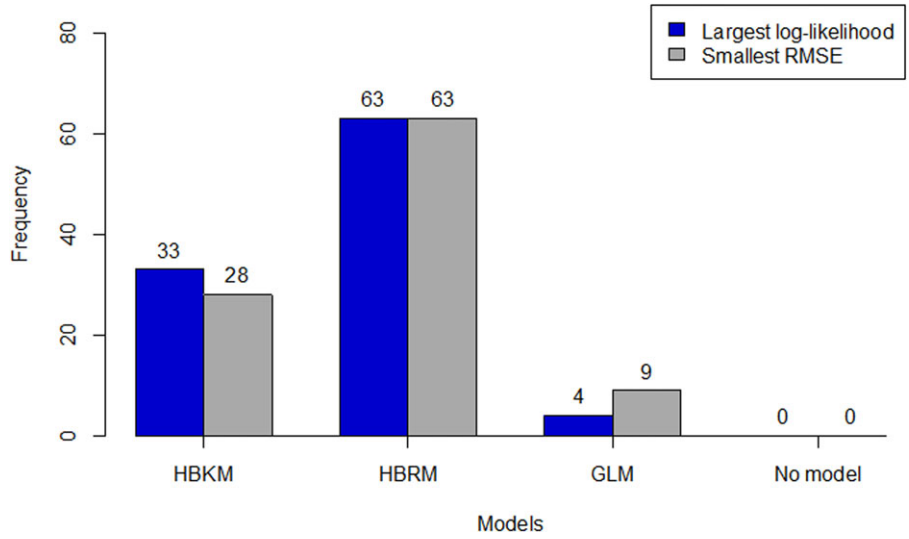


Fig. 10. Comparison of the frequency of achieving the best performance.

Table IV. Comparison of the Computation Time for Each Model

Model	Average (in Seconds)	*SD* (in Seconds)
HBKM	34.6	0.5
HBRM	75.9	14.7
GLM	0.0033	0.0043
No model	0.0004	0.0031

Note: The computation was performed on a desktop computer with Intel Core i7-6700 CPU @ 3.40 GHz and 16 GB installed memory.

5.2. HBKM in Risk Analysis

The outcome of HBKM allows for a comprehensive risk analysis of community resilience where the recovery rate at each point in time is evaluated as a point estimate or as a probability distribution. Fig. 11 provides an example of the predictions of the recovery rate using HBKM as a function of time after a storm. The recovery rate for each storm is calculated by dividing the number of people who regained power at a given time by the initial number of people without power. At each point in time, a probability distribution of the number of customers regaining power is generated that provides the decisionmaker with a range of values to consider for planning and restoration strategies. For illustration purposes, the predictions of storms S1 and S4 are presented. The predicted values follow a Poisson distribution with the predicted mean and observed values highlighted in the plot. From the distribution of predicted values, it can be observed that the predicted values have a relatively flat distribution at the early stages of recovery after a storm when the number of people without power is large, indicating a high uncertainty in the posterior distribution. As the community is recovering and the number of people without power decreases, the predicted values get closer to the observed values and the posterior distributions become narrower, indicting lower uncertainty in the predictions. Note that the reduction of uncertainty around the predicted recovery rate in Poisson HBKM may be due to the property of Poisson distribution that exhibits a higher variance for a higher mean. To analyze the source of variation in the prediction uncertainty, different versions of the model should be tested with non-Poisson likelihood functions. A generalized form of HBKM with nonconjugate priors constitutes the subject of future research and is, hence, outside the scope of this article.

Using hold-one-out cross-validation, the observed and predicted values are plotted for all five storms in Fig. 12, where it is shown that HBKM outperforms other models. Specifically, when compared to the results of other models, HBKM provides accurate predictions when the recovery rate is larger than 40%. For smaller values of the recovery rate, the model has a tendency to overestimate the true value due to an underestimation of the number of people without power. However, recall that HBKM outperforms all other models for larger values of RMSE, suggesting that in cases where there is a significant discrepancy between the observed and predicted value, the proposed model is still more

