INCORPORATING HURRICANE FORECAST UNCERTAINTY INTO A DECISION-SUPPORT APPLICATION FOR POWER OUTAGE MODELING

BY STEVEN M. QUIRING, ANDREA B. SCHUMACHER, AND SETH D. GUIKEMA

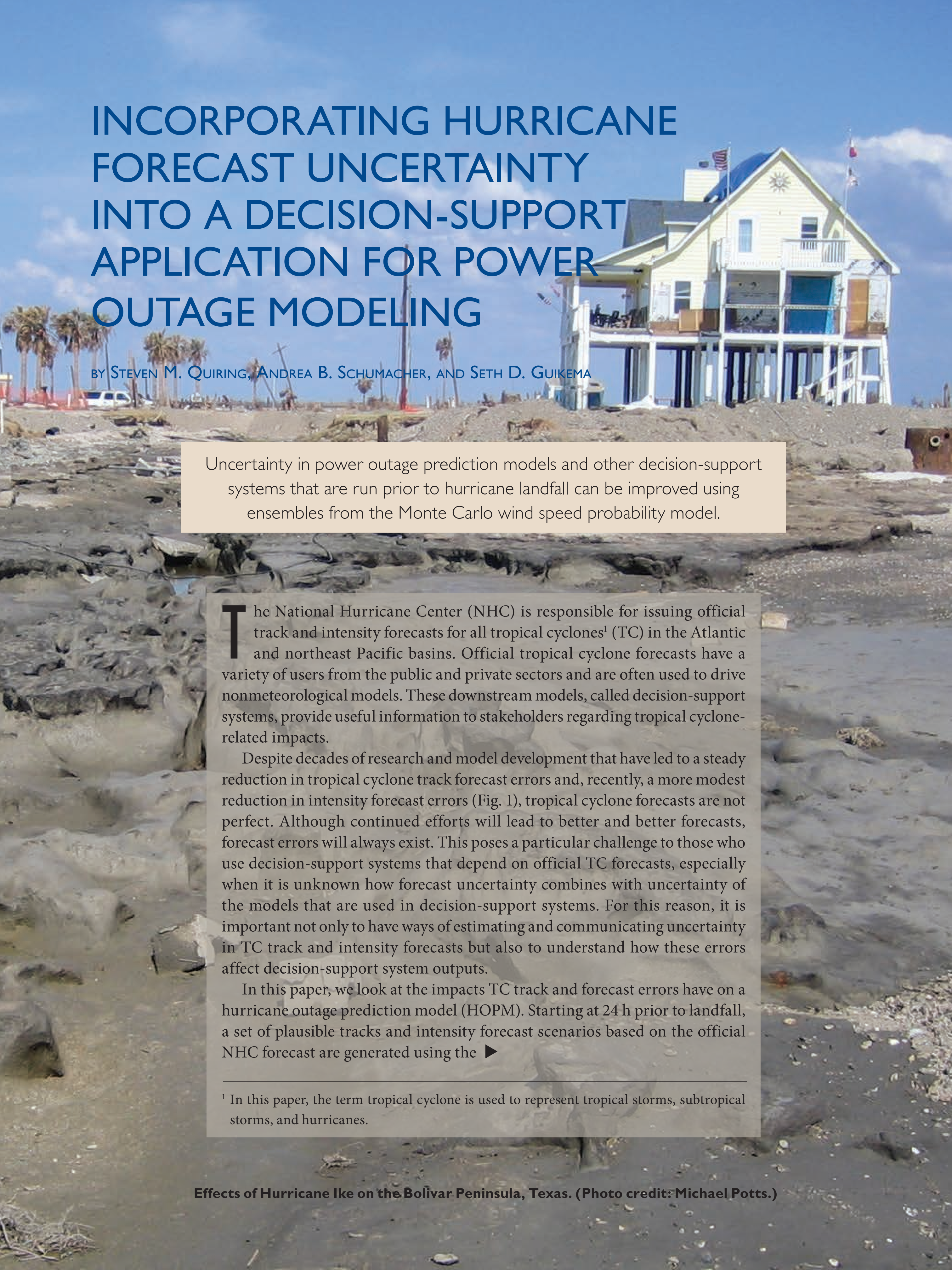
Uncertainty in power outage prediction models and other decision-support systems that are run prior to hurricane landfall can be improved using ensembles from the Monte Carlo wind speed probability model.

The National Hurricane Center (NHC) is responsible for issuing official track and intensity forecasts for all tropical cyclones[1] (TC) in the Atlantic and northeast Pacific basins. Official tropical cyclone forecasts have a variety of users from the public and private sectors and are often used to drive nonmeteorological models. These downstream models, called decision-support systems, provide useful information to stakeholders regarding tropical cyclone-related impacts.

Despite decades of research and model development that have led to a steady reduction in tropical cyclone track forecast errors and, recently, a more modest reduction in intensity forecast errors (Fig. 1), tropical cyclone forecasts are not perfect. Although continued efforts will lead to better and better forecasts, forecast errors will always exist. This poses a particular challenge to those who use decision-support systems that depend on official TC forecasts, especially when it is unknown how forecast uncertainty combines with uncertainty of the models that are used in decision-support systems. For this reason, it is important not only to have ways of estimating and communicating uncertainty in TC track and intensity forecasts but also to understand how these errors affect decision-support system outputs.

In this paper, we look at the impacts TC track and forecast errors have on a hurricane outage prediction model (HOPM). Starting at 24 h prior to landfall, a set of plausible tracks and intensity forecast scenarios based on the official NHC forecast are generated using the ▶

[1] In this paper, the term tropical cyclone is used to represent tropical storms, subtropical storms, and hurricanes.



Effects of Hurricane Ike on the Bolivar Peninsula, Texas. (Photo credit: Michael Potts.)

Monte Carlo wind speed probability (MCWSP) algorithm (DeMaria et al. 2009). Each of these TC forecast scenarios is then input into the HOPM, and the resulting outage predictions are analyzed.

OVERVIEW OF HOPM. Electric power utilities make many critical decisions in the days prior to hurricane landfall that are primarily based on the estimated impact to their service area. For example, utilities must determine how many repair crews to request from other utilities, the amount of material and equipment they will need to make repairs, and where in their geographically expansive service area to station crews and materials (DeGaetano et al. 2008). Obtaining extra crews and materials is expensive, with preparedness costs often in the tens of millions of dollars for a single storm for a single utility (Zhu et al. 2007; Cerruti and Decker 2012). At the same time, having too few resources available can significantly delay power restoration. Accurate forecasts of the impact of an approaching hurricane within their service area are critical for utilities in balancing the costs and benefits of different levels of resources.

