Risk Analysis, Vol. 36, No. 10, 2016 DOI: 10.1111/risa.12552

Hurricane Isaac: A Longitudinal Analysis of Storm Characteristics and Power Outage Risk

Gina L. Tonn,[1,*] Seth D. Guikema,[2] Celso M. Ferreira,[3] and Steven M. Quiring[4]

In August 2012, Hurricane Isaac, a Category 1 hurricane at landfall, caused extensive power outages in Louisiana. The storm brought high winds, storm surge, and flooding to Louisiana, and power outages were widespread and prolonged. Hourly power outage data for the state of Louisiana were collected during the storm and analyzed. This analysis included correlation of hourly power outage figures by zip code with storm conditions including wind, rainfall, and storm surge using a nonparametric ensemble data mining approach. Results were analyzed to understand how correlation of power outages with storm conditions differed geographically within the state. This analysis provided insight on how rainfall and storm surge, along with wind, contribute to power outages in hurricanes. By conducting a longitudinal study of outages at the zip code level, we were able to gain insight into the causal drivers of power outages during hurricanes. Our analysis showed that the statistical importance of storm characteristic covariates to power outages varies geographically. For Hurricane Isaac, wind speed, precipitation, and previous outages generally had high importance, whereas storm surge had lower importance, even in zip codes that experienced significant surge. The results of this analysis can inform the development of power outage forecasting models, which often focus strictly on wind-related covariates. Our study of Hurricane Isaac indicates that inclusion of other covariates, particularly precipitation, may improve model accuracy and robustness across a range of storm conditions and geography.

KEY WORDS: Hurricanes; power outages; random forest

1. INTRODUCTION

Hurricane Isaac hit Louisiana in August 2012 and caused substantial power outages. It was a Category 1 hurricane at landfall and 47% of the state's electric customers lost power. The storm was large, slow-moving, and had significant storm surge associated with it. In comparison with other hurricanes, Isaac ranks fourth in customer power outages, behind Hurricanes Katrina, Gustav, and Rita, for the Entergy service area in Louisiana, Mississippi, Texas, and Arkansas.(1) The track of the storm is illustrated in Fig. 1(2)

Power outages result in direct repair and restoration costs for utility companies, and can also result in loss of services from other types of critical infrastructure that rely on power service such as water, transportation, and communications systems. This can delay recovery times for a community that is impacted by a hurricane.(3) Accurate predictions of power outages prior to a storm can benefit both utility companies and government agencies by making planning and recovery more efficient.(4)

[1]Department of Geography and Environmental Engineering, Johns Hopkins University, Baltimore, MD, USA.
[2]Department of Industrial and Operations Engineering, University of Michigan, Ann Arbor, MI, USA.
[3]Department of Civil, Environmental, and Infrastructure Engineering, George Mason University, Fairfax, VA, USA.
[4]Department of Geography, Texas A&M University, College Station, TX, USA.
*Address correspondence to Gina Tonn, Department of Geography and Environmental Engineering, Johns Hopkins University, 3400 North Charles Street, Baltimore, MD 21218, USA; gtonn2@jhu.edu.

0272-4332/16/0100-1936$22.00/1 © 2016 Society for Risk Analysis

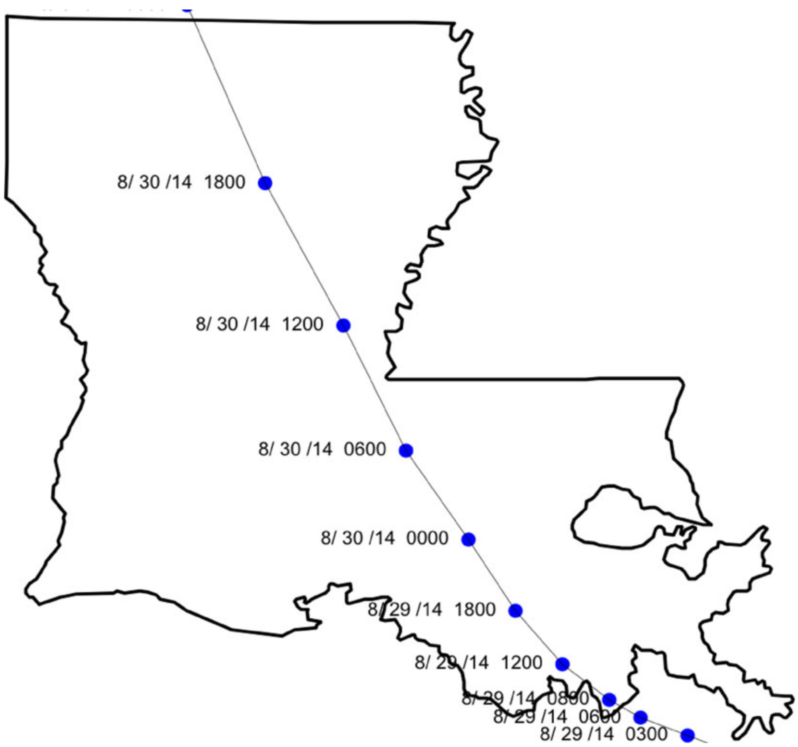

Fig. 1. Hurricane Isaac track.

Power outage prediction is often accomplished through the development of models based on wind field estimates, along with other covariates such as power system data, soil moisture levels, land use, and topographical indicators.[4] A number of such statistical models have been developed.[3–5] While these models can be very accurate for some storms, they are less accurate for others due to the differing characteristics of the storms.

In addition to accuracy of models varying from storm to storm, the causes of the outages can vary geographically across a region, and the existing models typically do not include some potential causes of power outages, particularly heavy rainfall. The main goal of this article is to obtain a better understanding of how storm characteristics correlate with power outages and how this correlation varies geographically. The purpose of this study is both to improve basic understanding of hurricane power outages and to provide a stronger basis for improving outage forecasting models. Are the outage drivers the same for a coastal area as an inland area? How important are rainfall and surge relative to wind? Damage to power systems is recorded by utilities, but good data on causes of outages are not generally available, making a longitudinal approach necessary. Statistical analysis of power outage data and covariate data is used in this analysis to provide a better understanding of how storm conditions correlate with power outages. Power outages were studied longitudinally across the state of Louisiana for Hurricane Isaac to identify how the importance of covariates changes geographically. The results of this analysis may inform power outage prediction models and help to build more resilient infrastructure through improved understanding of power outage risk.