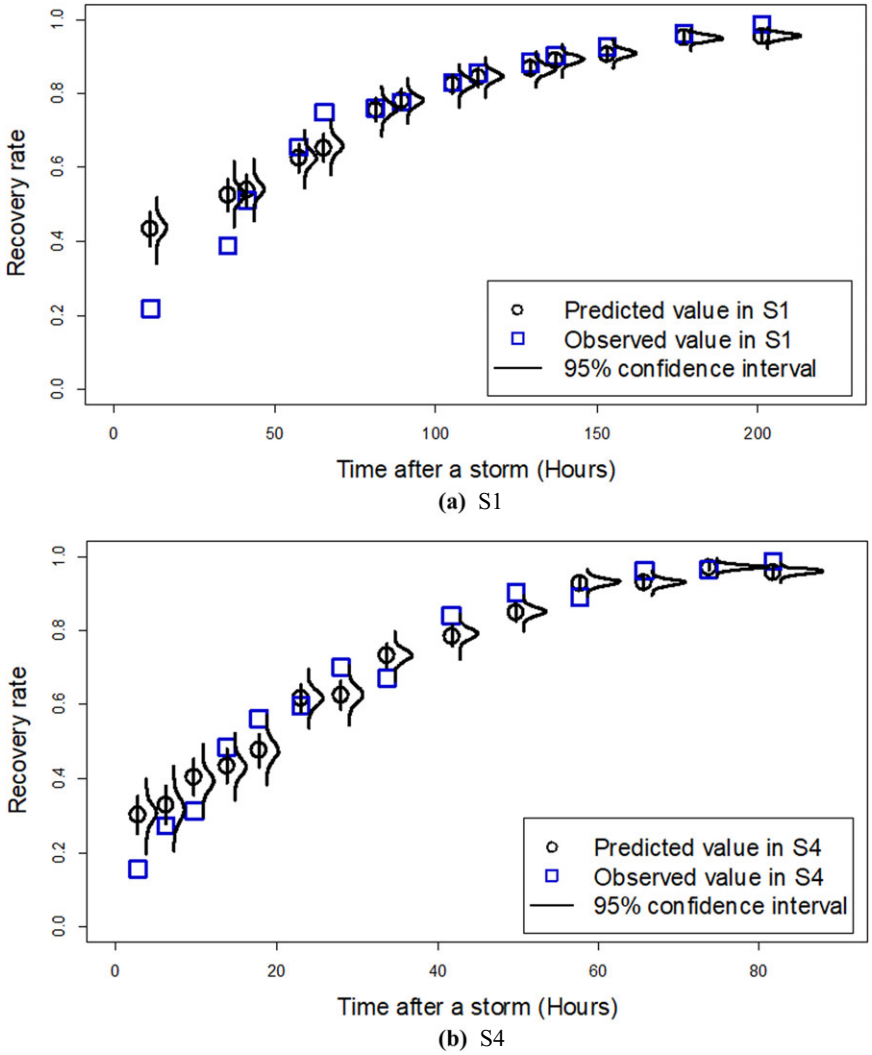


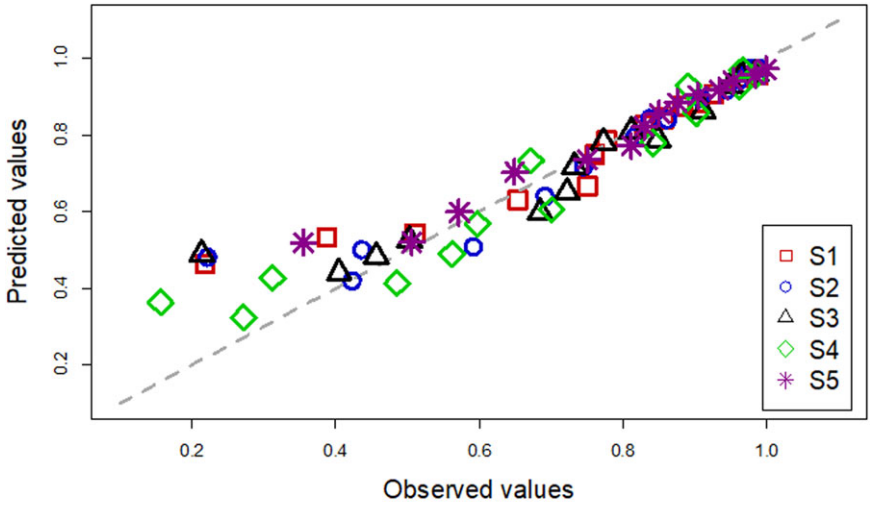
Fig. 11. Comparison between observed values and predicted values using HBKM given the time after a storm for (a) S1 and (b) S4. (The error bar represents the 95% confidence interval of a predicted value; Patil & Kulkarni, 2012.)

accurate than other statistical models considered in the analysis.

