A Case Study on Power Outage Impacts from Future Hurricane Sandy Scenarios

D. W. Wanik[1], E. N. Anagnostou[1], M. Astitha[1], B. M. Hartman[2],

G. M. Lackmann[3], J. Yang[1], D. Cerrai[1], J. He[4], M. E. B. Frediani[1]

[1] Department of Civil and Environmental Engineering, University of Connecticut

[2] Department of Statistics, Brigham Young University

[3]Department of Marine, Earth and Atmospheric Sciences, North Carolina State University

[4] Analytics and Research Department, Travelers Insurance

Resubmitted with Major Revision in September 2017

(Original manuscript JAMC-D-16-0184):

Journal of Applied Meteorology and Climatology (JAMC)

Corresponding Author: Dr. David Wanik, Civil and Environmental Engineering, University of Connecticut, Storrs, CT 06269 | Email: david.w.wanik@gmail.com | Tel.: 860-463-4265

Abstract

Hurricane Sandy (2012, referred to as "Current Sandy") was among the most devastating storms to impact Connecticut's overhead electric distribution network, resulting in over 15,000 outage locations that affected more than 500,000 customers. In this paper we estimate the severity of tree-caused outages in Connecticut under future-climate Hurricane Sandy simulations, each exhibiting strengthened winds and heavier rain accumulation over the study area from large-scale thermodynamic changes in the atmosphere and track changes in the year ~2100 ("Future Sandy"). Three machine learning models used five weather simulations and the ensemble mean of Current and Future Sandy, along with land use and overhead utility infrastructure data, to predict the frequency and spatial distribution of outages across the Eversource Energy-Connecticut service territory. To assess the influence of increased precipitation from Future Sandy, we compared two approaches: an outage model fit with a full set of variables accounting for both wind and precipitation, and a reduced set with only wind. Future Sandy displayed an outage increase of 42% - 64% when using the ensemble of WRF simulations fit with three different outage prediction models. This study is a proof-of-concept for the assessment of increased outage risk resulting from potential changes in tropical cyclone intensity associated with late-century thermodynamic changes driven by the IPCC AR4 A2 emissions scenario.

Capsule

"How many more power outages would occur if a storm like Hurricane Sandy impacted the Connecticut electric distribution utility in the future, in a warmer climate scenario?"

1. Introduction

Hurricane Sandy ("Sandy") was among the three major storms that affected Connecticut in the past decade (alongside Tropical Storm Irene and the October 2011 Nor'easter). Though technically classified as a post-tropical cyclone when it made landfall (Blake et al. 2012), Sandy was impactful to Connecticut's largest electric utility, The Connecticut Light & Power Company, doing business as Eversource Energy ("Eversource"). At the peak of the restoration over 500,000 customers were affected, with some customers without power for nine days. In addition, more than 15,000 outages were repaired by a workforce six times as large as Eversource's normal operating workforce (Caron et al. 2013). Figure 1 shows the spatial distribution of outages across the Eversource service territory. Most of the outages were concentrated in Fairfield County (southwestern Connecticut), where substantial overhead electric distribution infrastructure and population is present. Although storm surge was extensive during Sandy (Fanelli et al. 2013), the majority of outages in the Eversource service territory were caused by wind and trees affecting overhead lines (Personal Communication, Thomas Layton, Eversource Energy, 2015).

Weather is found to be responsible for nearly 44% of power outages in the United States (Campbell 2013), with hurricanes and tropical storms affecting an average of 782,695 customers per event (Hines et al. 2008). The annual cost of power outages (in 2012 USD) has been estimated between $28 billion to as much as $209 billion, with annual weather-related outages estimated to cost between $25 billion to $70 billion (Abraham et al. 2013). In addition to impacts of the economy, utilities can also incur direct costs from tens to hundreds of millions of dollars for labor and equipment due to the storm (Northeast Utilities 2013).

Given that Sandy was particularly impactful for utilities in the mid-Atlantic and New England (Henry and Ramirez-Marquez 2016), in this paper, we present a proof-of-concept for assessing the impacts of Sandy within a future climate scenario as it pertains to overhead electric distribution networks ("distribution networks"). A case study or storyline approach is consistent with the pseudo-global warming (PGW) approach taken here as a viable means to evaluate the impacts of climate warming on an observed weather event (Schär et al. 1996; Trenberth et al. 2015; Shepherd 2016). The present study complements existing long-term hurricane planning efforts in the United States and answers the following question: how many more outages would occur if Hurricane Sandy impacted Connecticut in the future, forced by a different large scale climate scenario? As noted by Staid et al. (2014), there is less consensus about whether the frequency of tropical cyclones will increase (Emanuel 2005) or decrease (Emanuel et al. 2008; Knutson et al. 2010). Nevertheless, there is consensus that the strongest tropical cyclones will strengthen to some degree (Pielke Jr. 2007; Knutson et al. 2010). Yates et al. (2014) examined the potential of substations being flooded under Future Sandy scenarios, and found that coastal flooding in Long Island, NY (close proximity to Connecticut) could nearly double in some areas.

This study is facilitated by the recent work of Lackmann (2015) who investigated Hurricane Sandy track scenarios under current ("Current Sandy"), future (~2100, "Future Sandy") and past climate (~1890, "Past Sandy") thermodynamic and sea surface conditions. Lackmann (2015) found that while past Sandy tracks are indistinguishable from the Current Sandy simulations, Future Sandy scenarios appear to be stronger and shifted further north towards New England. See Lackmann (2015) for plots of sea-level pressure; lower sea-level pressure is consistent with stronger storm-centric winds. While the Lackmann (2015) study has caveats, including whether or not Sandy would form under future conditions, the goal of his study was to isolate the influence of changes in the large-scale thermodynamic environment on the intensity and track of a system like Sandy.

Using Lackmann's technique (2015), we demonstrate a case study of how potential storm, a hypothetical future Hurricane Sandy, might affect the electrical grid in Connecticut. Generalized conclusions about how future tropical cyclones can affect the distribution network requires examination of many events, featuring a variety of tracks and intensities. Nevertheless, the added value of the presented methodology is that it can be implemented when data for future tropical cyclones becomes available. In order to assess hypothetical future changes in overhead electric distribution grid outages based on simulation of a single storm event, it is necessary to recognize that impact changes will be a function of (i) changes in the intensity and size of the storm itself, and (ii) changes in the track of the storm. This study combines these two aspects through Lackmann's (2015) simulations, which we believe provides a framework for emergency managers to evaluate the impacts of climate data on infrastructure networks they manage.

The paper is structured as follows: Section 2 discusses the weather simulation modeling framework, and a comparison of the Current and Future Sandy storms; Section 3 provides details on the outage prediction modeling, including an overview of the nonparametric models and our methodology; Section 4 contains the results and a discussion on how track and severity influenced the occurrence of power outages, as well as the limitations of the study; Section 5 contains major findings and future research directions.

2. Weather Data

2.1 Background

Within the IPCC Fourth Assessment Report (AR4) one can find several future emissions scenarios and the associated impact on global average temperature and sea level rise; these scenarios include keeping emissions at constant levels from the year 2000, and subsequent scenarios with increased emissions. In our study, we utilized the A2 emissions scenario which describes a heterogeneous world with increasing population and carbon emissions through the year 2100 (Nakicenovic and Swart 2000). It features the second-highest emission scenario of the scenarios used at that time, loosely corresponding to the RCP 8.5 scenario in the IPCC Fifth Assessment Report (AR5).

For the work presented here we relied on the Weather Research and Forecasting (WRF) model (Skamarock et al. 2008) simulations reported in Lackmann (2015). Two different five-member ensemble simulation sets of Sandy were used, one for the current and one for the future climate scenario. The model simulations included three gridded domains with 54, 18 and 6-km horizontal grid spacing using one-way nesting for the two inner grids. From the 17 members described by Lackmann (2015), five were selected to supply the outage prediction model input. The WRF members were selected based on the availability of the 6-km domain and the variations in the physical parameterization schemes. To achieve a sample of available WRF configurations, the variations included cumulus parameterization, microphysics, and planetary boundary layer schemes. A summary of the variations in the physical parameterizations for each WRF ensemble member is provided in Table 1 herein and in Lackmann (2015), and the 6-km domain is displayed in Figure 2a.

The initial and boundary conditions for the Current Sandy ensemble set were obtained from the European Center for Medium Range Weather Forecasting (ECMWF) interim reanalysis (Dee et al. 2011), with an approximate spatial grid of 0.7 deg. The pseudo-global warming (PGW) procedure used to generate future simulations of Sandy was described in detail by Lackmann (2015), and here we describe the essential aspects. Thermodynamic changes between the 1990s and 2090s were computed using a subset of general circulation model (GCM) projections from the CMIP3 project (Meehl et al. 2007) for the A2 emissions scenario. The GCM-based temperature change fields were applied to initial and lateral boundary conditions as well as to lower boundaries (sea-surface and soil temperatures) in the original ensemble. At constant relative humidity, warming was associated with an increase in specific humidity. A hydrostatically balanced geopotential field was then computed based on the modified virtual temperature. The digital filter initialization (DFI) procedure in WRF was used to ensure balance between the wind and mass fields in the model initial conditions. Thus, the future simulations essentially answer the question: "If the synoptic weather pattern preceding Sandy were to take place in a warmer, moister tropospheric environment, how would the track and intensity of the system change?"

The authors have much experience using gridded, numerical weather prediction (NWP) model outputs for predicting storm-related power outages (Wanik et al. 2015; He et al. 2016; Wanik et al. 2017). Similar to our previous work, the WRF simulations were processed into a set of parameters that serve as input to the outage model. Specifically, within the simulated hours enclosing the storm period across the study area, wind and precipitation variables were post-processed to summarize the storm temporal evolution. Wind speed at 10 meter height, precipitation accumulation, and surface gust (see Wanik et al. 2015 for computation) were reduced into the storm maxima and durations exceeding wind thresholds at each grid point within the area covering the Eversource service territory in Connecticut (see Table 2; a detailed post-process description is given by Wanik et al. 2015).

2.2 Evaluation of Current Sandy WRF Simulations

To evaluate the consistency of the Current Sandy runs, we compared the simulated wind speed and precipitation to available observations. We used wind speed observations from airport

(METAR) stations provided by the National Centers for Environmental Prediction (NCEP) ADP Global Upper Air and Surface Weather Observations (National Centers for Environmental Prediction et al. 1997) and precipitation from the NCEP Stage IV analysis data (radar and gauges; (Lin and Mitchell 2005). The statistical error metrics are listed in the Appendix.

We present a comparison of model-simulated temperature from the Current Sandy CNTRL simulation valid 18 UTC 28 October 2012, and the closely corresponding actual temperature as shown by a GOES-13 IR image at from 18:15 UTC 28 October 2012. The comparison demonstrates that the CNTRL simulation captured the asymmetrical structure of Current Sandy, and this builds confidence in the accuracy of the WRF simulations we use. The time series of 10-m wind speed (Figure 3) revealed a temporal bias, but overall the WRF model was able to depict the highest wind speeds across all simulations. The wind speed RMSE varied between 2.6 – 4.5 m s^{-1} and the mean bias (MB) between 0.01 – 3.2 m s^{-1}, depending on station and WRF simulation (Table 3).The model predicted precipitation exhibited low bias and errors compared to Stage-IV radar-gage data for the gridded domain over Connecticut for all WRF simulations (Table 3). The RMSE varied between 1.94 – 3.48 mm (2.62 mm for the ensemble mean) and the MB between 0.53-0.83 mm (0.69 mm for the ensemble mean). Spatial distribution and magnitude of the predicted accumulated precipitation agreed with the Stage IV data (Figure 4) in that all members depicted high accumulation at the southwest region of the domain, which was left of Sandy's landfall. Accumulated precipitation at the northeastern part of the domain exhibited the same spatial pattern and magnitude as shown in the Stage IV plot. Precipitation is believed to contribute to power outages by wetting the soil and allowing for easier uprooting of trees (Foster 1988; McRoberts et al. 2017).

2.3 Comparison between Current and Future Sandy Simulations

Changes in simulated future storm impacts in Connecticut may be attributable both to the more northward track and to the lower minimum sea level pressure of Future Sandy. The cause of Sandy's more northward future track was discussed in Lackmann (2015) and is also consistent with the simulations of Yates et al. (2014). Increased tropical cyclone intensity with warming has been analyzed by Hill and Lackmann (2011) and others, and can be interpreted as the result of increased condensational heating. In Section 2, we evaluate how the change in track drives the change in resulting wind and precipitation intensity across the Eversource service territory. In later sections, we will incorporate lessons learned about the consequence of each simulations' track into the results of the Outage Prediction Modeling.

2.3.1 Storm Track Comparison

The "best track" (thick, dashed black line on Figure 5), as defined by the National Hurricane Center (NHC), is a smoothed representation of the tropical cyclone's location and intensity (e.g., latitude, longitude, maximum sustained surface winds, and minimum sea-level pressure at 6-hourly intervals). The simulated storm tracks of Current Sandy agreed with the NHC best track, while the simulated Future Sandy tracks deviated towards the northeast rather than the Mid-Atlantic States, with the Future Sandy center passing considerably closer to the state of Connecticut. The ENS simulation made landfall closest to the NHC track compared to the other five WRF simulations in Current Sandy. We have confidence in the representativeness of the Sandy WRF runs from Lackmann (2015) because they accurately represented Sandy's track and intensity in the current climate, and they capture the asymmetry of the cloud and precipitation shield (Figures 2 and 4). See Table 1 for a list of all evaluated WRF model simulations.

All tracks except the NOTCFLX simulation made landfall below the NHC track in Current Sandy (Figures 4, 5). The NOTCFLX simulation made landfall farther northeast in both Current and Future Sandy, and the CNTRL simulation had the most southerly track of all members in Current Sandy. The GODDARD, MORRIS and WDM6 tracks were all very similar in Current Sandy, while the GODDARD, CNTRL and ENS tracks were very similar for Future Sandy. The MORRIS and NOTCFLX were the only simulations that had storm centers pass over Connecticut in Future Sandy. The GODDARD, CNTRL and ENS simulations had Future Sandy tracks that made landfall on Long Island, NY, and the WDM6 simulation made landfall in New Jersey.

2.3.2 Storm Magnitude Discussion

We evaluated the change in wind and precipitation magnitude by creating cumulative distribution function (CDF) plots of total accumulated precipitation (Figure 6), maximum gust (Figure 7), and maximum wind at 10 m (Figure 8) for Current and Future Sandy. Each CDF plot shows the distribution of the 2-km grid cells for the variable of interest strictly within the Eversource service territory. The shift to the right of Future Sandy in each plot relative to Current Sandy indicates an increase in the magnitude of wind and precipitation variables.

The GODDARD simulation shows that the maximum gust and wind at 10 m in some of the upper percentiles were decreased in Future Sandy relative to Current Sandy scenarios (Figures 7 -8). The GODDARD simulation exhibited very similar distributions of maximum wind at 10-m between Current and Future Sandy. Comparatively, the WDM6, CNTRL and ENS simulations had greater separation between the Current and Future Sandy distributions. The increase in the cumulative distributions of total precipitation, gust, and 10-m wind speed between Current and Future Sandy indicate that Future Sandy was more intense in most simulations.

Table 4 provides the average for each of the distributions in Figures 6 – 8. On average, the individual model maximum gust increased 3% - 10%, maximum 10-m wind speeds increased 6% - 13%, and total precipitation increased 60% - 187% when changing from Current to Future scenario in the Eversource service territory. In comparison, the ensemble mean increase of maximum 10 m wind speed (6%) and gust (4%) was lower relative to the individual ensemble members, and more similar to values from the GODDARD simulation.

Spatial distribution of changes in wind and precipitation variables show that the majority of WRF simulations exhibit an increase in magnitude of the evaluated weather variables across most of the Eversource service territory (Figure 9). Specifically, for each 2-km grid cell, we subtracted the Current Sandy value from the Future Sandy value, such that positive values on the map indicate an increase in magnitude and negative values indicate a decrease. The increase in the spatial distribution of total precipitation was mostly concentrated in southwest and central Connecticut, while the changes in gust and 10 m wind speed distribution varied depending on the simulation. Given that the tracks shifted northward towards Connecticut, we initially expected all variables to increase in southwest Connecticut, but this did not occur. While precipitation increased heavily in southwestern Connecticut, the majority of increased gust and wind at 10-m actually occurred over eastern Connecticut. The NOTCFLX and WDM6 simulations showed the greatest increases in total accumulated precipitation, up to 200 and 300 mm per grid cell in southwestern Connecticut. The WDM6 simulation was the most southerly of the Future Sandy tracks evaluated, and wind at 10-m height and gust were increased in eastern Connecticut while there were decreases in western Connecticut. Although each WRF simulation is a hypothetical scenario, each should be treated as equally plausible as each accurately captured the Current Sandy track (Figure 4), and wind (Figure 3) and precipitation magnitude (Figure 5).

3. Outage Prediction Model (OPM)

3.1 Background

There is history of research in the field of hurricane outage modeling (e.g., predicting locations needing repair), outage monitoring (e.g., detecting locations with power outages), and outage duration modeling (e.g., estimating time until power is restored) for electric distribution networks. Early research leveraged parametric models, such as generalized linear models (Li et al. 2010) and generalized linear mixed models (Guikema and Davidson 2006; Liu et al. 2008), and later researchers explored probabilistic methods (Mensah and Duenas-Osorio 2014) and non-parametric methods, including classification and regression trees (Quiring et al. 2011; Wanik et al. 2015), neural networks (Cole et al. 2017), Bayesian additive regression trees (Nateghi et al. 2011; He et al. 2016) and random forest (Nateghi et al. 2014; Wanik et al. 2017). Beyond building models for specific utilities (Nateghi et al. 2014; Wanik et al. 2015; He et al. 2016), outage models have been re-calibrated with publicly available data such that the models can be generalized to other geographic regions (Guikema et al. 2014). Other recent research has investigated how tropical cyclone risk would affect customer outages under different climate change scenarios (Staid et al. 2014).

