

Improving Hurricane Power Outage Prediction Models Through the Inclusion of Local Environmental Factors

D. Brent McRoberts¹, Steven M. Quiring², and Seth D. Guikema³

1. Department of Geography, Texas A&M University

2. Department of Geography, Ohio State University

3. Department of Industrial and Operations Engineering, University of Michigan

Corresponding Author Address:

D. Brent McRoberts

Department of Geography

Texas A&M University

MS 3147

College Station, TX 77843-3147

mcrobert@tamu.edu

p: 979-845-7141

f: 979-862-4487

ABSTRACT

This is the author manuscript accepted for publication and has undergone full peer review but has not been through the copyediting, typesetting, pagination and proofreading process, which may lead to differences between this version and the [Version of Record](#). Please cite this article as [doi: 10.1111/risa.12728](#).

This article is protected by copyright. All rights reserved.

Tropical cyclones can significantly damage the electrical power system, so an accurate spatiotemporal forecast of outages prior to landfall can help utilities to optimize the power restoration process. The purpose of this paper is enhance the predictive accuracy of the Spatially Generalized Hurricane Outage Prediction Model (SGHOPM) developed by Guikema et al. (2014). In this version of the SGHOPM, we introduce a new two-step prediction procedure and increase the number of predictor variables. The first model step predicts whether or not outages will occur in each location and the second step predicts the number of outages. The Guikema et al. (2014) SGHOPM environmental variables were limited to the wind characteristics (speed and duration of strong winds) of the tropical cyclones. This version of the model adds elevation, land cover, soil, precipitation, and vegetation characteristics in each location. Our results demonstrate that the use of a new two-step outage prediction model and inclusion of these additional environmental variables increases the overall accuracy of the SGHOPM by ~17%.

1. INTRODUCTION

Severe weather events, such as hurricanes, cause widespread and prolonged power outages. A key part of preparing for, and responding to, hurricanes is requesting and positioning resources in advance of a storm. This decision has a significant impact on how long it takes to restore power, as well as the cost of the restoration effort. Forecasts of the number and locations of outages can help utilities balance the high cost of external resources with the need to have enough resources to restore power quickly. Previous work has been done on power outage forecasting for high wind events, especially for hurricanes. Past efforts include *Reed*⁽¹⁾, where a linear regression model was developed for estimating power outages during wind storms in Seattle. *Liu et al.*^(2,3) used generalized linear models (GLMs) and generalized linear mixed models to estimate outages during hurricanes, and in follow-up work by *Han et al.*^(4,5) generalized additive models were used. *Nateghi et al.*⁽⁶⁾ built on this previous work by employing a Random Forest model⁽⁷⁾ to achieve higher predictive accuracy. All of these models used a wide range of input variables, including data about the power system, hurricane winds, land use, topographic information, soil moisture levels, and other geographic and climatological factors. These models provide strong predictive accuracy in the utility service area for which they were developed, supporting improved utility decision-making. However, these models do not provide outage estimates for areas outside of the utility service area and they are not available to emergency and risk management personnel in local, state, and federal governments.

Guikema et al.⁽⁸⁾ developed a spatially-generalized hurricane power outage prediction model, building on *Nateghi et al.*⁽⁶⁾. This model can be used anywhere along the U.S. coastline because it is based on publicly available information. However, one limitation of the *Guikema et al.*⁽⁸⁾ model is that it uses a greatly simplified set of input variables (population, maximum 3-sec wind gust, and duration of sustained winds exceeding 20 m sec⁻¹

¹⁾ as compared to the utility-specific models (e.g., *Nateghi et al.*⁽⁶⁾). This leads to lower predictive accuracy relative to the utility-specific models.

The purpose of this paper is to improve the predictive accuracy of the Spatially Generalized Hurricane Outage Prediction Model (SGHOPM) first introduced in *Guikema et al.*⁽⁸⁾ by adding more explanatory variables, specifically: elevation, land cover, root zone depth, precipitation, soil moisture, and tree characteristics. The value of these additional variables is evaluated by quantifying the reduction in error relative to the *Guikema et al.*⁽⁸⁾ version of the SGHOPM (hereafter called the baseline model). Several different combinations of explanatory variables were explored using a cross-validation procedure to determine the ideal covariate set for predicting outages in the Random Forest model.

In addition to adding more variables, this paper introduces a two-stage version of the SGHOPM. The two-stage approach to hurricane outage modeling was introduced by *Guikema and Quiring*⁽⁹⁾ to account for the zero-inflation of datasets that is characteristic of past outage events. Traditional statistical methods for dealing with zero-inflation, such as zero-inflated Poisson and zero-inflated negative binomial regressions, are unlikely to handle the complexity of factors contributing to non-outages in hurricanes⁽⁹⁾. Instead of traditional statistical models, *Guikema and Quiring*⁽⁹⁾ improved the accuracy of outage predictions using a two-stage classification tree/regression model approach. The first stage of the *Guikema and Quiring*⁽⁹⁾ “tree-GAM” model fits a classification tree to predict a zero or non-zero response and the second stage makes a quantitative prediction of outages using a generalized additive model (GAM). The use of classification trees provides an improved assessment of power system responses to explanatory variable thresholds (i.e., maximum wind speed at which are poles are likely to be blown over) compared to traditional models⁽⁹⁾. In this paper, the first stage of the two-stage SGHOPM makes a prediction with only two outcomes using a Random Forest classification model: (1) zero outages or (2) one or more outages. The second stage of

the model determines the number of outages using a Random Forest regression model, conditional on the first stage of the model predicting an outcome of one or more outages.

This paper is organized as follows. A description of the data is provided in Section 2. Section 3 provides an overview of the SGHOPM model and the methodology for selecting the best set of covariates is described in Section 4. Section 5 presents the results and Section 6 examines the influence of individual variables on outage predictions.

2. DATA AND METHODS

2.1. Data

Census tracts are used as the spatial unit of analysis in this paper. Population data are used as a proxy for the number of utility customers because the SGHOPM only uses publically available sources of data. The SGHOPM predicts the fraction of the population that will lose power (fractional outages) in each census tract.

The SGHOPM is developed and trained using the gridded outage data from a private utility company who wishes to remain anonymous. Their service area covers parts of three states in the southeastern U.S. Outage data from six tropical cyclones (TCs) are used in this paper: Opal (1995), Danny (1997), Georges (1998), Ivan (2004), Dennis (2005), and Katrina (2005).

The explanatory variables that are evaluated in this study can be divided into two categories. One category contains variables related to geographic and environmental characteristics that are invariant (static) in time. These variables include various measures of topography, land cover, tree characteristics, and soil characteristics in each census tract. The second category contains variables that are time-dependent (dynamic) and represent the antecedent conditions when a TC makes landfall. These variables include various measures of soil moisture and precipitation.

All of these variables originate from different sources and each has a different spatial resolution. The data for the static variables are available at a high resolution, so these data were re-scaled to census tracts by identifying the census tract centroid that is closest to the grid cell. The soil moisture and precipitation data used are at a coarser resolution than most census tracts, so they are re-scaled to the census tracts using spatial interpolation. The remainder of Section 2 provides a description of all of the predictor variables that are considered for inclusion in the SGHOPM (Table 1).

2.2. Response variable

The response variable \hat{Y}_i in the SGHOPM is the fraction of outages for a given census tract i ; fractional outages are more useful than the total number of outages due to the variability in census tract populations⁽¹⁰⁾. In training the SGHOPM, both the number of customers and number of customers without power (which will be referred to as outages in the rest of the paper¹) from the utility-specific data were used to compute fractional outages. In some instances, the number of customer outages reported in the utility's outage management system exceeded the number of customers in the utility dataset as discussed in *Guikema et al.*⁽¹¹⁾. When this occurred, the number of outages was set to the number of customers.

2.3. Baseline explanatory variables

Guikema et al.⁽⁸⁾ developed a SGHOPM that used only three variables: census tract population, the maximum 3-sec wind gust, and duration of sustained winds exceeding 20 m sec⁻¹. This 3-variable version of the SGHOPM is referred to as the baseline model. The only

¹ We define the term outages in this paper as the number of customer meters without power. Therefore a "customer" is really a customer meter.

difference is that the baseline model used in this paper uses population density rather than total population because a normalized population metric (density) is more appropriate for determining fractional outages.

2.4. Static explanatory variables

The static variables that are used include various measures of topography, land cover, tree characteristics, and soil characteristics. The value of including topographical characteristics in utility-specific power outages models was initially demonstrated by *Guikema et al.*⁽¹²⁾ and explored in more depth by *Quiring et al.*⁽¹³⁾. The topographical variables are derived from a global 30-arcsecond digital elevation model (DEM) produced by the United States Geological Survey (see *Danielson and Gesch*⁽¹⁴⁾ for details) and they include: mean elevation ($X_i, elev_mean$), median elevation ($X_i, elev_median$), standard deviation of elevation ($X_i, elev_stdev$), minimum elevation ($X_i, elev_min$), and maximum elevation ($X_i, elev_max$).