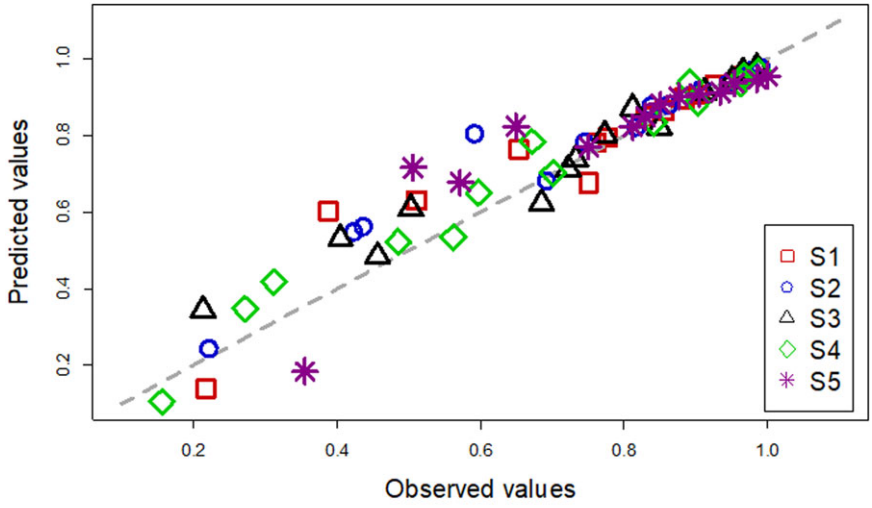
5.3. HBKM in Risk-Based Decision Making

By providing an accurate estimate of the recovery rate of a community, HBKM can be a powerful tool for utility companies and disaster responders in general. For utility companies, this model can help (i) reduce cost of investment in operations and maintenance by reducing call center volumes and managing vegetation along power lines, and (ii) optimize the number of crews required through "mutual aid agreements" to ensure rapid restoration of power to the affected areas (Quiring, Zhu, & Guikema, 2011). An underestimation of the required recovery crews may lead to prolonged recovery duration that increases the losses incurred by the utility and other infrastructure and economic sectors relying on electricity. An overestimation of the recovery rate will result in unnecessary cost of restoration, increasing the total loss incurred by the utility.

The importance of accurate estimation of the number of crews required is demonstrated numerically using an illustrative example (adapted from Figueroa-Candia, Felder, & Coit, 2018) that relates resilience predictive accuracy with cost of restoration. Suppose 30 components of an electric power



(a) HBKM



(b) HBRM

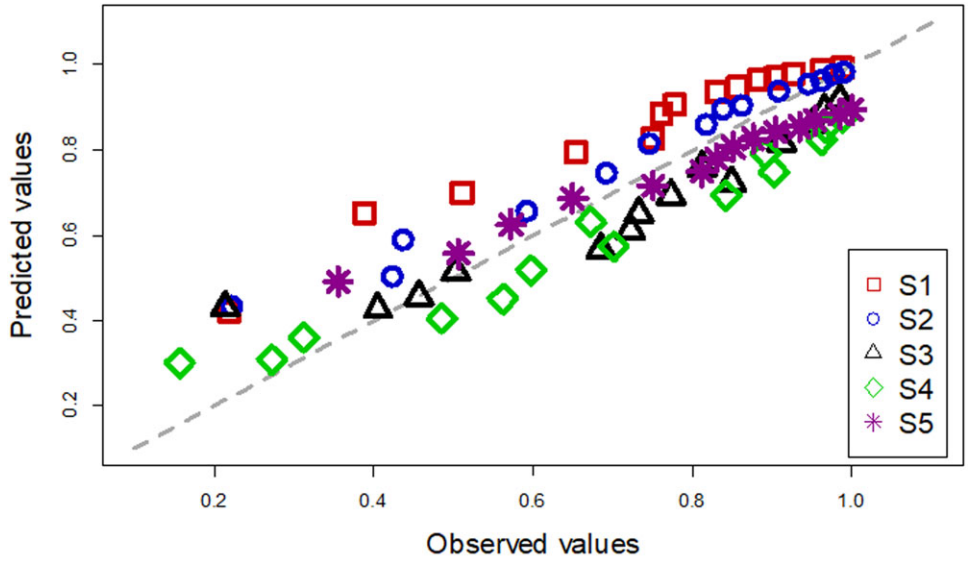
Fig. 12. Comparison between predicted values and observed values for all storms and all models (continued on next page).