This study is an extension of the outage prediction system previously created for Eversource (Wanik et al. 2015; He et al. 2016) to predict outages associated with synoptic scale weather systems. The response variable in both models was the count of outages per 2-km grid cell, defined as locations that require manual intervention to restore power. Given that some outage records were missing geographic coordinates (e.g., latitude and longitude), we used 15,251 of the 16,460 recorded outages from Current Sandy for modeling. The outage prediction modeling framework from the referenced works consisted of multiple machine learning models that used atmospheric conditions, infrastructure, and land use surrounding the overhead power lines to predict outages for upcoming weather events (Figure 10). Electric grid infrastructure was represented by the counts of isolating devices (i.e., transformers, switches, reclosers, and fuses) per 2-km grid cell. In this paper, land use and infrastructure variables were aggregated on the same 2-km grid by which outages were aggregated. Our research group previously demonstrated how including land use and infrastructure data contributed to improved spatial accuracy of outage predictions (Wanik et al., 2015), how results can be improved by including an indicator for tree-leaf condition (He et al., 2016), and how different machine learning models yielded more accurate point estimates and predictive intervals depending on the unit of aggregation (i.e., grid cells, towns, and service territories) (He et al., 2016). Key differences between the data used in this paper and the Wanik et al. (2015) and He et al. (2016) papers are (1) the use of different storms to train and validate the model, (2) the grid spacing of weather simulation data, (3) there is no tree-leaf condition indicator in this study as we assume the storm will have the same tree-leaf condition in Current and Future Sandy.

For this study, the infrastructure and land use are static variables, whereas the atmospheric conditions were obtained using numerical weather prediction (NWP) simulations. Atmospheric variables from each WRF simulation were then used as inputs for three machine learning models (see Section 3.2). In addition to the five individual WRF simulations, the ensemble mean of the five WRF simulations was used as input for the outage models. The atmospheric variables used for outage modeling were based on the 6-km nested domain of the WRF grid. In this study, we used the same 2-km aggregated land use and infrastructure data as in Wanik et al. (2015) and joined these data to the centroid of the nearest 6-km centroid of the atmospheric forcing data. A list of all data included in the outage models is presented in Table 2.

3.2 Nonparametric Models

Nonparametric models have been used in the power outage modeling community (Nateghi et al. 2014; Wanik et al. 2015; He et al. 2016) because they require fewer assumptions about the underlying relationship between the explanatory variables and the response variable than parametric models. In this study we used three nonparametric, machine learning (ML) models to evaluate each of the different Sandy scenarios: Bayesian additive regression trees (BART), boosted trees (BT) and random forest (RF). The model parameters were estimated using the R packages “bartMachine” (Kapelner and Bleich 2014), “gbm” (Ridgeway 2007), and “randomForest” (Liaw and Wiener 2002), respectfully. We previously used BT and RF in Wanik et al. (2015) and BART in He et al. (2016) to predict power outages in the Eversource service territory for a wide variety of storms (i.e., blizzards, thunderstorms, and hurricanes.)

The BART model is a derivation of the Bayesian classification and regression trees model (CART) that takes advantage of a back-fitting Markov chain Monte Carlo (MCMC) algorithm in generating the posterior sample of classification and regression trees (Chipman et al. 2012). BART as a Bayesian model utilizes a likelihood maximization procedure that benefits from well-selected prior distributions and parallel-grown decision trees. The BT model is a decision tree-based stochastic gradient boosting algorithm that fits a decision tree on the residuals of the previous tree so that overall fit becomes the cumulative effort of many “weak learners” (Friedman 2001). The RF model uses a random subset of the explanatory variables (with replacement) to build multiple decision trees, and the average of the predictions across all decision trees is used as the final prediction. The RF model was also ideal for our study because of its robustness to outliers and full use of the candidate variables (Breiman 2001).

Each nonparametric model evaluated has advantages and disadvantages, which is a function of how each handles the input data and relates it to the response variable (Mackinnon and Glick 1999; Vapnik 1999). An advantage of nonparametric models is that they are able to nonlinearly relate the input data to the response variable, which requires no assumptions from the analyst. Another advantage of these models is that one does not need to eliminate correlated explanatory variables - the correlated variables increase the time needed to train the models, but will not detract from predictive accuracy. A general disadvantage of nonparametric models is that they may not be good at extrapolating beyond the dynamic range of the independent or explanatory variables. For this reason we have evaluated a full and reduced weather data input with the machine learning models (see Section 3.3.1) and also explored the impacts of a limited dynamic range in Section 3.3.4. Also, while it is possible to explain the method by which each nonparametric model was fit to the data, it can be difficult to interpret the actual fitted model (e.g., there is no regression equation with coefficients for inference). For example, in the case of the RF model, the final model is the average of many decision tree models, and the average of the rules from the individual trees is incomprehensible. Therefore, we will rely on variable importance (Section 4.1.1) and partial dependence plots (Section 4.1.2) to analyze how these nonparametric models fit to the data. This is key to determining whether a nonparametric model has fit on an unusual pattern within the data (known as "overfitting".)

3.3 Methods

3.3.1 Full and Reduced OPM Data Inputs

Increased precipitation (Figures 6 and 9) in conjunction with attempting to address ML shortcomings (Section 3.2) are the reasons for employing full and reduced OPM data inputs. The full data input included both precipitation and wind variables, while the reduced model included only wind variables (Table 2). The combination of three machine learning models, with two data inputs (full and reduced) and six weather simulations (five WRF simulations and their calculated ensemble mean) yielded 36 scenarios each to be evaluated for Current and Future Sandy.

3.3.2 Outage Prediction Model for Current Sandy

We first establish that the WRF simulations could be used to predict Current Sandy outages with each of the three machine learning models. We refer to “model training” as using the tuned models as “in-sample” prediction on the Current Sandy data. The model training results were not included in this paper as they are not a true measure of model performance. Instead we present results from a leave-one-observation-out cross-validation (LOOCV) using the tuned models on Current Sandy to demonstrate their performance (we refer to this as “model validation”, “observations” are defined as 2-km grid cells). The following error metrics were calculated for each simulation and model across all grid cells: Pearson correlation (“correlation”, “r”) between actual and predicted outages per grid cell, mean absolute error (“MAE”) of predicted outages per grid cell, root-mean-square error (RMSE) of predicted outages per grid cell, and the sum of predicted outages over the service territory (to show estimation error of the predicted outages). Description of the calculation of these error metrics is provided in the Appendix.

3.3.3 Outage Prediction Model for Future Sandy

Once we established through model validation that the WRF simulations could predict outages for Current Sandy, we then performed an independent test to evaluate how the models would predict outages from a corresponding Future Sandy simulation. We refer to “model testing” as using our trained and validated models from Current Sandy that are used to predict Future Sandy outages. With the knowledge that some weather simulations may be inherently biased (Section 2.3), we assumed that any bias was consistent between the Current and Future Sandy simulations and absorbed these biases into our framework (Figure 10) by fitting pairwise outage models to account for the chosen configurations (i.e., an individual Current Sandy simulation from Table 1 is used to predict the corresponding Future Sandy outages), along with the ensemble mean.

In summary, the Current Sandy WRF simulation (joined with actual Current Sandy outages from Eversource) will be used for training and validated using LOOCV; and the Future Sandy simulation will be treated as an independent model test (e.g. holdout sample) of the trained model, respectively. In Section 4, we provide discussion on the validity of the predictions by examining the magnitude and distribution of predicted outages related to the input weather data.

3.3.4 Proof-of-Concept Results from Our Previous Research

As mentioned, in our previous work we have shown how storms of different types and magnitudes could be used to predict outages during each storm (Wanik et al. 2015; He et al. 2016). However, a technique that was not previously demonstrated was the use of a single hurricane to predict outages from another hurricane. To build confidence in our methodology, we used Hurricane Irene (2011) to train the outage prediction models (using BART, BT and RF), and used the trained models with full and reduced data inputs (see Section 3.3.1) to predict outages from Hurricane Sandy (2012) as an independent holdout, and vice-versa. These storms had similar storm outage totals despite differences in track and magnitude of wind and precipitation (Figure A.1 in the Appendix); Sandy had a more extreme distribution of wind-related variables, Irene had higher total accumulated precipitation over the Eversource-CT service territory.

The results from this proof-of-concept can be found in Table 5, and show that each ML model we investigated (BART, BT and RF) was able to predict the outages for each hurricane between -26% and +28% of the actual total outages for the full data input for both storms. However, using the reduced data input resulted in -25% to 2% of the actual total outages for Sandy, while Irene was overestimated 18% to 56%.

In addition to using a single hurricane to predict another hurricane's power outages, we also investigated whether an OPM trained on a large number of extratropical storms (n=76 storms) along with one hurricane could be used to improve the predictions for the other hurricane. For context, an extratropical cyclone is an asymmetric cyclone that usually occurs at the mid-latitudes, due to temperature and/or humidity gradients and wind shear. Their main characteristic is the presence of frontal systems slowly rotating counterclockwise (in the Northern Hemisphere) around the cyclone center. Their impact on the territory is usually manifested with long duration winds, gusts and precipitation, and outages ranged from 20 to 4,000 outages per storm (much less than the >15,000 outages in Irene and Sandy). In comparison, tropical cyclone have instead a symmetric structure, typically with an eyewall near the center, are not characterized by frontal structures. The results from this exercise were comparatively worse, with outage predictions typically underestimated by greater than 50% (Table A.1 in Appendix).

Given these additional proof-of-concept results, we note the uncertainty that can arise in the Future Sandy predictions when the forecasted weather has a different range than the historical storms. As shown in Figure A.1, Sandy and Irene were more similar to each other with respect to maximum gust and wind than to the extratropical storms, much like Current and Future Sandy (Figures 7 and 8). Therefore, we will proceed assuming that Current Sandy can be used to predict Future Sandy impacts, with the knowledge that these predictions may be underestimated given the dynamic range of the weather data input.

4. Results and Discussion

We will now show that although each nonparametric model was able to represent Current Sandy outage impacts for each WRF simulation (Section 4.1), there was a divergence in Future Sandy impacts owing to the non-linear response between the explanatory variables and power outages (Section 4.2). We also highlight how the inclusion of precipitation influenced outage prediction model accuracy for Current Sandy, and substantially altered the Future Sandy predictions. Note: from now on we will often refer to variable names as they appear in Table 1.

4.1 Outage Predictions for Current Sandy Scenarios (Model Validation)

BT and BART accurately predicted Current Sandy while the RF model tended to underestimate the outages and had poorer error metrics (Table 6). The BART and BT models had similar performance for both the full and reduced data inputs, with high correlation values between actual and predicted outages (0.85 - 0.87), low RMSE (4.58 - 4.77) and low MAE (2.38 – 2.5) per 2-km grid cell. In contrast, the RF model had comparatively lower correlation (0.54 – 0.8), higher RMSE (6.88-7.98) and higher MAE (3.3 – 4.07) values than the BART and BT models. Interestingly, the models calibrated on the full data input resulted in little change of MAE per grid cell (e.g., up to 1.7% improved MAE, or 3.3% worsened MAE) across all WRF simulations (Table 7).

The model validation results (Table 6) show that the BT and BART models were superior at predicting Current Sandy outages across all five individual weather simulations and the ensemble mean (e.g., high correlation, low error metrics). These low LOOCV error metrics provide confidence in the Outage Prediction Model and suggest that the Future Sandy outage predictions from these ML models will also be reliable. As previously discussed (Section 2), temporal lags between simulation and observation did not affect the outage model performance as the dependency was removed at the post-processing stage by converting the time series into variables representing the storm peak and severity (Table 2). Additionally, it is worth noting the MAE values for the BT and BART validation were improved compared to results from our previous studies on Current Sandy (Wanik et al. 2015; He et al. 2016) and comparable to others who also conducted hurricane outage modeling studies (Han et al. 2009a; Han et al. 2009b; Nateghi et al. 2014; McRoberts et al. 2017). However, comparison to these studies should be done with caution as each study referenced uses a different storm, outage data, geographic regions, aggregations and spatial resolutions. Given that RF consistently underestimated the storm total outages for Current Sandy, we expected an underestimation of Future Sandy predictions relative to the BT and BART models which more accurately captured the storm total outages (Table 6). However, this did not occur and we will motivate the Future Sandy results (Section 4.2) by analyzing the Current Sandy variable importance and partial dependence plots in the next subsections.

4.1.1 Variable Importance for Current Sandy

Variable importance refers to measuring the contribution of each variable in a ML model, and each ML model's corresponding R package had its own method for measuring variable importance. We will now provide high-level detail on the variable importance calculations for each ML model, the reader may refer to the R package documentation cited in Section 3.2 for a thorough description of how variable importance was calculated in the BART (e.g., inclusion proportion), BT (e.g. relative influence) and RF models (e.g. inclusion node purity). Generally speaking, the higher the variable importance, the more influential a variable will be in determining the predicted response variable. In the BART model, variable importance was the inclusion proportion for any given predictor, the proportion of times that variable is chosen as a splitting rule out of all splitting rules among the posterior draws of the sum-of-trees model. The importance score for a variable in the RF model was calculated by measuring the out-of-bag forecasting accuracy that occurs from shuffling the values for a particular predictor and dropping the out-of-bag observations down each tree. In the BT model, the reduction in the loss function attributed to each variable at each split was tabulated and the sum returned, which was then summed over each boosting iteration.

Though not shown here, there was moderate positive correlation between the count of assets and actual outages per grid cell during Current Sandy, and the count of assets was the most important variable across all combinations of ML model and WRF simulation evaluated. To facilitate comparison of variable importance, we computed the relative variable importance for each ML model and WRF simulation by normalizing variable importance values by the largest non-asset variable (as the assets had importance values that were generally double the next most important variable). Hence, a value of “100” means that this variable was the most important variable in the WRF simulation and ML model, and the importance of all other variables were scaled by this quantity excluding assets.

The variable importance of each ML model is presented in Figure 11 for the full data input (wind and rain variables), and the values are color-coded to help the reader discern which variables were most important (e.g., darker colors represent more important variables). One will notice that BART and RF have much more coloring than BT, indicating that many more variables had an impact closer to the impact of the assets.

The BART models had high variable importance for land use variables (PercConif, PercDecid, PercDev) followed by wind and precipitation variables. Similar to BART, the RF models were influenced from a comparatively larger subset of explanatory variables than BT. In comparison, the BT model had many variables at 0, which suggests only a subset of variables were used for prediction.

4.1.2 Partial Dependence for Current Sandy

Partial dependence plots were created for each of the Current Sandy simulations. Each plot visualizes how an explanatory variable of interest influences the response variable after accounting for all other variables. The X axis represents the explanatory variable of interest, and the Y axis shows the predicted outages per grid cell with all other variables at their mean. Note that the Y axis will change between ML models. A positive, increasing trend on each subpanel represents increased predicted outages and vice-versa. A flat line represents no change in the predicted Y values for given X values.

We present three groups of partial dependence plots that correspond to a subset of the most important variables listed in Figure 11. More specifically, Figure 12 contains partial dependence plots of geographic and land use variables, Figure 13 contains wind wind-related variables, and Figure 14 contains precipitation-related variables. Each figure is grouped by ML model, and each subpanel contains six lines which correspond to the WRF simulations (colors correspond to Figure 5). For brevity, we will focus on the most interesting observed patterns. Note that there is not the same amount of data in each section of the X-axis, and patterns observed at the extreme values of the X axis may be influenced by few data points not representative of the entire calibration data (i.e., see “Assets” for the BART model in Figure 12).

The Assets were not only most important in each ML model and WRF simulation, they resulted in the highest predicted outages (up to 25 outages per grid cell), holding all other variables held at their means (Figure 12). Similarly, all ML models and WRF simulations predicted higher outages for increased PercDecid. The wind-related variables in Figure 13 show a trade-off between calculate mean and maximum variables across ML models - variables of the same group (i.e., wind-related) may show an increase for one variable, and may show flat lines for other correlated variables. The presence of a flat line does not necessarily indicate that the variable was not "important" (i.e., see MEANWind10m for BART, Figure 13, and the variable importance was similar to the other wind-related variables). Within the RF model for wind-related variables (Figure 13), we see that most WRF simulations show a positive trend except the NOTCFLX and WDM6 simulations, which show a negative trend for MAXGust. For GODDARD and MORRIS, the MAXGust variable shows a large positive trend in BT and RF – and this agrees well with the variable importance listed in Figure 11, which confirms they are among the most important variables. The same is true for CNTRL and ENS simulations within the BT and RF models for MAXWind10m and MEANGust. The precipitation-related variables are shown in Figure 14, where most WRF simulations show a positive trend for TotPrec except for the ENS simulation. Within the BT and RF models, the TotPrec had a variable importance that was greatly less than the value of the most important wind-related variables, but the variable importance was similar to wind-related variables in the BART model. There was a mix of trends across all other ML models and WRF simulations for MAXPreRate and MEANPreRate.

4.2 Outage Predictions for Future Sandy Scenarios (Model Test)

Each Future Sandy WRF simulation was treated as an equally likely scenario, and each exhibited differences in the landfall location and magnitude of wind and precipitation within the

Eversource service territory (Figure 9). As previously mentioned, total precipitation increased drastically in some simulations so we compared models with full and reduced data inputs. Each Future Sandy simulation evaluated had differing predicted outage counts, but there were some consistent trends (Tables 8 and 9). The vast majority of Future Sandy scenarios evaluated show higher outages for Future Sandy, except for the following combinations: GODDARD (BT full, BT reduced, BART reduced) and MORRIS (BART full) simulations. Generally, the full model resulted in higher predicted outages than the reduced model, and these full model predictions are displayed in Figures 15 and 16 by machine learning model and simulation. Note how the change in outages were most pronounced in areas with the highest population density (Figure 1), which is inherently related to the amount of electric grid infrastructure.