There have been a number of methods developed for estimating power outages caused by weather events such as hurricanes. Some of these studies have focused on modeling how weather events affect power system reliability. Brown et al. (1997) developed methods to assess the reliability of the power distribution system to momentary interruptions and storms. They modeled potential storms events (defined as a period of high winds) using Monte Carlo simulation based on 20 years of historical wind data from Snohomish County, Washington. Brown et al. (1997) used both wind speed and duration of strong winds to determine power system reliability (i.e., average number of interruptions per year). Balijepalli et al. (2005) also used Monte Carlo simulation to evaluate the impact of lightning on power system reliability. They used data from 177 storms in Iowa and a bootstrapping method to calculate the average annual lightning fault rates. Zhou et al. (2006) used a Poisson regression model and a Bayesian network model to predict annual weather-related failure rates based on 6 years of data from Manhattan, Kansas. Although both models performed similarly, they identified the Bayesian approach as preferable because it is more informative, easier to implement, and can be updated more easily than the Poisson regression model. They considered 10 different weather conditions but determined that wind, ice, and lightning events had the largest influence on the power system. Like the previous studies, Zhou et al. (2006) used Monte Carlo simulation to quantify uncertainty in the predictions and determine confidence bands. These three studies focused on modeling how weather events influence power system reliability on an annual basis; they were not focused on forecasting outages for specific storms. In contrast, Reed (2008), focused on particular windstorms impacting the power system in the Seattle area, deriving empirical relationships between gust wind speeds and the likelihood of damage and duration of outages. Reed et al. (2010a) examined damage to energy infrastructure from Hurricane Rita (2005). While the focus was on understanding past damage, not predicting future damage, Reed et al. (2010a) did conclude that high gust wind speeds alone can lead directly to damage. In contrast, Reed et al. (2010b) investigated energy system performance during Hurricane Katrina (2005) and concluded that substantial damage can occur at lower wind speeds than wind-based models would predict because of the effects of storm surge and inland rainfall. Reed et al. (2010a,b) did not consider either predictive modeling or the uncertainty in hurricane forecasts, but their conclusions do indicate the complexity involved in predicting the impacts of hurricanes on power systems.

There has also been prior work on developing models to forecast power outages or damage resulting from weather events such as thunderstorms and ice storms. Zhu et al. (2007) developed models to predict outages for approaching storms. They trained their models using 10 years of historical data and identified six different types of storms. Their models are based on temperature, wind speed, and lightning. Zhu et al. (2007) found that lightning is the most important cause of outages (>50%).

AFFILIATIONS: Quiring—Department of Geography, Texas A&M University, College Station, Texas; Schumacher—Cooperative Institute for Research in the Atmosphere, Colorado State University, Fort Collins, Colorado; Guikema—Department of Geography and Environmental Engineering, The Johns Hopkins University, Baltimore, Maryland
CORRESPONDING AUTHOR: Steven M. Quiring, Department of Geography, Texas A&M University, College Station, TX 77843-3147
E-mail: squiring@tamu.edu

The abstract for this article can be found in this issue, following the table of contents.
DOI:10.1175/BAMS-D-12-00012.1

In final form 24 May 2013

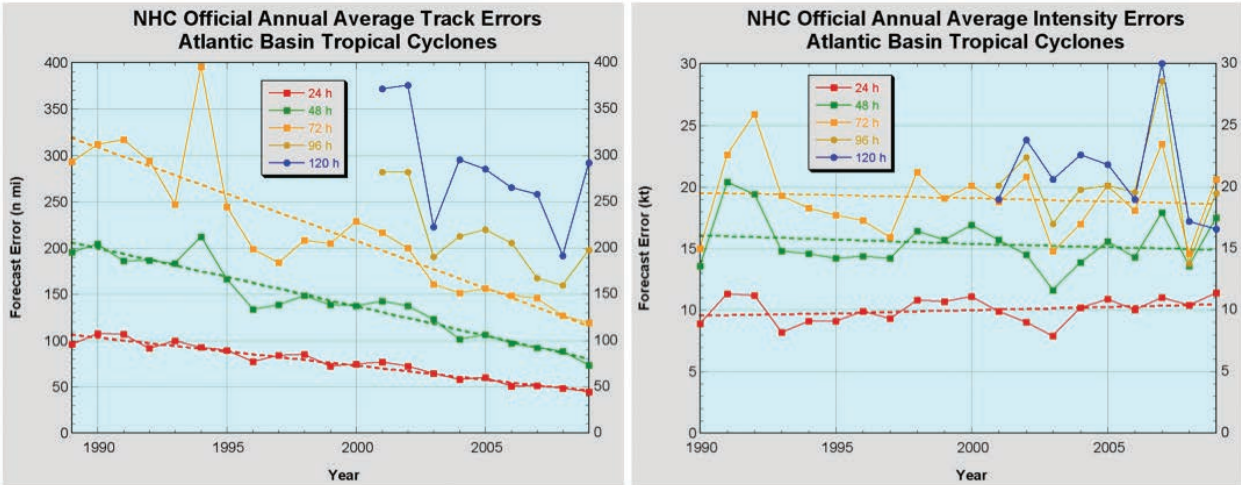


Fig. 1. NHC Verification Report Figs. 5.3 and 5.4 (www.nhc.noaa.gov/verification/). Annual average official (left) track (in n mi) and (right) intensity (kt) errors for Atlantic basin tropical cyclones for the period 1989–2009, with least-squares trend lines superimposed (National Hurricane Center 2010).

DeGaetano et al. (2008) developed an approach for forecasting ice accretion on electric distribution lines using the Weather Research and Forecasting model and an ice accretion model. The forecasts of ice accretion, made 6–12 h in advance, are then used to predict damage to the power system. DeGaetano et al. (2008) applied these models in two regions: one centered on the border of North Carolina and Virginia (23 ice storm events) and the second over central New York and northern Pennsylvania (15 ice storm events). Their approach provided relatively accurate short-term forecasts of ice accretion that are valuable for electric utilities.

Li et al. (2010) developed a Poisson regression model in a Bayesian hierarchical framework for predicting power outages with up to a 3-day lead time based on forecasts of severe weather events (e.g., hurricanes, tornados, and thunderstorms). The weather forecasts are based on the IBM Deep Thunder weather modeling system that provides forecasts at 1–2-km resolution. Their power outage model utilizes forecasts of wind gusts, gust frequency, rainfall, and temperature as well as rainfall in the two weeks preceding the storm event. They developed the power outage model based on data from a utility in southeastern New York for 19 storm events (2004–09). Unlike many of the other outage models, they explicitly incorporate uncertainties in the weather forecast and uncertainties in the damage model and provide estimates of confidence for their predictions of power outages.

Cerruti and Decker (2012) developed an infrastructure damage model for predicting weather-related damage to electric infrastructure using generalized linear models. They developed their models using damage data (2003–08) from the Public Service Electric and Gas Company service area in New Jersey. The models predicted weather-related damage using variables such as air temperature, dew-point temperature, precipitation, and maximum wind gust. They found that model accuracy was improved when days were stratified into six weather types (e.g., thunderstorms, heat, wind).