Davidson et al.⁽¹⁵⁾ found that the inclusion of land cover (LC) types improved the accuracy of hurricane-related outage predictions. *Quiring et al.*⁽¹³⁾ demonstrated that LC variables are particularly useful when utility-specific asset data (e.g., number of poles, transformers, etc.) is not available because certain LC variables can serve as proxies for these data. The LC data used in this paper are from the National Land Cover Database 2011 (NLCD 2011; see *Homer et al.*⁽¹⁶⁾ for details). There are 8 major land cover classes in the NLCD 2011, based on the *Anderson*⁽¹⁷⁾ classification system. The fractional coverage of these 8 LC types was determined for each census tract (i): water LC (X_i, LC_water), developed LC ($X_i, LC_developed$), barren LC (X_i, LC_barren), forest LC (X_i, LC_forest), scrub LC (X_i, LC_scrub), grassland LC ($X_i, LC_grassland$), pasture LC ($X_i, LC_pasture$), and wetlands LC ($X_i, LC_wetlands$).

A variety of tree/vegetation-related variables have been used in previous outage prediction modeling. For example, *Guikema et al.*⁽¹⁸⁾ and *Nateghi et al.*⁽⁶⁾ incorporated a

measure of tree trimming frequency, and *Lui et al.*⁽²⁾ and *Madeira*⁽¹⁹⁾ included variables related to specific tree species. All of these previous projects found that the inclusion of vegetation information can significantly enhance model performance.

This paper is the first to incorporate tree species data and tree characteristics such as type of root system, depth of the root system, tree diameter, tree height, density of the wood, hardness of the wood, and the crushing strength of the wood. The tree species data are from the 2012 National Insect and Disease Risk Map (NIDRM; *Krist et al.*⁽²⁰⁾). The NIDRM identifies a single, dominant tree species in each 240 m grid cell⁽¹⁹⁾. There are a total of 217 tree species in our region. The eight tree-related variables that are considered in this study are: fractional area of a census tract (i) covered by trees ($X_{i, treed}$), percentage of trees with a deep root system ($X_{i, deep}$), percentage of trees with a taproot system ($X_{i, taproot}$), maximum tree species height ($X_{i, hgt}$), maximum tree species diameter at breast height ($X_{i, dbh}$), tree density ($X_{i, density}$), Janka Hardness scale ($X_{i, janka}$), and crushing strength ($X_{i, crushing}$). Details on the tree species variables used in this study, and the source of the data, are provided in Table 2 (*Van Dersal et al.*⁽²¹⁾; *Burns and Honkala*⁽²²⁾; *Stoecklein*⁽²³⁾; *US Forest Service Tree List*⁽²⁴⁾).

Root zone (RZ) depths are derived from the USDA Gridded Soil Survey Geographic (gSSURGO) which is available at 30 m resolution. The root zone is characterized as the depth within the soil column that crop roots can effectively extract water (*Soil Survey Staff*⁽²⁵⁾). Based on *Dobos et al.*⁽²⁷⁾, the maximum RZ depth in gSSURGO is 1.5 m. *Quiring et al.*⁽¹³⁾ used a related metric, the depth to bedrock, to characterize the effective depth of the soil layer. The three RZ depth metrics used in this study (mean RZ depth (X_{i, RZ_mean}), majority RZ depth ($X_{i, RZ_majority}$), and median RZ depth (X_{i, RZ_median})) are based on a summary of all the 30 m gSSURGO grid cells that are within each census tract. Because of the high resolution of the gSSURGO dataset, it was first re-scaled to the same 240 m resolution as the tree species data, using the majority value from 64 gSSURGO grid cells.

2.4. Dynamic variables

The dynamic soil moisture and precipitation variables represent moisture conditions 3 days prior to landfall. The 3-day lag was chosen because when the SGHOPM is used for operational forecasts of power outages, these forecasts are typically initialized 3 to 5 days in advance of landfall. Soil moisture and precipitation have been shown to be important for power outage modeling because wetter soils can increase the likelihood that trees and utility poles will be uprooted/fail in strong winds (*Han et al.*⁽⁴⁾; *Quiring et al.*⁽⁶⁾).

Both the precipitation and soil moisture are from the North America Land Data Assimilation System Phase 2 (NLDAS-2) which contains modeled parameters (including soil moisture) and forcing variables (including precipitation). These data are available at 1/8th degree spatial resolution and at an hourly temporal resolution from January 1979-present. NLDAS-2 contains three different land-surface models (MOSAIC, NOAH, and VIC). In this study we use soil moisture data from the VIC model (*Liang et al.*⁽²⁸⁾) because of its past performance in accurately assessing variations in soil moisture (*Han et al.*^(4,5); *Guikema et al.*^(8,12)).

VIC-derived soil moisture is estimated for three layers: 0 – 10 cm ($X_{i, \text{soil1}}$), 10 – 40 cm ($X_{i, \text{soil2}}$), and 40 – 100 cm ($X_{i, \text{soil3}}$). When the soil layers did not match these standard depths, the volumetric moisture content was converted to these standard depths using a simple linear interpolation. Rather than using the fractional soil moisture values used in past studies (*Han et al.*^(4,5); *Guikema et al.*⁽¹²⁾), all soil moisture data in this paper were converted to percentiles based on the historical cumulative distribution function (CDF). The soil moisture CDFs used 36 years of historical data from 1979 through 2014 to compute non-parametric L-moment ratios (see *Hosking and Wallis*⁽²⁹⁾ for details) that were transformed to Pearson Type III distribution parameters. This approach differs from that used by the Climate Prediction

Center (CPC), which uses an empirically derived historical distribution (*Andreadis et al.*⁽³⁰⁾; *Mo et al.*⁽³¹⁾).

Precipitation is represented using the Standardized Precipitation Index (SPI). The SPI uses a CDF to normalize precipitation for a given time period using historical precipitation data (*McKee et al.*⁽³²⁾; *Guttman*⁽³³⁾; *Heim*⁽³⁴⁾). We use Pearson Type III to fit the CDF, which is consistent with *Han et al.*^(4,5). Five different SPI time-scales were used to represent short-term (1-month ($X_{i, SPI1}$), medium-term (3-month ($X_{i, SPI3}$), 6-month ($X_{i, SPI6}$), and long-term (12-month ($X_{i, SPI12}$), and 24-month ($X_{i, SPI24}$)) moisture conditions in each census tract.

The NLDAS-2 data were rescaled to census tracts by interpolating data from the four nearest NLDAS-2 grid cells to the census tract centroid using inverse distance weighting. NLDAS-2 soil moisture and precipitation were interpolated to census tracts prior to computing the soil moisture percentiles and SPI.

The 1200 UTC soil moisture data were used to fill a daily time series at each census tract from 1 Jan 1979 – 31 Dec 2014. The hourly precipitation were aggregated to daily data. Pearson Type III parameters were computed for all three soil layers and five SPI time scales for each census tract and calendar date (i.e., 30 Jul).

3. Model Background

3.1. Random Forest model

The SGHOPM is a non-parametric ensemble data mining model that is based on the Random Forest method created by *Breiman*⁽⁷⁾. Random Forest models are insensitive to outliers and noise (*Hastie et al.*⁽³⁵⁾) and have been proven to make accurate outage predictions (e.g., *Nateghi et al.*⁽⁶⁾). The SGHOPM is coded in the open-source R programming language, calling on the “randomForest” library for model computations and can be run on any modern computer. The Random Forest technique fits a large number (K) of regression trees, each

time using a bootstrapped sample from a training data set (*Guikema et al.*⁽⁸⁾). At each branch in the regression tree, the data are recursively partitioned into two groups based on a subset of the covariates until the number of terminal nodes reaches a specified value (*Hastie et al.*⁽³⁵⁾). The independence of each individual tree within the forest is improved by randomly sampling the training data before training a given tree and randomly selecting the set of covariates allowed to be used at each tree branch (*Nateghi et al.*⁽⁶⁾). Therefore, the set of K trees (in this paper $K = 500$) are approximately uncorrelated and unbiased (*Guikema et al.*⁽⁸⁾). For each tree, the predictive power of each variable is computed using the out-of-bag sample for that tree (*Hastie et al.*⁽³⁵⁾). The final prediction, given a set of input covariates, is the average of the individual tree predictions.

3.2. Cross-Validation Testing

We seek to minimize the prediction error of the SGHOPM model through the addition of explanatory variables to the baseline variables used by *Guikema et al.*⁽⁸⁾. Most regression models are built to optimize goodness-of-fit for the data used to train the model, and adding more variables increases model performance. However, this does not optimize the predictive power of the model. Random forest is powerful as a predictive tool because model development is based on the predictive performance of independent, out-of-bag samples *Hastie et al.*⁽³⁵⁾.

A holdout cross-validation analysis technique was used to assess the predictive power of each set of potential covariates in the SGHOPM, based on the same methodology as *Nateghi et al.*⁽⁶⁾. Given a set of n census tracts, we holdout 20% of the data, using random sampling without replacement, and train the model using the remaining 80%. We then test the out-of-sample predictive accuracy of the model using the 20% of data not used to train the model.

We repeat this process N times (in this paper $N = 30$), each time using different data in the training and validation groups.