4.2.1 Comparison between Full and Reduced Data Inputs on Future Sandy Outages

The change in Future Sandy predicted outages varied between the full and reduced data inputs depending on which WRF simulation and machine learning model was considered (Table 9). The BART model predicted -30% to +31% storm total outages for Future Sandy across the five individual WRF simulations when precipitation variables were included, while the BT model predicted -12% to +20% of total outages, and RF predicted +11% to +53% (Table 9). It was interesting to see that the full data input resulted in consistently increased outages over the reduced data input for the RF model (Table 9), even though the RF model had a slightly less accurate calibration by including precipitation variables (Table 7). In comparison BART and BT calibrations for Current Sandy were slightly improved (e.g., lower LOOCV error metrics) by including precipitation variables, and resulted in increased outage predictions for WDM6 and ENS, while all other WRF simulations had no discernable patterns.

Despite underestimating Current Sandy outages (Table 6), RF predicted the most outages from Future Sandy for both the arithmetic average of the five simulations (97% full; 74% reduced) and the ensemble mean (116% full; 75% reduced; Table 8). Figure 17 shows the quantile-quantile (QQ) plot relating the actual Current Sandy outages to the Future Sandy predicted outages for the full data input for all WRF simulations. Generally, RF had the highest change in outages, followed by BART and lastly BT. Decreases below the 45 degree line for the 95th percentile in the QQ plot shows that ML models did not merely predict the extremely large (and rare) values from the distribution of Current Sandy outages.

4.2.2 Influence of Storm Track on Future Sandy Outages

Three of the six Future Sandy tracks made landfall in the center of Long Island, New York. These include the GODDARD, CNTRL and ENS simulations. The WDM6 simulation made landfall in New Jersey, farther south than all other WRF simulations. The MORRIS and NOTCFLX models made landfall in eastern Long Island, New York and the centers of these tracks made land fall over southwestern Connecticut (Figures 4 and 5). From track alone, we would have expected MORRIS and NOTCFLX to have the highest outages, and WDM6 to have the lowest outages for Future Sandy, yet this did not occur. Further, though not shown here, we found that there was little correlation between the change in latitude or longitude at which a storm made landfall and the change in predicted outages, which supports that it is wind and precipitation magnitude and not track that influences power outages.

4.2.3 Influence of Storm Magnitude on Future Sandy Outages

We now focus our discussion on the behavior of machine learning models that used the full data input (both wind and precipitation variables), and how they used changes in wind and precipitation magnitude to predict Future Sandy outages. To support this analysis, the reader can check Figure 1 for a labeled map of Connecticut counties. Much of our analysis will focus on the most populated county in the territory, Fairfield County (southwest Connecticut, population of ~945,000 residents), which also had the most outages of any other county in the service territory during Current Sandy (Figure 1).

Qualitatively, one can compare the colors on Figure 9 (wind and precipitation magnitude changes) and Figure 16 (outage magnitude changes) to see the changes between Current and Future Sandy. To quantify this relationship, the Spearman rank correlation coefficient (ρ) was computed between changes in wind and precipitation magnitude per grid cell and the change in predicted outages from Current Sandy to Future Sandy for the full and reduced data input (Figure 18), and select results were presented for the service territory and Fairfield County (note: service territory results include grid cells from Fairfield County). Spearman's rank correlation coefficients that are close to one have a strong positive relationship, values close to 0 have no relationship, and values close to -1 have a strong negative relationship.

There were minor differences between Spearman correlations in the full and reduced data set (Figure 18). This supports that inclusion of precipitation-related variables did not substantially alter the Future Sandy outage predictions despite the increased accumulated precipitation (Tables 8 and 9). Also worth noting is that the correlations within Fairfield County were generally stronger than the entire Eversource-Connecticut service territory, which we suspect may be related to the vast amount of electric infrastructure present compared to other parts of the territory. The BART and BT models generally had weak to moderate positive correlation between changes in wind-related variables and changes in outage magnitude for Fairfield County (up to ρ=0.63), and the Eversource service territory (up to ρ=0.38). In comparison, the RF model generally had a negative correlation for wind-related variables (except for WDM6) and moderate positive correlation for total precipitation. This suggests that most outage changes within RF were driven by precipitation while BART and BT were driven by wind-related variables.

As mentioned, the correlations listed in Figure 18 can also be verified with visual inspection between Figure 9 and 16. The decreases in gust and wind in the southwest (Fairfield County) during the CTNRL, GODDARD, MORRIS and ENS simulations appear to match the corresponding outage decreases in BART and BT, and not for the RF model. The GODDARD simulation had nearly unchanged winds in Hartford County and decreased winds in Fairfield County, which may explain why the GODDARD simulation generally had the lowest predicted outage impacts for Future Sandy across ML models. The MORRIS simulation had increased winds and gusts in central and coastal Connecticut, with unchanged or decreased gusts in the southwest and northwest. MORRIS also had a region of increased precipitation in eastern Connecticut. The BART model correspondingly had decreased outages in both southwest (Fairfield County) and eastern Connecticut (Windham County and Fairfield County), resulting in this Future Sandy simulation with the full data input having less outages than Current Sandy input (14,595 outages). Interestingly, the reduced data input for BART gave increased outages (20,735 outages), which may indicate that BART with the full data input was overfit on the precipitation data. Another example of potential overfitting on the precipitation data was the BART and BT models which predicted decreased outages in Fairfield County for the NOTCFLX simulation, despite this region having most accumulated precipitation with unchanged gusts and winds. The WDM6 simulation had lowest increases in precipitation and large increases in gust and wind across Connecticut. There was moderate positive correlation between gust and predicted outages (ρ >0.5) in Fairfield County, and weak correlations for wind. The ML models had consistently higher outage predictions with the full data input for WDM6, even though the increases in accumulated precipitation were low compared to other WRF simulations.

4.2.4 Comparison of Machine Learning Models

While the BT model had excellent calibration metrics for Current Sandy, it usually gave the lowest outages for Future Sandy. However, BT had stronger correlation between changes in wind, gust, and changes in predicted outages than RF (see Section 4.2.3). We suspect the reason for divergence between BART and BT for Future Sandy was caused by the BT model being more influenced by the assets per grid cell than the weather variables; hence increases in weather variable severity across the simulations did not result in increased outages. In comparison, the RF model uses all explanatory variables in the input data while the BT and BART models only use variables that improve model accuracy (see Section 3.2 for details). For the same five weather simulations, BT had the smallest average increased outages, while RF and BART had comparatively larger increased outages (Table 8). The partial dependence plots from Current Sandy showed that increased assets per grid cell led to increased outage predictions across all ML models, but it is worth noting the dynamic range of the Y axis in the partial dependence plots for the BT model were typically much smaller than the BART and RF models, further suggesting it could have been overfit on the assets. For the sake of this paper, we treat all nonparametric models as equally valid, highlighting the model dependence of results.

4.3 Limitations

Changes in the severity of outages from future tropical cyclones can be caused by several factors, but here, we isolated the meteorological component. Those are: (i) the future storm was more intense, and (ii) it made landfall much closer to the study area. As mentioned earlier, we have accounted for both changes in track and intensity by making direct use of Lackmann (2015)

WRF simulations. Differentiating between those two mechanisms requires shifting the track of Current Sandy to match the future one, while keeping the same intensity. This scenario is challenging to achieve due to changes in storm-relative coastline orientation that would not allow us to simply translate future scenario model wind speeds relative to the present cyclone track moves. Therefore this scenario would require further model simulations, which is beyond the scope of the current study.

Our ability to generalize these results is limited by the use of a single case and we are aware that for this particular storm, the track changed in a way that helped to potentially maximize winds and precipitation of Future Sandy across the Eversource service territory. Simulations of many cases, with various tracks and intensities, are necessary to reach more general conclusions about how future tropical cyclone impacts could change with warming. How the frequency and intensity of tropical cyclones will change with climate warming remains an area of active research in the atmospheric sciences community. While we do not believe that we have modeled the worst case scenario for utilities to prepare for, the case study presented in this paper serves as a proof-of-concept method that can be readily implemented when weather data for many additional cases of future tropical cyclones becomes available.

There are many other factors that may play a role in modifying how an electrical distribution system responds to adverse weather. Utilities invest in structural and electrical hardening initiatives which may increase resilience to extreme weather events – depending on the level and type of investment, the grid may respond differently to severe weather (Kuntz et al. 2002). The level of foliage (Ennos 1999; James et al. 2014), which is a function of the day a storm would hit in the future (Fahey 2016; Carter et al. 2017) would also alter the relationship between wind, trees and resultant outages. On a broader level, tree species mixes may also change as a function of altered temperature and precipitation (Rustad et al. 2012). Further, if a utility were to alter the tree conditions such that the trees were less prone to impact through vegetation management activities, future outages may be limited (Guikema et al. 2006; Wanik et al. 2017). However, the presence of invasive species, such as Emerald ash borer (Poland and McCullough 2006), will weaken roadside trees and forests and may lead to greater outage counts in select regions. The electric distribution network typically follows population by necessity, thereby increasing the exposure of the network and contributing to potential outages by virtue of simply having more infrastructure where the system is overhead as population increases (Larsen et al. 2016). Should future population growth occur in cities (Heath 2001; Dawson 2007) rather than rural communities, infrastructure exposure and associated risk may be comparatively lowered as distribution infrastructure tends to be underground in urban areas.

5. Conclusions

This case study was based on a scientific question about the change in severity of power outages if a storm similar to Hurricane Sandy was to impact Connecticut in the future, taking place with warmer atmospheric conditions. We have presented a case study of how we would expect future outages to occur under different future Hurricane Sandy scenarios given end-of-century atmospheric thermodynamic conditions informed by numerical weather prediction simulations from a recent published work (Lackmann 2015). We acknowledge that changes in both track and intensity affect changes in outage impacts. We did not attempt to separate these effects as our purpose here was to provide a case study of potential power outages owing to a stronger storm and altered track induced by future climate conditions and to illustrate a technique that could be used with a more complete set of future tropical cyclone scenarios. For example, applying this technique to multi-season simulations of future climate scenario (or historic) hurricane tracks and severities could provide a more thorough treatment of the problem.

These simulations between Current Sandy and Future Sandy were shown to increase power outages in Connecticut by an amount ranging between 42% (reduced data input) and 64% (full data input) using the ensemble mean of each atmospheric variable from the five WRF simulations to run the outage models, and 55% (reduced data input) and 64% (full data input) using the arithmetic average of the five ensemble member outage simulations (Table 8).

To limit the weather-related outages, many utilities are investing in multimillion-dollar grid resilience projects to address substation flooding, vegetation management, and pole integrity improvements (Consolidated Edison 2013; Eversource Energy 2013; Public Service Enterprise Group 2016). The study did not account for electric grid hardening activities, which would likely moderate future storm impacts. Storm surge and inland flooding, while not evaluated in our model, are also be expected to contribute to increased outages in future hurricane scenarios, due to Future Sandy's stronger winds and closer track to the Eversource service territory relative to Current Sandy generate a higher storm surge. Soil moisture may increase power outages in both drought and saturated soil conditions by making tree branches more likely to break (Meir et al. 2015) or making tree roots more likely to be uprooted (James et al. 2014; Vogel 1996). Past research has explored the use of soil moisture data for improving the accuracy of hurricane outage prediction models (Han et al. 2009a; Han et al. 2009b), and were recently demonstrated to be useful for predicting outages during Hurricane Matthew (Gorder 2016).

Although we have only analyzed impacts on the electric distribution network by tree-caused damages, there are many other types of infrastructures that would likely be informed by an analysis of this type (i.e., water supply, wastewater, and telecommunications). A future extension of this analysis that includes simulations of many tropical cyclones (not just Hurricane Sandy), with various tracks and intensities, will allow us to reach more general conclusions about how future tropical cyclone impacts could change in a warming climate.

Acknowledgements

We gratefully acknowledge the support of Eversource and the Eversource Energy Center at the University of Connecticut, which provided funding and data for this research. WRF is developed and maintained by the National Center for Atmospheric Research, funded by NSF. Precipitation NCEP/EMC 4KM Gridded Stage IV data is provided by NCAR/EOL under sponsorship of the National Science Foundation http://data.eol.ucar.edu/. We acknowledge the ECMWF for provision of the ERA-Interim Reanalysis, and NCAR for making the ERA-Interim available.

Appendixes

Error Metrics

The statistical metrics used in the model evaluation analyses are presented below. The modelled variable (i.e., wind, precipitation, outages etc.) is represented by Y, the observed variable by X and N is the total number of data points used in the calculations.

- Root Mean Square Error (RMSE):

$$RMSE = \sqrt{\frac{1}{N}\sum_{N}(Y-X)^2}$$

- Mean Bias (MB):

$$MB = \frac{1}{N}\sum_{N}(Y-X)$$

- Mean Absolute Error (MAE):

$$MAE = \frac{1}{N}\sum_{N}|Y-X|$$

- Pearson Correlation (r):

$$r = \frac{\sum XY - \frac{(\sum X)(\sum Y)}{N}}{\sqrt{\left(\sum X^2 - \frac{(\sum X)^2}{n}\right)\left(\sum Y - \frac{(\sum Y)^2}{n}\right)}}$$

Proof-of-Concept Using Prior Research Data

The following plot and tables support analysis in Section 3.3.4.

[Table A.1 and Figure A.1 to be placed here by AMS technical editors]

References

Abraham K., J. Stock J and B. Stevenson, 2013: Economic benefits of increasing electric grid resilience to weather outages. Accessed 1 May 2017. [Available online at https://energy.gov/sites/prod/files/2013/08/f2/Grid%20Resiliency%20Report_FINAL.pdf]

Blake, E. S., T. B. Kimberlain, R. J. Berg, J. P. Cangialosi, and J. L. Beven, 2012: Tropical Cyclone Report: Hurricane Sandy (AL182012) 22-29 October 2012. 157 pp. [Available online at: http://www.nhc.noaa.gov/data/tcr/AL182012_Sandy.pdf]

Breiman, L., 2001: Random forests. Accessed 1 October 2014. [Available online at: https://www.stat.berkeley.edu/~breiman/randomforest2001.pdf]

Campbell, R.J., 2013: Weather-related power outages and electric system resiliency. Electricity Reliability in the United States, 18 pp. [Available online at https://fas.org/sgp/crs/misc/R42696.pdf]

Caron, M. A., A. H. House, and J. W. Betkowski, 2013: PURA investigates into the performance of Connecticut's electric distribution companies and gas companies in restoring service following Storm Sandy. CT Public Utilities Regulatory Authority Docket No. 12-11-07. 64 pp. [Available online at http://www.ct.gov/pura/lib/pura/pressreleases/2012/2012march26nudraftdecision.pdf]

Carter, J. M., M. E. Orive, L. M. Gerhart, J. H. Stern, R. M. Marchin, J. Nagel, and J. K. Ward, 2017: Warmest extreme year in U.S. history alters thermal requirements for tree phenology. *Oecologia*, **183,** 1197-1210, doi:10.1007/s00442-017-3838-z

Chipman, H. A., George, E. I., McCulloch, R. E., 2012: BART: Bayesian additive regression trees. *Ann Appl Statistics*, **4,** 266-298, doi: 10.1214/09-AOAS285

Cole T., Wanik D. W., Molthan A., Roman M., Griffin E., 2017: Synergistic Use of Nighttime Satellite Data, Electric Utility Infrastructure, and Ambient Population to Improve Power Outage Detections in Urban Areas. *Remote Sens,* **9**, 286, doi:10.3390/rs9030286

Consolidated Edison, 2013: Post Sandy Enhancement Plan. Accessed 1 May 2017. [Available online at https://www.coned.com/-/media/files/coned/documents/services-outages/post_sandy_enhancement_plan.pdf]

Dawson, R., 2007: Re-engineering cities: A framework for adaptation to global change. *Philos Trans R Soc A Math Phys Eng Sci*., **365,** 3085-3098, doi: 10.1098/rsta.2007.0008

Dee, D.P., and Coauthors, 2011: The ERA-Interim reanalysis: configuration and performance of the data assimilation system. *Q J R Meteorol Soc*., **137,** 553–597, doi: 10.1002/qj.828

Emanuel, K., R. Sundararajan, J. Williams, 2008: Hurricanes and global warming: Results from downscaling IPCC AR4 simulations. *Bull. Am. Meteorol. Soc.*, **89**, 347–367, doi: 10.1175/BAMS-89-3-347.