In Section 2, the data used for the analysis as well as the statistical analysis methods are presented. Results and discussion are included in Section 3, and conclusions in Section 4.

2. METHODS AND DATA

We focused on covariates related to three key physical hazards associated with hurricanes—wind, storm surge, and rainfall—in order to gain a better understanding of the relative contribution of these three storm characteristics. We analyzed all covariates on an hourly basis, and so included covariates that change over time as the storm progresses. We obtained data for the covariates of interest from publicly available sources or modeled them based on

Table I. Summary of Covariates

Category	Covariate	Source	Description
Precipitation	Cumulative Precipitation	National Climatic Data Center (NCDC)	Total precipitation amount during storm duration to hour of analysis in centimeters
	Hourly Precipitation Total	NCDC	Precipitation amount in hour of analysis in centimeters
Wind	Wind Speed	Wind model	Wind speed in meters/second for zip code in hour of analysis
	Wind Gust Duration	Wind model	Duration of wind gust >20 m/sec for zip code in hour of analysis
Outages	Previous Outages	Entergy website	Number of outages in previous hour of analysis for zip code
Population	Population	U.S. Census Bureau	Population estimate for zip code
Surge	Average Surge	ADCIRC+SWAN models	Average storm surge depth for zip code in hour of analysis in meters
	Minimum Surge	ADCIRC+SWAN models	Minimum storm surge depth for zip code in hour of analysis in meters
	Maximum Surge	ADCIRC+SWAN models	Maximum storm surge depth for zip code in hour of analysis in meters
	Surge Variance	ADCIRC+SWAN models	Variance of storm surge depth for zip code in hour of analysis in meters

publicly available data. A summary of the covariates, data sources, and a description of each covariate are provided in Table I. While data were available in varying time increments for each covariate, we performed interpolation to obtain hourly estimates. We chose the hourly change in outages as the response variable, and hours that did not have a positive increase in outages were removed from the analysis to focus the analysis on only the outage occurrence portion of the storm, not the outage restoration part of the storm. A more detailed description of each category of data and the data interpolation are provided in Sections 2.1–2.5.

Initial analysis was done on a statewide basis, with the remainder of the analysis done on a zip code basis. After completing the data collection and interpolation, we generated a random forest model for the entire data set including all zip codes. The most important covariates were identified through random-forest-based importance measures for use in additional analysis as described further below. Using this reduced set of covariates, we trained a random forest model for each zip code separately so that impacts could be analyzed spatially. We plotted the results in map format for analysis of spatial trends. We used a quantile regression forest model for selected zip codes to gain insight into model accuracy. The modeling and analysis methods are described in more detail in Sections 2.6 and 2.7.

2.1. Outage Data

Power outage data were harvested from the Entergy Louisiana website during the duration of the storm from August 27 to September 5, 2012.[(6)] The data were collected on a half-hourly basis during periods of peak outages, and were collected less frequently during nonpeak outage periods. Data collected included the number of current customer outages by zip code. In order to standardize the data for use in analysis, we performed linear interpolation to estimate the number of outages for each zip code at the top of each hour for the duration of the storm. Some areas of Louisiana are not serviced by Entergy and were not included in this analysis.

We chose the change in outages (termed delta outages in this article) for each hour of analysis for each zip code as the response variable for this analysis. Total power outages for the previous hour of analysis (Previous Outages) for each zip code was included as a covariate to account for the fact that the number of customers already without power impacts the number of power outages occurring in a given hour.

2.2. Precipitation Data

Precipitation data were obtained from the National Climatic Data Center (NCDC) website. Data were available for 36 rainfall stations in Louisiana. The time intervals at which the precipitation data were recorded varied by station, but were typically hourly or half-hourly. The data obtained were the hourly total rainfall.[7] In order to standardize the data for use in analysis, we interpolated the data set to estimate the hourly precipitation (precipitation that occurred in the previous 60-minute period) at the top of the hour for each station. Because our analysis was performed on a zip code basis, we needed rainfall estimates for each zip code. Based on the geographic coordinates of the zip code centroids and on the locations of the stations, we generated hourly rainfall estimates for each zip code using inverse distance weighted interpolation based on the spatially sparser set of rainfall stations that were available.

2.3. Storm Surge Model

We used the coupled version of the 2-Dimensional Depth Integrated version of the Advanced Circulation (ADCIRC) model and the wave model SWAN[8] to simulate hurricane storm surge. The ADCIRC model[9] is a finite element, shallow water model that solves for water levels and currents at a range of scales and is widely used for storm surge modeling[10] This version of the program solves the Generalized Wave Continuity Equation (GWCE) and the vertically integrated momentum equations. SWAN is a third-generation spectral wave model[11] that computes the time and spatial variation of directional wave spectra. We used the prevalidated numerical mesh SL15 presented in Bunya *et al.*[12] and validated by Dietrich *et al.*[13] with resolution up to 30 m in some areas. The hurricane surge model was forced by wind and pressure fields developed by a parametric asymmetric wind model[14] that computes wind stress, average wind speed, and direction inside the Planetary Boundary Layer (PBL) based on the National Hurricane Center (NHC) best track data[15] meteorological conditions (e.g., central pressure, forward speed, and radius to maximum wind). The simulations for Hurricane Isaac included tides (tidal potential components M2, S2, N2, K2, K1, O1, and Q1) and neglected rivers inflows. Simulation results were recorded at 15-minute intervals for every model node in the study region. The water levels for each model node within each zip code were extracted from the entire model domain and inundation levels were converted to the NAVD88 vertical datum. Covariates based on the storm surge model include average storm surge, maximum storm surge, minimum storm surge, and storm surge variance.