Each training set uses 5,023 grid-cell–storm combinations (performance of the power system in one grid cell in one TC), which is 80% of the overall data set. Both the training and holdout groups can have data from the same TC (or the same census tract). This allows for a robust evaluation of each set of covariates. Model performance for a given set of covariates is measured by the percent improvement in mean absolute error (MAE) as compared to the *Guikema et al.*⁽⁸⁾ version of the SGHOPM. It is important to note that population density, maximum wind gust, and duration of strong winds are included in every set of covariates that are evaluated.

4. Model Development and Testing

4.1. Overview

The overall procedure for choice of explanatory variables began with testing covariate sets within each variable type (topography, land cover, tree characteristics, soil characteristics, SPI, and soil moisture). Each group of variables were evaluated using a 30-fold cross-validation procedure to determine the three explanatory variables with the lowest predictive error. A second 30-fold cross-validation was used to evaluate whether adding these three variables reduced the model error as compared to the baseline covariate set.

For each variable type (i.e., topography), numerous covariate sets were tested to determine the combination of variables that provided the largest improvement over the baseline model. The result was six different combinations of variables, one for each of the six variable types, which minimized MAE. These six variable combinations were then blended, with each blend containing variables of two or more types (i.e., a blend of dynamic variables contains SPI and soil moisture variables). Three types of blends used were (1) a blend of all

the static variable types, (2) a blend of the two dynamic variables, and (3) a blend of all six variable types. Additionally, “reduced” blends were formed using no more than three individual variables for each variable type. Overall, there were 12 main covariate sets tested, six using a single variable type and six using blends of two or more different variable types.

4.2. Random Forest Response Variable

One of the challenges in predicting power outages is the large number of locations with no outages (i.e., zero inflation; *Guikema and Quiring*⁽⁹⁾). For example, approximately 30% of the census tracts in the training dataset have zero outages (Fig. 2), a percentage that remains consistent even in the cross-validation training datasets. However, the predicted fractional outage is rarely zero. Therefore, in this paper the prediction is done in two phases. First, we train a Random Forest model to predict whether or not outages will occur in a census tract ($\hat{c}_i = 0$ or $\hat{c}_i = 1$), this is called the Binary Classification (BC) model. Second, we predict the fractional outage prediction ($0 < \hat{f}_i \leq 1$) if $\hat{c}_i = 1$ using a separate Non-Zero Outage (NOZE) model. If $\hat{c}_i = 0$, then $\hat{f}_i = 0$, regardless of the NOZE prediction \hat{Y}_i . This is the same approach that was used in *Guikema and Quiring*⁽⁹⁾, except they used a classification tree for the BC portion and a Generalized Additive Model (GAM) for the NOZE portion. In this paper we have use a Random Forest model for both steps.

There can be a different set of optimal covariates in each model, and the BC model was tested independently of the NOZE model. The BC Random Forest model response (\hat{c}_i) is either a zero or one. Therefore, census tracts in the training dataset with observed non-zero outage fractions were assigned c_i values of 1. The first choice in the model development was determining if the training dataset used to build the NOZE Random Forest model should include census tracts with zero outages. A testing procedure deemed that it was better to use

all the training data ($f_i \geq 0$) in the NOZE model rather than just the subset of training data with observed fractional outages ($f_i > 0$).

4.3. Binary Classification (BC) Model

Adopting a two-stage modeling process, without adding additional variables, improves model performance by almost 4% (standard deviation of 0.78% in the 30 holdouts) as compared to *Guikema et al.*⁽⁸⁾. Using the reduction in MAE relative to the *Guikema et al.*⁽⁸⁾ model as the performance metric, the BC model with an optimized covariate set improves the model by over 9% (standard deviation of 1.19%; Table 3). Several sets of covariates were evaluated to determine the optimal set of covariates for the BC model. After determining the set of variables within each group having the most predictive power, the best variables within each variable group were combined in an additional cross-validation procedure testing procedure. Performance of the six covariate set blends used in the BC model were compared to simpler covariate sets (Table 3); these included one variable of each type and the baseline variables. In general, using a larger number of covariates improves the performance of the BC model. Our results indicate that the BC performs best when using all of the static variable types (Table 3). In 20 of the 30 repetitions, the static covariate set (Table 2) outperformed the other 13 covariate sets.

4.4. Non-zero Outage (NOZE) Model

The purpose of the NOZE model is to predict how many outages will occur in the census tracks that were identified as experiencing storm impacts by the BC model. As with the BC model, we evaluate each of the variable types to determine which variables should be included in the final, optimal set of covariates (Table 4). The test results indicate that the optimal NOZE model covariate set contains 23 variables and that multiple variables from

each group are included in the optimal model (Table 2). Overall, there is a 17.3% improvement in the performance of the two-step SGHOPM with an enhanced set of covariates as compared to the *Guikema et al.*⁽⁸⁾ model.

5. Model application to previous tropical cyclones

After completing the holdout testing procedures, the SGHOPM was used to forecast outages for each of the six TCs with available data in a storm-specific holdout procedure. Rather than randomly withholding 20% of the data across all storms as done in Section 4, a storm-specific holdout was used (i.e., withhold all data from a single storm and use data from the remaining storms to train the model).

The two-step SGHOPM developed in this paper does poorly in comparison to other models in the storm-specific holdout testing (Table 5). Using two different NOZE model covariate sets, each prediction of number of outages was plotted as a function of the observed number of outages, using data from all six storms (Fig. 3). It is apparent that the storm-specific model predictions made using the covariate set found optimal in Section 4 systematically underestimates total outages when both the observed and predicted values are non-zero (Fig. 3a). Removing the dynamic variables from that covariate set improves the accuracy of the storm-specific predictions (Fig. 3b), however, there is still a systematic underestimation of predicted outages. In addition, the slope of the non-zero outage pairs (f_i and \hat{f}_i both non-zero) is less than one, indicating that the variance in the predicted values is less than the observed variance. The ratio of the variance of the predicted outages to the variance of the actual outages can be denoted as $R_{var} = \text{var}(\hat{f}_i \times P_i) / \text{var}(f_i \times P_i)$, where P_i is the census tract population. $R_{var} = 0.52$ for NOZE model using baseline plus the set of static variables and $R_{var} = 0.37$ for the *Guikema et al.*⁽⁸⁾ model (Table 5). These R_{var} values include the locations with no outages and there are fewer locations with no outages in the predictions

than in the observations. The BC model using only baseline variables predicted that no outages would occur in 15.7% of census tracts, compared to 18.2% using the covariate set with additional variables; both of these values are well below the observed frequency of 30.0%. This is one of the reasons that variance of the predictions is much less than the observed.

For each covariate set used to build a NOZE model, the ratio of the model prediction variance relative to the observed outage variance (R_{var}) was also computed using only non-zero outage pairs. A comparison using non-zero outage pairs gives a better idea of the NOZE model influence on variance and the results indicate a reduction in model variance relative to the observed outages for both the baseline ($R_{var} = 0.34$) and static ($R_{var} = 0.48$) versions of the NOZE model (Table 5). Although the predicted variance is much lower than the observed, the two-step SGHOPM used in this paper is an improvement over *Guikema et al.*⁽⁸⁾.

The total number of outages (summed over all census tracts) predicted over all six storms compares favorably to the observed outages (Fig. 4a). The inclusion of additional covariates generally increases the number of model outages predicted, with the exception of the antecedent precipitation variables (e.g., SPI). Model performance is highly variable when viewed on a storm-by-storm basis (Fig. 4b). The model performed better for the three storms (Opal, Ivan, and Katrina) with the largest values of observed outages. This agrees with our previous work which shows that the SGHOPM does well when there are strong storms that have large impacts on the power system, but the performance is more uneven for smaller storms *Quiring et al.*⁽¹³⁾.

Figure 5 compares the *Guikema et al.*⁽⁸⁾ model to the two-step version of the SGHOPM developed in this paper. At each census tract in our domain, the total number of model-predicted outages was summed over all six storms and this quantity was compared to the summation of observed outages (Fig. 5). In over 71% of the census tracts, the two-step

SGHOPM outperformed the *Guikema et al.*⁽⁸⁾ model. The mean accuracy of the outage predictions across the six storms improved by >25% in more than two-fifths of the census tracts; only 10% of census tracts had a decrease in accuracy of >25%. Much of this improvement is due to adopting a two-phase modeling approach.

In 113 of the 994 census tracts, there were zero total outages for all six storms. The BC model correctly predicted zero outages for 78 of these census tracts (nearly 70%); in these same 78 census tracts, the *Guikema et al.*⁽⁸⁾ model predicted over 315,000 outages. Fig. 5 also shows that outage predictions tend to be more accurate in census tracts that experience more outages (i.e., tracts with more than 10,000 observed outages over the six storms). This shows that the model performs better in places that experience more outages and for stronger storms.

6. Partial dependence of response variable to covariates

6.1 Variable Importance

Every Random Forest model measures variable importance (VI), which is computed as the data are recursively split into two groups at each node using a subset of the explanatory variables. For a given covariate set, each explanatory variable VI value represents its usefulness in splitting the data at each node. A “pure” node does a better job partitioning the response variable than an “impure” node. For classification models (e.g., BC model), VI is measured by the average decrease in the Gini index g ($Liaw^{(36)}$), which is a measure of node impurity. For a regression model (e.g., NOZE model), VI is measured by the average decrease in the residual sum of squares.