Emanuel, K., 2005: Increasing destructiveness of tropical cyclones over the past 30 years. *Nature*, **436,** 686-689, doi:10.1038/nature03906

Ennos, A. R., 1999: The aerodynamics and hydrodynamics of plants. *J. Exp. Biol.,* **202,** 3281-3284

Eversource Energy, 2013: Improving Electric Reliability: Eversource's System Resiliency Program. Accessed 1 May 2017. [Available online at https://www.eversource.com/Content/docs/default-source/ct---pdfs/system-reliability-project.pdf?sfvrsn=0]

Fahey, R. T., 2016: Variation in responsiveness of woody plant leaf out phenology to anomalous spring onset. *Ecosphere*, **7,** doi: 10.1002/ecs2.1209

Fanelli C., P. Fanelli and D. Wolcott, 2013: Hurricane Sandy 2012 Water Level and Meteorological Data Report. 62 pp. [Available online at https://tidesandcurrents.noaa.gov/publications/Hurricane_Sandy_2012_Water_Level_and_Meteorological_Data_Report.pdf]

Foster, D. R., 1988: Species and stand response to catastrophic wind in central New England, USA. *J. Ecol.*, **76,** 135-151, doi: 10.2307/2260458

Friedman, J. H., 2001: Greedy function approximation: A gradient boosting machine. *Ann. Stat.*, **29,** 1189-1232

Gorder, P., 2016: How soil moisture can help predict power outages caused by hurricanes. Accessed 1 May 2017. [Available online at https://phys.org/news/2016-12-soil-moisture-power-outages-hurricanes.html]

Guikema, S. D., R. Nateghi, S. M. Quiring, A. Staid, A. C. Reilly, M. Gao, 2014: Predicting Hurricane Power Outages to Support Storm Response Planning. *IEEE Access*, **2,** 1364-1373, doi: 10.1109/ACCESS.2014.2365716

Guikema, S. D., and R. A. Davidson, 2006: Modelling critical infrastructure reliability with generalized linear (mixed) models. *8th International Conference on Probabilistic Safety Assessment and Management*, New Orleans, LA, PSAM 2006. [Available online at https://experts.umich.edu/en/publications/modelling-critical-infrastructure-reliability-with-generalized-li]

Guikema, S. D., R. A. Davidson and H. Liu, 2006: Statistical models of the effects of tree trimming on power system outages. *IEEE Trans Power Delivery*, **21,** 1549 – 1557, doi: 10.1109/TPWRD.2005.860238

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 Analysis,* **29,** 1443-1453, doi: 10.1111/j.1539-6924.2009.01280.x

Han, S-R, S. D. Guikema, S. M. Quiring, K. Lee, D. 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, doi: 10.1016/j.ress.2008.02.018

He, J., D. W. Wanik, B. M. Hartman, E. N. Anagnostou, M. Astitha, and M. E. B. Frediani, 2016: Nonparametric Tree-Based Predictive Modeling of Storm Outages on an Electric Distribution Network. *Risk Anal.,* **37,** 441-458, doi: 10.1111/risa.12652

Heath, T., 2001: Revitalizing cities: Attitudes toward city-center living in the United Kingdom. *J. Plann. Educ. Res.,* **20,** 464-475, doi: 10.1177/0739456X0102000410

Henry, D. and J. E. Ramirez-Marquez, 2016: On the Impacts of Power Outages during Hurricane Sandy - A Resilience-Based Analysis. *Syst Eng.,* **19,** 59-75, doi: 10.1002/sys.21338

Hill, K.A. and G. M. Lackmann, 2011: The Impact of Future Climate Change on TC Intensity and Structure: A Downscaling Approach. *J. Clim.,* **24,** 4644-4461, doi: 10.1175/2011JCLI3761.1

Hines, P., J. Apt and S. Talukdar, 2008: Trends in the history of large blackouts in the United States. *IEEE Power Energy Society Gen Meeting*, Pittsburgh, PA, IEEE. doi: 10.1109/PES.2008.4596715

James, K. R., G. A. Dahle, J. Grabosky, B. Kane and A. Detter, 2014: Tree biomechanics literature review: Dynamics. *Arboric Urban. For.,* **40,** 1-15

Kapelner, A., and J. Bleich, 2014: bartMachine: A Powerful Tool for Machine Learning. *Journal of Statistical Software*, **70,** 1-40, doi: 10.18637/jss.v070.i04

Knutson, T. R. and Coauthors, 2010: Tropical cyclones and climate change. *Nature Geoscience,* **3,** 157-163, doi: 10.1038/ngeo779

Kuntz, P. A., R. D. Christie and S. S. Venkata, 2002: Optimal vegetation maintenance scheduling of overhead electric power distribution systems. *IEEE Trans Power Delivery*, **17,** 1164-1169, doi: 10.1109/TPWRD.2002.804007

Lackmann, G. M., 2015: Hurricane Sandy before 1900, and after 2100. *Bull. Amer. Meteor. Soc.,* **96,** 547-560, doi: 10.1175/BAMS-D-14-00123.1

Larsen, P. H., K. H. LaCommare, J. H. Eto, and J. L. Sweeney, 2016: Recent trends in power system reliability and implications for evaluating future investments in resiliency. *Energy*, **117,** 29-46, doi: 10.1016/j.energy.2016.10.063

Li, H., L. A. Treinish, and J. Hosking, 2010: A statistical model for risk management of electric outage forecasts. *IBM Journal of Research and Development*, **54,** 1-11, doi: 10.1147/JRD.2010.2044836

Liaw A. and M. Wiener, 2002: Classification and Regression by randomForest. *R News,* **2**, 18-22

Lin, Y., and K. E. Mitchell, 2005: The NCEP stage II/IV hourly precipitation analyses: Development and applications, in: *19th Conf. Hydrology*, San Diego, CA, American

Meteorological Society, USA [Available online at https://ams.confex.com/ams/Annual2005/techprogram/paper_83847.htm]

Liu, H., R. A. Davidson and T. V. Apanasovich, 2008: Spatial generalized linear mixed models of electric power outages due to hurricanes and ice storms. *Reliab Eng Syst Saf.*, **93,** 897-912, doi: 10.1016/j.ress.2007.03.038

Mackinnon, M. J., and N. Glick, 1999: Data mining and knowledge discovery in databases - An overview. *Austral. & New Zealand J. Statist.*, **41,** 255–275, doi: 10.1111/1467-842X.00081

McRoberts, D. B., S. M. Quiring, and S. D. Guikema, 2017: Improving Hurricane Power Outage Prediction Models Through the Inclusion of Local Environmental Factors. *Risk Anal.*, **In Press,** doi: 10.1111/risa.12728

Meehl, G. A, and Coauthors, 2007: The WCRP CMIP3 multimodel dataset: A new era in climatic change research. *Bull. Am. Meteorol. Soc.*, **88,** 1383-1394, doi: 10.1175/BAMS-88-9-1383

Meir P., M. Mencuccini, and R. C. Dewar, 2015: Drought-related tree mortality: Addressing the gaps in understanding and prediction. *New Phytol.*, **207,** 28-33, doi: 10.1111/nph.13382

Mensah, A. F. and L. Duenas-Osorio, 2014: Outage predictions of electric power systems under Hurricane winds by Bayesian networks. *Int Conf Probab Methods Appl Power Syst*, Durham, U.K., PMAPS - Conf Proc, doi: 10.1109/PMAPS.2014.6960677

Nakicenovic N and R. Swart, 2000: Intergovernmental panel on climate change special report on emissions scenarios. [Available at http://www.ipcc.ch/ipccreports/sres/emission/index.php?idp=0]

Nateghi, R., S. D. Guikema and S. M. Quiring, 2014: Power Outage Estimation for Tropical Cyclones: Improved Accuracy with Simpler Models. *Risk Anal.*, **34,** 981-1159, doi: 10.1111/risa.12131

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,** 1897-1907, doi: 10.1111/j.1539-6924.2011.01618.x

National Centers for Environmental Prediction, NOAA, U.S. Department of Commerce, 1997: NCEP ADP global upper air and surface weather observations (PREPBUFR format): research data archive at the National Center for Atmospheric Research, Computational and Information Systems Laboratory. In: Research data archive at the National Center for Atmospheric Research, Computational and Information Systems Laboratory. [Available at http://rda.ucar.edu/datasets/ds337.0/] Accessed 3/1 2015.

Northeast Utilities, 2013: Annual Report. Accessed 1 April 2017. [Available online at https://www.eversource.com/content/docs/default-source/pdfs/2013_nu_annual_report.pdf?sfvrsn=2. Accessed 5/1 2017]

Pielke Jr., R. A., 2007: Future economic damage from tropical cyclones: Sensitivities to societal and climate changes. *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences*, **365,** 2717-2729, doi: 10.1098/rsta.2007.2086

Poland, T.M. and D. G. McCullough, 2006: Emerald ash borer: Invasion of the urban forest and the threat to North America's ash resource. *J. For.*, **104,** 118-124

Public Service Enterprise Group, 2016: Gov. Christie Highlights PSE&G Post-Sandy Investments to Improve Reliability with Visit to Hackensack: $1.2 billion infrastructure upgrades are making New Jersey Energy Strong. Accessed 1 April 2017. [Available online at https://www.pseg.com/info/media/newsreleases/2016/2016-10-28.jsp#.WR8s7evysXM.]

Quiring, S. M., L. Zhu, S. D. Guikema, 2011: Importance of soil and elevation characteristics for modeling hurricane-induced power outages. *Nat Hazards.*, **58,** 365-390, doi: 10.1007/s11069-010-9672-9

Ridgeway, G., 2007: Generalized Boosted Models: A guide to the gbm package. [Available at http://ftp.auckland.ac.nz/software/CRAN/doc/vignettes/gbm/gbm.pdf] Accessed 3/1 2014.

Rustad, L., J. Campbell, J. Dukes, T. Huntington, K. Lambert, J. Mohan, and N. Rodenhouse, 2012: Changing Climate, Changing Forests: The Impacts of Climate Change on Forests of the Northeastern United States and Eastern Canada. United States Department of Agriculture: General Technical Report NRS-99:36

Schär, C., C. Frei, D. Lüthi, and H. C. Davies, 1996: Surrogate climate-change scenarios for regional climate models. *Geophys. Res. Lett.,* **23,** 669-672, doi: 10.1029/96GL00265

Shepherd T. G., 2016: A Common Framework for Approaches to Extreme Event Attribution. *Current Climate Change Reports,* **2**, 28-38, doi: 10.1007/s40641-016-0033-y

Skamarock, W. C., and Coauthors, 2008: A description of the Advanced Research WRF version 3. NCAR Tech. Note NCAR/TN-475+STR, 113 pp, doi: 10.5065/D68S4MVH.

Staid A., S. D. Guikema, R. Nateghi, S. M. Quiring, and M. Z. Gao, 2014: Simulation of tropical cyclone impacts to the U.S. power system under climate change scenarios. *Clim. Change,* **127,** 535-546, doi: 10.1007/s10584-014-1272-3

Trenberth K. E., J. T. Fasullo, and T. G. Shepherd, 2015: Attribution of climate extreme events. *Nat. Clim. Change,* **5,** 725-730, doi: 10.1038/nclimate2657

Vapnik, V. N., 1999: An overview of statistical learning theory. *IEEE Trans Neural Networks,* **10,** 988-999, doi: 10.1109/72.788640

Vogel, S., 1996: Blowing in the wind: Storm-resisting features of the design of trees. *J. Aboric.* **22,** 92-98

Wanik, D. W., E. N. Anagnostou, B. M. Hartman, M. E. B. Frediani and M. Astitha, 2015: Storm outage modeling for an electric distribution network in Northeastern USA. *Natural Hazards,* **79,** 1359-1384, doi: 10.1007/s11069-015-1908-2

Wanik, D. W., J. R. Parent, E. N. Anagnostou, and B. M. Hartman, 2017: Using vegetation management and LiDAR-derived tree height data to improve outage predictions for electric utilities. *Electr Power Syst Res.*, **146,** 236-245, doi: 10.1016/j.epsr.2017.01.039

Yates, D. and Coauthors, 2014: Stormy weather: Assessing climate change hazards to electric power infrastructure: A Sandy case study. *IEEE Power Energ Mag.*, **12,** 66-75, doi: 10.1109/MPE.2014.2331901

List of Figures

line represents the National Hurricane Center (NHC) “best track” for Current Sandy. (b) Zoomed in to highlight storm landfall location.

Figure 6: Cumulative distributions of total accumulated precipitation for Current and Future Sandy simulations in the sub-region of the model domain enclosing the Eversource service territory. Colors correspond to WRF simulations in Figure 3.

Figure 7: Cumulative distributions of maximum gust for Current and Future Sandy simulations in the sub-region of the model domain enclosing the Eversource service territory. Colors correspond to WRF simulations in Figure 3.

Figure 8: Cumulative distributions of maximum wind at 10-m for Current and Future Sandy simulations in the sub-region of the model domain enclosing the Eversource service territory. Colors correspond to WRF simulations in Figure 3.

Figure 9: Changes in select wind and precipitation magnitude from Current to Future Sandy. Positive values indicate an increase in intensity during Future Sandy.

Figure 10: Modeling framework that combines weather, land use and infrastructure into outage predictions for Current Sandy scenarios. Calibrated models were then applied to Future Sandy scenarios.

Figure 11: Relative variable importance for the BART, BT and RF models, with full data input (normalized by highest value in column – does not include assets per grid cell). Darker colors indicate higher relative importance.

Figure 12: Partial dependence plots related to select geographic variables. Y axis represent change in predicted outages per 2-km grid cell. Colors are related to WRF simulations in Figure 3.

Figure 13: Partial dependence plots related to select wind variables. Y axis represent change in predicted outages per 2-km grid cell. Colors are related to WRF simulations in Figure 3.

Figure 14: Partial dependence plots related to select precipitation variables. Y axis represent change in predicted outages per 2-km grid cell. Colors are related to WRF simulations in Figure 3.

Figure 15: Distribution of predicted outages for Future Sandy by simulation and machine learning models for the full model forcing (wind and precipitation variables). Legend matches Current Sandy actual outages in Figure 1.

Figure 16: Change in predicted outages from Current to Future Sandy for the full data input (positive numbers indicate an increase in Future Sandy).

Figure 17: Quantile-quantile plot showing the increase in predicted outages per grid cells for Future Sandy (Y axis) compared to actual Current Sandy outages per grid cell (X axis) for BART, BT and RF models with the full data input. Quantiles represent the 5, 10, 20, 30, 40, 50, 60, 70, 80, 90, and 95th percentiles.

Figure 18: Correlation between increased outages and weather magnitude using Spearman correlation for Fairfield County and the Eversource Connecticut service territory for the full and reduced data input. Red cells indicate positive correlation, blues cells indicate negative correlation, and white cells indicate a lack of correlation. Difference between Spearman

correlations for full and reduced data input are also tabulated in right third (positive values indicate an improvement in the full model).

Figure A.1: Comparison of CDF plots for select weather variables for 76 extratropical storms (occurred between 2005 and 2017), Hurricane Irene (2011), and Hurricane Sandy (2012).

Table 1: Configuration used in the WRF model simulations from Lackmann (2015). The convective parameterization (CP) choices included Kain-Fritsch (KF) and none. The microphysics choices include the WRF single-moment 6-class microphysics scheme (WSM6), the Goddard scheme, the WRF double-moment 6-class microphysics scheme (WDM6), and the Morrison scheme. The Planetary Boundary Layer (PBL) and Tropical Cyclone (TC) flux column includes use of the Yonsei University (YSU) scheme, and all but run NOTCFLX utilized the TC flux correction option. All simulations used vertical motion damping, 50 dry-air sigma model levels, and a model top at 50 hPa. All simulations have the same initialization time: 26 Oct 2012, 0000 UTC. For more detail on individual microphysics schemes, please refer to Skamarock et al. (2008).

Simulation	Grid length (km)	CP scheme by grid	Microphysics	PBL/TC flux
CNTRL	54, 18, 6	KF, KF, none	WSM6	YSU/Yes
GODDARD	54, 18, 6	KF, KF, none	Goddard	YSU/Yes
MORRIS	54, 18, 6	KF, KF, none	Morrison	YSU/Yes
NOTCFLX	54, 18, 6	KF, KF, none	WSM6	YSU/No
WDM6	54, 18, 6	KF, KF, none	WDM6	YSU/Yes

Table 2: Summary of variables used in the models; variables with an asterisk have both mean and maximum values calculated (as described in Section 2). The full and reduced models are as discussed in Section 3.3.1.

Variable	Abbreviation	Description	Type	Units	Full	Reduced
Duration of wind at 10 meters above 5 m s^{-1}	wgt5	Weather	Continuous	hr	X	X
Duration of wind at 10 meters above 9 m s^{-1}	wgt9	Weather	Continuous	hr	X	X
Duration of wind at 10 meters above 13 m s^{-1}	wgt13	Weather	Continuous	hr	X	X
Duration of wind at 10 meters above 18 m s^{-1}	wgt18	Weather	Continuous	hr	X	X
Duration of wind gusts above 13 m s^{-1}	ggt13	Weather	Continuous	hr	X	X
Duration of wind gusts above 18 m s^{-1}	ggt18	Weather	Continuous	hr	X	X
Duration of wind gusts above 22 m s^{-1}	ggt22	Weather	Continuous	hr	X	X
Duration of wind gusts above 27 m s^{-1}	ggt27	Weather	Continuous	hr	X	X
Duration of wind gusts above 35 m s^{-1}	ggt36	Weather	Continuous	hr	X	X
Duration of wind gusts above 44 m s^{-1}	ggt45	Weather	Continuous	hr	X	X
Continuous duration of wind at 10 meters above 5 m s^{-1}	Cowgt5	Weather	Continuous	hr	X	X
Continuous duration of wind at 10 meters above 9 m s^{-1}	Cowgt9	Weather	Continuous	hr	X	X
Continuous duration of wind at 10 meters above 13 m s^{-1}	Cowgt13	Weather	Continuous	hr	X	X
Continuous duration of wind at 10 meters above 18 m s^{-1}	Cowgt18	Weather	Continuous	hr	X	X
Total precipitation	TotPrec	Weather	Continuous	mm	X	
Wind gust*	Gust	Weather	Continuous	m s^{-1}	X	X
Wind at 10-m height*	Wind10m	Weather	Continuous	m s^{-1}	X	X
Soil moisture*	SoilMst	Weather	Continuous	mm mm^{-1}	X	X
Precipitation Rate*	PreRate	Weather	Continuous	mm hr^{-1}	X	
Count of Assets	Assets	Infrastructure	Continuous	count	X	X
Percent Developed	PercDeveloped	Land Cover	Continuous	%	X	X
Percent Coniferous	PercConif	Land Cover	Continuous	%	X	X
Percent Deciduous	PercDecid	Land Cover	Continuous	%	X	X

Table 3: Current Sandy simulation evaluation metrics; the formulas for RMSE, MB and ME are listed in the Appendix. The wind speed observations are taken from METAR airport stations and the precipitation from Stage IV data for Connecticut. Times with missing data were not included in the error metrics. The location of these airport stations are denoted with red circles on Figure 1.

