Risk Analysis, Vol. 30, No. 12, 2010 DOI: 10.1111/j.1539-6924.2010.01510.x

Prestorm Estimation of Hurricane Damage to Electric Power Distribution Systems

Seth D. Guikema,[1,*] Steven M. Quiring,[2] and Seung-Ryong Han[3]

Hurricanes frequently cause damage to electric power systems in the United States, leading to widespread and prolonged loss of electric service. Restoring service quickly requires the use of repair crews and materials that must be requested, at considerable cost, prior to the storm. U.S. utilities have struggled to strike a good balance between over- and under-preparation largely because of a lack of methods for rigorously estimating the impacts of an approaching hurricane on their systems. Previous work developed methods for estimating the risk of power outages and customer loss of power, with an outage defined as nontransitory activation of a protective device. In this article, we move beyond these previous approaches to directly estimate damage to the electric power system. Our approach is based on damage data from past storms together with regression and data mining techniques to estimate the number of utility poles that will need to be replaced. Because restoration times and resource needs are more closely tied to the number of poles and transformers that need to be replaced than to the number of outages, this pole-based assessment provides a much stronger basis for prestorm planning by utilities. Our results show that damage to poles during hurricanes can be assessed accurately, provided that adequate past damage data are available. However, the availability of data can, and currently often is, the limiting factor in developing these types of models in practice. Opportunities for further enhancing the damage data recorded during hurricanes are also discussed.

KEY WORDS: Damage; data mining; distribution system; hurricane; regression; risk

1. INTRODUCTION

Hurricanes cause substantial disruptions to electric power service and significant amounts of damage to electric power distribution systems. This damage is costly and time-consuming to repair, leading to significant burdens on both utilities and consumers. For example, during Hurricane Katrina in 2005, 1.3 million customers (82% of the customers) of one of the largest utilities in the U.S. Gulf Coast area lost power, and some of these customers were without power for up to 12 days. A given utility company does not have sufficient personnel or material on hand to rapidly restore power after a major hurricane. Instead, it relies on assistance from other utilities and material suppliers through mutual aid agreements and prestorm requests for crews and materials. In making these requests, utilities must balance the substantial costs associated with requesting external resources with the need to restore power quickly. Requesting more external resources than needed incurs excess salary, logistic, and supply costs that can, together, run into millions of dollars for a major hurricane. Requesting fewer

[1]Johns Hopkins University, Baltimore, MD, USA.
[2]Department of Geography, Texas A&M University, College Station, TX, USA.
[3]Department of Civil Engineering, Korea University, Seoul, Korea.
*Address correspondence to S.D. Guikema, Johns Hopkins University, 313 Ames Hall, 3400 N. Charles St., Baltimore, MD 21218, USA; tel: 410-516-6042; fax: 410-516-8996; sguikema@jhu.edu.

external resources than needed can lead to significant delays in restoring electric power. At the same time, the crews and materials that are requested must be positioned as close as possible to the areas of highest damage in order to minimize restoration times. Being able to accurately plan crew needs and placement in advance of an approaching hurricane is essential to striking the proper balance between cost and rapid power restoration. This prestorm planning requires estimates of the impacts of an approaching hurricane. These estimates must represent the spatial pattern of impacts accurately and provide an accurate overall estimate of the degree of impact to the system.

Statistical modeling provides a useful approach for estimating the impacts of an approaching hurricane on electric power systems prior to landfall. This approach uses data about the performance of a power system during past hurricanes to develop a model for estimating performance during the approaching hurricane. However, the previous research on developing statistical models for estimating the impacts hurricanes have on power systems has not dealt directly with damage.