Table 6 shows VI for all the explanatory variables used in the optimal BC model covariate set. It is common practice to normalize VI by giving the most important variable a score of 100 (in our case maximum wind gust). As expected based on our previous work, the

duration of strong winds is important. The variable with the second highest VI score in the BC model is the density of wood (i.e., a trait that is tree-species specific). Further analysis indicates that the mean of the average density $\bar{\rho}_0$ for census tracts with no observed outages is 604.0, whereas $\bar{\rho}_1 = 614.0$ in census tracts with at least one observed outage. Based on a difference of means test, this is a statistically significant difference (p -value < 0.001). The wood density variable may be indicative of the relative mixture of softwoods and hardwoods within each census tract. It appears that softwood species tend to be associated with locations that experience lower fractional outages. This may be because softwoods are more flexible and therefore bend, but do not break, under strong winds.

Other important variables are those related to topography, which include the mean, median, maximum, and minimum elevation of census tracts. Elevation may serve as a proxy for the proximity to the coastline. Census tracts that are located closer to the coast are likely to experience higher wind speeds and therefore are associated with more outages. An additional consideration is that regions with higher elevation may leave trees more exposed to the impacts of high winds (*Chapman*⁽³⁷⁾), thus leading to more outages.

In the NOZE model, the two baseline variables related to the TC winds have the highest VI scores (Table 7), with population density ranked the fourth most important. Aside from the three baseline variables, the six highest VI scores belong to variables describing antecedent SPI and soil moisture. The VI scores are confirmation of the importance of the dynamic variables in making fractional outage predictions.

6.2 Partial Dependence

Partial dependence plots are used to illustrate the marginal contribution of a single explanatory variable to the Random Forest model response with the rest of the explanatory variables averaged out *Nateghi et al.*⁽⁶⁾. Whereas VI is a single metric, partial dependence

plots show changes in the response variable as a continuous function of a single explanatory variable value. Partial dependence plots for the eight explanatory variables that were identified as important in the holdout testing are shown in Fig. 6.

The partial dependence plots for the three baseline variables (Figs. 6a-c) have similar shape, with a sudden increase in predicted outages at lower values that asymptotes at larger values. More specifically, there is a large spike in outages when maximum winds are above 20 m sec^{-1} (Fig. 6a). There is a significant increase in predicted outages when strong winds ($\geq 20 \text{ m sec}^{-1}$) last 4 hours or more relative to shorter durations (Fig. 6b). The fractional outage prediction increases sharply with population density (Fig. 6c) at values $< 800 \text{ customers km}^{-2}$, with a more gradual increase at larger values of population density.

Partial dependence plots are shown for three dynamic variables, which are the layer 1 (Fig. 6d) and layer 2 (Fig. 6e) soil moisture percentiles and the 12-month SPI (Fig. 6f). The two soil moisture plots show that soil moisture has little explanatory power in drier soils. However, the partial dependence increases in wetter soils. In general, the stability of soil decreases with increasing wetness *Quiring et al.*⁽¹³⁾, leading to an increased susceptibility of trees in saturated soils being uprooted. The 12-month SPI plot (Fig. 6f) seems to contradict these findings because it shows that there is a dramatic increase in outages when $X_{i, SPI12} \leq -1$ (which should be associated with drier soils). However, negative 12-month SPI values are indicative of longer-term drought stress on trees, which can lead to weakening and increased susceptibility to being damaged (*Guikema et al.*⁽⁸⁾). Near-surface soil moisture (layer 1 and 2) is not necessarily strongly correlated with long-term drought conditions. Based on analysis of the partial dependence plots for the dynamic variables, trees are most susceptible to being blown down by strong winds when there are very wet shallow soils in regions where trees have been weakened by long-term drought conditions.

Average wood density (Fig. 6g) was the static variable with the highest VI score in the NOZE model and, as mentioned previously, it was important in the BC model. In general, predicted outages increase as census tracts have tree species with higher wood density, particularly when $X_{i, \text{density}} > 650 \text{ kg m}^{-3}$. Loblolly pine (570 kg m^{-3}) is by far the most prevalent species in our testing region and despite its lower wood density, it is more susceptible to being wind-thrown than other pines. This likely accounts for the increase in predicted outages with decreasing wood density values when $X_{i, \text{density}}$ is below 600 kg m^{-3} .

Topographic variables were relatively unimportant in the NOZE model compared to the other baseline and dynamic variables (Table 7). However, the partial dependence plot for maximum elevation (Fig. 6h) shows that at very low elevations ($< 50 \text{ m}$), there is an increase in outages followed by a sharp decrease to a minimum around 100 m . This suggests that topographic variables may be proxies for distance to the coast. At elevations $> 100 \text{ m}$, there is a gradual increase in outages, consistent with the findings of *Chapman*⁽³⁷⁾.

7. Summary and Conclusions

Our results demonstrate that the inclusion of more variables and the use of a new two-step outage prediction model increases the overall accuracy of the SGHOPM by $\sim 17\%$. Approximately, half of this improvement ($\sim 9\%$) is due to adopting the two-step outage prediction model. In the first step, the BC model makes a categorical outage occurrence/non-occurrence prediction and then the NOZE model is used to predict the number of outages. The addition of the BC model helps address the zero inflation issue in the outage data. These improvements are important because power outages due to landfalling tropical cyclones can be expensive and difficult for utility providers to handle. Providing an accurate spatial forecast of outages prior to landfall can help utility companies with resource allocation and power restoration.

One way to assess the value of the different types of variables that we considered for inclusion in the SGHOPM is to compare the improvement in model performance when that variable type is added to the baseline model. The mean improvement in model performance shows that for the BC model the most valuable variables (Table 3), in order of importance, are: tree characteristics, SPI, soil moisture, land cover, topography, and root zone depth. For the NOZE model the most valuable variables (Table 4), in order of importance, are: SPI, soil moisture, tree characteristics, land cover, topography, and root zone depth. Therefore, it is clear that the inclusion of tree characteristics and antecedent meteorological conditions (SPI and soil moisture) are valuable and significantly improve the accuracy of the SGHOPM. Land cover and topography variables are also useful, but to a lesser extent, and the inclusion of root zone depth had a minimal impact on the model.

We have demonstrated that while the three variables used in the *Guikema et al.*⁽⁸⁾ version of the SGHOPM are the most important for modeling power outages, the inclusion of information on elevation, land cover, soil, precipitation, and vegetation characteristics improves the predictive accuracy. The static variables (elevation, land cover, soil, and vegetation characteristics) provide a general and invariant assessment of power failure vulnerability for each census tract. In particular, the tree characteristics help to identify locations that are susceptible to outages due to the density and type of trees present. The dynamic variables (precipitation and soil moisture) determine how antecedent meteorological conditions either increase or decrease the vulnerability of trees being wind-thrown and leading to outages. The results suggest that wetter soils reduce soil stability, which in turn increases the probability of trees being blow over. In addition, long-term precipitation deficits associated with drought conditions also weaken the trees and therefore are associated with increased outages.

ACKNOWLEDGMENTS

This research was funded by the DOE's Energy Infrastructure Modeling and Analysis (EIMA) office (Contract No. DE-AC02-05CH11231).

REFERENCES

1. Reed DA. Electric utility distribution analysis for extreme winds. *Journal of Wind Engineering and Industrial Aerodynamics*, 2008; 96(1): 123–140.
2. 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.
3. Liu H, Davidson RA, Apanasovich TV. Spatial generalized linear mixed models of electric power outages due to hurricanes and ice storms. *Reliability Engineering*, 2008; 93: 875–890.
4. Han SR, Guikema SD, Quiring SM. Improving the predictive accuracy of hurricane power outage forecasts using generalized additive models. *Risk Analysis*, 2009; 29(10): 1443–1453.
5. Han SR, Guikema SD, Quiring SM, Lee K, Rosowsky D, Davidson RA. Estimating the spatial distribution of power outages during hurricanes in the Gulf Coast region. *Reliability Engineering and System Safety*, 2009; 94(2): 199–210.
6. Nateghi R, Guikema SD, Quiring, SM. Power outage estimation for tropical cyclones: Improved accuracy with simpler models. *Risk Analysis*, 2014; 34(6): 1069–1078.
7. Breiman L, Random forests. *Machine Learning*, 2001; 45(1): 5–32.
8. Guikema SD, Nateghi R, Quiring SM, Reilly A, Gao M. Predicting hurricane power outages to support storm response planning. *IEEE Access*, 2014; 2: 1364–1373.
9. Guikema SD, Quiring, SM. Hybrid Data Mining-Regression for Infrastructure Risk Assessment Based on Zero-Inflated Data. *Reliability Engineering & System Safety*, 2012; 99: 178–182.