2.4. Wind Model

The parametric wind field model of Willoughby *et al.*[16] was used to generate wind estimates for the duration of the hurricane at the zip code level for Hurricane Isaac. Parametric hurricane models are formulated from a physical understanding of hurricane wind fields. That is, winds are calm in the eye of the hurricane and they are typically at a maximum in the eyewall. Outside the eyewall the wind decreases with radius, although not always monotonically, and becomes near zero at some distance from the center of circulation. This wind field model was previously used in Han *et al.*[3,17] Two of the covariates are based on output from this model. The first is maximum wind speed in meters per second in the previous hour. The second is wind gust duration greater than 20 m/sec, with duration being taken cumulatively over the life of the storm for each zip code. Both of these covariates are simulated for the centroid of each zip code polygon based on running the wind field model every 60 minutes over the duration of the storm.

2.5. Other Data

Population estimates for each zip code were obtained from the U.S. Census Bureau American Community Survey. These estimates were based on the U.S. Census Bureau data for the year 2011. Because the U.S. Census Bureau does not track population on a zip code basis, the population data are estimates based on census tract data.[18]

2.6. Random Forest and Quantile Regression Forest Methods Overview

A random forest is a nonparametric ensemble data mining method.[19] In the method, a large number of regression trees are developed, with each tree based on a bootstrapped sample of the data set. Random forest models are good for data sets with nonlinear data, outliers, and noise. Two types of output from the random forest model fit very nicely with the objectives of this analysis. The first is variable importance, which is a measure of the contribution of a given covariate to the model prediction accuracy. The second is the partial dependence plot. These plots show the marginal effect of a covariate on the response variable. The randomForest package in R was used for this analysis.[20]

Quantile regression forests provide a nonparametric way of estimating conditional quantiles based on an underlying random forest model. Quantiles give more information about the distribution of the response variable as a function of the covariates than just using the conditional mean as a standard random forest model does. In this method, regression trees are grown as in the random forest method. Then the weighted distribution of the observed response variables is used to estimate a conditional distribution. The difference between random forest models and quantile regression forest models is that random forest models keep only the mean predictions and disregard other information. Quantile regression forests estimate the quantiles of the predictions based on the trained forest.[21] The quantregForest package in R was used for this analysis.[22] Predictions made using this package are based on out-of-bag data generated through the standard random forest bootstrapping process.[20]

2.7. Statistical Analysis

A statewide random forest model was run using the data for all covariates and zip codes. Only positive delta outages were included, to limit the analysis to the occurrence of power outages, not the restoration of power. In order to better understand the predictive accuracy of the random forest model, a quantile regression forest model was run on 10 selected zip codes. The zip codes were chosen so that different geographic areas in the state were represented.

Variable importance was reviewed to identify the variables that were most significant for predictive accuracy. Based on the variable importance, one covariate from each category of covariates (precipitation, wind, storm surge, and outages) was retained for individual zip code analysis in order to better understand the influences of the different variables. Partial dependence plots were generated for each of these covariates, and were reviewed to understand the marginal effects of these covariates on the response variable.

In order to understand the relative importance of the four covariates, and how that importance varied geographically, plots of importance for each of the covariates were generated. Because the magnitude of variable importance was not the same for each random forest run, comparing the variable importance between zip codes would not be useful. Instead, we calculated a percent variable importance for each zip code. The variable importance for the four covariates (wind speed, cumulative precipitation, maximum storm surge, and previous outages) was summed to calculate the total importance value for each zip code. Then the percent of total importance accounted for by each covariate was calculated. For each of the four covariates, we plotted the percent variable importance by zip code. We visually reviewed these plots to identify how the percent importance for each covariate differed geographically. The plots were also evaluated in light of the plots of the covariate values, so that the magnitude of the covariates was accounted for in evaluating the percent importance trends.

3. RESULTS AND DISCUSSION

3.1. Quantile Regression Forest

We ran a quantile regression forest model on 10 selected zip codes in order to better understand the predictive accuracy of the random forest model. These zip codes were selected to cover the geographical range of the state and to include zip codes with varying numbers of outages. Plots of the quantile regression forest results for three zip codes are shown in Fig. 2. These plots show the 90% prediction confidence intervals and whether predictions using out-of-bag data fall inside or outside of the prediction intervals. As shown on the plots, the majority of the predictions fall within the prediction intervals.

Table II shows the percent of predictions that fall between the 10% and 90% quantiles for the 10 zip codes analyzed using the quantile regression forest model. The percent coverage (percent of predictions within the 80% confidence interval) was calculated for three ranges of delta outages: low (0–2), medium (2–75), and high (75 and above), so that we could understand how the predictive accuracy varied across a range of values. In some cases, no prediction values fell within the low or high range, and this is indicated with an N/A in Table II. The model predictive accuracy is poor within the low range, except for in one zip code. In the medium and high range, the predictive accuracy is generally good, with the exception of predictions for two zip codes in each range. None of the zip codes have a high coverage of the 80% interval throughout the low, medium, and high ranges. However, six of the zip codes have high coverage (75% or greater) in two of the ranges.

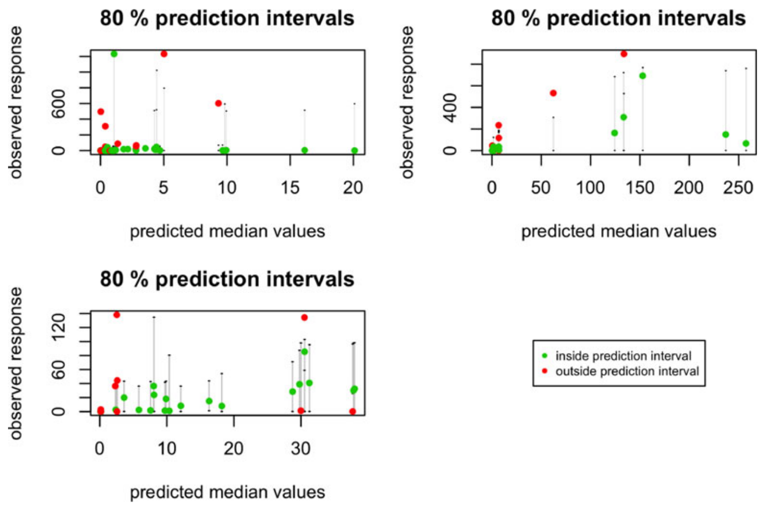

Fig. 2. Quantile regression forest plots for (a) zip code 70129, (b) zip code 71220, and (c) zip code 70546.