Past research in estimating the impacts of hurricanes on power systems in advance of an approaching hurricane has focused on modeling power outages, where a power outage is defined as a nontransient activation of a protective device.[1–4] This previous work focused on outages largely because outage data are automatically collected by utilities through an automated outage management system (OMS), providing a convenient basis for model development. However, a single outage could affect a small number of customers or hundreds of customers and it could be associated with minimal damage or damage to a significant fraction of the poles on a feeder. Direct estimates of the amount of actual damage (e.g., the number of poles to be replaced) would be more closely aligned with the methods utility companies use for prehurricane deployment of repair crews and materials. They would allow restoration durations to be more easily estimated through the use of piece-rate duration estimates (e.g., the crew hours required to replace a single pole). However, spatially detailed damage data from past hurricanes are much less commonly recorded and saved in practice than outage data. In this article, we show that actual damage to the system, as measured by the number of utility poles that need to be replaced, can be modeled accurately provided that sufficient data about past performance of the system are available, and we then discuss ways in which damage data collection could be improved. We do this by examining detailed pole replacement data for parts of Mississippi from Hurricane Katrina. The results from utilizing these detailed data from a single hurricane are compared with earlier results from Han[5] based on less spatially detailed damage data from three hurricanes. The opportunities for significantly improving the basis for damage estimation are discussed within this modeling context.

2. BACKGROUND

This article presents regression and data mining models developed for estimating the number of poles that need to be replaced during an approaching hurricane. This work builds from the outage models of Liu *et al.*[1,2] and Han *et al.*[3,4] These papers used generalized linear models (GLMs),[1,3] generalized linear mixed models (GLMMs),[2] and generalized additive models (GAMs).[4] We utilize these same types of models but in addition utilize tree-based data mining techniques. This section provides an overview of these different methods.

2.1. Generalized Linear Models

A GLM is a generalization of standard linear regression that allows regression analysis of count data.[6,7] A GLM model consists of three components: (1) a conditional distribution for the count events given the parameter(s) of the distribution, (2) a "link" equation relating the distribution parameter(s) to a function of the explanatory variables, and (3) an equation specifying the function of the explanatory variables to be used in the link function. For example, a negative binomial GLM, a widely used GLM for count data, is given by Equations (1) and (2) where the vector x_i contains the explanatory variables, the vector β is the regression parameters to be estimated, and α is the overdispersion parameter of the negative binomial distribution:

$$f_Y(y \mid \alpha, \lambda) \sim \frac{\Gamma(y_i + \alpha)}{\Gamma(y_i + 1)\Gamma(\alpha^{-1})} \times \left(\frac{\alpha^{-1}}{\alpha^{-1} + \lambda_i}\right)^{\alpha^{-1}} \left(\frac{\lambda_i}{\alpha^{-1} + \lambda_i}\right) \quad (1)$$

$$\log(\lambda_i) = \beta_0 + \sum_i \beta_i x_i. \quad (2)$$

The negative binomial GLM is particularly appealing for modeling power system failures during hurricanes because it allows the variance of the counts to be higher than the mean of the counts. This "overdispersion" is typically observed in power system performance data. The presence of overdispersion makes the use of the more standard Poisson GLM inappropriate because the Poisson GLM assumes that the mean and variance of the counts based on the fitted model are equal. If this is not the case, using a Poisson GLM can lead to misestimation of the standard errors of the regression parameters, leading to erroneous conclusions about the statistical significance of covariates. Liu *et al.*[1] and Han *et al.*[3] also used a negative binomial GLM to model outages during hurricanes. Other types of GLMs are available if different variance structures must be represented (see, for example, Guikema and Coffelt[8]).

2.2. Generalized Additive Models

The structure of a GAM differs from the structure of a GLM only in how the parameters of the conditional distribution are related to the covariates.[9] Unlike a GLM that assumes linearity, a GAM allows for nonlinearity in this relationship. Equation (2) is replaced by a nonparametric smoothing function over the covariates. A simple example of such a GAM link function is shown in Equation (3) where each function $s(x_i)$ is a nonparametric smoothing function over the covariate x_i:[9]

$$\log(\lambda_i) = \beta_0 + \sum_i s(x_i). \tag{3}$$

Different types of smoothing functions can be used, but the standard approach, and the one we use in this article, is to use a cubic regression spline with a potential "knot point" (segment at which the function can change) at each unique value in the vector x_i. These smoothing splines are iteratively estimated as the overall model is fit to the data. Han *et al.*[4] showed that a GAM can fit hurricane power outage data better and can yield significantly more accurate power outage predictions for hurricanes not included in the fitting data set relative to a GLM. In this article, we use a Poisson GAM based on preliminary model results that showed that overdispersion was not a substantial issue once the link function was changed to the form of Equation (3) from the form of Equation (2). If overdispersion had been present even with the nonlinear link function, a negative binomial GAM would have been needed.