Some studies have focused specifically on modeling power outages or damage resulting from hurricanes. Huang et al. (2001a) used event-based Monte Carlo hurricane simulation techniques to determine the 50-yr recurrence interval for hurricane winds in the southeastern United States. They used 112 years of historical hurricane data to characterize the central pressure, radius of maximum winds, approach angle, translation velocity, and annual occurrence rate of hurricanes that had influenced this region. Hurricane winds were simulated with the gradient wind model developed by Georgiou (1985), and the wind field model was combined with a damage model, constructed using actual insurance loss data, to evaluate long-term risk (i.e., expected annual loss ratio). They found that the expected annual damage ratio drops very quickly with distance from the coast. They also concluded that the design load standard [American Society of Civil Engineers (ASCE) standard 7-95], may overestimate the design wind speeds in coastal areas. Huang et al. (2001b) applied the approach of Huang et al. (2001a) to estimate damage to residential structures and to estimate annualized losses due to hurricanes in North Carolina, South Carolina, and Florida. They found that estimated annualized loses in Florida are much greater than in North Carolina or South Carolina. Davidson et al. (2003) analyzed power outages from five hurricanes that affected Duke Power and Carolina Power and Light (now Progress Energy). They found a statistically significant relationship between the maximum wind gust and the number of outages. They concluded that, while wind speed is an important predictor, it is not sufficient for explaining the pattern of outages and that other explanatory variables are necessary (Davidson et al. 2003). Liu et al. (2005) extended the work of Davidson et al. (2003) by developing Poisson and negative binomial generalized linear models (GLMs) for the same study region. They found that the most important variables for explaining variations in power outages were maximum wind gust, the number of transformers, the power company affected, and the hurricane indicator (Liu et al. 2005).

Liu et al. (2007) developed accelerated failure time (AFT) models for power restoration times from outages caused by hurricanes and ice storms. Many of the same variables (e.g., maximum wind gust, hurricane indicator) found to be important in previous studies were also statistically significant for predicting restoration times. One of the primary limitations identified by Liu et al. (2007) was the lack of potentially important tree-related explanatory variables (e.g., number, type, age of trees, and tree trimming frequency) and infrastructure variables (e.g., age and condition of the poles). Liu et al. (2008) developed spatial generalized linear mixed models (GLMMs) using the same study area and hurricanes as Liu et al. (2007) but at a higher spatial resolution. They found that the results from the spatial GLMMs are not significantly better than the negative binomial GLM or simpler (nonspatial) GLMM. Consistent with their prior work, the best models considered maximum wind gust, the number of protective devices in each grid cell, and the hurricane and company indicator variables. Winkler et al. (2010) examined the effect of network topology on power system reliability during hurricanes. They estimated the probability of damage of each utility pole based solely on gust wind speed and then estimated power outages based on the topology of the power system. Their predictions power outages for Hurricane Ike had an error of ~16% for the overall service area.

A conceptually different approach from that of Winkler et al. (2010) is the HOPM of Han et al. (2009a,b), Guikema and Quiring (2012), and Nateghi et al. (2011). The HOPM are a family of statistical models that utilize predictions of tropical cyclone wind speed and duration of strong winds, along with power system and environmental variables (e.g., soil moisture, long-term precipitation), to forecast the number and location of power outages and, in the case of Nateghi et al. (2011), the duration of outages prior to hurricane landfall. An outage is defined as a nontransitory activation of a protective device (e.g., fuses, circuit breakers, automatic circuit reclosers). Utility companies are often most interested in physical damage to the electric power system since this directly affects restoration times and costs. However, much of the past modeling work has focused on predicting power outages because outage data are more readily available for model validation. The modeling framework and variables used in the HOPM have evolved over time based on our conversations and meetings with utility company personnel and our experience with applying the HOPM in real time.

The first generation of the HOPM utilized negative binomial GLMs to predict power outages (Han et al. 2009b). The predictions were based on maximum wind gust and duration of strong winds, the time since the last hurricane, radius of maximum winds, and central pressure deficit. Han et al. (2009b) also used soil moisture levels from three soil layers, mean annual precipitation, and antecedent precipitation conditions over time scales ranging from 1 month to 2 years (e.g., standardized precipitation index). Soil moisture and antecedent precipitation data provide information about the stability of the soil, since saturated soils increase the likelihood of trees being uprooted or poles being blown over when subjected to strong winds.

Han et al. (2009a) improved upon the accuracy of Han et al. (2009b) by utilizing generalized additive models (GAMs) for predicting hurricane-related outages. They found that GAMs more accurately predict the spatial distribution of power outages and GAMs overcome some of the overprediction problems associated with GLMs. Guikema et al. (2010) developed a set of models for predicting physical damage to the power system using both the regression-based models employed in previous studies and two data mining approaches [e.g., classification and regression trees (CART) and Bayesian regression trees (BART)]. The data mining approaches outperformed the regression-based approaches. One limitation of all of these statistical models is that they are not sufficiently flexible and robust to capture the complexities of the infrastructure data. Specifically, they have difficulty with zero inflated data (i.e., datasets that contain many lots of zeros). Therefore, Guikema and Quiring (2012) developed a two-stage model that uses CART and a Poisson GAM to predict the number of power outages. The CART is used to model whether a location will experience zero power outages or more than zero power outages and the Poisson GAM then estimates the number of power outages in all of the locations where CART predicts outages will occur. This approach has strong predictive accuracy and it provides predictions that are responsive to hurricane characteristics and local conditions (Guikema and Quiring 2012). This hybrid model is the HOPM used in this paper because it provides better out of sample predictive accuracy than the Han et al. (2009a,b) models. A detailed discussion of the modeling approach and an evaluation of model accuracy are provided in Guikema and Quiring (2012).

The occurrence of power outages during hurricanes depends on a number of factors that influence the vulnerability of electric power systems to outages including power system exposure to falling trees, windborne debris, and high winds. We use the length of distribution line, the number of poles, switches, transformers, and customers in different geographic areas as proxy measures for exposure of the power system. Other factors include spatially varying conditions such as land use/land cover, long-term precipitation patterns in the area, and soil moisture levels before a hurricane makes landfall. In our models, we utilize data from the service area of a major utility in the central Gulf Coast region. The service area is divided into 3.66 km (12,000 ft) by 2.44 km (8,000 ft) grid cells. These rectangular grid cells are defined by the utility company and used in their internal inventory and monitoring systems. This is the resolution at which the outage data and data on the power system (e.g., poles, transformers) are available. The 6,681 grid cells that cover the service area are the unit of analysis for our statistical modeling. The models are trained and validated using power outage data at the grid cell level from five past hurricanes [Danny (1997; 627 outages), Dennis (2005; 4,840 outages), Georges (1998; 1,075 outages), Ivan (2004; 13,568 outages), and Katrina (2005; 10,105 outages)]. The data used in the HOPM are described in more detail in Han et al. (2009a,b).