system break down and 3,000 residential customers lose power. The utility managers need to decide on the number of additional repair crews required, apart from the three crews of their own. The cost of the three crews of their own is assumed to be $5,000 per day and the cost of additional crews and logistics combined per day is given by Equation (14), where n_a represents the number of additional crews, which can be a maximum of 10.

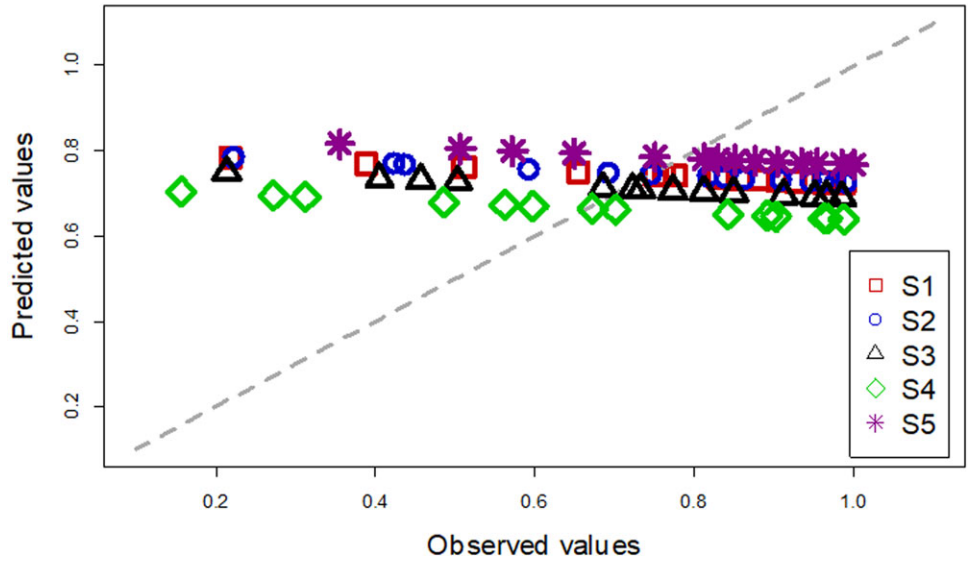
$$f_c(n_a) = \begin{cases} \$10{,}000 \text{ per day}, & n_a = 1, 2, 3 \\ \$10{,}000 + \$5{,}000\,(n_a - 3) \text{ per day}, & 4 \le n_a \le 10 \end{cases} \quad (14)$$

Next, for simplicity, full restoration is assumed to be achieved in 10 days with the three original crews. That is, one crew can restore power at an average of 100 customers per day. However, the efficiency of additional crews is lower than the original crews and will decline with the increase in the number of additional crews. The efficiency of one additional crew f_e, defined by the number of customers regaining power, is given by Equation (15).

$$f_e(n_a) = 100 - 5n_a,\ 1 \le n_a \le 10 \quad (15)$$



(c) GLM



(d) "No model"

Fig. 12. *Continued.*

With $(n_a + 3)$ crews, the total time of full recovery is T days (Equation (16)).

$$T(n_a) = \left\lceil \frac{3{,}000}{3 \times 100 + n_a f_e(n_a)} \right\rceil, \quad 0 \le n_a \le 10 \quad (16)$$

The expected loss per day from losing power for a residential customer is assumed to be \$100. Customers' loss L is calculated as the sum of customers' loss each day during the entire recovery process. Customers' loss on the first day is $100 \times 3{,}000$. The loss on the second day is $\{3{,}000 - [3 \times 100 + n_a(f_a)]\}$. The loss incurred each day afterward can be computed similarly to the second day and the total customers' loss for T days is given by Equation (17).