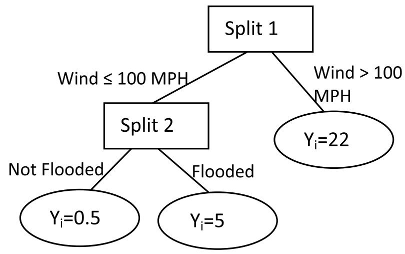

Fig. 1. Hypothetical CART tree.

2.3. Tree-Based Data Mining

In this article, we also use two tree-based data mining approaches. A tree-based data mining method represents the relationship between the response variable of interest (here the number of damaged poles) and the explanatory variables through the use of recursive binary partitioning of the data set. This provides a flexible representation of the relationships in the data set.

The first tree-based method we use is the classification and regression trees (CART) approach of Brieman *et al.*[10] CART is the classic tree-based regression model. The CART approach develops a single tree to try to capture the relationship between the response variable and the explanatory variables. Fig. 1 gives a simple, hypothetical CART tree. The first node splits the data set according to whether the gust wind speed forecast for a given location is above or below 100 mph. Those locations (geographic grid cells, for example) with forecast wind speeds above 100 mph are predicted to have 22 damaged poles. For those areas with forecast gust wind speeds below 100 mph, the example tree then partitions the data set on whether or not flooding is forecast for that grid cell. If flooding is forecast, the area is predicted to have five damaged poles. If not, it is predicted to have less than one damaged pole (on average). While the trees developed based on real data are significantly more complex and involve a much larger set of explanatory variables, this hypothetical tree serves to illustrate the idea of a CART model.

A CART model is fit in a two-stage process. First, a large tree is grown recursively. At each step, the algorithm chooses the variable to split the data set on and the value of that variable at which to split the data to maximize the decrease in the sum of squared error (SSE) of the fitted values at that node. This process is continued until each end node reaches purity (SSE = 0) or until a minimum number of records (e.g., five data records) are reached in a given node. This large tree is then "pruned" to remove nodes in order to yield a tree offering good predictive ability. This is done by using cross-validation to recursively remove nodes, with the node contributing the least to predictive accuracy in the cross-validation being removed at each step. However, in the trees fit in our study, the largest tree was the optimal tree in terms of predictive ability so no pruning was conducted.

The second tree-based method used in this study is Bayesian additive regression trees (BART).[11] A BART model consists of a large number of small trees, none of which models the response variable fully on its own. Letting $g(\mu_j \,|\, T_j, x_i)$ be the prediction from the jth BART tree T_j for record x_i, the overall form of the prediction from a BART model is given by Equation (4) where ε is an unknown random error term that is assumed to have a standard normal distribution with a variance to be estimated as a part of the model fitting process:

$$\hat{y}_i = \sum_j g(\mu_j \mid T_j, x_i) + \varepsilon. \tag{4}$$

Equation (4) shows that a BART model estimates a value for the response variable for a given set of explanatory variable values as the sum of the predictions from a large number of small trees.

BART models are estimated using a Markov chain Monte Carlo (MCMC) sampling method to draw replications from the posteriors for the parameters of the individual trees in the BART model with the number of trees being fixed prior to model estimation.[11] We used 200 trees, the default value, in our BART models. Chipman *et al.*[11] have shown that BART models yield very strong predictive accuracy in a wide range of settings including both simulated and real data.