Table II. Percent of Predictions Within 80% Confidence Interval

Zip Code	Region	Maximum Outages	Delta Outage Range Low (0–2)	Medium (2–75)	High (75+)
70129	Southeast	3,364	0%	88%	57%
70454	Southeast	11,314	N/A	100%	96%
70546	Southwest	356	0%	100%	N/A
70560	South	1,684	3%	39%	50%
70607	Southwest	415	N/A	75%	100%
70806	South	11,616	N/A	100%	100%
71055	Northwest	367	N/A	100%	83%
71070	West	437	88%	69%	N/A
71220	North	2,896	3%	76%	100%
71351	East	3,314	100%	20%	96%

For low values of delta outages (0–2), four of the zip codes did not have values in this range. With the exception of two zip codes, the coverage of the 80% interval is very low; the model has little reliability at the lowest level of delta outages. For middle of the range values of delta outages (2–75), the model confidence interval coverage is fairly high for eight of the zip codes, ranging from 69% to 100%. However, the other two zip codes had only 20% and 39% of predictions within the 80% confidence interval. At the high end of the delta outages range (75+), the coverage accuracy varies significantly. This makes sense given the nature of power outages and the covariates used in the model. Very low increases in power outages are not likely well correlated to storm characteristics, and are more likely caused by random events occurring at individual houses. Very high increases in power outages can sometimes be correlated with high precipitation or wind, but could also occur due to sudden problems in the power grid.

Given the low percentage of predictions within the 80% confidence interval for several analyzed zip codes, we decided to investigate whether changing the data set from including all positive delta outages to only delta outages greater than one would increase predictive accuracy. Table III shows this comparison. Increased percent predictions within the 80% confidence interval occurred for nine of the zip codes, while a slight decrease was observed in zip code 71351. Based on this marked improvement, we decided to include only delta outages greater than one for the remainder of the analysis. This created a more accurate model, without reducing functionality, since prediction of very low delta outages (<1) is unnecessary.

3.2. Variable Importance

The variable importance results for the random forest model with all covariates included are shown

Table III. Percent of Predictions Within 80% Confidence Interval, Delta Outages Greater Than 0 Versus Greater Than 1

Zip Code	Percent Predictions Within 80% Confidence Interval	
	Delta Outages 0+	Delta Outages 1+
70129	59%	85%
70454	96%	96%
70546	65%	100%
70560	24%	74%
70607	82%	100%
70806	100%	100%
71055	96%	100%
71070	73%	91%
71220	75%	100%
71351	69%	67%

in Fig. 3. Variable importance is a measure of the contribution of a given covariate to the model prediction accuracy, and the magnitude of the importance is based on the data set. In Fig. 3, the variable importance is presented as the increase in node purity resulting from splitting over each variable, averaged over all trees. Cumulative precipitation, wind speed, and previous outages are the most important variables, followed by population and hourly precipitation. All of the surge variables, along with wind gust duration, had considerably lower variable importance. This differs from some previous work where wind gust duration was shown to be an important variable[17] and may be specific to this hurricane, for which wind speeds were lower than in the hurricanes included in the Han *et al.*[17] work.

Based on these results, four covariates were selected as part of a reduced covariate set to be used for the remainder of the analysis. These covariates were: cumulative precipitation, wind speed, previous outages, and maximum surge. Maximum surge depth was selected over average surge depth in each zip code because it had a clearer physical interpretation than the average surge depth yet had nearly the same importance score. Population was not included because the remainder of the analysis was done on an individual zip code basis wherein population is constant. The random forest model for the entire state was rerun with this reduced set of covariates. The resulting variable importance plot is included as Fig. 4. In this model, the cumulative precipitation covariate has the highest variable importance, followed closely by previous outages and wind speed. Maximum surge

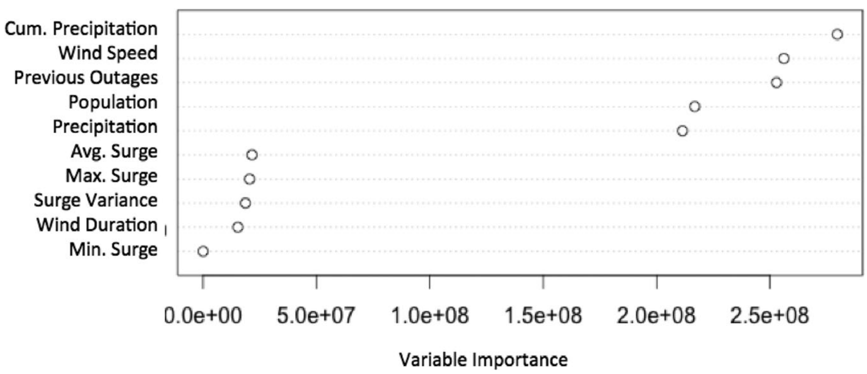

Fig. 3. Variable importance, all covariates included.

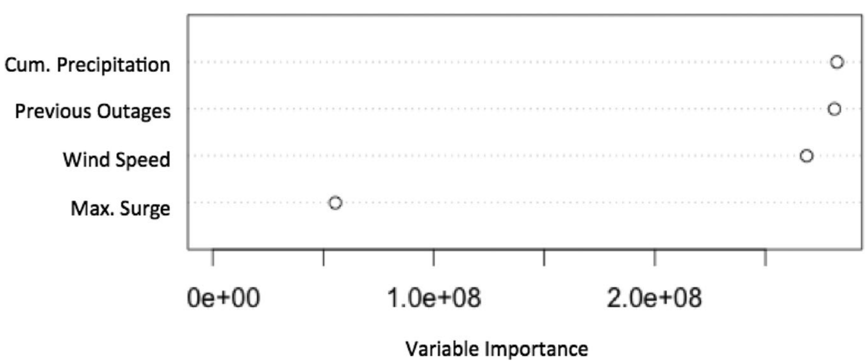

Fig. 4. Variable importance, reduced covariate set.