10. Reilly AC, Guikema SD. Bayesian multiscale modeling of spatial infrastructure performance predictions. *Journal of Infrastructure Systems*, 2014; 21(2): 04014036.
11. Guikema SD, Nateghi R, Quiring SM. Predicting infrastructure loss of service from natural hazards with statistical models: Experiences and advances with hurricane power outage prediction. *Proceedings of the 22nd European Safety and Reliability Conference*, Amsterdam, 2013.
12. Guikema SD, Quiring SM, Han SR. Prestorm Estimation of Hurricane Damage to Electric Power Distribution Systems. *Risk Analysis*, 2010; 30(12): 1744-1752.
13. Quiring SM, Zhu L, Guikema SD. Importance of soil and elevation characteristics for modeling hurricane-induced power outages. *Natural Hazard*, 2011; 58: 365–390.
14. Danielson JJ, Gesch, DB. Global multi-resolution terrain elevation data 2010 (GMTED2010). No. 2011-1073. US Geological Survey, 2011; 26 pp.
15. Davidson RA, Haibin L, Sarpong IK, Sparks P, Rosowsky DV. Electric power distribution system performance in Carolina hurricanes. *Natural Hazards Review*, 2003; 8: 36-45.
16. Homer C, et al. Completion of the 2011 National Land Cover Database for the Conterminous United States – Representing a decade of land cover change information. *Photogrammetric Engineering and Remote Sensing*, 2015; 81(5): 345-354.
17. Anderson, JR. A land use and land cover classification system for use with remote sensor data. US Government Printing Office, 1976; 964: 28 pp.
18. Guikema SD, Davidson RA, Liu H. Statistical models of the effects of tree trimming on power system outages. *IEEE Transactions on Power Delivery*, 2006; 21(3):1549–1557.
19. Maderia CM. Importance of tree species and precipitation for modeling hurricane-induced power outages, Master's Thesis, Department of Geography, Texas A&M University, August 2015.
20. Krist Jr. FJ, et al. 2013–2027 National Insect and Disease Forest Risk Assessment. US Forest Service, 2012; 199 pp.

21. Van Dersal WR, Mulford FL, Thornthwaite CW. Native woody plants of the United States: their erosion-control and wildlife values. USDA Miscellaneous Publication 303, 1938; 362 pp.
22. Burns RM, Honkala BH. Silvics of North America: 1. Conifers; 2. Hardwoods. US Forest Service Agriculture Handbook 654, 1990; 877 pp.
23. Stoecklein, Marc C. The complete plant selection guide for landscape design. Purdue University Press, 2001; 747 pp.
24. US Forest Service Tree List, 2015; <http://www.fs.fed.us/database/feis/plants/tree/>.
25. Wood Database, 2015; <http://www.wood-database.com>.
26. Soil Survey Staff. Gridded Soil Survey Geographic (gSSURGO) Database for the United States of America and the Territories, Commonwealths, and Island Nations served by the USDA-NRCS. United States Department of Agriculture, Natural Resources Conservation Service, 2015; <https://gdg.sc.egov.usda.gov>.
27. Dobos R, Sinclair H, Hipple, K. User Guide National Commodity Crop Productivity Index (NCCPI) Version 1.0. US Department of Agriculture, Natural Resources Conservation Service, 2012.
28. Liang X, Wood EF, Lettenmaier DP. Surface soil moisture parameterization of the VIC-2L model: Evaluation and modification. *Global and Planetary Change*, 1996; 13(1): 195-206.
29. Hosking JRM, Wallis JR. Regional frequency analysis: an approach based on L-moments. Cambridge University Press, 2005; 244 pp.
30. Andreadis KM, Clark EA, Wood AW, Hamlet AF, Lettenmaier DP. Twentieth-century drought in the conterminous United States. *Journal of Hydrometeorology*, 2005; 6(6): 985-1001.
31. Mo KC, Long LN, Xia Y, Yang SK, Schemm, JE, Ek M. Drought indices based on the Climate Forecast System Reanalysis and ensemble NLDAS. *Journal of Hydrometeorology*, 2011; 12(2): 181-205.

32. McKee TB, Doesken NJ, Kleist J. The relationship of drought frequency and duration to time scales. Proceedings of the 8th Conference on Applied Climatology, Anaheim, CA, American Meteorological Society, 1993; 179-184.
33. Guttman NB. Comparing the Palmer drought index and the standardized precipitation index. Journal of the American Water Resources Association, 1998; 34(1): 113-121.
34. Heim Jr. RR. A review of twentieth-century drought indices used in the United States. Bulletin of the American Meteorological Society, 2002; 83(8): 1149-1165.
35. Hastie T, JRM, Tibshirani R, Friedman J. The Elements of Statistical Learning: Data Mining, Inference, and Prediction. Springer Science & Business Media, 2013; 536 pp.
36. Liaw A, Wiener M. Classification and Regression by RandomForest. R News, 2002; 2(3): 18-22.
37. Chapman L. Assessing topographic exposure. Meteorological Applications, 2000; 7(4): 335-340.

TABLES

Table 1. Summary of the variables used in the SGHOPM. Variables shown in bold-italics are used in the optimal BC model covariate set (26 total) and the variables that are underlined are used in the optimal NOZE model covariate set (25 total).

Variable	Response	Baseline	Static				Dynamic	
			Elevation Census Tract Statistics	Land Cover (%)	Root Zone Depth Census Tract Statistics	Tree Characteristics	Soil Moisture (CDF)	Precipitation (SPI)
1	Outage Fraction	<u>Total Population</u>	<i>Mean</i>	<i>Water</i>	<u>Mean</u>	<i>Pct. treed</i>	<u>Layer 1 (0 -</u>	<u>1- month</u>

							10 cm)	
2		<i><u>Max Wind Gust</u></i>	<i><u>Median</u></i>	<i><u>Developed</u></i>	<i><u>Majority</u></i>	<i><u>Pct. Deep-Rooted</u></i>	Layer 2 (10 – 40 cm)	3– month
3		<i><u>Strong Winds Duration</u></i>	<i><u>Standard Deviation</u></i>	<i><u>Barren</u></i>	Median	<i><u>Pct. Taproot</u></i>	Layer 3 (40 – 100 cm)	6– month
4			<i><u>Minimum</u></i>	<i><u>Scrub</u></i>		<i><u>Max Tree Height</u></i>		12– month
5			<i><u>Maximum</u></i>	<i><u>Forest</u></i>		<i><u>Maximum DBH</u></i>		24– month
6				<i><u>Grassland</u></i>		<i><u>Avg. Wood density</u></i>		
7				<i><u>Pasture</u></i>		<i><u>Avg. Janka Hardness</u></i>		
8				<i><u>Wetlands</u></i>		<i><u>Avg. crushing strength</u></i>		

Table 2. Eight tree characteristics from the National Insect and Disease Risk Map (NIDRM) dataset (light gray background). The NIDRM data were re-scaled to census tracts by aggregating all NIDRM grid cells within each census tract (dark gray background). Information about each individual tree species was compiled from the following sources: (1) *Van Dersal 1938*⁽²⁰⁾, (2) *Burns and Honkala 1990a*⁽²¹⁾, (3) *Burns and Honkala 1990b*⁽²²⁾, (4) *Stoecklein 2001*⁽²³⁾, (5) *US Forest Service Tree List 2015*⁽²⁴⁾, and (6) *Wood Database 2015*⁽²⁵⁾.

Variable		NIDRM level (240 m resolution)		Census Tract Level		
		Metric	Data Source	Model Variable	Model Variable Description	
1	Trees	Trees or no trees?	1,2,3,4,5	$X_{treed, i}$	Percentage covered by trees	
2	Taproot	Taproot or not?		$X_{taproot, i}$	Percentage with taprooted species	
3	Rooting depth	Deep or shallow rooted species?		$X_{deep, i}$	Percentage with deep-rooted species	
4	Tree Height	Species Maximum	6	$X_{hgt, i}$	Census tract average	
5	Diameter at Breast Height			$X_{dbh, i}$		
6	Wood Density	$X_{density, i}$				
7	Janka Hardness	$X_{janka, i}$				
8	Crushing Strength	$X_{crushing, i}$				
		Species Average				

Table 3. Performance metrics for the 14 different covariate sets that were evaluated to determine the optimal set of explanatory variables in the BC model. Improvement the reduction in MAE relative to the *Guikema et al.*⁽⁸⁾ NOZE-only model and is based on 30 replicates. Decrease in root mean square error is in parentheses next to model improvement.

BC Covariate Sets	Rank	Frequency Most Improved	Mean Improvement	Standard Error
No BC Model	14	0	0.00% (0.00%)	0.00%
Baseline	13	0	3.74% (-1.69%)	0.70%
Elevation	11	0	5.84% (-0.80%)	1.16%
Land Cover	10	0	6.62% (-0.29%)	1.33%
Root Zone	12	0	3.88% (-1.14%)	0.67%
Soil Moisture	9	0	6.66% (-0.84%)	0.92%
SPI	5	2	7.73% (-0.08%)	1.20%
Trees	6	0	7.60% (0.26%)	1.20%
<i>Static</i>	<i>1</i>	<i>18</i>	<i>9.04%</i> (0.44%)	1.36%
Static Reduced	2	2	8.47% (0.11%)	1.51%
Dynamic	4	2	7.88% (-0.14%)	1.08%
Dynamic Reduced	7	1	7.56% (-0.29%)	0.98%
Full	3	4	8.24% (0.58%)	1.03%
Full Reduced	8	1	7.53% (0.03%)	1.12%

Table 4. Performance metrics for the 15 different covariate sets that were evaluated to determine the optimal set of explanatory variables in the NOZE model. Improvement is the reduction in MAE relative to the *Guikema et al.*⁽⁸⁾ NOZE-only model and is based on 30 replicates. The BC covariate set uses variables that are in bold-italics in Table 1 unless otherwise noted. Decrease in root mean square error is in parentheses next to model improvement.