Uncertainties in the hurricane track and intensity forecasts are believed to be the largest source of uncertainty in the HOPM predictions. Shifts in the track or changes in intensity that are not accurately forecast by numerical weather prediction models can lead to large differences between the predicted and the observed power outages. This uncertainty can only be reduced through improved track and intensity forecasting. Other sources of uncertainty in modeling power outages include the accuracy of the input variables (e.g., wind speed, soil moisture) and the form (e.g., type of model used) and fit of the hurricane outage prediction models. Prior studies have assessed the accuracy of the hurricane wind field model used in this study (Willoughby and Rahn 2004; Willoughby et al. 2006) and quantified the uncertainty associated with the form of the HOPM (Guikema et al. 2010; Guikema and Quiring 2012). This paper focuses solely on quantifying how errors in the hurricane track and intensity forecasts influence the accuracy of power outage predictions.

OVERVIEW OF MCWSP ALGORITHM. The MCWSP model estimates the probabilities of wind speeds exceeding 17 m s^{-1} [39 miles per hour (mph); tropical storm–force winds], 26 m s^{-1} (58 mph), and 33 m s^{-1} (74 mph; hurricane-force winds) at a given point within the next 12, 24, 36, …, 120 h (DeMaria et al. 2009). For each tropical cyclone, the MCWSP model generates 1,000 forecast realizations by sampling from track and intensity forecast errors from the last 5 years and determines the wind radii of each realization using a simple climatology and persistence scheme. Because the realizations are based on track and intensity forecast errors, the MCWSP model accounts for the improvements in track and intensity forecasts that are shown in Fig. 1. An example of the 1,000 track realizations for the 1200 UTC 25 August 2011 forecast for Hurricane Irene is shown in Fig. 2 (left). Wind speed probabilities are then derived at each point in the model domain by counting the number of realizations where the wind speed exceeds the threshold of interest relative to the total number of realizations (Fig. 2, right).

Forecasters at the National Hurricane Center use data from various sources, including statistical and dynamical models, to make tropical cyclone track and intensity forecasts (Rappaport et al. 2009). In addition to using individual model outputs, forecasters also typically consider the average of predictions made by groups of different models initialized at the same time, which is referred to as a consensus forecast. Strong agreement between models is one factor that may indicate that there is limited forecast uncertainty. Goerss (2007), for example, found consensus model spread to be one of the leading predictors of consensus track forecast errors in the Atlantic. Goerss (2007) examined whether errors in a track forecast consensus (CONU) could be predicted ahead of time to provide forecasters with a measure of confidence in the forecast. It was found that consensus model spread (i.e., the average distance of each of the member forecasts from the consensus forecast) and tropical cyclone intensity were important predictors of CONU forecast errors in the Atlantic. Using CONU tropical cyclone track forecasts data from 2001 to 2003 and a pool of potential predictors, a stepwise linear regression was used to develop models for predicting errors. The predicted CONU tropical cyclone track forecast errors are often referred to as the Goerss predicted consensus error (GPCE).

DeMaria et al. (2013) demonstrated that past track forecast errors can be separated into terciles based on their corresponding GPCE value and that track forecast errors in the low (high) terciles tend to correspond to less (more) spread in forecast errors. As such, this methodology has been incorporated into the MCWSP model (DeMaria et al. 2013). Including GPCE modifies the wind speed probabilities such that they are more confined to the actual forecast track when the GPCE value is low: that is, when track forecast uncertainty is low. Conversely, when the GPCE value is high (e.g., track forecast uncertainty is high), the wind speed probabilities are modified such that they are less confined to the actual forecast track. Since the GPCE MCWSP model has shown progress in refining estimates of track forecast errors and represents the operational MCWSP, it is used in this paper.

INCORPORATING HURRICANE FORECAST ERROR INFORMATION INTO THE HOPM. The MCWSP model is well suited for use in this paper because it can generate a large number of possible scenarios for any given tropical cyclone track and intensity forecast. Each of these scenarios can be used as input for the HOPM to estimate outages. Three cases were chosen to demonstrate how the MCWSP can be used to drive the HOPM: Hurricane Ivan (2004), Hurricane Dennis (2005), and Hurricane Katrina (2005). These cases were chosen because they fall within the domain of the HOPM. This study focused on running the HOPM when the hurricane was 24 h prior to landfall. Although most utility companies will begin to evaluate the potential impact of a tropical storm and run decision-support models starting 3–5 days prior to landfall, we chose to utilize forecasts 24 h prior to landfall since this is the minimum lead time that utility companies need to plan and prepare for a hurricane. Once there is less than 24 h until landfall, it becomes very difficult to prepare and position personnel and equipment. The amount of uncertainty in the forecast track and intensity is also much lower at 24 h than it is for 48, 72, or 96 h. Therefore, the errors reported in this paper likely represent the best-case scenario in terms of the accuracy of the power outage predictions. Official forecast tracks for each of these three cases at 24 h prior to landfall are shown in Fig. 3. For each case, the MCWSP model was run to generate 1,000 forecast realizations by sampling from track and intensity forecast errors from the last 5 years and adding these errors directly to the official forecast of track and intensity issued 24 h prior to landfall.

The track and intensity from each forecast realization were used as input for the Willoughby et al. (2006) wind field model to simulate the maximum wind gust and the duration of strong winds [length of time that wind speeds are in excess of 20 m s^{-1} (45 mph)] for each grid cell in the HOPM domain. The threshold used for calculating the duration of strong

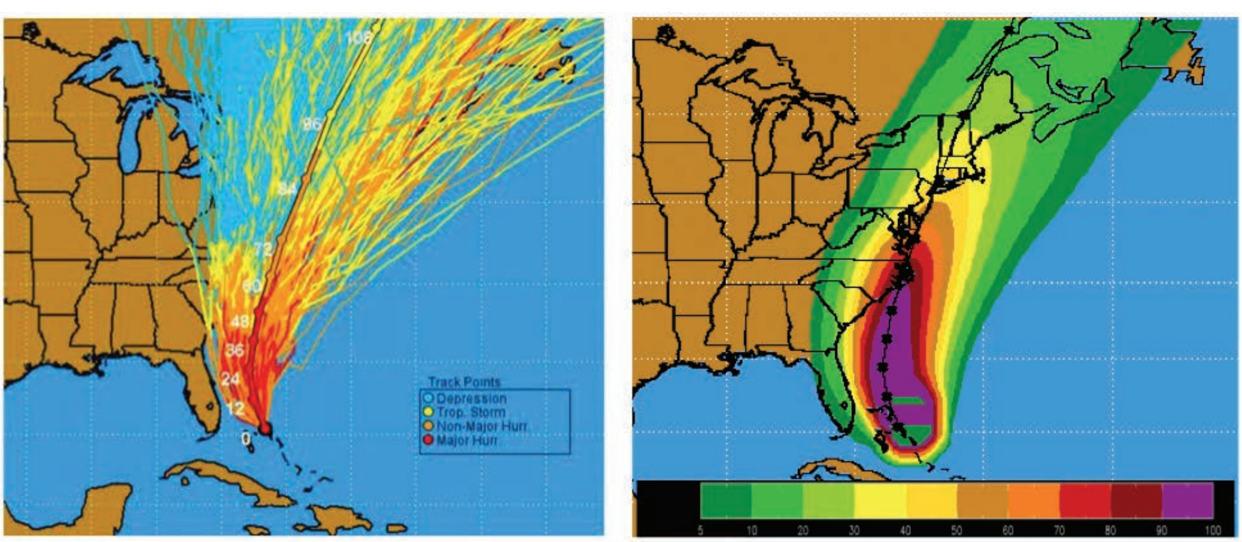


Fig. 2. (left) 1,000 track realizations for Hurricane Irene on 25 Aug 2011 and (right) the corresponding 34-kt 0–120-h cumulative wind speed probabilities.

winds was based on previous research (Huang et al. 2001a,b). Finally, these data are used as input to the HOPM and the location and number of outages are computed for each forecast realization.