2.4. Evaluating Predictive Accuracy

In developing a model for estimating damages associated with an approaching hurricane, the predictive accuracy of a model for a storm not in the fitting data set is the primary measure of success. To assess this predictive accuracy, we used holdout validation analysis. In our primary data set, we randomly partitioned the data set into training and validation sets with the training data set containing 80% of the records and the validation set containing the other 20% of the records. We then developed each of the four models above for the training data. The developed models were then used to predict the damage for the withheld 20% in the validation data. This process was repeated 150 times with a different random partitioning of the data conducted each time. The mean absolute error (MAE) and mean squared error (MSE) were calculated for each of these 150 holdout samples, with MAE and MSE defined as in Equations (5) and (6) where $\hat{y}_i$ is the estimate of the ith response measure, $n = 150$, and y_i is the actual value of the ith response measure:

$$\text{MAE} = \frac{1}{n}\sum_i^n |\hat{y}_i - y_i| \tag{5}$$

$$\text{MSE} = \frac{1}{n}\sum_i^n (\hat{y}_i - y_i)^2. \tag{6}$$

We then did pair-wise hypothesis tests comparing the means of the vectors of 150 MAE and MSE values for each possible pairing of the models using two-sided, paired Student's t tests with unequal means. Together with the difference in mean values for the MAE and MSE across the 150 random holdouts, this yields a sound basis for concluding which of the models yields the best predictive performance.

3. DATA

The pole replacement data that formed the basis for our model consisted of the number of poles that were replaced in 456 grid cells (12,000 feet by 8,000 feet) due to damage in parts of Mississippi during Hurricane Katrina. There were 8,698 total pole replacements in this data set with 2,308 of these being poles owned by a telephone company but used by the power company and 6,390 being poles owned by the power company. These pole replacement data were combined with the following explanatory variables.

- The number of transformers
- The number of poles
- The miles of primary lines
- The number of switches
- The maximum three-second gust windspeed in meters per second
- The duration of strong winds (three-second gust winds above 20 m/s) in hours
- The fractional soil moisture at three different levels each of the two days before the hurricane made landfall and the day of landfall (nine explanatory variables)
- The mean annual precipitation
- The standardized precipitation index (SPI) for periods of 1, 2, 3, 6, 12, and 24 months (six covariates total)

- The minimum, maximum, mean, median, and standard deviation of elevation (in meters) (five covariates total)
- The minimum, maximum, mean, median, and standard deviation of the compound topographic index (five covariates total)
- The minimum, maximum, mean, median, and standard deviation of slope (five covariates total)
- The fraction classified as each of 11 standard landcover types defined by the USGS

4. RESULTS

We fit each of the four basic models described above, conducting repeated random holdout validation tests. We also composed model average predictions of BART and CART together, BART, CART, and the GAM together, and all four basic models (BART, CART, GAM, and GLM) together and examined these model average predictions as well. These combinations were based on preliminary model runs that suggested that BART and CART offered substantially better predictions than the GLM and perhaps the GAM models. For the model averages, we took a simple unweighted average of predictions from the individual models. This yielded the following seven models that we tested:

1. BART
2. CART
3. Negative binomial GLM
4. Poisson GAM
5. BART/CART average
6. BART/CART/GAM average
7. BART/CART/GAM/GLM average

We then also compared the predictions from these seven models with the predictions obtained by using the mean in the fitting data set as the prediction for all values in the holdout data set. This provides an estimate of the degree to which each of the seven models above offers an improvement over using the historic mean as the future prediction. Note that this yielded 28 simultaneous hypothesis tests, necessitating the use of the Bonferroni correction for multiple hypothesis tests.[12] In order to maintain an overall confidence level of 95% for the combined set of hypothesis tests, each test can be considered significant if its p-value is below 1.8×10^{-3}.

As shown in Table I, the BART/CART and BART/CART/GAM model averages have the lowest MAE values, and these MAEs are lower than the MAEs of all other models except for the BART/CART/GLM/GAM average model by a statistically significant amount ($P < 4 \times 10^{-4}$ in all cases). However, the difference in MAE between the BART/CART and BART/CART/GAM models is not statistically significant. All models outperform the model using the average in the fitting data set by a statistically significant amount ($P \leq 3.03 \times 10^{-19}$) except for the GLM. The BART/CART and BART/CART/GAM model MAE values are statistically indistinguishable. The BART/CART and BART/CART/GAM models are the models with the best predictive accuracy and both models offer a sizable and statistically significant improvement in prediction accuracy relative to basing the predictions on the past mean damage rate.