$$L(n_a) = 100 \times \{3{,}000T(n_a) - [3 \times 100 + n_a f_e(n_a)] [1 + 2 + \cdots + (T(n_a) - 1)]\} \quad (17)$$

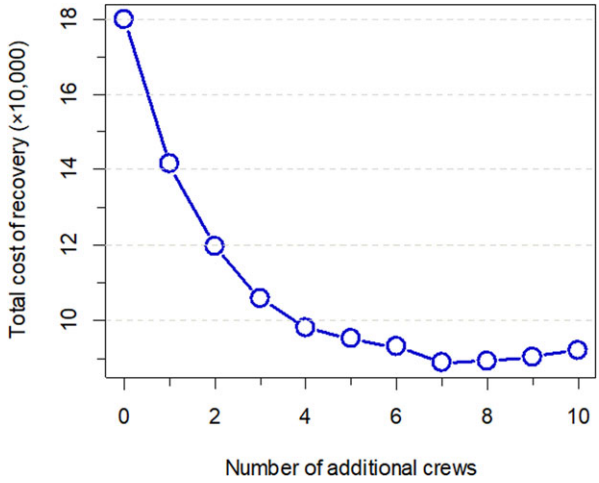


Fig. 13. The relationship between the total cost of recovery and the number of additional crews.

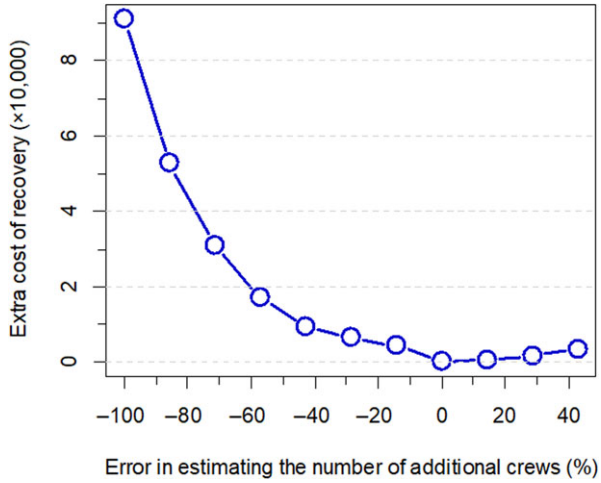


Fig. 14. The relationship between the extra cost of recovery and the error in estimating the number of additional crews.

Adding customers' loss and the cost of crews and logistics yields the total cost of recovery C_t, which is given by Equation (18).

$$C_t(n_a) = [f_c(n_a) + 5{,}000]\, T(n_a) + L(n_a) \quad (18)$$

The change in total cost relative to the change in the number of additional crews is depicted in Fig. 13. It is clear that the optimal number of additional crews is seven in this example.

After identifying the optimal number of additional crews, the error in estimating the optimal number of additional crews can be calculated using Equation (19), where n_e represents the estimated number of additional crews and n_{optm} represents the optimal number of additional crews.

$$\text{error}(\%) = \frac{n_e - n_{\text{optm}}}{n_{\text{optm}}} \times 100 \quad (19)$$

The extra cost of recovery is calculated as the difference between the real cost and the minimum cost. Fig. 14 displays the change in the extra cost of recovery relative to the change in the estimation error. Note that the negative error means underestimation while the positive error means overestimation. It can be observed from Fig. 14 that accurate estimation of the number of additional crews can significantly influence the total cost of recovery. If the number is underestimated by almost 40%, the extra cost will be approximately $10,000. Therefore, the ability to accurately predict the recovery rate can help minimize the overall cost of recovery by dynamically estimating the number of recovery crews needed as a function of time after a storm.

The accurate estimation of the recovery rate is of great concern to disaster management decisionmakers due to the wide variation in possible economic loss and inconvenience to communities. Hurricane Ike caused a power outage that lasted between a few days to several weeks for some communities, affecting more than 2.8 million customers in the Greater Houston area (Arab, Khodaei, Han, & Khator, 2015). During Hurricane Irma, prolonged power failure impacted the air conditioning system of a nursing home in Florida, which led to the deaths of eight people (Shapiro & Jacobo, 2017). The ability to understand and predict the potential impact of power outages on community resilience and their corresponding recovery trajectory can help reduce uncertainty and improve resource allocation for planning, preparedness, and recovery.