The resulting power outage distributions are shown in Fig. 4. In addition to the outage distributions, Fig. 5 also shows the tracks and intensities of the five forecast scenarios generated by the MCWSP model producing the smallest number of model-predicted outages (i.e., best-case scenarios) and the five forecast scenarios producing the largest number of model-predicted outages (i.e., worst-case scenarios) at 24 h prior to landfall.

The HOPM predictions in Fig. 4 show considerable spread, confirming that errors in the official tropical cyclone forecast track and intensity have a significant impact on the number of outages predicted by the HOPM. Hurricane Ivan appears to be the most extreme of the three cases tested, with HOPM power outages estimates ranging from 175 to almost 34,000 (standard deviation of 8,241 outages). This is in agreement with the experiences of utility personnel in the area impacted by the storm who reported that Hurricane Ivan was a particularly difficult storm for them to estimate impacts for (C. Wallis, personal communication). Hurricane Dennis also was quite sensitive to hurricane track/intensity forecast errors, ranging from 1,330 to almost 30,000 predicted outages (standard deviation of 6,348). The variation in the number of outages predicted for Hurricane Katrina was less sensitive to errors in the forecast track and intensity yet still range from 3,377 to almost 34,000 outages (standard deviation of 3,682). Note that Hurricanes Ivan and Dennis had similar tracks, with landfall locations in less densely populated eastern Alabama and the Florida Panhandle, areas where population density is also more variable. In contrast, Hurricane Katrina made landfall over a more densely

Fig. 3. Official National Hurricane Center forecast tracks for (left) Hurricane Ivan, (middle) Hurricane Dennis, and (right) Hurricane Katrina 24 h prior to landfall.

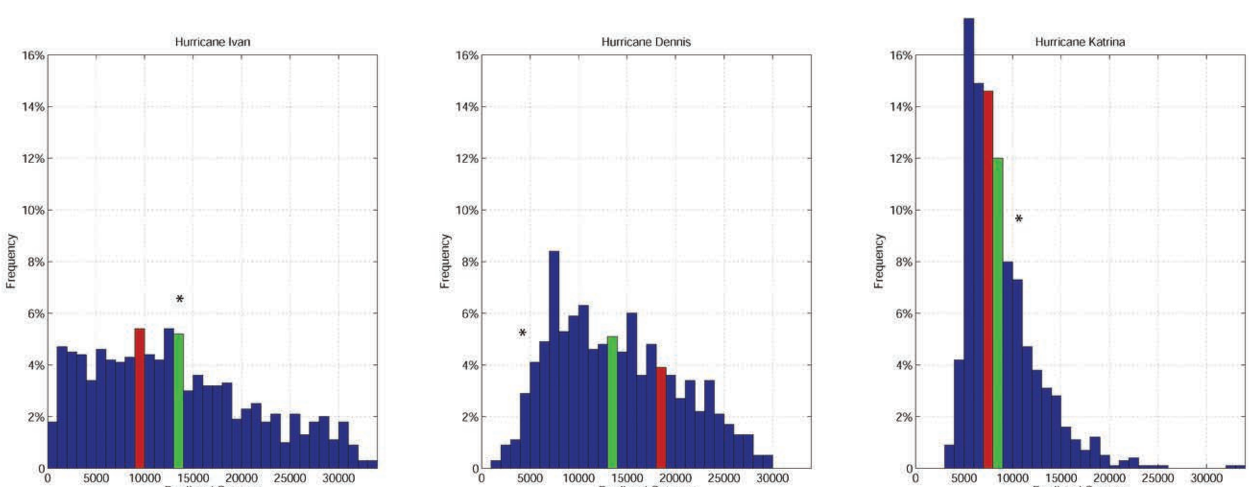


Fig. 4. Distribution of the HOPM-predicted power outages for (left) Hurricane Ivan, (middle) Hurricane Dennis, and (right) Hurricane Katrina made 24 h prior to landfall based on the 1,000 possible forecast track and intensity scenarios generated by the MCWSP. The bin corresponding to the number of outages based on the NHC official forecast is red, the bin corresponding to the mean of the 1,000 scenarios is green, and the bin corresponding to the actual number of outages has an asterisk above it.

populated area with a more uniform population density. Because of this, power outage predictions for Hurricane Katrina should be less sensitive to relatively small track changes. Small track changes for the other two storms could bring the areas of highest impact more directly over the more scattered areas of higher population density, but for Hurricane Katrina shifts in the track would keep the areas of high impact over more uniformly—and densely—populated areas.

In addition, it is not surprising that the power outage predictions for Hurricane Katrina had less spread than for Ivan and Dennis because Katrina had the lowest GPCE value. That is, the track forecast uncertainty for Katrina was relatively low (in the lowest tercile) and therefore the realizations produced by the GPCE MCWSP model were more constrained. The NHC track forecasts for Katrina that were issued up to 2½ days before landfall were very accurate. The official track forecast issued 24 h prior to 1200 UTC 29 August had an error of only 24 n mi (1 n mi = 1.15 mi), less than half the corresponding 10-yr average (1995–2004) calculated from all Atlantic basin forecasts (Knabb et al. 2005). On the other hand, the average NHC intensity forecast errors during Katrina were 17 kt (1 kt = 1.15 mph) for the 24-h forecasts, much larger than the corresponding Atlantic 10-yr average of 10 kt (Knabb et al. 2005).