When predictive accuracy is judged based on MSE, the BART, BART/CART, and BART/CART/GAM models have the lowest MSE values by substantial margins (see Table II). The differences in MSE values of the BART, BART/CART, and

Table I. Comparison of Holdout Mean Absolute Errors (MAEs) Based on Detailed Pole-Level Damage Data on the Basis of 150 Random Holdout Samples

Model	Mean MAE	p-Value: (2)	p-Value: (3)	p-Value: (4)	p-Value: (5)	p-Value: (6)	p-Value: (7)	p-Value: (8)
(1) BART	11.5	0.69	$\mathbf{1.01 \times 10^{-5}}$	$\mathbf{2.27 \times 10^{-4}}$	$\mathbf{2.82 \times 10^{-7}}$	$\mathbf{2.87 \times 10^{-5}}$	0.43	$\mathbf{5.68 \times 10^{-21}}$
(2) CART	11.7		$\mathbf{4.70 \times 10^{-6}}$	$\mathbf{1.39 \times 10^{-4}}$	$\mathbf{2.65 \times 10^{-9}}$	$\mathbf{3.03 \times 10^{-6}}$	0.52	$\mathbf{4.38 \times 10^{-25}}$
(3) GLM	21.4			$\mathbf{8.23 \times 10^{-5}}$	$\mathbf{6.17 \times 10^{-7}}$	$\mathbf{1.14 \times 10^{-6}}$	$\mathbf{8.46 \times 10^{-8}}$	0.46
(4) GAM	13.6				$\mathbf{6.01 \times 10^{-9}}$	$\mathbf{2.02 \times 10^{-11}}$	$\mathbf{5.08 \times 10^{-4}}$	$\mathbf{3.03 \times 10^{-15}}$
(5) BART \ CART	10.3					0.61	$\mathbf{1.23 \times 10^{-3}}$	$\mathbf{1.41 \times 10^{-27}}$
(6) BART \ CART \ GAM	10.4						3.40×10^{-3}	$\mathbf{2.58 \times 10^{-26}}$
(7) BART \ CART \ GAM \ GLM	12.0							$\mathbf{1.84 \times 10^{-19}}$
(8) Prediction by the Mean	20.0							

Note: Bold values are statistically significant comparisons at an overall 5% significance level for the family of simultaneous tests.

Table II. Comparison of Holdout Mean Squared Errors (MSEs) Based on Detailed Pole-Level Damage Data on the Basis of 150 Random Holdout Samples

Model	Mean MSE	*p*-Value: (2)	*p*-Value: (3)	*p*-Value: (4)	*p*-Value: (5)	*p*-Value: (6)	*p*-Value: (7)	*p*-Value: (8)
(1) BART	476	0.003	**0.0004**	$\mathbf{1.64 \times 10^{-6}}$	0.61	0.58	**0.001**	$\mathbf{2.59 \times 10^{-9}}$
(2) CART	640		**0.0004**	$\mathbf{1.174 \times 10^{-4}}$	$\mathbf{2.91 \times 10^{-7}}$	$\mathbf{2.61 \times 10^{-4}}$	0.009	8.7×10^{-7}
(3) GLM	11,872			$\mathbf{6.90 \times 10^{-4}}$	$\mathbf{3.76 \times 10^{-4}}$	$\mathbf{3.94 \times 10^{-4}}$	**0.0004**	$\mathbf{6.21 \times 10^{-4}}$
(4) GAM	1,161				$\mathbf{7.29 \times 10^{-7}}$	$\mathbf{1.95 \times 10^{-7}}$	0.85	0.50
(5) BART \ CART	463					0.14	**0.0007**	$\mathbf{7.86 \times 10^{-11}}$
(6) BART \ CART \ GAM	493						**0.001**	$\mathbf{2.1 \times 10^{-9}}$
(7) BART \ CART \ GAM \ GLM	1,198							0.50
(8) Prediction by the mean	1,069							

Note: Bold values are statistically significant comparisons at an overall 5% significance level for the family of simultaneous tests.