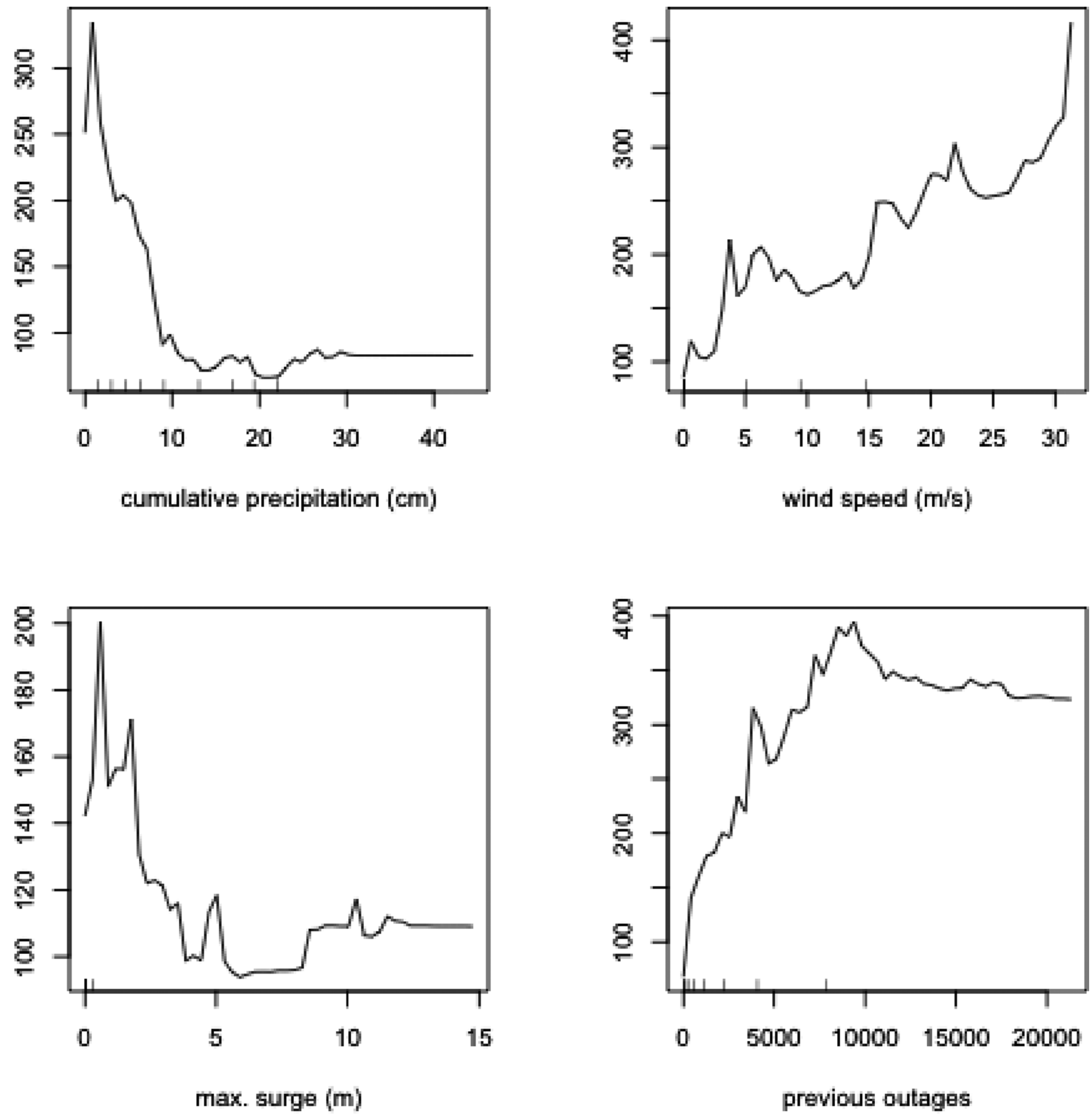

Fig. 5. Partial dependence plots: (a) partial dependence on cumulative precipitation, (b) partial dependence on wind speed, (c) partial dependence on maximum surge, and (d) partial dependence on number of previous outages. The x-axis represents the value of the covariate and the y-axis represents the marginal influence of the covariate on delta outages.

has a lower importance, as should be expected since only a small portion of the state was impacted by storm surge.

3.3. Partial Dependence

Partial dependence plots were generated for the four covariates in the reduced set, and are provided as Fig. 5. Partial dependence provides insight into the marginal impact of the covariate on the response variable, increase in outages.

The marginal influence of the cumulative precipitation covariate is highest for about 0–10 cm of precipitation. This is primarily due to the timing of the storm, with the highest values of delta outages generally occurring in the earlier part of the storm. Cumulative precipitation continued for days after the initial power outages occurred, with limited number of power outages occurring later in the storm. This resulted in a higher marginal influence for lower values of cumulative precipitation. Additionally, only a small percentage of zip codes experienced the