The track and intensity of the hurricane scenarios that had the highest (i.e., worst case) and lowest (i.e., best case) number of predicted outages are shown in Fig. 5. Not surprisingly, the worst-case scenarios are associated with hurricanes that intensify during the 24 h prior to landfall and with storms that track directly over (or just to the west of) the utility company's service area. In particular, the worst-case scenarios for Hurricane Ivan and Hurricane Katrina are associated with storms that remained major hurricanes as they moved through the service area (i.e., they weakened very slowly after landfall). The

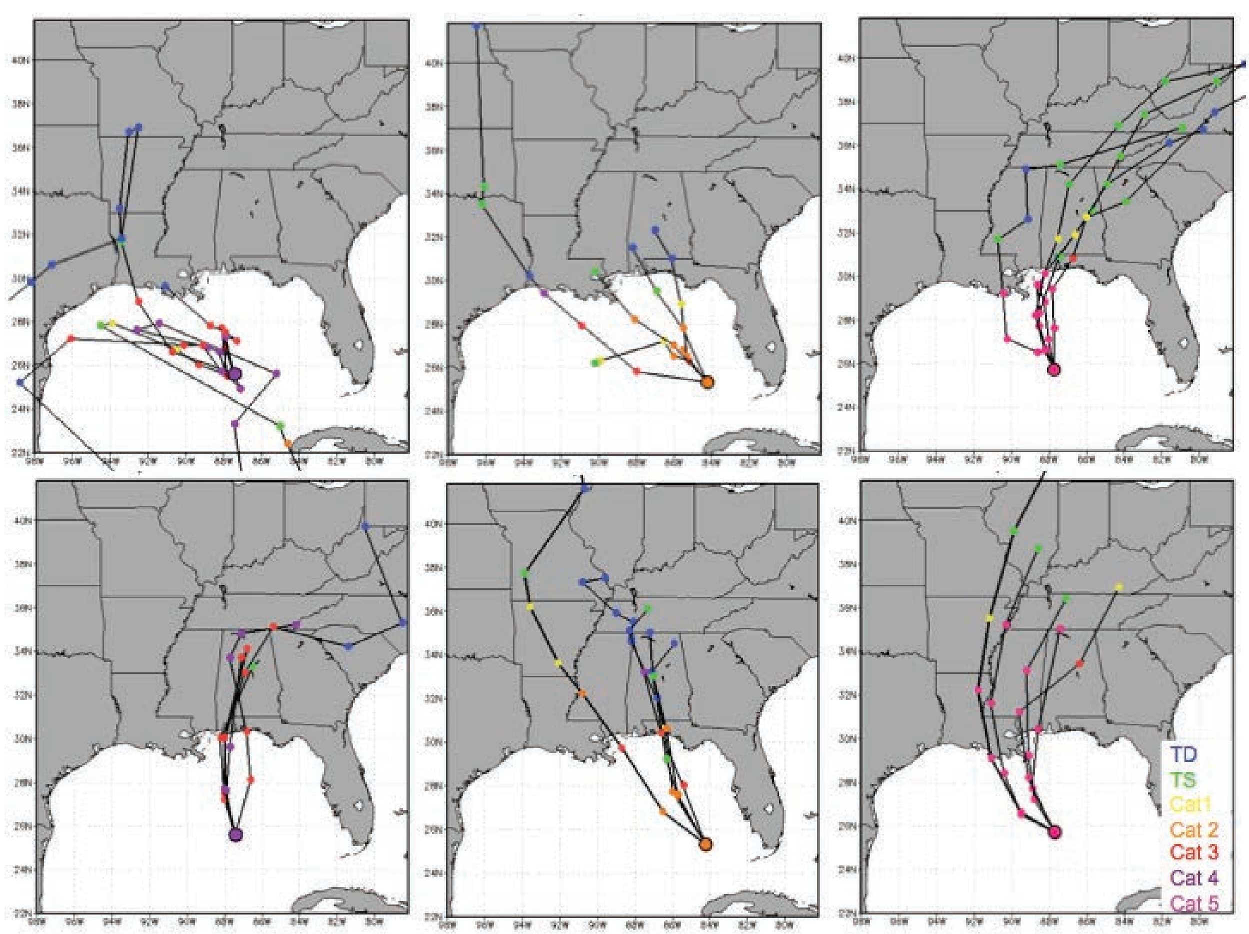


Fig. 5. (top) The five best-case scenarios (i.e., fewest predicted outages) and (bottom) the five worst-case scenarios (i.e., most predicted outages) based on HOPM-predicted power outages for (left) Hurricane Ivan, (middle) Hurricane Dennis, and (right) Hurricane Katrina using 1,000 possible forecast track and intensity scenarios generated by the MCWSP 24 h prior to landfall.

location and strength of these storms result in very high wind speeds in the service area. Conversely, all of the best-case scenarios are associated with hurricanes that weakened during the 24 h prior to landfall and with storms that shifted away from the service area or into the less densely populated areas. For example, for Hurricane Katrina the best-case scenarios were storms whose track was shifted to the east of Mobile and whose intensity decreased to tropical storm strength at (or shortly after) landfall. The best- and worst-case scenarios represent the extremes and they are much different than most of the 1,000 scenarios generated by the MCWSP model. However, these cases illustrate how changes in track and intensity (even 24 h prior to landfall) can have a significant impact on power outage predictions.

One method for quantifying the sensitivity of the HOPM predictions to variations in track and intensity is to calculate the coefficient of variation (coefficient of variation = standard deviation/ensemble mean) for each hurricane. The coefficient of variation provides a measure of the relative variability in the HOPM predictions and it allows us to directly compare the three cases. The coefficient of variation is 42% for Hurricane Katrina, 46% for Hurricane Dennis, and 63% for Hurricane Ivan. This suggests that more confidence (less uncertainty) could be placed in the outage predictions for Hurricane Katrina and Hurricane Dennis than for Hurricane Ivan.

While it is instructive to look at the extremes from all 1,000 scenarios to demonstrate how variations in hurricane track and intensity can influence the HOPM, it is better to use the ensemble average (i.e., the average number of outages based on all 1,000 simulations) to predict the number of outages. In all three cases that we examined, the ensemble average based on the 1,000 simulations is closer to the actual number of outages than predictions based on only the National Hurricane Center official forecast (OFCL) (Table 1). The HOPM prediction for Hurricane Ivan based on the OFCL 24 h prior to landfall had a relative error of 32.4% and the MCWSP ensemble average had a relative error of 3.9%. For Hurricane Dennis, the HOPM prediction based on the OFCL had a relative error of 288.4% and the MCWSP ensemble average had a relative error of 185.3%. The HOPM prediction for Hurricane Katrina based on the OFCL had relative error of 30.2% and the MCWSP ensemble average had a relative error of 12.9%. For these three cases, the ensemble average based on running the HOPM 1,000 times provides a better estimate of the actual number of outages than running the HOPM once using the National Hurricane Center official forecast. Based on the three cases that we examined, the ensemble average provides a reasonable estimate of the expected number of outages. Of course these three cases are not enough for us to conclude whether this finding is robust and would hold up if we evaluated more hurricanes.

The ensemble average appears to be a reasonable predictor of the power outages, but it is a single number. The 1,000 hurricane scenarios can also be used to estimate the uncertainty in this prediction by constructing confidence bounds around the ensemble average. For consistency with the forecast cone, the range that captures 67% of the forecast outages will be calculated. If the HOPM outages were normally distributed, this range could be calculated using mean and standard deviation from the 1,000 scenarios (e.g., lower bound = ensemble mean – 1 standard deviation; upper bound = ensemble mean + 1 standard deviation). However, it is evident from the histograms show in Fig. 4 that the outage predictions are not normally distributed and therefore the confidence bounds will be constructed using the empirical distribution by rank ordering the outage predictions and identifying the range that captures 67% of the forecast outages. Based on this approach, the outage range for Hurricane

Table 1. Summary of the HOPM-predicted power outages based on the National Hurricane Center official forecast and the 1,000 possible forecast track and intensity scenarios generated by the MCWSP (ensemble). The ensemble mean and standard deviation are based on the 1,000 track/intensity scenarios. The observed outages are the actual number of outages reported by the utility. The relative error is based on the difference between the forecast and observed number of outages divided by the observed number of outages.