BART/CART/GAM models are not statistically significant at an overall significance level of 95%. However, the MSE value for BART is lower (better) than the MSE values of all of the other models except the CART, BART/CART, and BART/CART/GAM models by a statistically significant amount. Similarly, the MSE values of the BART/CART and BART/CART/GAM models are lower than all of the other models except for the BART model by a statistically significant amount. The MSE values for all models except the GAM and BART/CART/GAM/GLM models are lower than the MSE of the mean model by a statistically significant amount; all other models outperform the mean-only model.

Combining the results of the MAE- and MSE-based model comparisons, the BART/CART average and the BART/CART/GAM average models provide the best predictive accuracy based on our data set. However, there is not a statistically distinguishable difference between the predictive accuracy of these two types of models in our analysis. While either or both could be used, the simpler BART/CART average model would be preferred in most situations due to decreased implementation burden.

5. COMPARISON WITH PREVIOUS RESULTS

We previously developed a statistical model for estimating the number of poles and transformers damaged during three past hurricanes (Dennis, Ivan, and Katrina) based on a different data set utilizing only GLMs (Han,[5] chapter 6). The data set used by Han[5] contained the number of poles and transformers damaged in past hurricanes for a limited fraction of a different portion of the service area of the same utility company providing the Hurricane Katrina damage data. These data are aggregated over much larger areas than the grid cells used in our models. These damage aggregation areas were irregularly shaped and overlapped a number of smaller grid cells. In some cases, these areas contain as many as 224 smaller grid cells, making geographically detailed estimation difficult. However, unlike the spatially more detailed Hurricane Katrina data, the damage data of Han[5] included damage during three hurricanes, potentially enhancing the ability to extrapolate to future hurricanes not in the training data set.

Due to the level of aggregation of the damage data, we assumed that the *rate* of damage of poles and transformers, calculated as the number of damaged poles or transformers divided by the total number of poles or transformers, was constant within each of the large grid cells. This allowed us to scale down to the smaller grid cells, making use of all of the detailed explanatory variables available at the level of the small grid cells. This scaling was done by first calculating the total number of poles and transformers in each of the data aggregation areas by summing over the smaller grid cells that were included in each of the larger damage aggregation areas. Then the *rate* of pole and transformer damage was calculated by dividing the total number of damaged poles or transformers by the total number of poles or transformers in the data aggregation areas. Then these two damage rates were assumed to be constant, and the number of damaged poles and transformers in each of the original grid cells was estimated by multiplying the pole or transformer damage rate by the number of poles or transformers in each of the smaller original grid cells. A negative binomial GLM was then developed for predicting the rate of damaged poles and damaged transformers at the level of the small grid cells using the same methods used for developing the outage models in Han *et al.*[2,4] However, unlike previous models of outages,[1−4] an offset was included in the link function to obtain the mean estimates of the number of poles and transformers damaged based on the estimated damage rates and the total number of poles or transformers in each grid cell. The predictions from the damage models are the number of poles and the number of transformers damaged in each of the small grid cells. This leads to variability in the predictions that is not present in the original data set, which included information only at the level of the larger data aggregation areas.

To examine the predictive accuracy of the pole and transformer damage models, we conducted holdout analysis at the hurricane level as in Han.[5] We first removed the data for one of the three hurricanes (Dennis, Ivan, and Katrina) from the data set, fit the model to the remaining two hurricanes, and then used the fitted model to predict the number of poles and transformers damaged in each grid cell during the withheld hurricane. We repeated the same process to obtain an estimate of the relative error in the damage predictions. Table III shows the results of this process.

We see that the average per-cell errors (i.e., MAE) are approximately equal to the average number of damaged poles and damaged transformers for Hurricane Dennis, substantially higher for Hurricane Ivan, and lower for Hurricane Katrina. The predictive accuracy of these models is poor for Hurricane Ivan, marginal for Hurricane Dennis, and reasonable for Hurricane Katrina.

In comparing the results in Table III with the results in Tables I and II, we see the ratio of MAE to the mean number of damaged poles in the data set is substantially higher for the models of Han[5] (Table III) than for any of the models in Tables I and II. The data underlying Han[5] are substantially less precise because of the need to assume a constant damage rate in order to downscale to the smaller grid cells from the aggregated damage data provided. These results demonstrate the importance of collecting and recording geographically detailed damage information after a hurricane if accurate, geographically detailed damage estimates are desired for future events.