NOZE Covariate Sets	Rank	Frequency Most Improved	Mean Improvement	Standard Error
Baseline (No BC Model)	14	0	0.00% (0.00%)	0.00%
Baseline	13	0	8.84% (0.37%)	1.23%
Elevation	11	0	11.35% (3.51%)	1.57%
Land Cover	10	0	11.80% (3.96%)	1.73%
Root Zone	12	0	9.90% (1.73%)	1.39%
Soil Moisture	8	0	12.78% (3.25%)	1.78%
SPI	5	2	15.74% (4.17%)	1.71%
Trees	9	0	12.08% (4.53%)	1.79%
Static	7	0	13.66% (5.47%)	1.72%
Static Reduced	6	0	13.99% (5.78%)	1.71%
Dynamic	3	1	15.91% (4.33%)	1.95%
Dynamic Reduced	4	2	15.85% (4.55%)	1.96%
<i>Full</i>	<i>1</i>	<i>15</i>	<i>16.75%</i> (6.88%)	1.72%
Full Reduced	2	10	16.62% (6.77%)	1.78%

Table 5. Performance metrics for the 16 different NOZE model covariate sets based on the storm-specific holdout testing. Improvement is the reduction in MAE relative to the *Guikema et al.*⁽⁸⁾ NOZE-only model. The BC covariate set uses variables that are in bold-italics in Table 1 unless otherwise noted.

NOZE Covariate Sets	Model Improvement	Ratio of Variance to Observed Variance (R_{var})

	Rank	Value	Rank	Value
Baseline (No BC Model)	9	0.00%	10	0.36
Baseline (BC Model w/ only Baseline)	12	-1.32%	7	0.40
Baseline	7	8.36%	8	0.39
Elevation	1	11.76%	4	0.42
Land Cover	3	10.15%	5	0.41
Root Zone Depth	2	11.42%	9	0.38
Soil Moisture	14	-3.84%	6	0.40
SPI	13	-2.48%	15	0.27
Trees	6	9.72%	3	0.48
Static	5	9.95%	1	0.51
Static Reduced	4	10.03%	2	0.49
Dynamic	8	0.68%	14	0.28
Dynamic Reduced	15	-4.79%	11	0.33
Full	11	-0.87%	13	0.32
Full Reduced	10	-0.84%	12	0.33

Table 6. Variable importance in the BC model, originally measured as the decrease in the Gini index (g).

Rank	Type	Variable	VI
1	Baseline	Max Wind Speed	100.00
2	Tree	Average Wood Density	89.63
3	Baseline	Strong Winds Duration	87.19
4	Elevation	Mean Elevation	85.77
5	Elevation	Max Elevation	76.54
6	Elevation	Median Elevation	71.91
7	Tree	Average Crushing Strength	59.14
8	Elevation	Min Elevation	57.99
9	Land Cover	Wetlands Land Cover	55.14
10	Tree	Percentage Taproot	51.66
11	Baseline	Population Density	47.85
12	Tree	Percentage Deep	44.85
13	Elevation	Elevation Stdev	42.04
14	Land Cover	Developed Land Cover	41.83
15	Tree	Average Max Tree Height	39.09
16	Land Cover	Scrub Land Cover	37.78

17	Tree	Average Janka Hardness	36.99
18	Land Cover	Barren Land Cover	31.23
19	Land Cover	Pasture Land Cover	30.28
20	Land Cover	Forest Land Cover	29.15
21	Tree	Average Maximum DBH	28.26
22	Tree	Grassland Land Cover	26.84
23	Root Zone Depth	Root Zone Mean Depth	26.70
24	Land Cover	Water Land Cover	25.98
25	Tree	Percentage Treed	24.61
26	Root Zone Depth	Root Zone Majority Depth	1.10

Table 7. Variable importance in the NOZE model, originally measured as the residual sum of squares.

Rank	Type	Variable	VI
1	Baseline	Max Wind Speed	100.00
2	Baseline	Strong Winds Duration	87.45
3	SPI	SPI12	70.16
4	Baseline	Population Density	41.05
5	Soil Moisture	Soil CDF 2	38.54
6	Soil Moisture	Soil CDF 1	38.42
7	SPI	SPI3	37.75
8	SPI	SPI24	35.69
9	SPI	SPI6	33.33
10	Tree	Average Wood Density	33.01
11	Soil Moisture	Soil CDF 3	30.38
12	SPI	SPI1	30.32
13	Land Cover	Wetlands Land Cover	28.69
14	Elevation	Max Elevation	27.63
15	Tree	Percentage Deep	26.83
16	Tree	Percentage Taproot	26.83
17	Root Zone Depth	Root Zone Mean Depth	26.58
18	Tree	Average Janka Hardness	26.20
19	Tree	Average Max Tree Height	25.73
20	Land Cover	Forest Land Cover	24.52
21	Tree	Percentage Treed	23.64
22	Land Cover	Grassland Land Cover	22.89

23	Elevation	Median Elevation	22.19
24	Tree	Average Crushing Strength	22.14
25	Tree	Average Maximum DBH	6.76

FIGURES

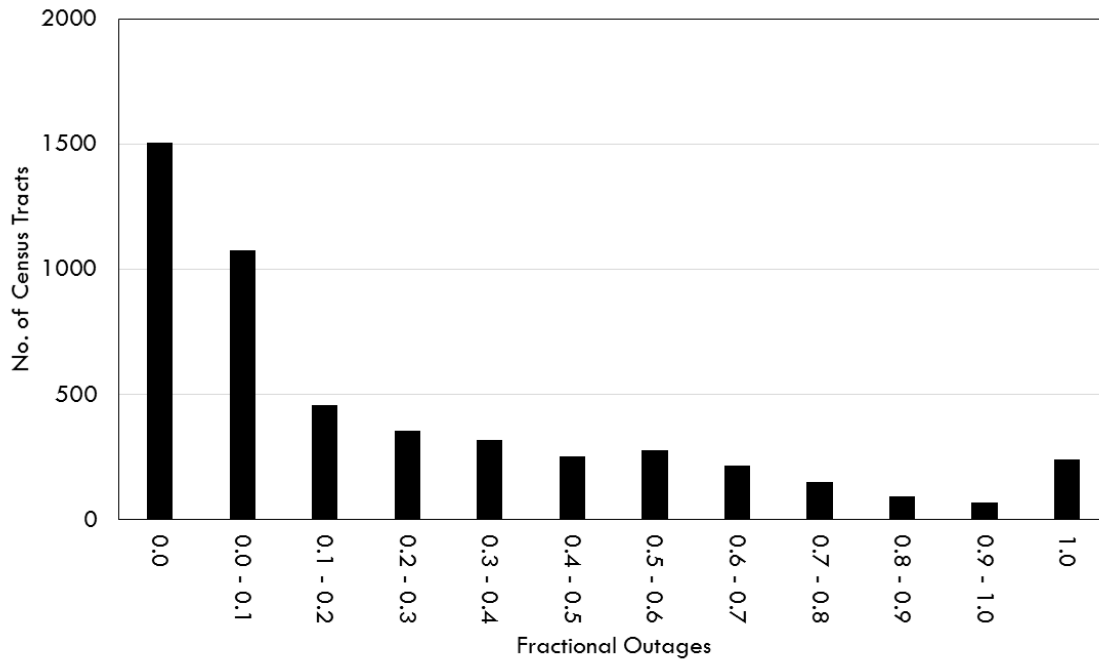
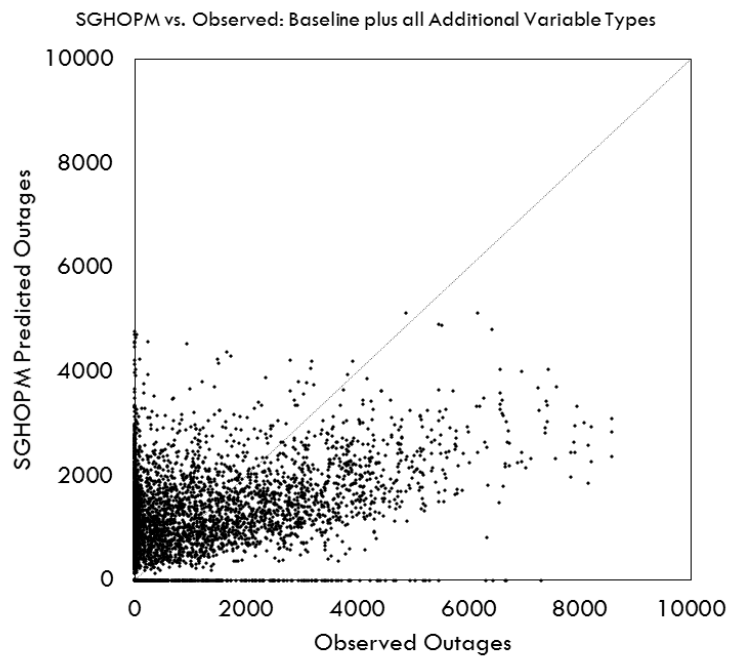
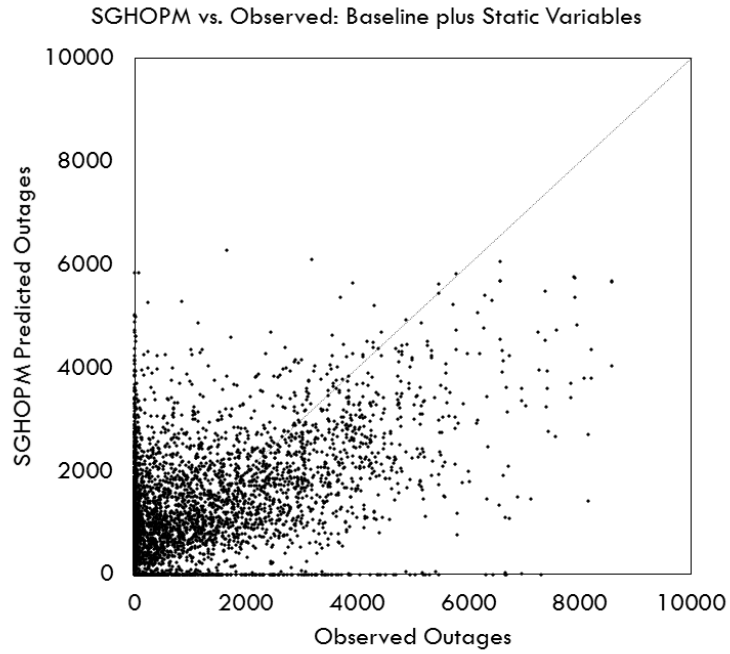