	Hurricane Ivan	Hurricane Dennis	Hurricane Katrina
OFCL forecast	9,167	18,798	7,053
OFCL forecast relative error (%)	32.4	288.4	30.2
Ensemble mean (standard deviation)	13,035 (8241)	13,808 (6348)	8,800 (3628)
Ensemble mean relative error (%)	3.9	185.3	12.9
Observed outages	13,568	4,840	10,105

Ivan is 4,318–21,793; the outage range for Hurricane Dennis is 7,296–20,834; and the outage range for Hurricane Katrina is 5,631–11,793. In other words, the model estimates that there is a 67% chance that the actual number of outages for Hurricane Katrina will be between 5,631 and 11,793. The outage ranges for these three storms match the sensitivity results described above. Hurricane Ivan has the largest range (and therefore the greatest uncertainty), and Hurricane Katrina has the smallest range. The actual number of outages was within these ranges for Hurricane Ivan and Hurricane Katrina, but the actual number of outages for Hurricane Dennis (4840) was below the forecast range. We expect that actual outages will be outside this range 33% of the time.

Of course, the 1,000 scenarios can be used to construct outage ranges for any confidence level of interest (20%, 50%, 90%, etc.). These data can also be used to answer other questions, such as the probability that the number of outages will exceed 5,000 (or 10,000) or whatever other threshold is seen as important from a planning standpoint. This ability to estimate quantiles of the probability distribution for outages is particularly valuable for utility decision makers, who in many cases may be risk averse. A risk-averse decision maker is particularly concerned about uncertainty in impacts, weighting impacts nonlinearly (Clemen and Reilly 1999). If a model such as HOPM provides only a point estimate, it is not providing the information needed by risk-averse decision makers, limiting their ability to make well-supported decisions about prestorm preparation activities.

The ensemble-based approach used in this paper provides a significant advance over previous work for hurricane power outage prediction. Most previous methods for predicting power outages due to hurricanes (e.g., Liu et al. 2005, 2007, 2008; Han et al. 2009a,b; Guikema et al. 2010; Winkler et al. 2010) have not adequately modeled the uncertainty in the hurricane forecasts or incorporated it into the outage forecasts. Li et al. (2010) is the only prior study that has explicitly accounted for weather forecast uncertainty (i.e., differences between forecasted and observed weather variables). They provided confidence intervals on their damage predictions as well as worst-case scenarios. However, Li et al. (2010) focused on predicting damage to the power system from all weather events. Therefore, they did not explicitly incorporate uncertainty due to variations in the track and intensity of tropical storms and so their approach is not appropriate for quantify uncertainty associated with hurricanes. We have demonstrated that incorporating the uncertainty can both improve the predictions and provide a strong characterization of the uncertainties in the forecasts. Although this study has focused on the needs of electric utilities, the approach and methods used for quantifying and representing uncertainty can be applied to model how other weather-sensitive enterprises such as agriculture, transportation, emergency management, insurance, and water resources will be effected by adverse weather events such as hurricanes (Li et al. 2010). Accounting for the uncertain nature of hurricane forecasts in decision-support systems will provide better information that can enhance risk management and business continuity.

SUMMARY AND CONCLUSIONS. The main results of this study can be summarized as follows:

1) Small errors in the official track and/or intensity forecast can lead to large errors in the resulting HOPM outage prediction. Table 2 shows the official 24-h forecast position at approximately 24 h prior to landfall for Hurricanes Ivan, Katrina, and Dennis versus the observed positions and intensities at the forecast validation time (obtained from best-track data). The official 24-h track (intensity) forecasts for Hurricanes Ivan and Katrina 24 h prior to landfall had below (above) average errors (Franklin 2005, 2006), resulting in HOPM outage predictions with a relative error of about 30%. However, the official 24-h track forecast for Hurricane Dennis had below average

Table 2. Official 24-h forecast positions and wind speeds valid at approximately the time of landfall and the corresponding observed positions and wind speeds.

	Ivan	Dennis	Katrina
Forecast time	0600 UTC 15 Sep 2004	1800 UTC 9 Jul 2005	1200 UTC 28 Aug 2005
Valid time	0600 UTC 16 Sep 2004	1800 UTC 10 Jul 2005	1200 UTC 29 Aug 2005
Forecast (obs) lat	29.6°N (30.0°N)	29.9 N (29.9°N)	29.1 N (29.5°N)
Forecast (obs) lon	88.2°W (87.9°W)	87.2 W (86.9°W)	89.6 W (89.6°W)
Forecast (obs) wind speed	120 kt (105 kt)	120 kt (110 kt)	140 kt (110 kt)

error, the intensity forecast had approximately average error (Franklin 2006), and the resulting HOPM outage prediction had a relative error of almost 290%. This suggests that the HOPM is sensitive to errors in the official forecast track and intensity.

2) The hurricane realizations generated by the MCWSP produced a large range in the HOPM-predicted outages. The MCWSP algorithm creates realizations by sampling from the last 5 years of official forecast track and intensity errors. As such, the MCWSP realizations used to run the HOPM in this study can be considered to be relatively plausible TC track and intensity scenarios. Yet Fig. 4 shows the large range in HOPM-predicted outages that resulted from this ensemble of plausible TC scenarios. In general, HOPM-predicted outages were found to range from a few thousand to 30,000. This result confirms that the HOPM is sensitive to reasonable (expected) errors in the official forecast track and intensity despite HOPM giving accurate outage forecasts when run poststorm using the best-track data. Stated differently, the uncertainty in track and intensity forecasts accuracy has a major influence on the accuracy of the outage forecasts.
3) The ensemble methodology used in this study yielded supplemental information about the HOPM prediction that may be useful to end users. For Hurricanes Ivan, Katrina, and Dennis, the ensemble mean provided a more accurate prediction of the observed number of outages than the HOPM outage prediction generated using the official TC forecast. However, the small sample size prevents us from determining whether this finding is robust. The primary advantage of using an ensemble approach is that it provides a means to communicate uncertainty to decision makers. For example, confidence bounds can be constructed from the resulting outage distributions to help support decision making by end users concerned about forecast uncertainty. Future work should examine what elements of this information are most useful to end users and the best way to communicate that information.

As is the situation with most case studies, the above results cannot be generalized too broadly. Not all models used for decision support are necessarily as sensitive to errors in their meteorological inputs as the HOPM. Yet we believe these results do emphasize the importance of understanding and estimating how uncertainty and errors compound when forecasts models are integrated. This may be especially challenging when integrating models from different disciplines such as when meteorological predictions are used in a socioeconomic or infrastructure model. The development and use of decision-support models brings together people with different backgrounds who may not fully understand the intricacies of the other models and inputs.