6. DISCUSSION AND DATA AVAILABILITY ISSUES

Overall, the results of this article show that damage to electric power distribution systems can be accurately modeled with statistical models if appropriate data are available and appropriate methods are used. In our study, the predictive accuracy of traditional parametric regression models such as GLMs was not as good as that of the nonparametric regression (GAMs) and data mining (BART and CART) approaches in repeated, random holdout validation testing. In addition, model averaging generally improved the predictive accuracy by statistically significant and practically noticeable amounts. This suggests that more flexible regression and data mining methods together with model averaging should be examined for future hurricane damage modeling.

While selecting the best model is important, no model will provide accurate, useful damage estimates if it is not developed based on sound, representative information. Comparing the results of Han[5] with those in Tables I and II clearly demonstrates that geographically aggregated data are insufficient for developing models for producing accurate, geographically detailed damage estimates. Utility companies must begin to gather accurate damage data at a geographically detailed scale if they wish to use these data to develop damage prediction models for future events. However, gathering accurate, geographically detailed damage data after a hurricane is difficult in practice, and innovative approaches are needed.

There are many challenges in gathering posthurricane damage data. In particular, the restoration

Table III. Predictive Accuracy of the GLM-Based Damage Estimation Models of Han[6] by the Comparisons of Holdout MAEs and MSEs

	Ivan (2004)	Dennis (2005)	Katrina (2005)
Actual number of damaged poles (mean/grid cell)	2,364 (2.99)	144 (0.17)	834 (1.47)
Actual number of damaged transformers (mean/grid cell)	709 (0.90)	107 (0.13)	602 (1.06)
MAE, poles	243	148	64
MSE, poles	69,862	117,116	10,371
MAE, transformers	73	61	57
MSE, transformers	6,285	17,460	11,024

process used at most utilities is not conducive to data gathering. The primary goal of restoration crews is and should be to restore electrical power as quickly as possible. Gathering damage data and preserving it accurately imposes an additional task that can slow the progress of repair crews. In addition, crews from many different utilities are involved in the restoration process after a major storm. These crews generally lack detailed knowledge of the system they are working on and any data management hardware and software utilized by the host utility. Furthermore, not all utilities realize the importance of gathering detailed damage data after a hurricane. They can meet their immediate goal of rapid service restoration without gathering the data, and they are under little pressure from regulatory bodies to gather damage data, making it seem to be an unjustified expense. The utility providing the data for this analysis was an exception in that it had gathered useful damage data from recent hurricanes.

Anecdotal discussions with a number of utilities have revealed a number of different attempts to gather damage data. Some have attempted to have the restoration crews use personal digital assistants (PDAs) with preloaded data collection software to record the location of each repair they made. These data would then be loaded into a central database for future use. However, practical experience suggested that crew compliance was low due to the added burden on their time. In addition, mutual aid crews from other utilities did not have training with the PDAs and were reportedly unlikely to utilize them. Another approach reported by a different utility was to have experienced personnel make a prerepair damage assessment in advance of the crews in a limited number of "statistical" sampling grid cells covering small portions of the system. This approach does yield damage data, but the areas that are sampled are generally constrained to relatively easily accessible portions of the system, risking a biased sample of the overall system. In addition, this approach takes experienced personnel away from the restoration process, possibly slowing the overall restoration process. A third approach, and the one utilized to collect the data used in Tables I and II, is to use billing records for replacement of third-party poles to determine damage. Many utilities share poles with other organizations such as telephone companies. If the utility replaces these poles as part of its restoration process, records must be kept in enough detail to enable the utility to bill the third party. These records would generally be kept based on a paper or PDA-based recording process by field personnel, but the added importance of billing for the work helps to ensure higher compliance. The limitation is that this is currently done primarily *only* for third-party poles. There may be opportunities to expand this type of data-gathering activity and rely more on utility billing records to extract more detailed damage data for past events.