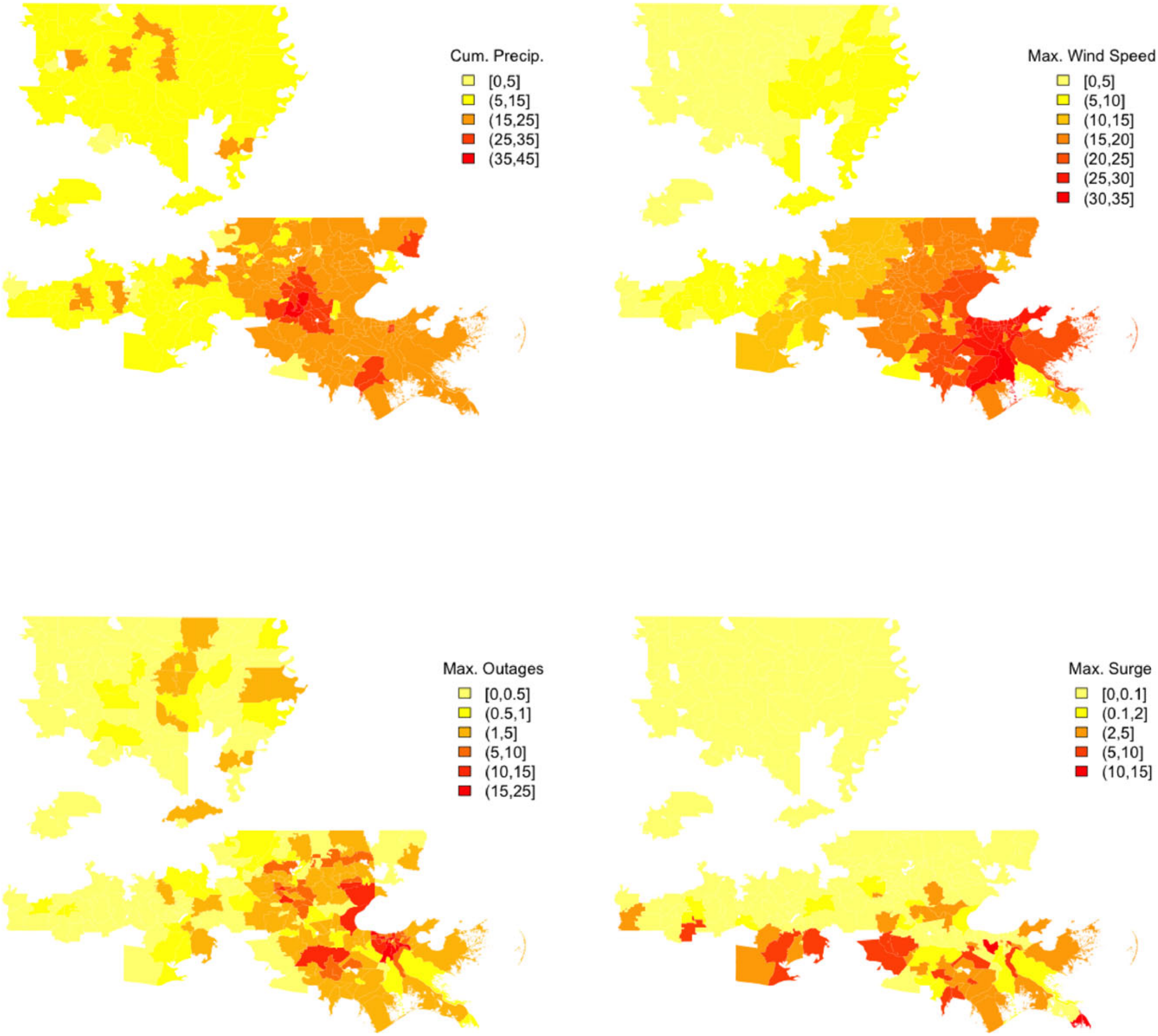

Fig. 6. Covariate values for (a) cumulative precipitation (cm), (b) maximum wind speed (m/sec), (c) maximum surge (m), and (d) maximum number of outages 1,000s. Zip codes not colored are not part of the utility's service area (colors visible in on-line version).

highest cumulative precipitation totals (30+ cm). The marginal influence of wind speed generally increases with increasing wind speed, which is intuitive. The influence of maximum surge is more variable, which may be due to the fairly low number of zip codes that experienced storm surge. The influence is higher at lower values of surge, likely because few zip codes experienced maximum surge values above 5 m. The marginal influence of the previous outages covariate increases up to around 10,000 outages, and then slightly decreases, since once a high number of outages occurs in a zip code, additional outages may be small in magnitude, as most customers have already lost power.

3.4. Geospatial Analysis

In order to analyze spatial trends across the state, we generated plots to get a sense of the magnitude of precipitation, wind speed, storm surge, and outages, and how the magnitude varied across the state. These plots are presented as Fig. 6. Total precipitation (cumulative precipitation) was highest in the southeastern part of the state, with more than 30 cm of