Station ID/Location	Variable	Simulation	RMSE	MB	MAE
KBDL: North/Central CT 41.94° N, 72.68° W	Wind speed at 10-m (m s^{-1})	CNTRL	2.66	0.09	2.16
		GODDARD	3.09	0.13	2.52
		MORRIS	2.38	0.39	2.13
		NOTCFLX	2.83	0.01	2.21
		WDM6	2.86	0.01	2.35
		ENS MEAN	2.61	0.12	2.19
KDXR: Southwestern CT 41.37° N, 73.48° W	Wind speed at 10-m (m s^{-1})	CNTRL	3.84	2.49	3.14
		GODDARD	4.62	2.76	3.64
		MORRIS	3.92	3.14	3.25
		NOTCFLX	3.91	2.57	3.08
		WDM6	4.29	2.67	3.45
		ENS MEAN	3.96	2.73	3.22
KGON: Southeastern CT 41.33° N, 72.045° W	Wind speed at 10-m (m s^{-1})	CNTRL	3.71	2.53	3.00
		GODDARD	4.54	2.98	3.66
		MORRIS	3.75	2.53	3.08
		NOTCFLX	3.75	2.75	3.17
		WDM6	3.68	1.88	2.78
		ENS MEAN	3.61	2.53	2.88
KIJD: Eastern CT 41.74° N, 72.18° W	Wind speed at 10-m (m s^{-1})	CNTRL	3.76	2.53	3.32
		GODDARD	4.30	2.69	3.67
		MORRIS	3.64	2.90	3.14
		NOTCFLX	4.16	3.20	3.72
		WDM6	3.77	2.32	3.20
		ENS MEAN	3.81	2.73	3.37
Averaged over Connecticut	Precipitation (mm)	CNTRL	2.78	0.83	1.60
		GODDARD	2.74	0.53	1.71
		MORRIS	1.94	0.55	1.25
		NOTCFLX	3.48	0.97	2.02
		WDM6	2.97	0.55	1.59
		ENS MEAN	2.62	0.69	1.57

Table 4: Average maximum gust, maximum wind at 10-m and total precipitation for the Current ("Mean(Current)) and Future (Mean(Future)) Sandy runs in the Eversource service territory. The last column (% Increase) represents the percentage increase from Current Sandy to Future Sandy by each simulation for each weather variable.

Variable	Simulation	Unit	Mean(Current)	Mean(Future)	% Increase
MAXGust	CNTRL	m s^{-1}	28.4	31.3	10.3%
MAXGust	GODDARD	m s^{-1}	29	29.8	3.0%
MAXGust	MORRIS	m s^{-1}	28.1	29.9	6.3%
MAXGust	NOTCFLX	m s^{-1}	27.3	29	6.2%
MAXGust	WDM6	m s^{-1}	28.3	31.1	9.9%
MAXGust	ENS MEAN	m s^{-1}	27.7	28.9	4.3%
MAXWind10m	CNTRL	m s^{-1}	15.9	17.8	12.0%
MAXWind10m	GODDARD	m s^{-1}	16.3	17.2	5.5%
MAXWind10m	MORRIS	m s^{-1}	15.6	17.4	11.0%
MAXWind10m	NOTCFLX	m s^{-1}	15.2	16.4	8.0%
MAXWind10m	WDM6	m s^{-1}	15.6	17.6	12.5%
MAXWind10m	ENS MEAN	m s^{-1}	15.4	16.29	5.8%
TotPrec	CNTRL	mm	51.4	103.6	101.6%
TotPrec	GODDARD	mm	42.8	80.1	87.4%
TotPrec	MORRIS	mm	43.8	125.5	186.8%
TotPrec	NOTCFLX	mm	51	116.2	127.6%
TotPrec	WDM6	mm	43.5	69.5	59.7%
TotPrec	ENS MEAN	mm	46.5	98.9	112.7%

Table 5: Proof-of-concept results for predicting Hurricane Irene from Hurricane Sandy, and Hurricane Sandy from Irene using data from Wanik et al. (2015) and He et al. (2016).

		Full Model					**Reduced Model**				
Holdout Sample	**ML**	**r**	**MAE**	**RMSE**	**sum(Pred)**	**PercErr**	**r**	**MAE**	**RMSE**	**sum(Pred)**	**PercErr**
Irene (2011)	BART	0.7	3.59	6.85	11,658	-26%	0.71	4.81	7.15	23,819	56%
	BT	0.66	4.22	6.88	17,831	13%	0.65	4.26	6.94	17,990	18%
	RF	0.69	4.69	7.01	20,234	28%	0.67	5.11	7.25	22,510	48%
Sandy (2012)	BART	0.59	3.75	6	16,823	11%	0.63	3.59	5.72	13,094	-17%
	BT	0.59	3.98	6.12	18,014	18%	0.69	3.38	5.47	11,844	-25%
	RF	0.57	4.02	6.02	18,692	23%	0.65	3.59	5.43	16,185	2%

Table 6: Error metrics for Current Sandy calibration from model validation (using leave-one-out cross-validation). The correlation (r), root-mean-square error (RMSE) and mean absolute error (MAE) were calculated for each grid cell; the sum of predicted outages for Current Sandy (sum(Pred)) was calculated for the entire service territory. The actual outages that occurred during Current Sandy were15,251.

		BART				**BT**				**RF**			
		r	**RMSE**	**MAE**	**sum(Pred)**	**r**	**RMSE**	**MAE**	**sum(Pred)**	**r**	**RMSE**	**MAE**	**sum(Pred)**
FULL MODEL (WIND AND PRECIP)	CNTRL	0.86	4.68	2.44	15,357	0.87	4.58	2.38	14,721	0.56	7.84	3.76	10,856
	GODDARD	0.86	4.75	2.49	15,300	0.86	4.62	2.41	14,862	0.60	7.53	3.58	11,937
	MORRIS	0.86	4.68	2.44	15,386	0.86	4.64	2.42	14,867	0.75	6.88	3.30	11,334
	NOTCFLX	0.85	4.83	2.46	15,213	0.87	4.59	2.38	14,817	0.52	7.98	4.07	12,669
	WDM6	0.86	4.64	2.47	15,305	0.86	4.73	2.42	14,794	0.58	7.77	3.70	10,399
	ENS MEAN	0.85	4.78	2.47	15,428	0.86	4.65	2.41	14,845	0.62	7.67	3.73	10,767
REDUCED MODEL (WIND ONLY)	CNTRL	0.86	4.73	2.48	15,433	0.86	4.71	2.40	14,793	0.62	7.53	3.80	12,933
	GODDARD	0.85	4.77	2.50	15,273	0.86	4.66	2.41	14,769	0.64	7.47	3.51	11,500
	MORRIS	0.86	4.68	2.44	15,357	0.86	4.65	2.40	14,804	0.8	6.88	3.30	11,499
	NOTCFLX	0.86	4.64	2.43	15,372	0.86	4.66	2.42	14,840	0.54	7.76	3.98	13,397
	WDM6	0.86	4.72	2.49	15,456	0.86	4.69	2.42	14,675	0.60	7.69	3.62	10,514
	ENS MEAN	0.86	4.72	2.47	15,242	0.86	4.66	2.41	14,803	0.66	7.65	3.61	10,105

Table 7: Comparison of model improvement in MAE per grid cell when precipitation variables were included in the outage prediction model for Current Sandy (full data input). Positive difference ("Diff") and percent difference ("PercDiff") values indicate an increase in predicted Future Sandy outages for machine learning (ML) models using the full data input (MAE is lower in a more accurate model).

	BART				BT				RF			
WRF	**Reduced**	**Full**	**Diff**	**PercDiff**	**Reduced**	**Full**	**Diff**	**PercDiff**	**Reduced**	**Full**	**Diff**	**PercDiff**
CNTRL	2.48	2.44	0.04	1.6%	2.40	2.38	0.02	0.8%	3.80	3.76	0.04	1.1%
GODDARD	2.50	2.49	0.01	0.4%	2.41	2.41	0.00	0.0%	3.51	3.58	-0.07	-2.0%
MORRIS	2.44	2.44	0.00	0.0%	2.40	2.42	-0.02	-0.8%	3.30	3.30	0.00	0.0%
NOTCFLX	2.43	2.46	-0.03	-1.2%	2.42	2.38	0.04	1.7%	3.98	4.07	-0.09	-2.3%
WDM6	2.49	2.47	0.02	0.8%	2.42	2.42	0.00	0.0%	3.62	3.70	-0.08	-2.2%
ENS	2.47	2.47	0.00	0.0%	2.41	2.41	0.00	0.0%	3.61	3.73	-0.12	-3.3%

Table 8: Actual and relative percentage increase or decrease of predicted Future Sandy outages compared to Current Sandy for each scenario by weather simulation and machine learning model. The AVG column and row represent the arithmetic average of the weather simulations and machine learning models, respectively.

		BART Outages	BT Outages	RF Outages	AVG Outages	BART %Δ	BT %Δ	RF %Δ	AVG %Δ
WIND AND RAIN	CNTRL	36,207	23,185	35,383	31,592	137%	52%	132%	107%
	GODDARD	15,674	13,313	19,164	16,050	3%	-13%	26%	5%
	MORRIS	14,595	20,239	32,650	22,495	-4%	33%	114%	48%
	NOTCFLX	25,800	19,319	35,177	26,765	69%	27%	131%	76%
	WDM6	34,629	23,408	27,431	28,489	127%	53%	80%	87%
	ENS	20,076	22,176	32,988	25,080	32%	45%	116%	64%
WIND ONLY	CNTRL	41,680	19,323	34,658	31,887	173%	27%	127%	109%
	GODDARD	15,216	13,749	18,363	15,776	0%	-10%	20%	3%
	MORRIS	20,735	17,718	27,090	21,848	36%	16%	78%	43%
	NOTCFLX	23,830	22,073	27,996	24,633	56%	45%	84%	62%
	WDM6	26,354	21,687	24,559	24,200	73%	42%	61%	59%
	ENS	17,770	20,376	26,693	21,613	17%	34%	75%	42%

Table 9: Comparison of Future Sandy outage predictions when precipitation variables were included (full data input). Positive difference ("Diff.") and percent difference ("Perc. Diff.") values indicate an increase in predicted Future Sandy outages for ML models using the full data input.

	BART				BT				RF			
WRF	**Reduced**	**Full**	**Diff**	**PercDiff**	**Reduced**	**Full**	**Diff**	**PercDiff**	**Reduced**	**Full**	**Diff**	**PercDiff**
CNTRL	41,680	36,207	-5,473	-13%	19,323	23,185	3,862	20%	31,887	35,383	3,496	11%
GODDARD	15,216	15,674	458	3%	13,749	13,313	-436	-3%	15,776	19,164	3,388	21%
MORRIS	20,735	14,595	-6,140	-30%	17,718	20,239	2,521	14%	21,848	32,650	10,802	49%
NOTCFLX	23,830	25,800	1,970	8%	22,073	19,319	-2,754	-12%	24,633	35,177	10,544	43%
WDM6	26,354	34,629	8,275	31%	21,687	23,408	1,721	8%	24,200	27,431	3,231	13%
ENS	17,770	20,076	2,306	13%	20,376	22,176	1,800	9%	21,613	32,988	11,375	53%

Table A.1: Outage model predictions for Sandy and Irene using BART, BT and RF models with full and reduced data input, trained using 76 extratropical storms and the other tropical storm.

Holdout Sample		ML	r	MAE	RMSE	sum(Pred)	PercErr
Full Model	Irene (2011)	BART	0.25	4.88	8.07	5,411	-65%
		BT	0.24	5.24	8.73	1,164	-92%
		RF	0.31	4.44	7.64	5,933	-62%
	Sandy (2012)	BART	0.4	4.89	9.59	2,370	-84%
		BT	0.36	4.37	9.13	6,668	-56%
		RF	0.54	4.38	9.21	4,922	-67%
Reduced Model	Irene (2011)	BART	0.48	4.91	8.19	2,186	-86%
		BT	0.3	5.21	8.68	1,234	-92%
		RF	0.6	4.82	8.1	2,523	-84%
	Sandy (2012)	BART	0.57	4.03	8.03	5,286	-66%
		BT	0.54	4.15	8.1	10,024	-34%
		RF	0.7	3.93	8.31	6,745	-56%

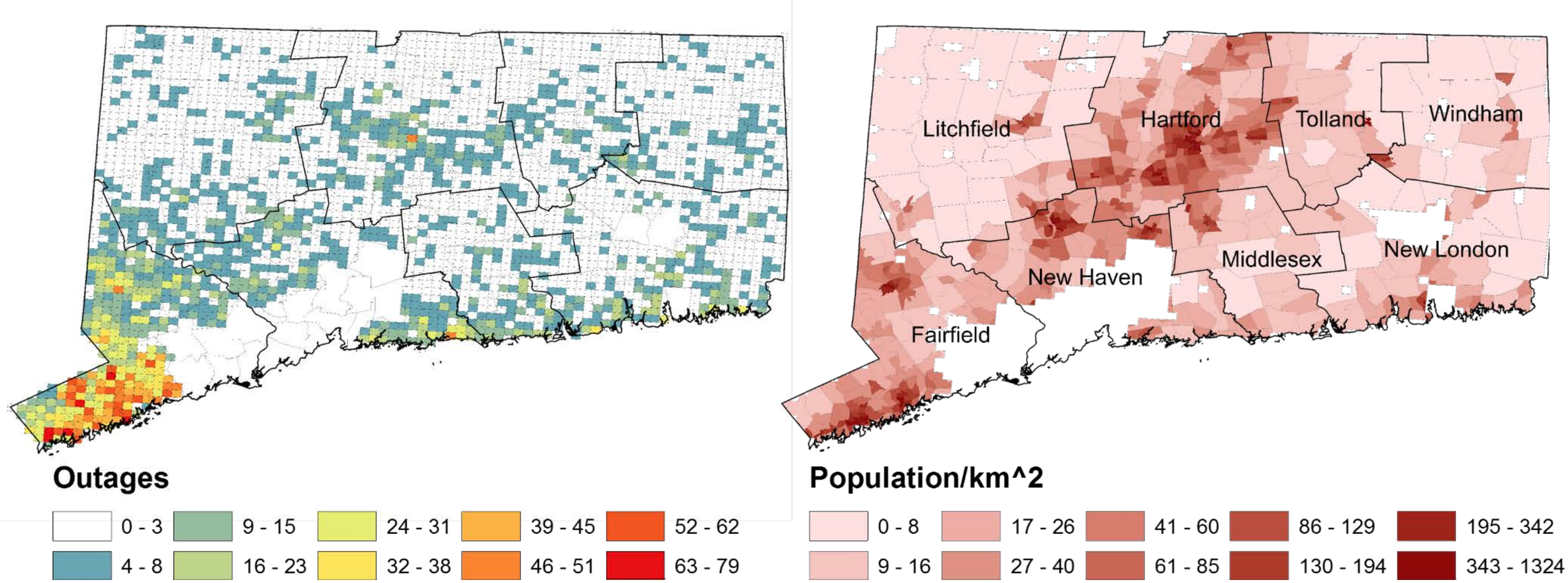

Figure 1: (left) Distribution of actual outages per 2-km grid cells that cover the Eversource service territory during Current Sandy (2012). White areas without grid cells represent regions served by other utility companies. (right) Population density per census tract (source: 2000 Census data.) Counties denoted with thick black lines and labels.

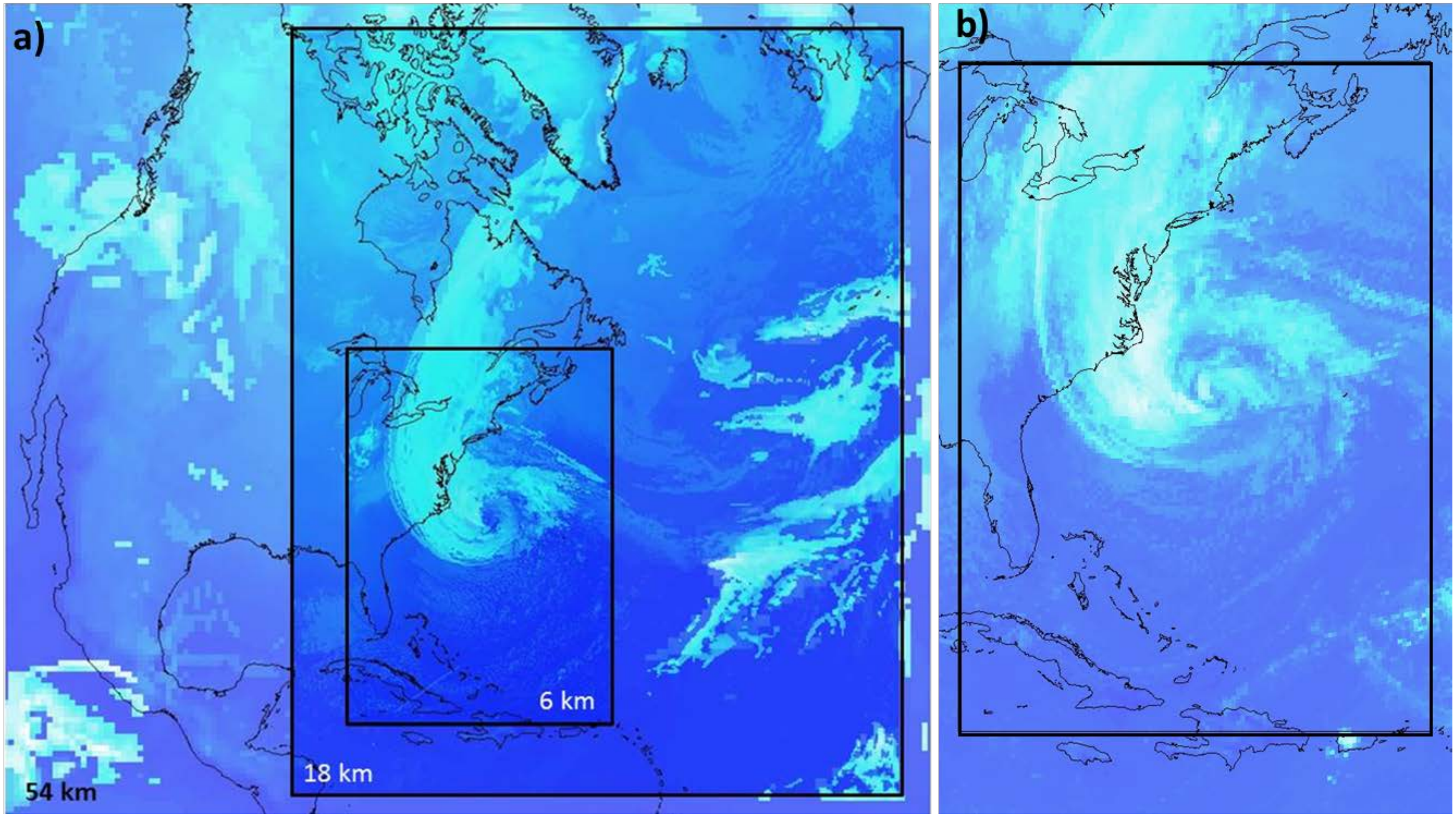

Figure 2: (a) Map denoting the approximate location of the 54, 18 and 6 km WRF domains used for weather simulations. The field displayed is model-simulated brightness temperature at hour 66 of the CNTRL simulation, valid 18 UTC 28 October 2012; with resolution corresponding to the grid length in each domain. (b) Corresponding GOES-13 Infrared image from 18:15 UTC 28 October 2012, black box corresponds to 6km domain in subpanel (a), and color palettes are approximate but not exact.