The comparison of the results of Han[5] (Table III) with the results in Tables I and II demonstrate the importance of having data from multiple hurricanes. The holdout results in Tables I and II are only for Hurricane Katrina. We do not know how well these models generalize to other hurricanes because the data are not available to test this. On the other hand, the holdout results in Table III explicitly deal with cross-hurricane generalization, albeit with only three hurricanes. Accurate, geographically detailed data are needed from additional hurricanes in the future to better support the development of detailed damage models.

Finally, damage models such as those developed in this article, in conjunction with outage models such as those of Han *et al.*,[4] could be coupled with a model estimating long-term hurricane occurrence rates and intensities. This would provide a strong basis on which utilities could make decisions about system reinforcement ("hardening") and about the level of emergency response materials (e.g., poles and line) to keep on hand. Having sound, long-term predictive models as the basis for balancing the long-term costs with the short-term power outage impacts in this type of decision making would then provide an improved basis on which the reasoning for utility decisions could be explained to regulators and the public.

7. CONCLUSION

This article has shown that accurate, geographically detailed models can be developed for estimating posthurricane damage to electric power systems in advance of an approaching hurricane. However, the accuracy of these models is dependent on both the type of models used and the availability of appropriate information on which to develop the models. The results of this study suggest that nonparametric regression and data mining models (e.g., GAM, BART, and CART) may provide a better basis for accurate prediction of hurricane damage. This study also shows that accurate, geographically detailed damage data from multiple hurricanes is needed to have a strong basis for developing damage models. This presents practical challenges to utilities, but these challenges can be overcome provided appropriate prestorm development of the needed procedures and approaches.

ACKNOWLEDGMENTS

This work is part of a multiyear agreement between Johns Hopkins University and a large investor-owned utility in the United States. Part of that agreement specifies that the sponsor will not be acknowledged by name.

We gratefully acknowledge its support. All opinions expressed in this article are those of the authors and do not necessarily reflect the position of this sponsor.

REFERENCES

1. Liu H, Davidson RA, Rosowsky DV, Stedinger JR. Negative binomial regression of electric power outages in hurricanes. Journal of Infrastructure Systems, 2005; 11(4):258–267.
2. Liu H, Davidson RA, Apanasovich TV. Spatial generalized linear mixed models of electric power outages due to hurricanes and ice storms. Reliability Engineering and System Safety, 2008; 93(6):897–912.
3. Han SR, Guikema S, Quiring S, Lee KH, Davidson R, Rosowsky D. Estimating the spatial distribution of power outages during hurricanes in the Gulf Coast region. Reliability Engineering and System Safety, 2009; 94(2):199–210.
4. Han SR, Guikema S, Quiring S. Improving the predictive accuracy of hurricane power outage forecasts using generalized additive models. Risk Analysis, 2009; 29(10):1443–1453.
5. Han SR. Estimating Hurricane Outage and Damage Risk in Power Distribution Systems, Ph.D. Dissertation, Department of Civil Engineering, Texas A&M University, August 2008.
6. Nelder JA, Wedderburn RWM. Generalized linear models. Journal of the Royal Statistical Society, Series A (General), 1972; 135(3):370–384.
7. Cameron AC, Trivedi PK. Regression Analysis of Count Data. Econometric Society Monographs No. 30, Cambridge, UK:Cambridge University Press, 1998.
8. Guikema S, Coffelt JP. Modeling count data for non-linear, complex infrastructure systems. ASCE Journal of Infrastructure Systems, 2009; 15(3):172–178.
9. Hastie T, Tibshirani R. Generalized additive models. Statistical Science, 1986; 1(3):297–310.
10. Breiman L, Friedman JH, Olshen RA, Stone CJ. Classification and Regression Trees. Boca Raton, FL: CRC, 1998.
11. Chipman H, George E, McCulloch R. BART: Bayesian Additive Regression Trees, 2006. Available at: http://www-stat.wharton.upenn.edu/~edgeorge/Research_papers/BART%20June%2008.pdf, Accessed on July 20, 2009.
12. Devore JL. Probability and Statistics for Engineering and the Sciences, 3rd ed. Pacific Grove, CA: Brooks/Cole, 1995.