Fig. 1. Histogram showing of observed fractional outages in the utility company dataset.

Author



(a)



(b)

Fig. 2. Scatter plots of total predicted versus observed outages based on the two-step SGHOPM:

(a) NOZE model using baseline + all additional variable types, and (b) NOZE model using baseline + all static variables. The perfect prediction line (1:1) is plotted for reference.

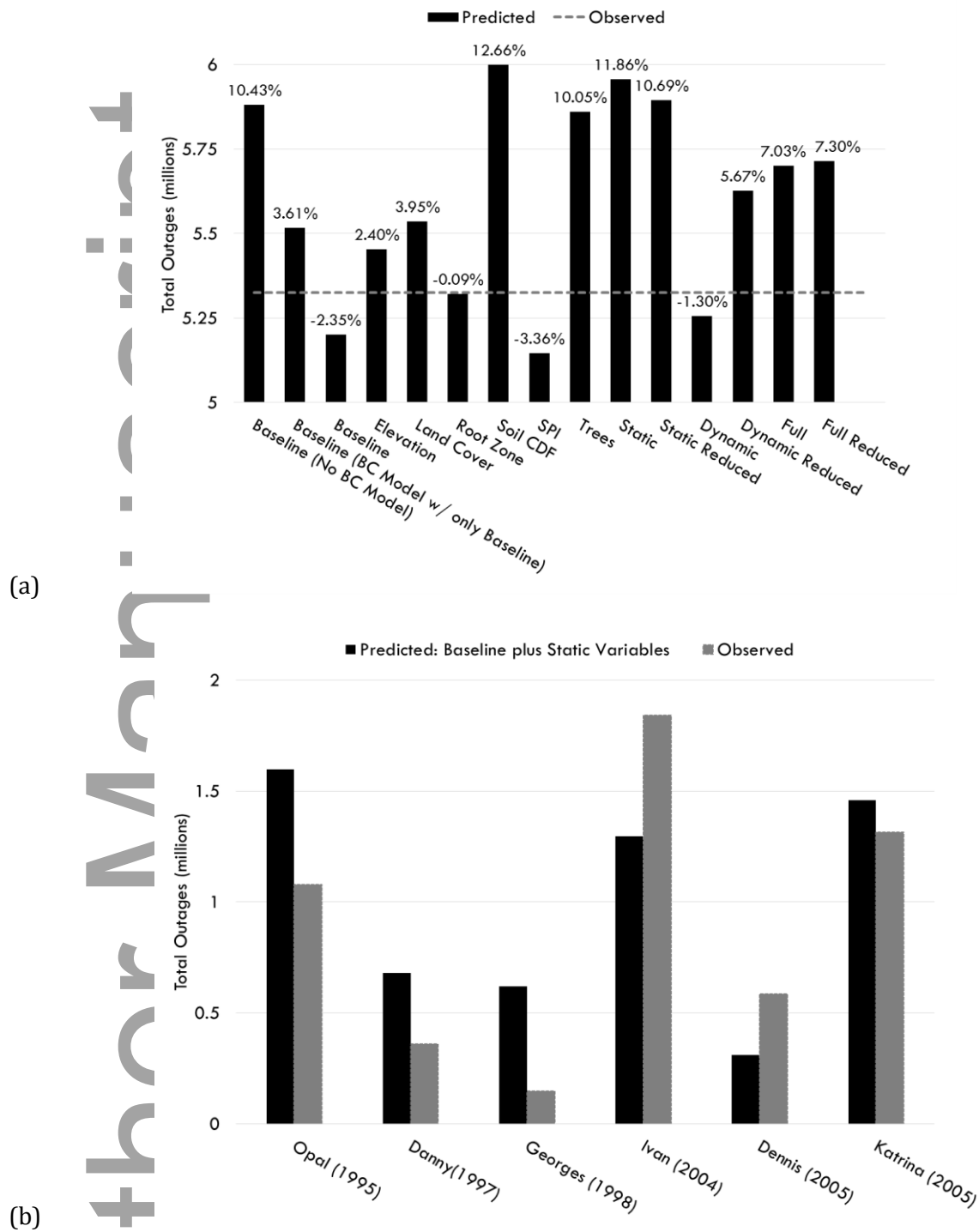


Fig. 3. (a) Predicted number of outages (summed for all census tracts) for different covariate sets in the NOZE model. (b) Predicted number of outages using the NOZE model with the covariate set including baseline plus static variables for each storm-specific holdout.

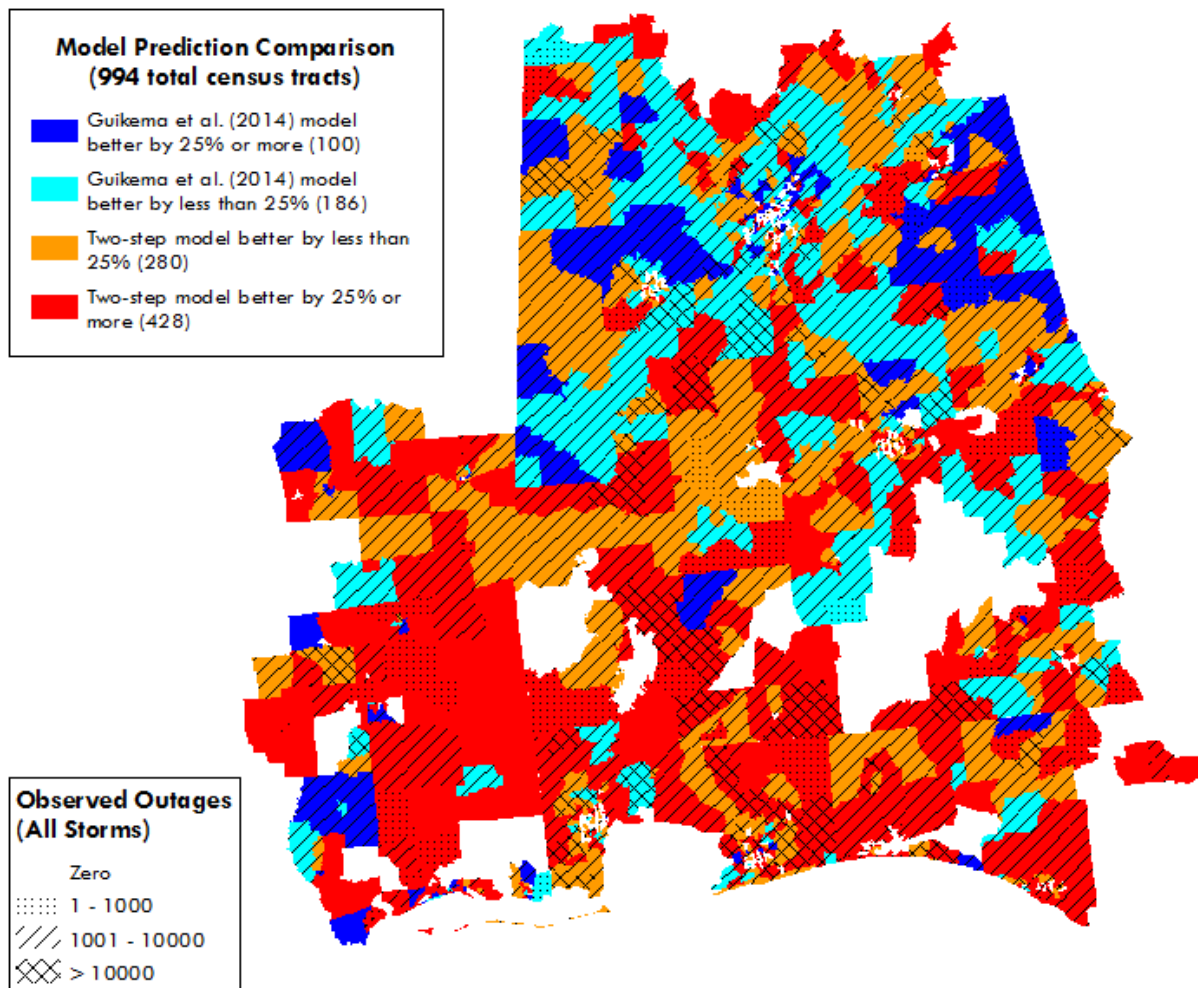
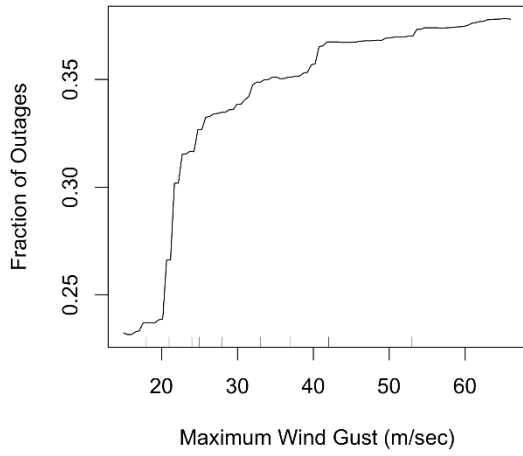