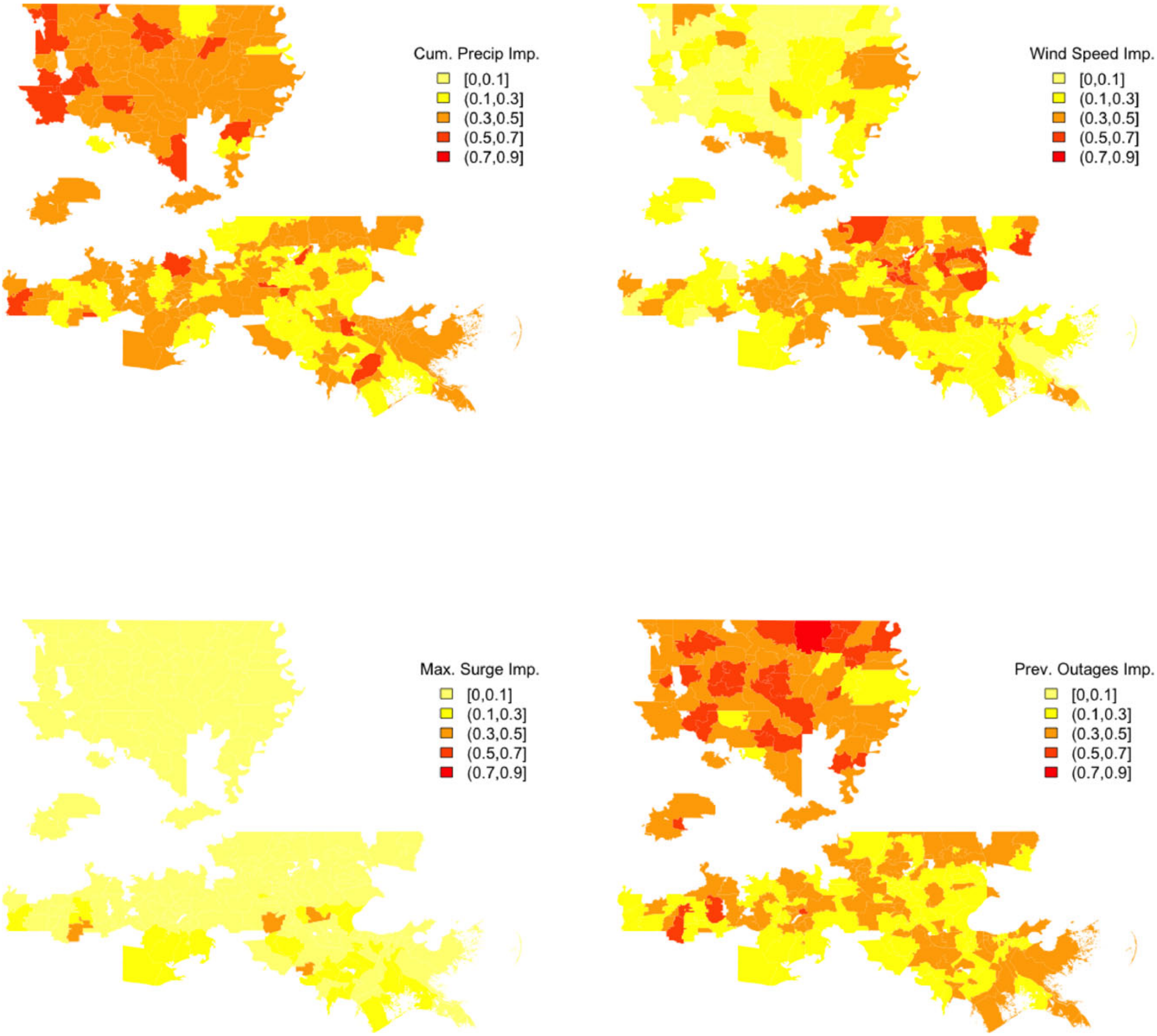

Fig. 7. Percent importance plots for (a) cumulative precipitation, (b) wind speed, (c) maximum surge, and (d) previous outages.

precipitation recorded in some locations. Maximum wind speed was also highest in the southeastern part of the state, where the hurricane made landfall. Maximum storm surge was highest in zip codes bordering the Gulf of Mexico, as well as in several zip codes bordering the Mississippi River. The maximum numbers of power outages were observed in zip codes in the southeast, around New Orleans, where the population is greatest and the storm impacts were more pronounced.

Fig. 7 illustrates the relative importance of cumulative precipitation, wind speed, maximum storm surge, and previous outages for all zip codes analyzed in Louisiana. In the northern part of the state, both cumulative precipitation and previous outages had high relative importance. Moderate amounts of precipitation occurred in this area, while wind speeds and total number of outages in the northern zip codes were lower than in other parts of the state. In the east central part of the state (Baton Rouge area), moderate to high precipitation, winds, and outages were experienced. Wind speed generally had the highest importance in this region, but precipitation and previous outages were also important. In the southeast

(New Orleans area), high wind speeds, precipitation, and outages occurred. Precipitation and previous outages had the highest importance in this region, and wind speed was also of importance. In the southwest and south central portions of the state, low to moderate precipitation and winds were experienced. High storm surge occurred in some coastal zip codes. The overall number of outages was low in most zip codes in this region, and the relative importance of each covariate varied considerably by zip code.

Cumulative precipitation was of moderate to high importance in most zip codes throughout the state, including those with relatively low precipitation. Conversely, wind speed generally only had high importance in areas that experienced high wind speeds. With the exception of a few zip codes, the percent importance for maximum storm surge was less than 30%, even in coastal areas. The relative importance of previous outages was moderate to high in most zip codes, and the maximum number of outages in a zip code does not seem directly related to the importance of previous outages in that zip code.

These results indicate that the importance of covariates varies geographically. This is due to the storm's track and characteristics, but also potentially due to the interaction of other factors pertaining to the topography and power system. Both wind speed and cumulative precipitation were highest in the east central and southeastern part of the state, due to the storm's track; however, wind speed generally had greater importance in those areas than precipitation. In the northern part of the state, where precipitation was moderately high, but wind speeds were low, precipitation was of greater importance. The previous outages covariate was generally more important in areas that had a low to moderate maximum outages value.

4. CONCLUSIONS

The purpose of this analysis was to provide insight on how rainfall and storm surge, along with wind, contribute to risk of power outages in hurricanes. By conducting a longitudinal study of outages at the zip code level, we were able to gain insight into the causal drivers of power outages during hurricanes. Our analysis showed that the correlation of storm characteristics with power outages and the importance of the covariates can vary geographically. In Louisiana, during Hurricane Isaac, rainfall and previous outages were the most important covariates in the north, while both rainfall and wind were important in the southeast. Rainfall, wind, and previous outages were all relatively important in the southwest. With the exception of a few zip codes, storm surge was generally not an important variable in predicting power outages, reinforcing the findings of Guikema *et al.*,[23] which also found that hurricane storm surge was not a particularly important variable in predicting power outages from hurricanes. The geographical variation of the correlation between storm characteristics and power outages is likely due to physical characteristics of the location and of the storm. In areas where the highest wind speeds are experienced, wind is likely to be the most important covariate. Elsewhere, the importance of covariates differs geographically.

While a random forest model proved to offer good out-of-sample predictive accuracy for this data set, a quantile regression forest provided additional information about the uncertainty in and accuracy of the estimates. We found that modeling only hours with delta outages greater than one resulted in improved predictive accuracy. The low-outage periods proved to be difficult to model accurately, as one would expect. Hours with small but positive increases in outage counts at the zip code level are more likely associated with random events than the types of larger-scale system damage that cause higher magnitude outages.

Based on previously published modeling efforts that focused on wind-related covariates to predict power outages, one might expect that wind speed would be the most significant covariate in our model, particularly in areas that experienced high wind speeds. Wind speed was of high importance in areas with high wind speeds, but cumulative precipitation was of moderate to high importance in more parts of the state, and was also important in the areas that experienced high winds. Storm surge was of limited importance in most areas, including those that experienced storm surge. These results point to the conclusion that the use of only wind-related variables in power outage forecasting models may result in a less accurate model than one that includes additional variables such as precipitation and perhaps surge inundation, especially in areas outside of the highest wind areas. Storm characteristics and their importance vary from storm to storm, and while many outages may be driven by wind, power outage modelers should include other covariates, particularly precipitation, to improve their model's robustness to differing storm conditions.