REFERENCES

Balijepalli, N., S. Venkata, C. Richter Jr., R. Christie, and V. Longo, 2005: Distribution system reliability assessment due to lightning storms. *IEEE Trans. Power Delivery,* **20,** 2153–2159.

Brown, R. E., S. Gupta, R. D. Christie, S. S. Venkata, and R. Fletcher, 1997: Distribution system reliability assessment: Momentary interruptions and storms. *IEEE Trans. Power Deliv.,* **12,** 1569–1575.

Cerruti, B. J., and S. G. Decker, 2012: A statistical forecast model of weather-related damage to a major electric utility. *J. Appl. Meteor. Climatol.,* **51,** 191–204.

Clemen, R. T., and T. Reilly, 1999: *Making Hard Decisions.* Duxbury Press, 733 pp.

Davidson, R. A., H. Liu, I. K. Sarpong, P. R. Sparks, and D. V. Rosowsky, 2003: Electric power distribution system performance in Carolina hurricanes. *Nat. Hazards Rev.,* **4,** 36–45.

DeGaetano, A. T., B. N. Belcher, and P. L. Spier, 2008: Short-term ice accretion forecasts for electric utilities using the Weather Research and Forecasting model and a modified precipitation-type algorithm. *Wea. Forecasting,* **23,** 838–853.

DeMaria, M., J. A. Knaff, R. Knabb, C. Lauer, C. R. Sampson, and R. T. DeMaria, 2009: A new method for estimating tropical cyclone wind speed probabilities. *Wea. Forecasting,* **24,** 1573–1591.

——, and Coauthors, 2013: Improvements to the operational tropical cyclone wind speed probability model. *Wea. Forecasting,* **28,** 586–602.

Franklin, J., 2005: 2004 National Hurricane Center forecast verification report. National Hurricane Center Rep., 46 pp. [Available online at www.nhc.noaa.gov/verification/pdfs/Verification_2004.pdf.]

——, 2006: 2005 National Hurricane Center forecast verification report. National Hurricane Center Rep., 52 pp. [Available online at www.nhc.noaa.gov/verification/pdfs/Verification_2005.pdf.]

Georgiou, P. N., 1985: Design wind speeds in tropical cyclone-prone regions. Ph.D dissertation, University of Western Ontario, 295 pp.

Goerss, J. S., 2007: Prediction of consensus tropical cyclone track forecast error. *Mon. Wea. Rev.,* **135,** 1985–1993.

Guikema, S. D., and S. M. Quiring, 2012: Hybrid data mining-regression for infrastructure risk assessment based on zero-inflated data. *Reliab. Eng. Syst. Saf.,* **99,** 178–182.

——, ——, and S. D. Han, 2010: Prestorm estimation of hurricane damage to electric power distribution systems. *Risk Anal.,* **30,** 1744–1752.

Han, S. R., S. D. Guikema, and S. M. Quiring, 2009a: Improving the predictive accuracy of hurricane power outage forecasts using generalized additive models. *Risk Anal.,* **29,** 1443–1453.

——, ——, ——, K. H. Lee, D. V. Rosowsky, and R. A. Davidson, 2009b: Estimating the spatial distribution of power outages during hurricanes in the Gulf Coast region. *Reliab. Eng. Syst. Saf.,* **94,** 199–210.

Huang, Z., D. V. Rosowsky, and P. R. Sparks, 2001a: Hurricane simulation techniques for the evaluation of wind-speeds and expected insurance losses. *J. Wind Eng. Ind. Aerodyn.,* **89,** 605–617.

——, ——, and ——, 2001b: Long-term hurricane risk assessment and expected damage to residential structures. *Reliab. Eng. Syst. Saf.,* **74,** 239–249.

Knabb, R. D., J. R. Rhome, and D. P. Brown, 2005: Tropical cyclone report: Hurricane Katrina, 23-30 August 2005. National Hurricane Center Rep., 43 pp. [Available online at www.nhc.noaa.gov/pdf/TCR-AL122005_Katrina.pdf.]

Li, H., L. A. Treinish, and J. R. M. Hosking, 2010: A statistical model for risk management of electric outage forecasts. *IBM J. Res. Dev.,* **54,** 8:1–8:11, doi:10.1147/jrd.2010.2044836.

Liu, H., R. A. Davidson, D. V. Rosowsky, and J. R. Stedinger, 2005: Negative binomial regression of electric power outages in hurricanes. *J. Infrastruct. Syst.,* **11,** 258–267.

——, ——, and T. V. Apanasovich, 2007: Statistical forecasting of electric power restoration times in hurricanes and ice storms. *IEEE Trans. Power Syst.,* **22,** 2270–2279.

——, ——, and ——, 2008: Spatial generalized linear mixed models of electric power outages due to hurricanes and ice storms. *Reliab. Eng. Syst. Saf.,* **93,** 875–890.

Nateghi, R., S. D. Guikema, and S. M. Quiring, 2011: Comparison and validation of statistical methods for predicting power outage durations in the event of hurricanes. *Risk Anal.,* **31,** 1897–1906.

National Hurricane Center, cited 2010: National Hurricane Center forecast verification. [Available online at www.nhc.noaa.gov/verification/.]

Rappaport, E. N., and Coauthors, 2009: Advances and challenges at the National Hurricane Center. *Wea. Forecasting,* **24,** 395–419.

Reed, D. A., 2008: Electric utility disruption analysis for extreme winds. *J. Wind Eng. Ind. Aerodyn.,* **96,** 123–140.

——, M. D. Powell, and J. Westerman, 2010a: Energy infrastructure damage for Hurricane Rita. *Nat. Hazards Rev.,* **11,** 102–109.

——, ——, and ——, 2010b: Energy supply system performance for Hurricane Katrina. *J. Energy Eng.,* **136,** 95–102.

Willoughby, H. E., and M. E. Rahn, 2004: Parametric representation of the primary hurricane vortex. Part I: Observations and evaluation of the Holland (1980) model. *Mon. Wea. Rev.,* **132,** 3033–3048.

——, R. W. R. Darling, and M. E. Rahn, 2006: Parametric representation of the primary hurricane vortex. Part II: A new family of sectionally continuous profiles. *Mon. Wea. Rev.,* **134,** 1102–1120.

Winkler, J., L. Dueñas-Osorio, R. Stein, and D. Subramaniam, 2010: Performance assessment of topologically diverse power systems subjected to hurricane events. *Reliab. Eng. Syst. Saf.,* **95,** 323–336.

Zhou, Y., A. Pahwa, and S. Yang, 2006: Modeling weather-related failures of overhead distribution lines. *IEEE Trans. Power Syst.,* **21,** 1683–1690.

Zhu, D., D. Cheng, R. P. Broadwater, and C. Scirbona, 2007: Storm modeling for prediction of power distribution system outages. *Electr. Power Syst. Res.,* **77,** 973–979.