Fig. 4. Comparison of model performance for the Guikema et al. (2014) and two-step model versions of the SGHOPM (colors) for all six storms in our study. The comparison metric is the percentage decrease in MAE of the “better” model relative to the other. The total number of observed outages for all six storms is denoted by the given patterns.

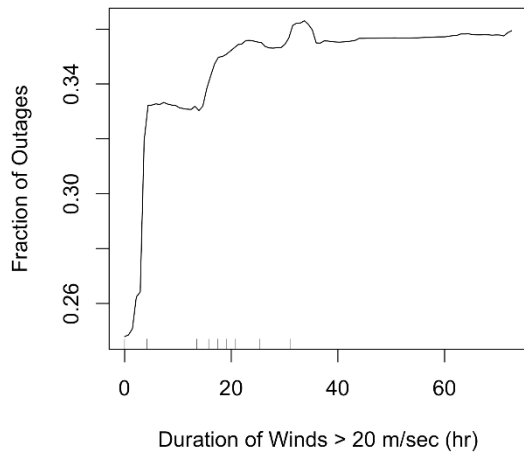
Maximum Wind Gust



(a)

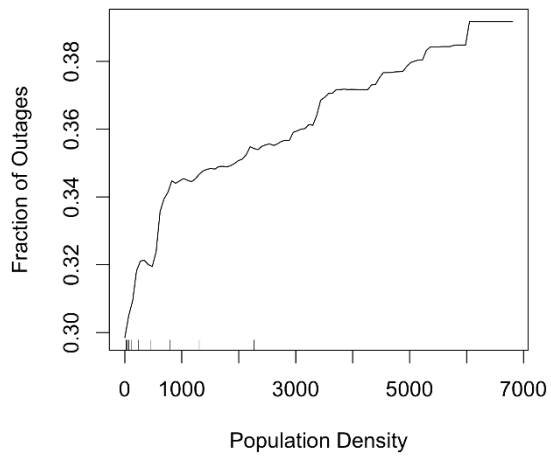
(b)

Duration of Strong Winds



Autho

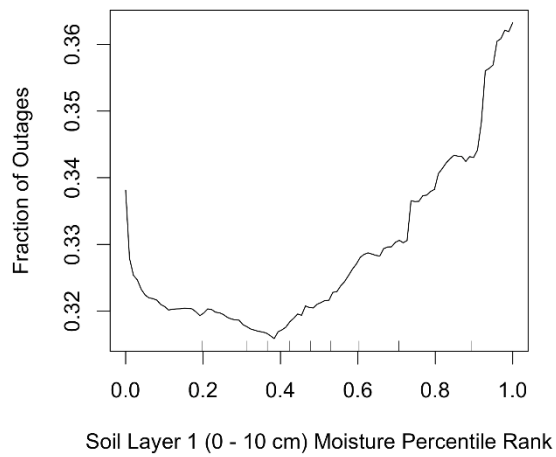
Population Density



(c)

(d)

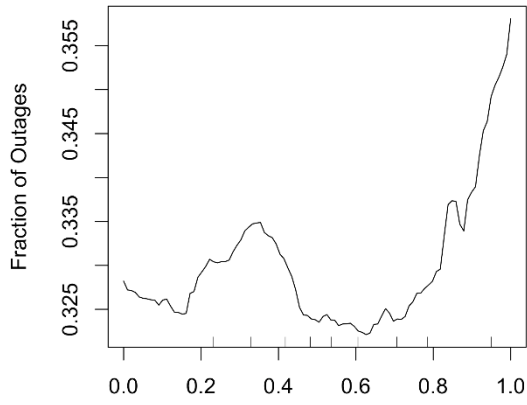
Soil Layer 1 Moisture Percentile Rank



Soil Layer 1 (0 - 10 cm) Moisture Percentile Rank

Autho

Soil Layer 2 Moisture Percentile Rank

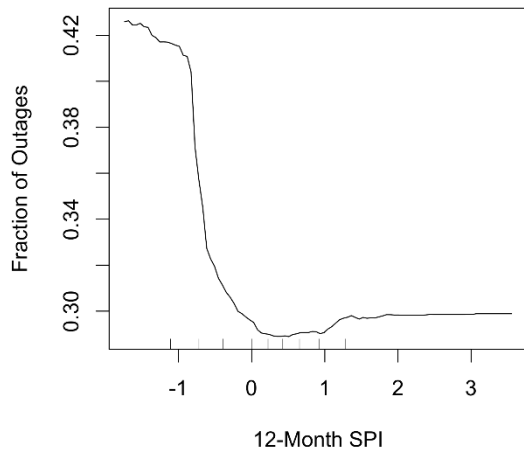


(e)



(f)

12-Month Standardized Precipitation Index



Autho

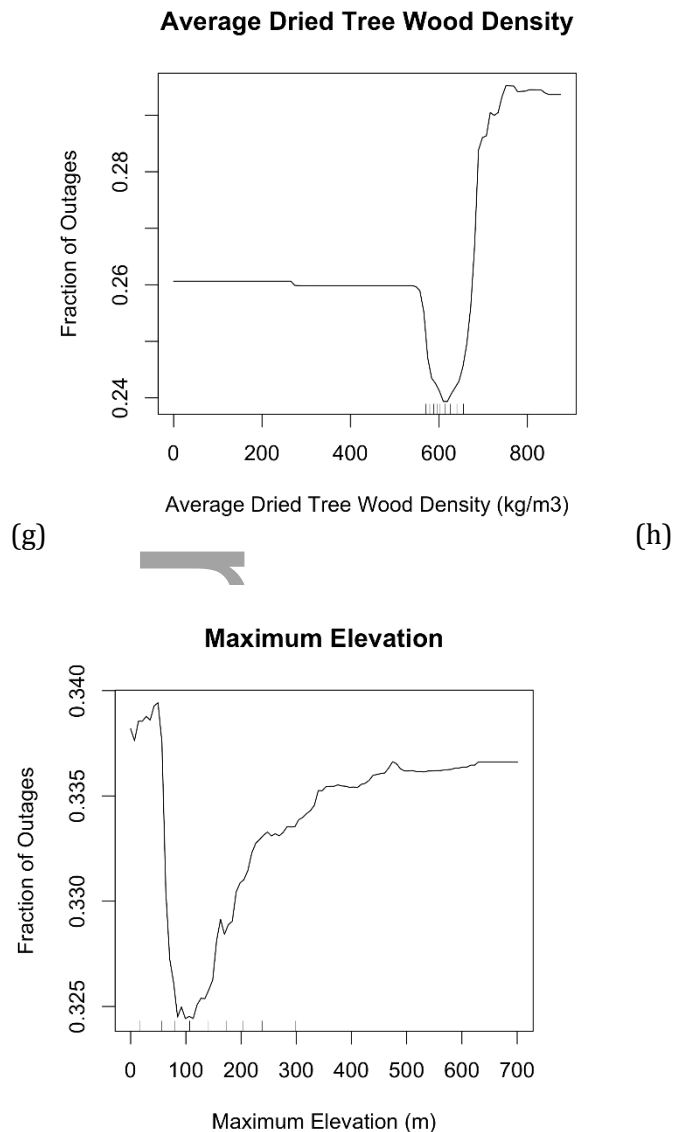


Fig. 5. Partial dependent plots for variables identified as being important in the holdout evaluations: (a) maximum wind gust (m s^{-1}), (b) duration of strong winds (hours), (c) population density (people km^{-2}), (d) soil moisture percentile in layer 1 (0 to 10 cm), (e) soil moisture percentile in layer 2 (10 to 40 cm), (f) 12-month Standardized Precipitation Index (SPI), (g) Average dried wood density (kg m^{-3}), and (h) maximum elevation (m).