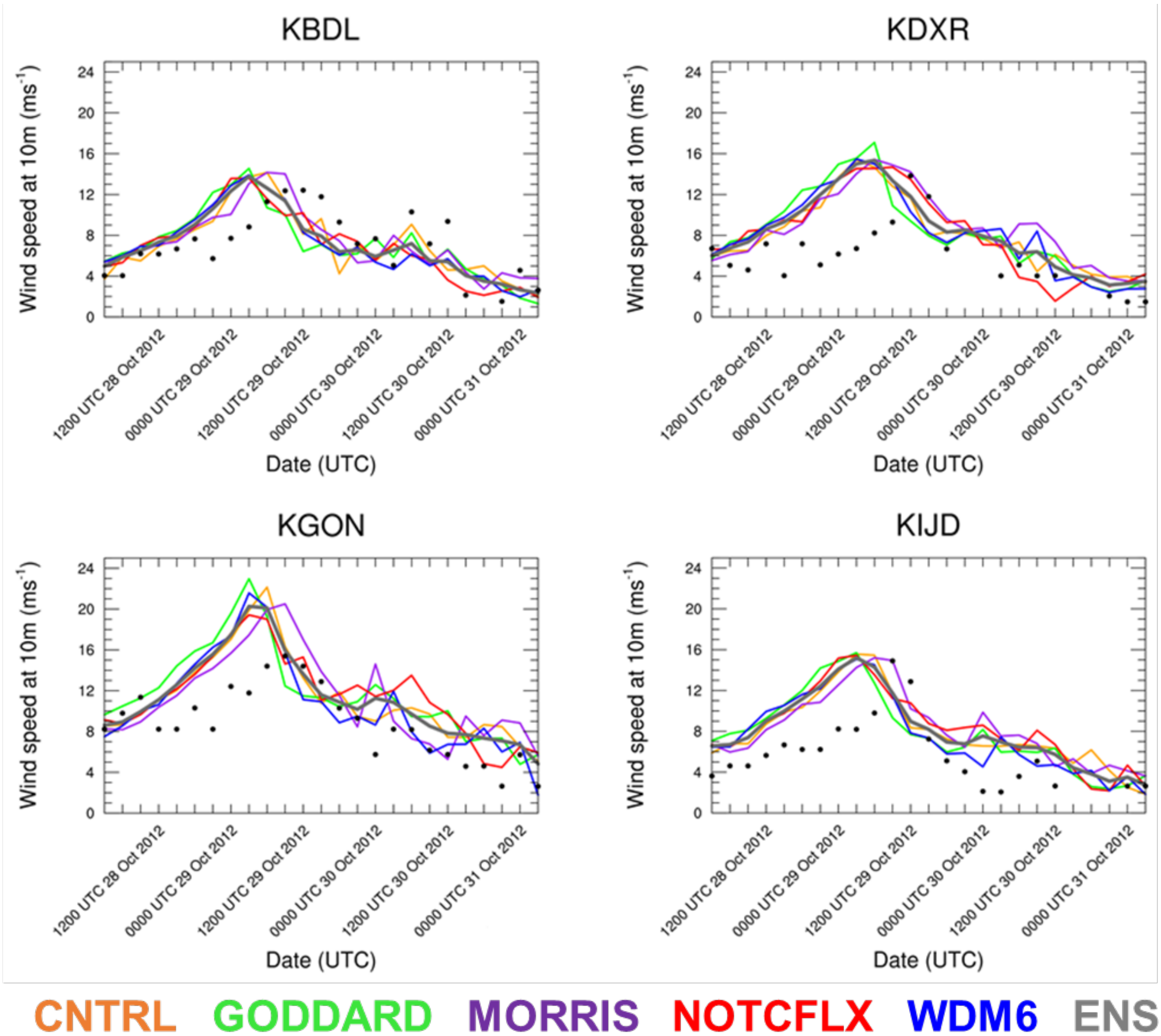

Figure 3: Wind speed at 10-m for the Current Sandy simulations compared to observations (black dots) in four Connecticut stations. Each colored line is an individual WRF simulation. There may be missing values in the observations at different time steps depending on the evaluated airport station.

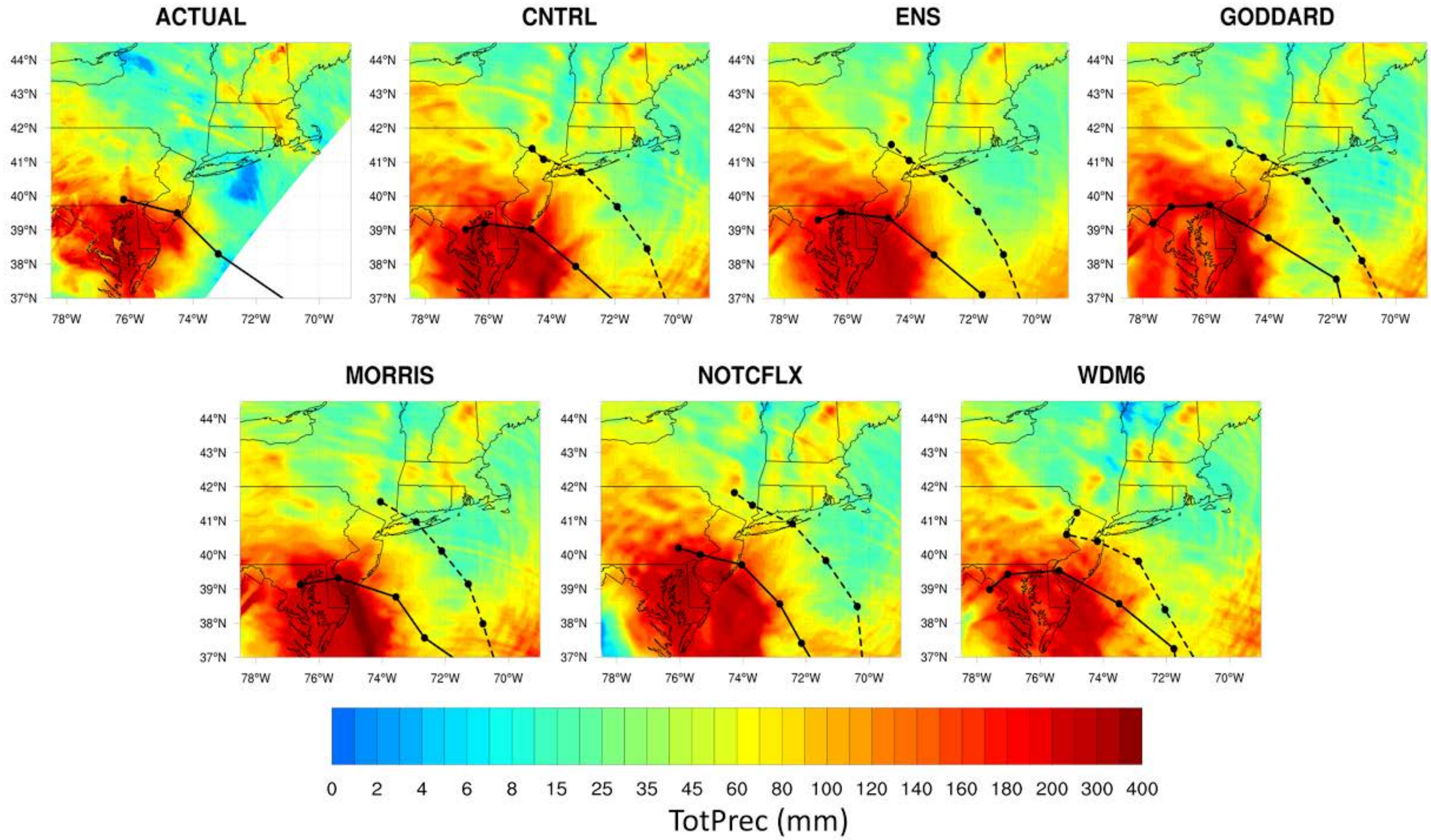

Figure 4: Accumulated precipitation from each WRF simulation ensemble member (name conventions correspond to Table 1), the ensemble mean of the five members (ENS), and Stage IV radar data (ACTUAL) which represents precipitation observations. The Current Sandy tracks are added in thick black lines, with Future Sandy tracks in dashed lines.

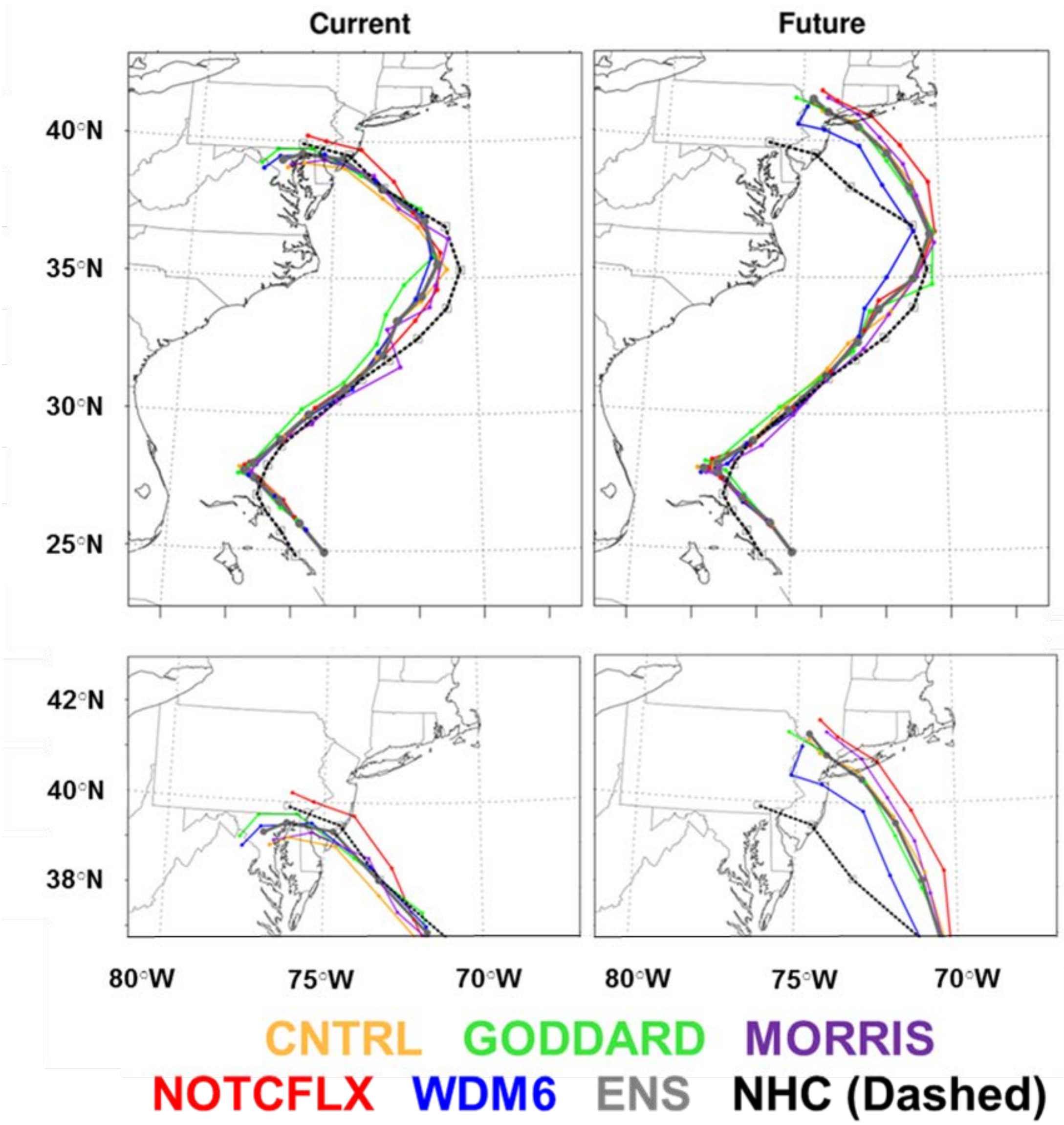

Figure 5: (a) Current and Future Sandy storm tracks. Colored lines correspond to individual WRF simulations, the grey line indicates the ensemble mean track (ENS), and the dashed black line represents the National Hurricane Center (NHC) "best track" for Current Sandy. (b) Zoomed in to highlight storm landfall location.

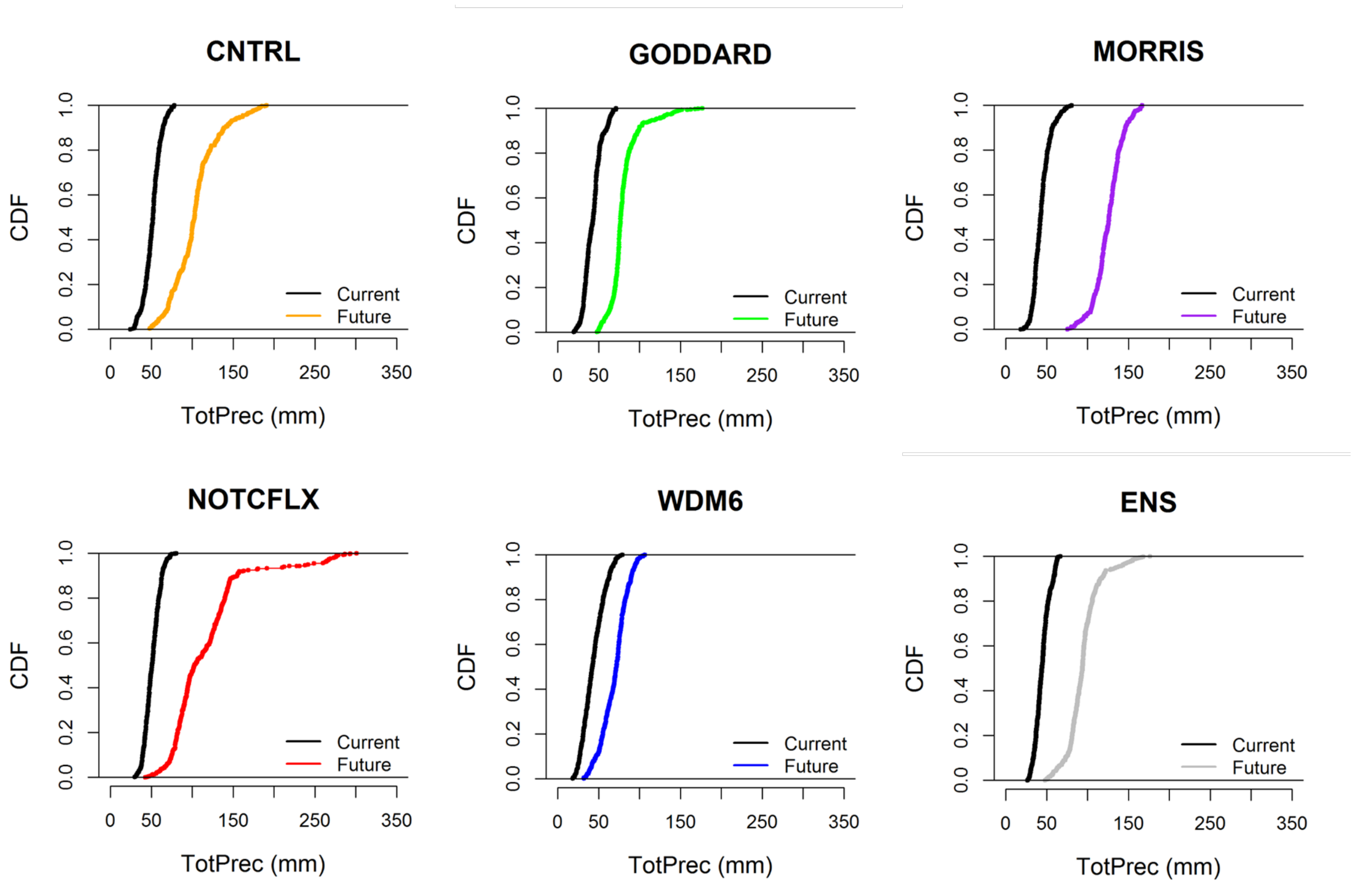

Figure 6: Cumulative distributions of total accumulated precipitation for Current and Future Sandy simulations in the sub-region of the model domain enclosing the Eversource service territory. Colors correspond to WRF simulations in Figure 3.