6. CONCLUSION

The accurate estimation of a community's recovery rate is important to help responders and planners optimize their allocation of resources during the recovery process and inform planning investments to improve resilience to future events. In this article, a new model, the HBKM, is proposed to predict the recovery rate of communities when data are scarce. This method addresses the problem of data scarcity by borrowing information from other similar subsets of the data and improves the estimation of parameters by assigning a distribution to the prior parameter that is updated using the entire data set. The HBKM

is compared to HBRM and GLM using recovery data on power outages from five of the strongest storms that have hit Shelby County in Tennessee between 2007 and 2017. HBKM yields on average the highest out-of-sample predictive accuracy. Although HBRM achieves the largest frequency of yielding the highest log-likelihood and lowest RMSE, the proposed HBKM requires less computation time while offering higher overall predictive accuracy. Accurate predictive measures of community resilience can significantly improve restoration of infrastructure; a lower estimation error can help minimize the cost of recovery by avoiding a prolonged period of power outage and unnecessary costs of restoration by dynamically updating the number of crews needed to restore power.

The proposed model can be a useful tool for utilities and disaster responders to understand the response mechanism of a certain community and plan accordingly. The ultimate goal of this work is to be able to model the recovery process of communities as a function of different infrastructure and socioeconomic characteristics. As such, future work will incorporate additional information such as demographic data to evaluate the impact of social vulnerability on the recovery process.

ACKNOWLEDGMENTS

This work is funded by the National Science Foundation under Grant No. 1635717.

REFERENCES

Allenby, G. M., Rossi, P. E., & McCulloch, R. E. (2005). *Hierarchical Bayes models: A practitioners guide*. SSRN Scholarly Paper ID 655541. Rochester, NY: Social Science Research Network.

Andrade, A. R., & Teixeira, P. F. (2015). Statistical modelling of railway track geometry degradation using hierarchical Bayesian models. *Reliability Engineering & System Safety*, *142*, 169–183.

Arab, A., Khodaei, A., Han, Z., & Khator, S. K. (2015). Proactive recovery of electric power assets for resiliency enhancement. *IEEE Access*, *3*, 99–109.

Baroud, H. (2015). *Bayesian kernel methods for the risk analysis and resilience modeling of critical infrastructure systems*. Norman, OK: University of Oklahoma.

Baroud, H., & Barker, K. (2018). A Bayesian kernel approach to modeling resilience-based network component importance. *Reliability Engineering & System Safety*, *170*, 10–19.

Baroud, H., Barker, K., & Lurvey, R. (2013). Bayesian kernel model for disruptive event data. *Proceedings of IIE Annual Conference* (p. 1777), Institute of Industrial and Systems Engineers (IISE), Peachtree Corners, GA.

Baroud, H., Francis, R., & Barker, K. (2016). Data-driven methods for the risk analysis of global supply chains. In *13th International Conference on Probabilistic Safety Assessment and Management (PSAM 13)*. Seoul, Korea.

Browne, W. J., & Draper, D. (2006). A comparison of Bayesian and likelihood-based methods for fitting multilevel models. *Bayesian Analysis*, *1*(3), 473–514.

Chang, S. E. (2010). Urban disaster recovery: A measurement framework and its application to the 1995 Kobe earthquake. *Disasters*, *34*(2), 303–327.

Coull, B. A., Mezzetti, M., & Ryan, L. M. (2003). A Bayesian hierarchical model for risk assessment of methylmercury. *Journal of Agricultural, Biological, and Environmental Statistics*, *8*(3), 253.

Faes, C., Geys, H., Aerts, M., & Molenberghs, G. (2006). A hierarchical modeling approach for risk assessment in developmental toxicity studies. *Computational Statistics & Data Analysis*, *51*(3), 1848–1861.

Figueroa-Candia, M., Felder, F. A., & Coit, D. W. (2018). Resiliency-based optimization of restoration policies for electric power distribution systems. *Electric Power Systems Research*, *161*, 188–198.

Friedman, M., & Savage, L. J. (1968). *The utility analysis of choices involving risk*. Indianapolis, IN: Bobbs-Merrill Company.