In addition to storm characteristics differing from storm to storm, our findings indicate that correlation of storm characteristics with power outages can vary geographically. It is unclear if this variation is due to characteristics of the storm, or other geographic considerations such as topography, power system characteristics, vegetation, and soil types.[24] Completing this type of analysis over multiple storms might clarify the reasons for this variation. Analysis of multiple hurricanes would also help assess the robustness of this analysis, and would be useful in informing the development of a power outage model for a state or region. This type of longitudinal analysis could result in a better understanding of the drivers of power outages and in better predictive models.

ACKNOWLEDGMENTS

The authors would like to acknowledge Michael Gao, a Johns Hopkins University undergraduate researcher, for his GIS assistance and members of Seth Guikema's research group for their assistance with collecting power outage data during Hurricane Isaac. This work was supported by a fellowship to Gina Tonn by the Environment, Energy, Sustainability and Health Institute (E2SHI) and the NSF IGERT Water, Climate, and Health program at Johns Hopkins University as well as by NSF grant number 1149460 from CMMI. The views in this article are those of the authors and do not necessarily reflect those of the sponsors.

REFERENCES

1. UPDATE 1—Entergy estimates Hurricane Isaac damage at $500 mln. Chicago Tribune, September 18, 2012.
2. NHC. NHC GIS Archive — Tropical Cyclone Best Track, 2015. Available at: http://www.nhc.noaa.gov/gis/archive_best track_results.php. Accessed January 6, 2015.
3. Han S, Guikema S, Quiring S. Improving the predictive accuracy of hurricane power outage forecasts using generalized additive models. Risk Analysis, 2009; 29(10):1443–1453.
4. Nateghi R, Guikema S, Quiring S. Power outage estimation for tropical cyclones: Improved accuracy with simpler models. Risk Analysis, 2014; 34(6):1069–1078.
5. Guikema S, Nateghi R, Quiring S. Predicting infrastructure loss of service from natural hazards with statistical models: Experiences and advances with hurricane power outage prediction. Proceedings ESREL, 2013.
6. Entergy. Entergy View Outages, 2012. Available at: http://www.entergy.com/storm_center/outages.aspx, Accessed August 27–September 5, 2012.
7. NOAA. National Climatic Data Center (NCDC), 2012. Available at: http://www.ncdc.noaa.gov, Accessed November 16, 2012.
8. Dietrich JC, Zijlema M, Westerink JJ, Holthuijsen LH, Dawson C, Luettich RA, Jensen R, Smith JM, Stelling GS, Stone GW. Modeling hurricane waves and storm surge using integrally-coupled, scalable computations. Coastal Engineering, 2011; 58(1):45–65.
9. Luettich R, Westerink J. Formulation and numerical implementation of a 2D/3D ADCIRC Finite Element Model Version 4.46, 2010. Available at: http://adcirc.org/adcirc_theory_2004_12_08.pdf, Accessed November 13, 2010.
10. Ferreira CM, Irish J, Olivera F. Uncertainty in hurricane surge simulation due to land cover specification. Journal of Geophysical Research: Oceans, 119(3):1812–1827.
11. Booij N, Ris RC, Holthuijsen LH. A third generation wave model for coastal regions. Model description and validation. Journal of Geophysical Research, 1999; 104(C4):649–7666.
12. Bunya S, Dietrich J, Westerink J, Ebersole B, Smith J, Atkinson J, Jensen R, Resio D, Luettich R, Dawson C, Cardone V, Cox A, Powell M, Westerink J, Roberts H. A high-resolution coupled riverine flow, tide, wind, wind wave, and storm surge model for southern Louisiana and Mississippi. Part I: Model development and validation. Monthly Weather Review, 2010; 138(2):345–377.
13. Dietrich J, Bunya S, Westerink J, Ebersole B, Smith J, Atkinson J, Jensen R, Resio D, Luettich R, Dawson C, Cardone V, Cox A, Powell M, Westerink J, Roberts H. A high-resolution coupled riverine flow, tide, wind, wind wave, and storm surge model for southern Louisiana and Mississippi. Part II: Synoptic description and analysis of Hurricanes Katrina and Rita. Monthly Weather Review, 2010; 138(2):378–404.
14. Mattocks C, Forbes C. A real-time, event-triggered storm surge forecasting system for the state of North Carolina. Ocean Modelling, 2008; 25(3):95–119.
15. NOAA. Atlantic basin hurricane database (HURDAT), 2013. Available at: http://www.aoml.noaa.gov/hrd/hurdat/, Accessed July 8, 2013.
16. Willoughby HE, Darling RWR, Rahn ME. Parametric representation of the primary hurricane vortex. Part II: A new family of sectionally continuous profiles. Monthly Weather Review, 2006; 134(4):1102–1120.
17. Han S, Guikema S, Quiring S, Lee KH, Rosowsky D, Davidson R. Estimating the spatial distribution of power outages during hurricanes in the Gulf Coast region. Reliability Engineering & System Safety, 2009; 94(2):199–210.
18. US Census Bureau. American Community Survey, 2013. Available at: http://www.census.gov/acs/www, Accessed July 7, 2013.
19. Hastie T, Tibshirani R, Friedman J. The elements of statistical learning: Data mining. P. 763 in Inference and Prediction, 1st ed. New York: Springer, 2001.
20. Liaw A, Wiener M. Classification and regression by randomForest. R News, 2002; 2(3):18–22.
21. Meinshausen N. Quantile regression forests. Journal of Machine Learning Research, 2006; 7:983–999.
22. Meinshausen N. quantregForest: Quantile regression forests. R package version 0. 2–3, 2012.
23. Guikema SD, Udo I, Irish J, Nateghi R. The effects of hurricane surge in power system outage risk models. Submitted to Natural Hazards [under first review], 2014.
24. Quiring S, Zhu L, Guikema S. Importance of soil and elevation characteristics for modeling power system outages due to hurricanes. Natural Hazards, 2011; 58(1):365–390.