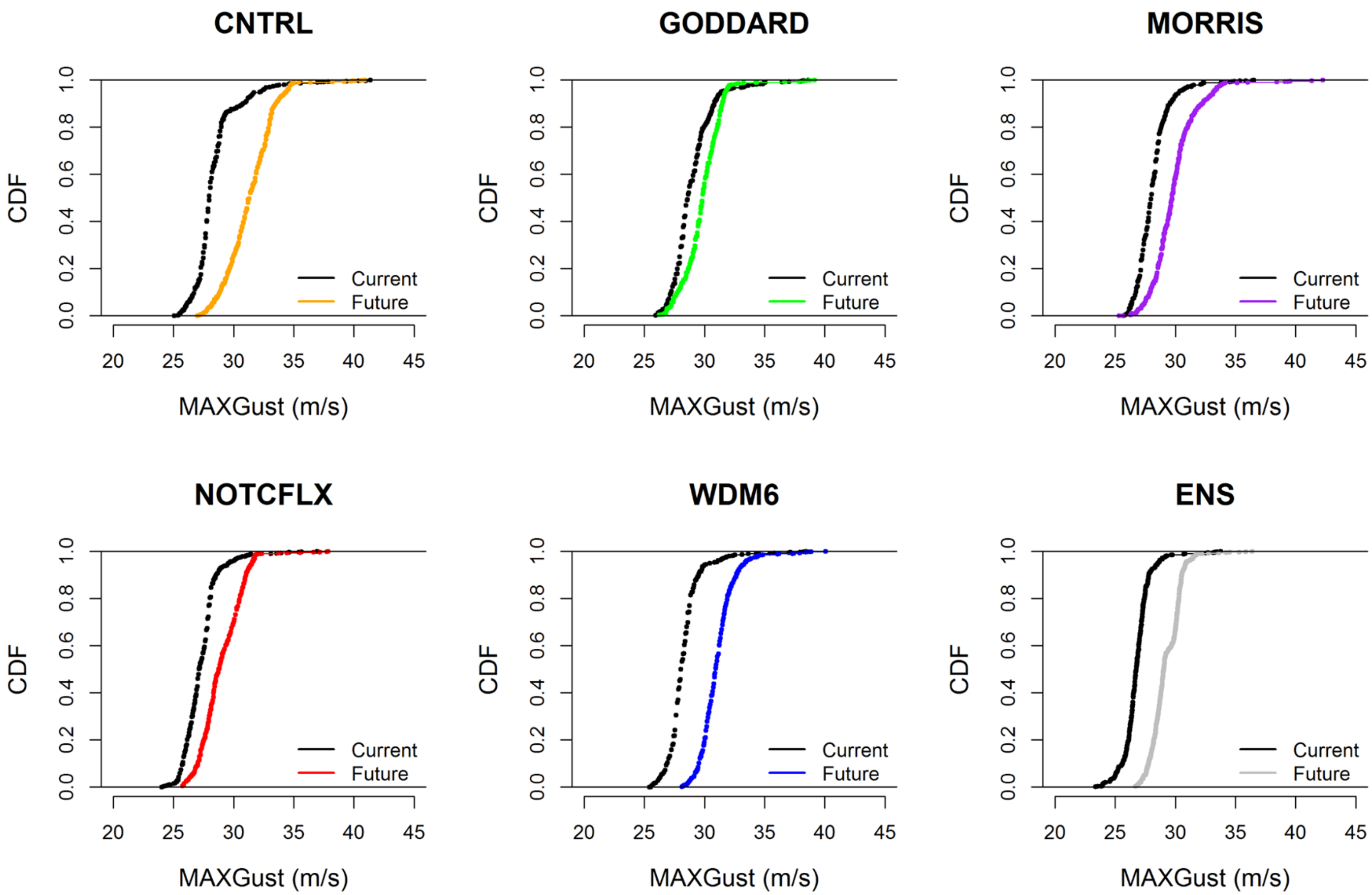

Figure 7: Cumulative distributions of maximum gust for Current and Future Sandy simulations in the sub-region of the model domain enclosing the Eversource service territory. Colors correspond to WRF simulations in Figure 3.

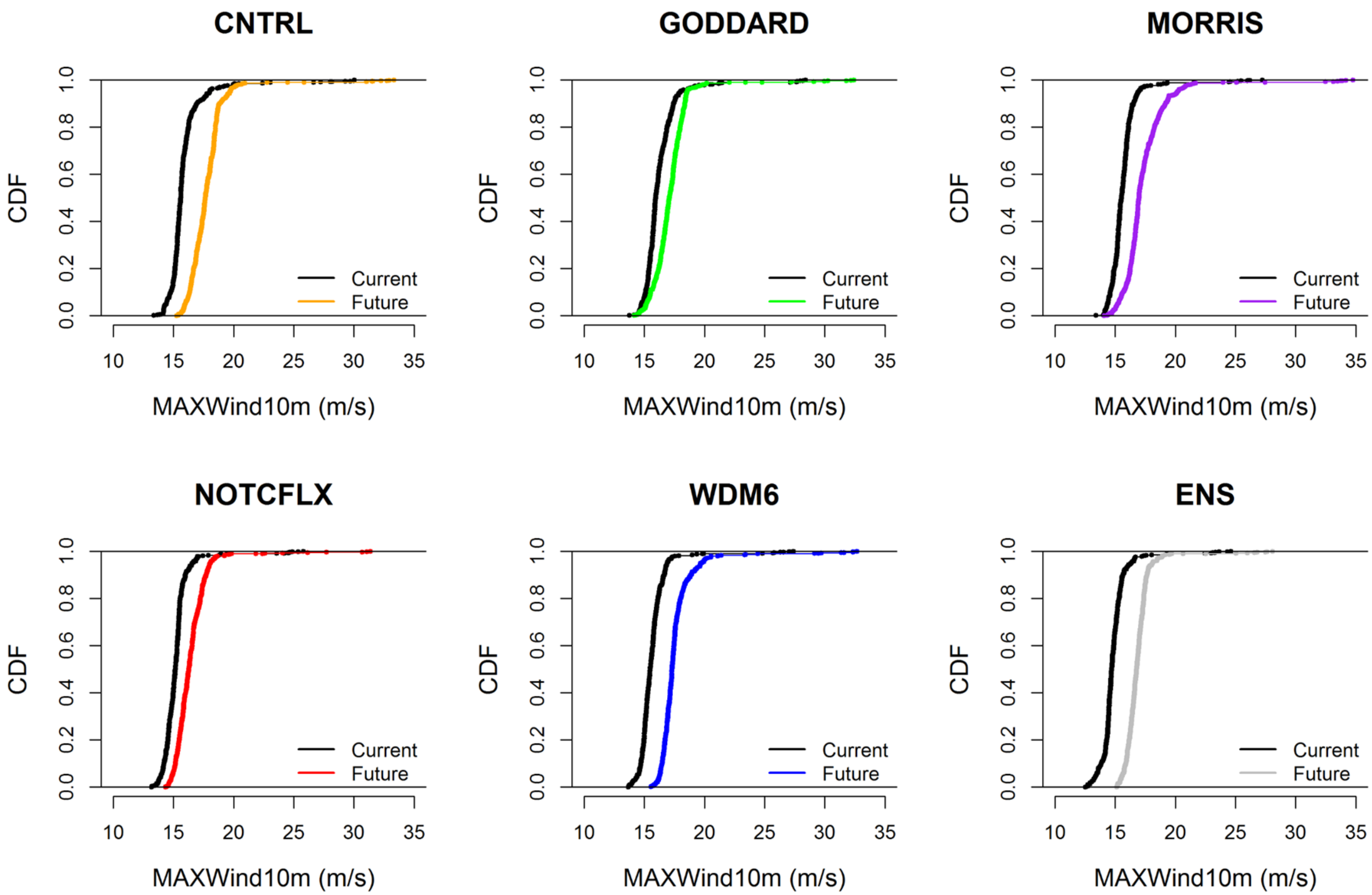

Figure 8: Cumulative distributions of maximum wind at 10-m for Current and Future Sandy simulations in the sub-region of the model domain enclosing the Eversource service territory. Colors correspond to WRF simulations in Figure 3.

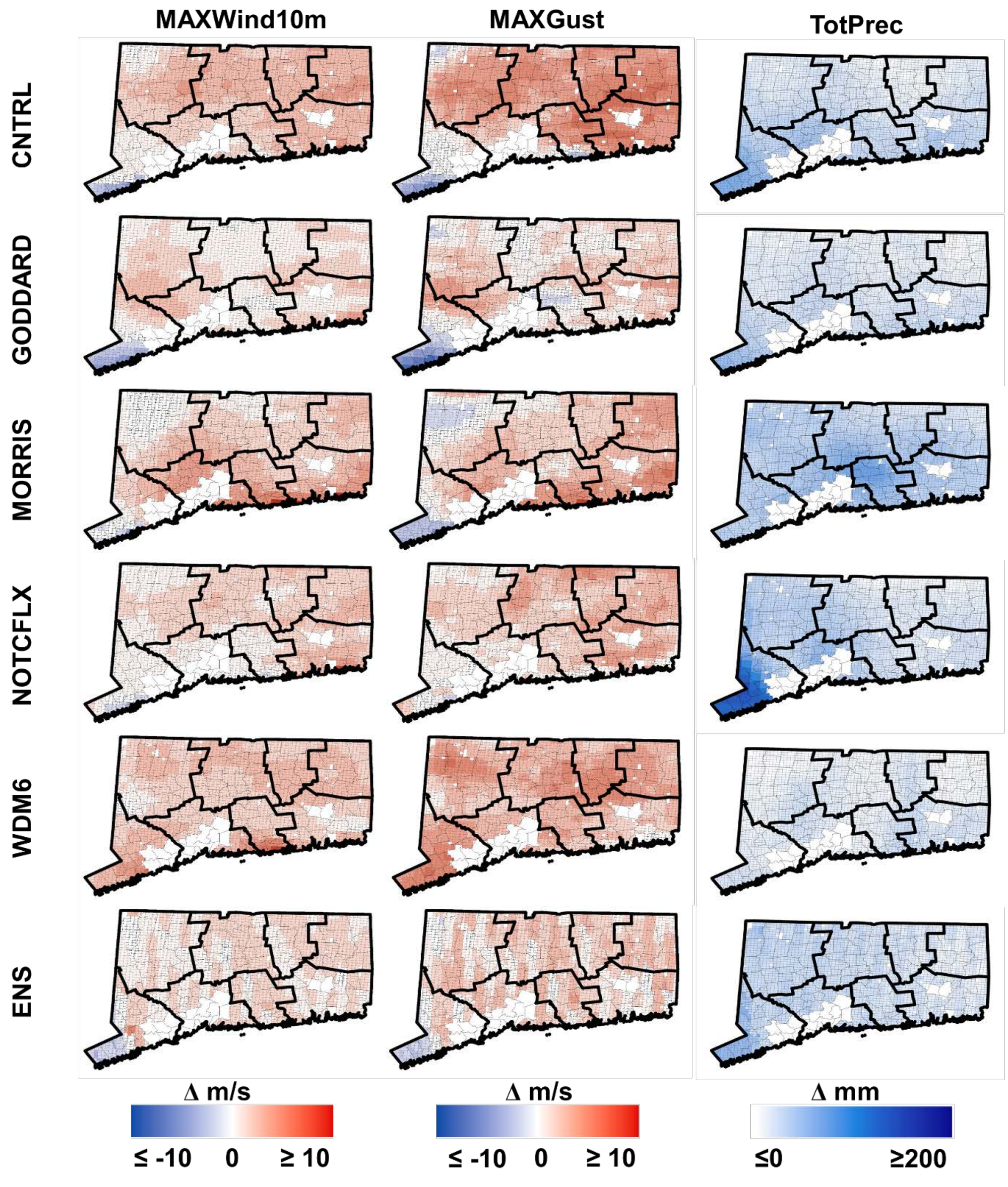

Figure 9: Changes in select wind and precipitation magnitude from Current to Future Sandy. Positive values indicate an increase in intensity during Future Sandy.

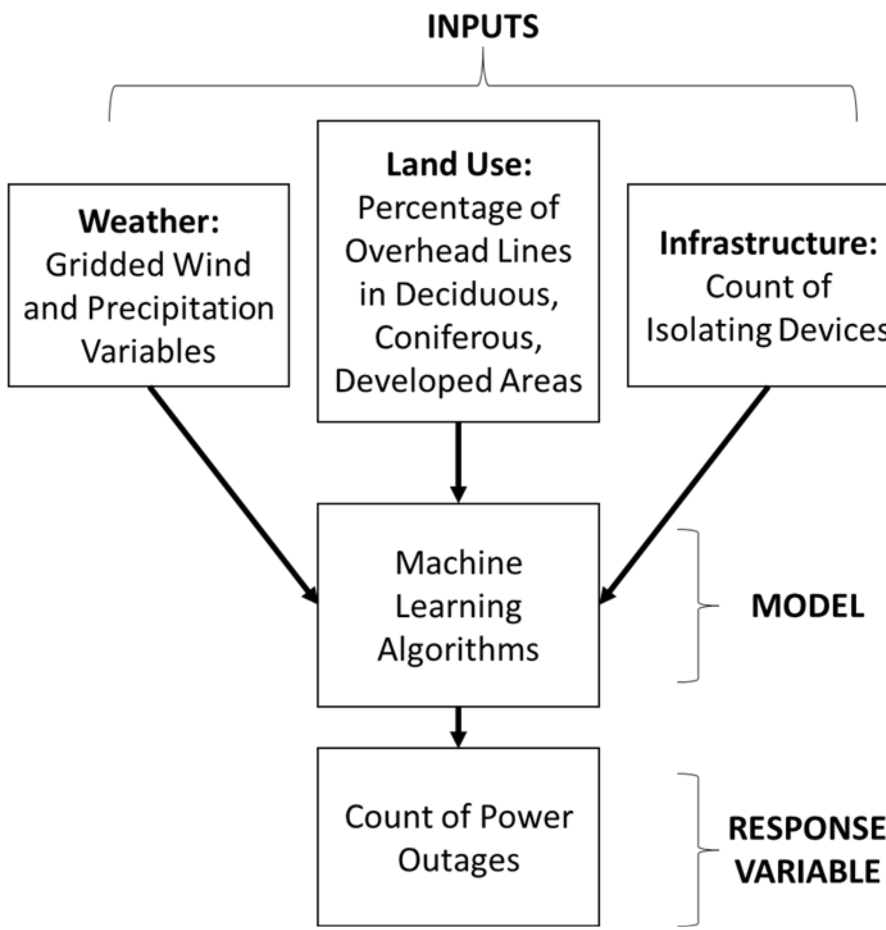

Figure 10: Modeling framework that combines weather, land use and infrastructure into outage predictions for Current Sandy scenarios. Calibrated models were then applied to Future Sandy scenarios.

Figure 11: Relative variable importance for the BART, BT and RF models, with full data input (normalized by highest value in column – does not include assets per grid cell). Darker colors indicate higher relative importance.

	BART						BT						RF					
Variable	CNTRL	GODDARD	MORRIS	NOTCFLX	WDM6	ENS	CNTRL	GODDARD	MORRIS	NOTCFLX	WDM6	ENS	CNTRL	GODDARD	MORRIS	NOTCFLX	WDM6	ENS
Cowgt13	27	42	47	36	36	20	0	0	0	38	0	0	10	11	48	25	33	2
Cowgt18	16	21	18	6	10	19	0	0	0	0	0	0	3	1	0	0	0	0
Cowgt5	44	47	58	52	59	33	3	1	4	53	1	0	18	14	38	83	33	5
Cowgt9	19	27	39	46	32	32	0	0	0	79	0	1	5	9	12	62	8	4
ggt13	51	32	47	35	51	33	1	1	3	0	52	0	21	14	28	13	29	10
ggt17	46	31	33	32	44	39	1	0	0	2	0	1	10	4	7	10	8	4
ggt22	45	24	44	42	23	29	0	0	0	12	0	0	13	2	8	14	3	5
ggt27	30	45	42	30	35	51	0	13	3	0	4	0	2	32	22	3	10	12
ggt35	8	16	7	6	8	0	0	0	0	0	0	0	0	0	0	0	0	0
MAXGust	60	47	52	33	46	57	3	100	100	0	0	1	46	100	100	12	22	27
MAXPreRate	62	61	45	38	34	37	1	0	0	0	0	0	33	32	29	59	21	13
MAXSoilMst	51	38	47	41	48	36	1	1	0	1	4	1	26	17	24	19	22	8
TotPrec	58	77	56	43	49	33	1	11	3	4	8	0	42	39	79	33	43	16
MAXWind10m	40	58	53	30	41	43	68	1	0	0	1	34	47	34	55	23	36	69
MEANGust	44	48	49	41	41	53	13	21	12	1	15	100	49	70	65	18	35	100
MEANPreRate	64	43	58	45	46	35	100	1	0	4	48	1	100	15	45	16	56	12
MEANSoilMst	37	37	47	36	41	32	0	0	0	0	0	0	35	15	23	15	26	7
MEANWind10m	47	43	41	29	48	52	0	1	0	2	9	25	42	38	40	29	42	56
PercConif	70	72	100	90	81	74	0	0	0	0	0	0	33	27	37	29	36	16
PercDecid	100	100	99	100	100	100	17	5	6	6	7	2	61	32	45	32	44	17
PercDev	81	81	91	79	75	67	28	6	14	15	33	1	91	49	63	44	81	22
wgt13	24	34	40	38	42	19	0	1	40	4	49	0	7	15	45	22	33	2
wgt18	24	24	13	7	9	19	0	1	0	0	0	0	3	1	0	0	0	0
wgt5	42	39	62	58	74	39	10	1	11	100	100	0	43	32	42	100	100	12
wgt9	23	34	36	42	52	34	0	0	0	1	15	0	9	15	19	17	15	6

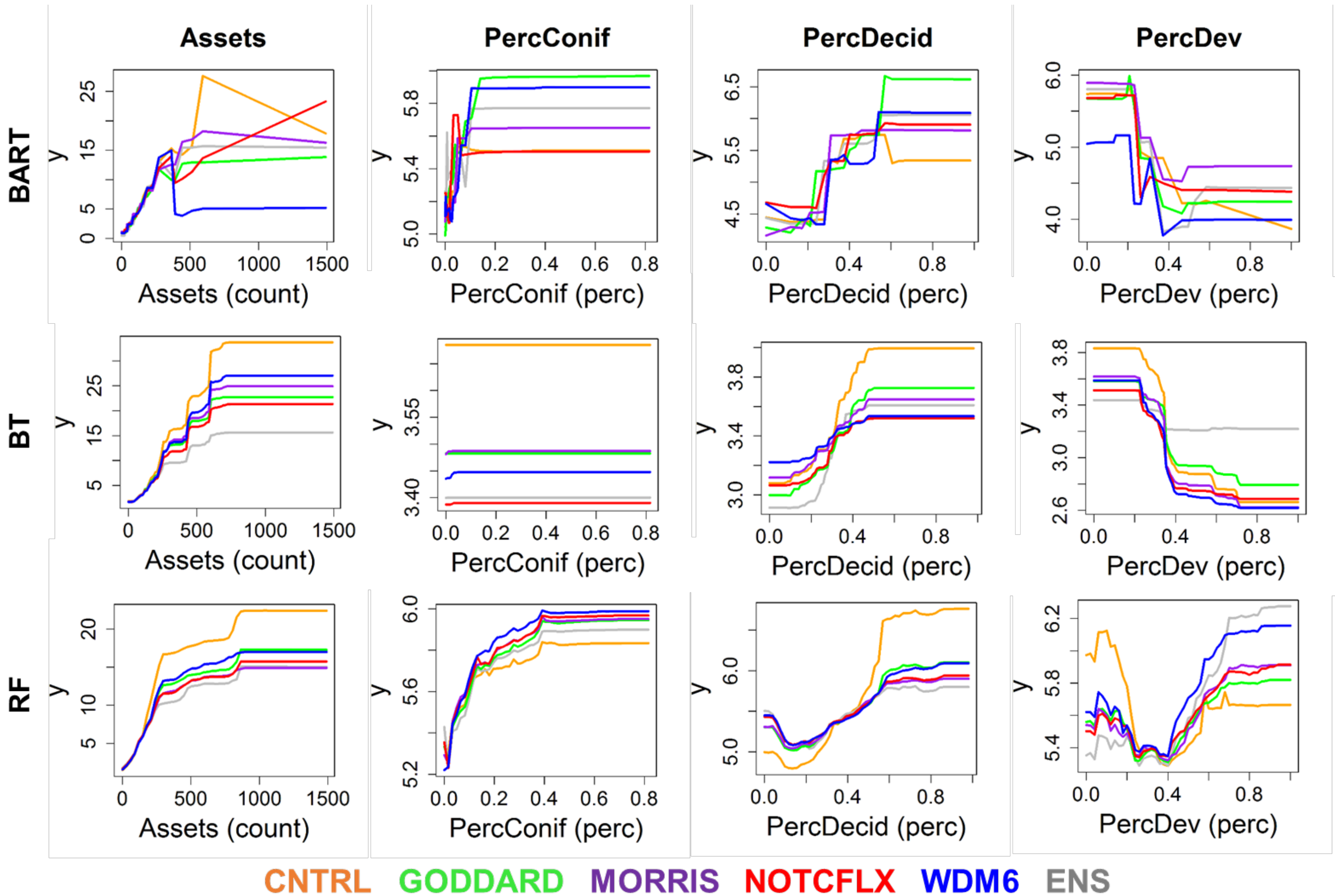

Figure 12: Partial dependence plots related to select geographic variables. Y axis represent change in predicted outages per 2-km grid cell. Colors are related to WRF simulations in Figure 3.