Gelman, A. (2006). Prior distributions for variance parameters in hierarchical models (comment on article by Browne and Draper). *Bayesian Analysis*, *1*(3), 515–534.

Gelman, A., Stern, H. S., Carlin, J. B., Dunson, D. B., Vehtari, A., & Rubin, D. B. (2013). *Bayesian data analysis*. London: Chapman and Hall/CRC.

Gillespie-Marthaler, L., Nelson, K., Baroud, H., & Abkowitz, M. (2019). Selecting indicators for assessing community sustainable resilience. *Risk Analysis*.

Guo, J., Lee, D., Sakrejda, K., Gabry, J., Goodrich, B., De Guzman, J., … Fletcher, J. (2016). rstan: R Interface to Stan. *R, 534*, 0–3.

Horney, J., Dwyer, C., Aminto, M., Berke, P., & Smith, G. (2017). Developing indicators to measure post-disaster community recovery in the United States. *Disasters*, *41*(1), 124–149.

Liu, J., Shi, Z., Lu, D., & Wang, Y. (2017). *Measuring and characterizing community recovery to earthquake: The case of 2008 Wenchuan Earthquake, China*. Discussion paper, Natural Hazards and Earth System Sciences, Germany.

MacKenzie, C. A., Trafalis, T. B., & Barker, K. (2014). A Bayesian beta kernel model for binary classification and online learning problems. *Statistical Analysis and Data Mining: The ASA Data Science Journal*, *7*(6), 434–449.

MacNab, Y. C. (2003). A Bayesian hierarchical model for accident and injury surveillance. *Accident Analysis & Prevention*, *35*(1), 91–102.

Migon, H. S., & Moura, F. A. (2005). Hierarchical Bayesian collective risk model: An application to health insurance. *Insurance: Mathematics and Economics*, *36*(2), 119–135.

Miles, S. B., & Chang, S. E. (2003). *Urban disaster recovery: A framework and simulation model*. MCEER Technical Reports (public). Buffalo, NY University at Buffalo.

Montesano, L., & Lopes, M. (2009). Learning grasping affordances from local visual descriptors. In *2009 IEEE 8th International Conference on Development and Learning* (pp. 1–6), Piscataway, NJ.

Nie, Y. M., Wu, X., & Homem-de-Mello, T. (2012). Optimal path problems with second-order stochastic dominance constraints. *Networks and Spatial Economics*, *12*(4), 561–587.

Ovaskainen, O., & Soininen, J. (2011). Making more out of sparse data: Hierarchical modeling of species communities. *Ecology*, *92*(2), 289–295.

Patil, V. V., & Kulkarni, H. V. (2012). Comparison of confidence intervals for the Poisson mean: Some new aspects. *REVSTAT—Statistical Journal*, *10*(2), 211–227.

Polettini, S., & Stander, J. (2004). A Bayesian hierarchical model approach to risk estimation in statistical disclosure limitation.

In *Proceedings of the International Workshop on Privacy in Statistical Databases* (pp. 247–261), Barcelona, Spain.

Qin, X., Ivan, J. N., Ravishanker, N., & Liu, J. (2005). Hierarchical Bayesian estimation of safety performance functions for two-lane highways using Markov chain Monte Carlo modeling. *Journal of Transportation Engineering*, *131*(5), 345–351.

Quiring, S. M., Zhu, L., & Guikema, S. D. (2011). Importance of soil and elevation characteristics for modeling hurricane-induced power outages. *Natural Hazards*, *58*(1), 365–390.

Schölkopf, B., & Smola, A. J. (2001). *Learning with kernels: Support vector machines, regularization, optimization, and beyond*. Cambridge, MA: MIT Press.

Shapiro, E., & Jacobo, J. (2017). 8 dead after Irma knocks out air conditioning at Florida nursing home. Retrieved from http://abcnews.go.com/US/dead-florida-nursing-home-irma-tore-state/story?id=49817477.

U.S. Census Bureau. (2016). *Quick facts*. Retrieved from https://www.census.gov/quickfacts/fact/table/shelbycountytennessee/PST045216.

Yan, Z., & Haimes, Y. Y. (2010). Cross-classified hierarchical Bayesian models for risk-based analysis of complex systems under sparse data. *Reliability Engineering & System Safety*, *95*(7), 764–776.