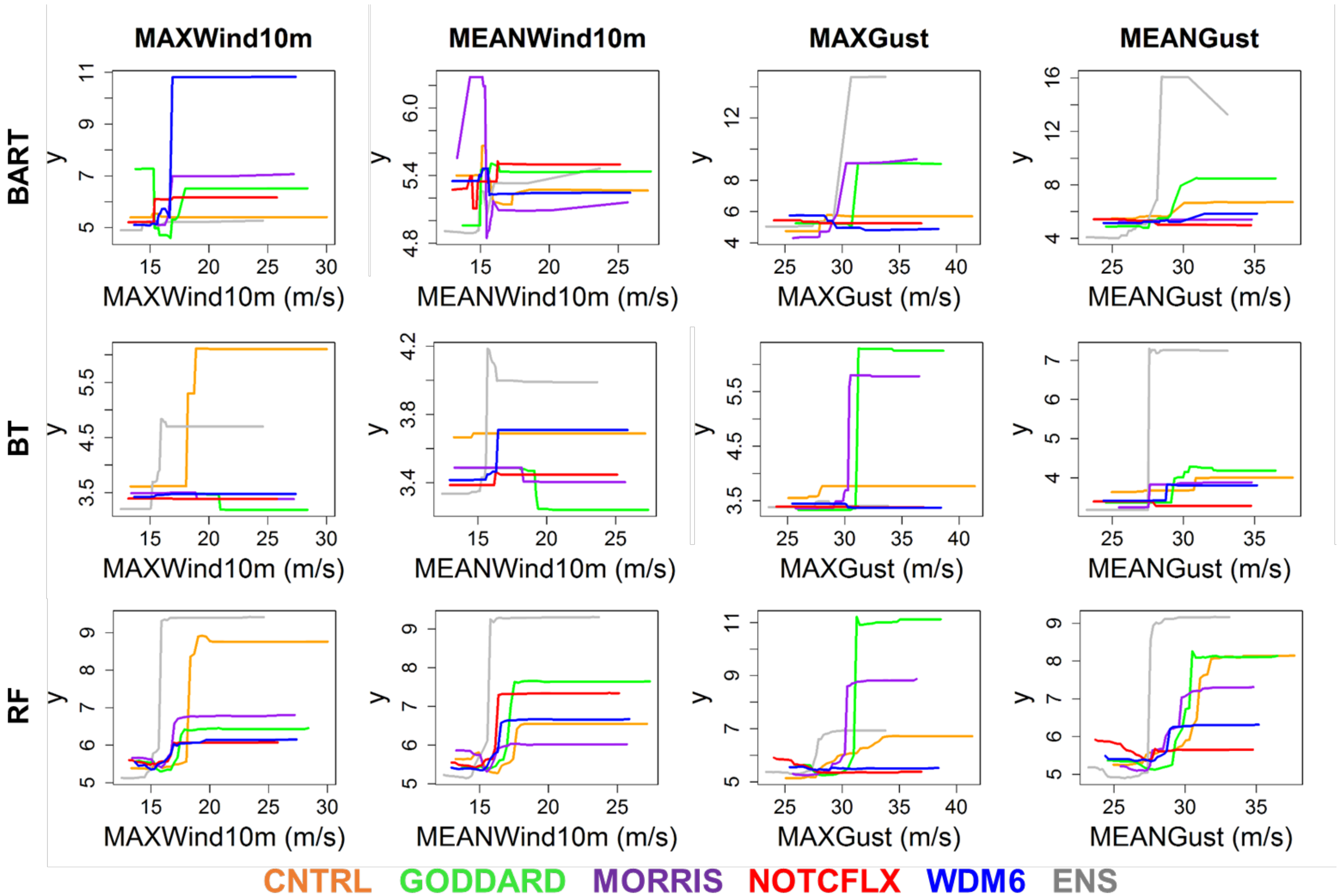

Figure 13: Partial dependence plots related to select wind variables. Y axis represent change in predicted outages per 2-km grid cell. Colors are related to WRF simulations in Figure 3.

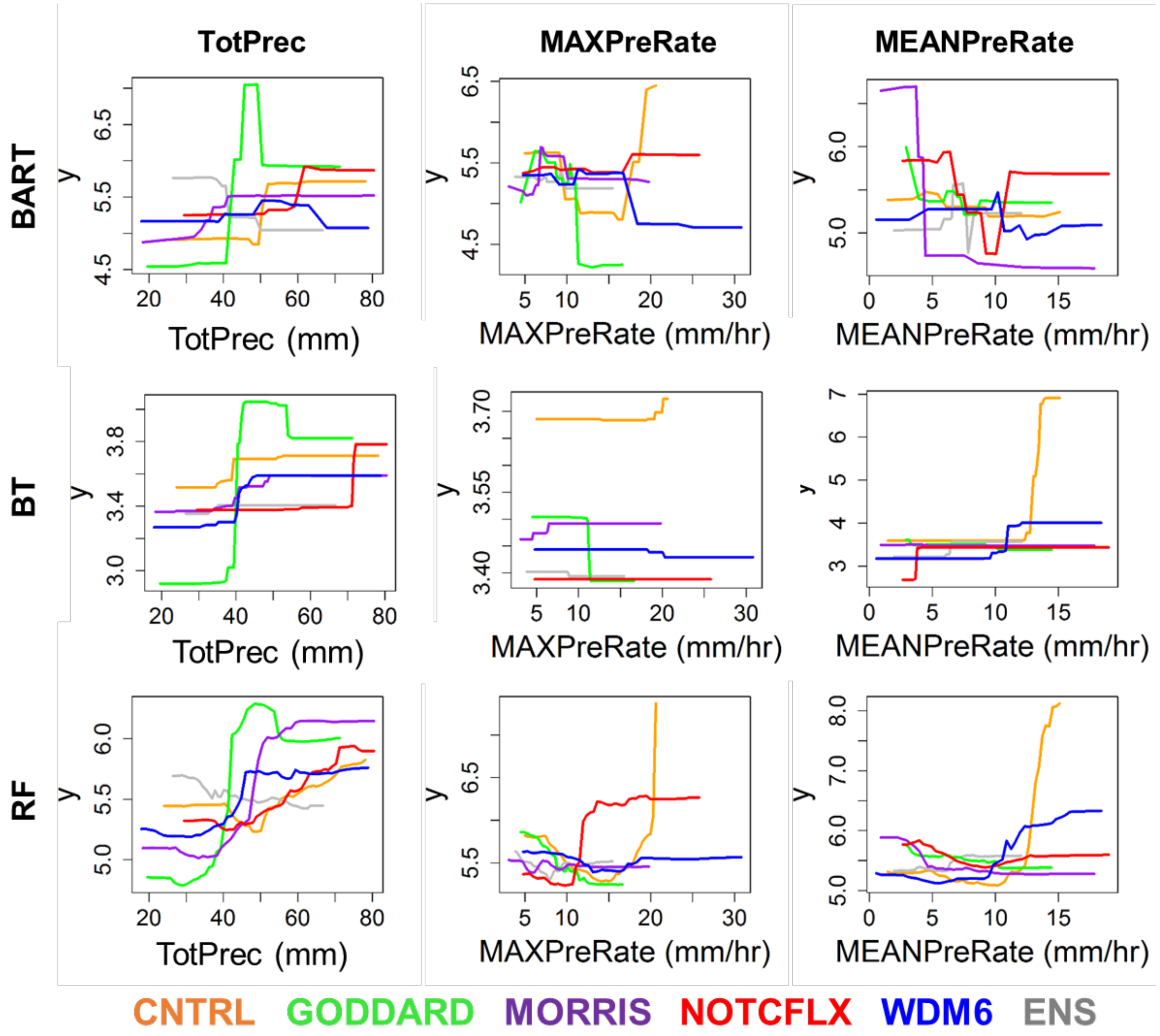

Figure 14: Partial dependence plots related to select precipitation variables. Y axis represent change in predicted outages per 2-km grid cell. Colors are related to WRF simulations in Figure 3.

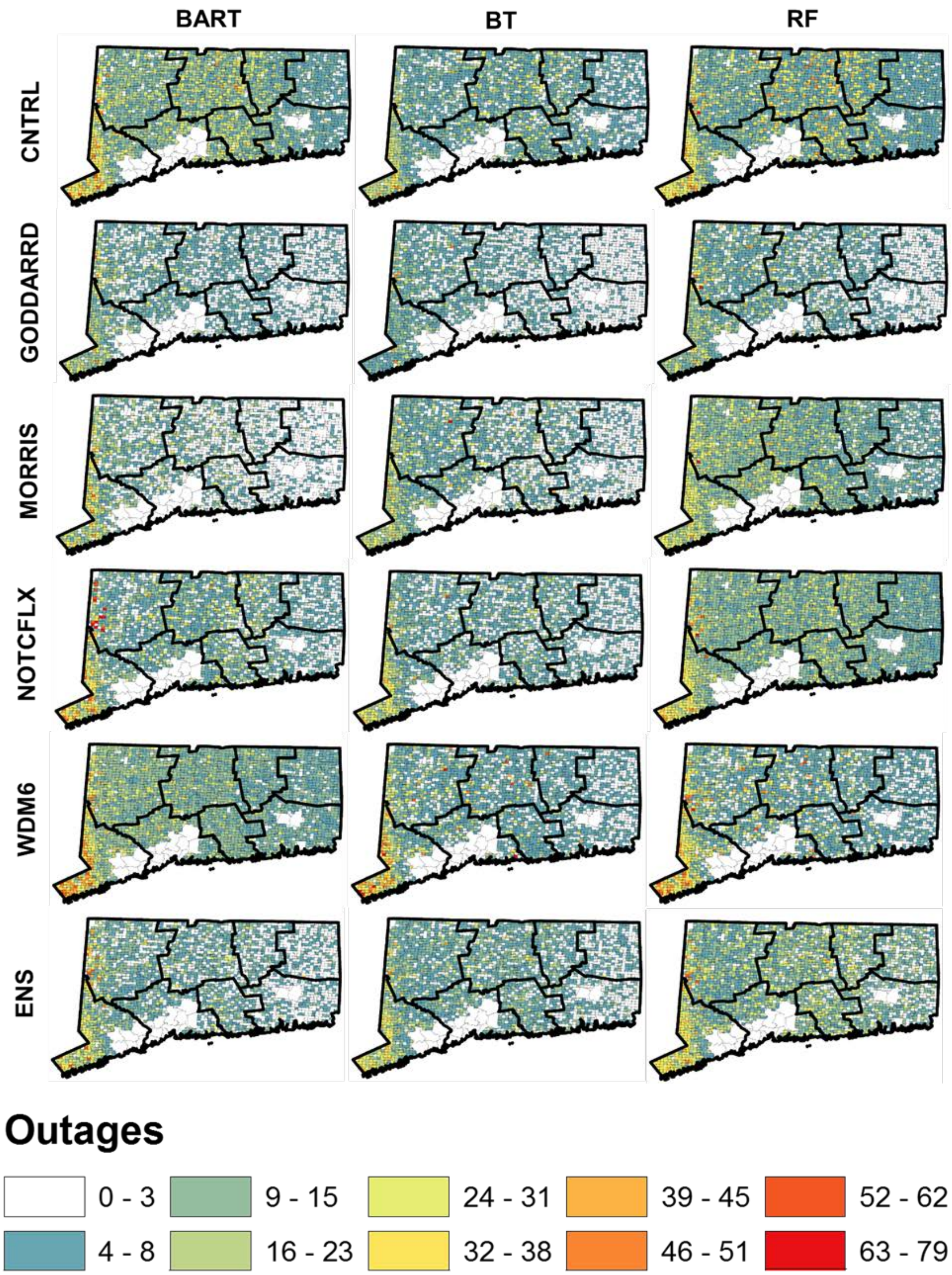

Figure 15: Distribution of predicted outages for Future Sandy by simulation and machine learning models for the full model forcing (wind and precipitation variables). Legend matches Current Sandy actual outages in Figure 1.

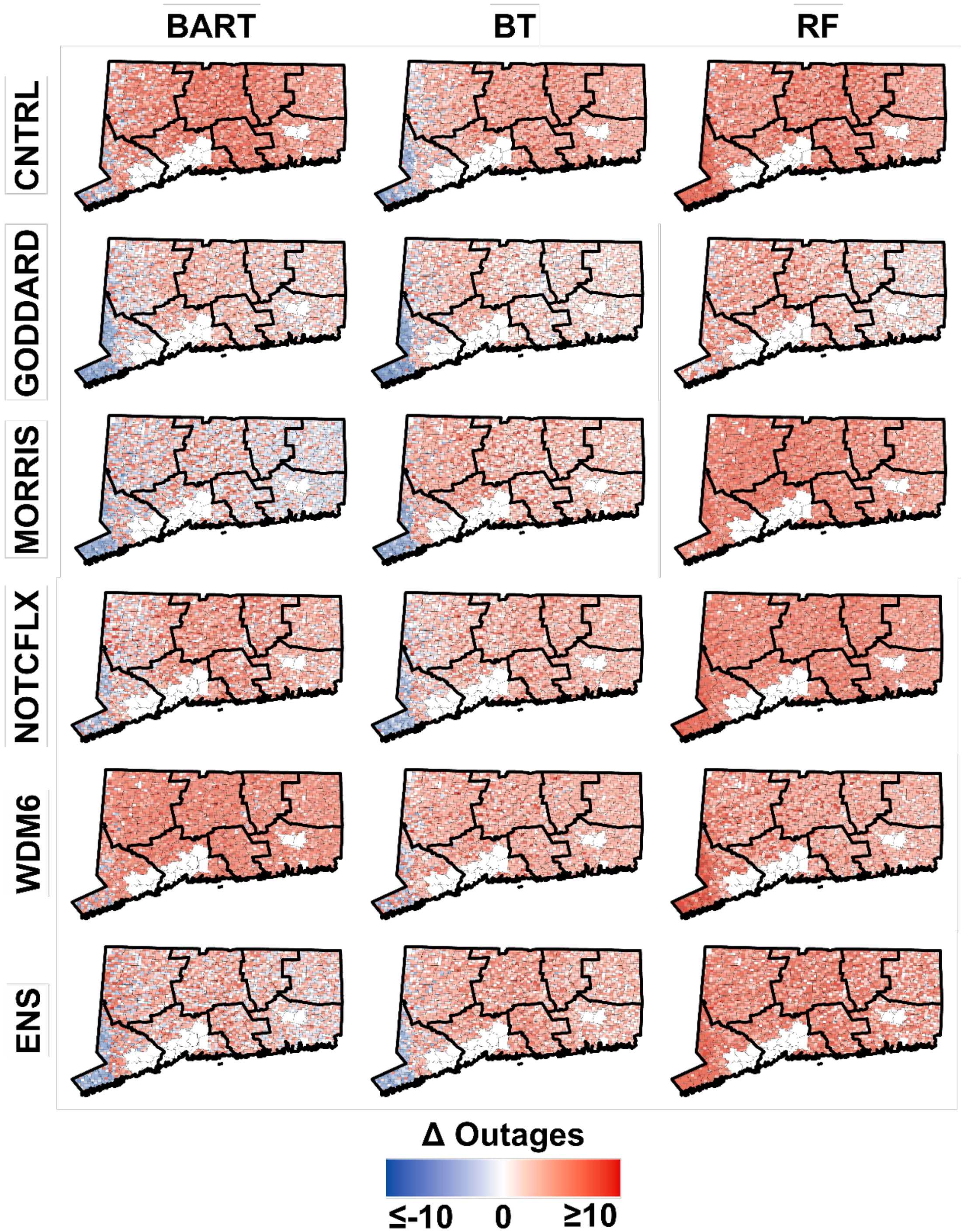

Figure 16: Change in predicted outages from Current to Future Sandy for the full data input (positive numbers indicate an increase in Future Sandy).

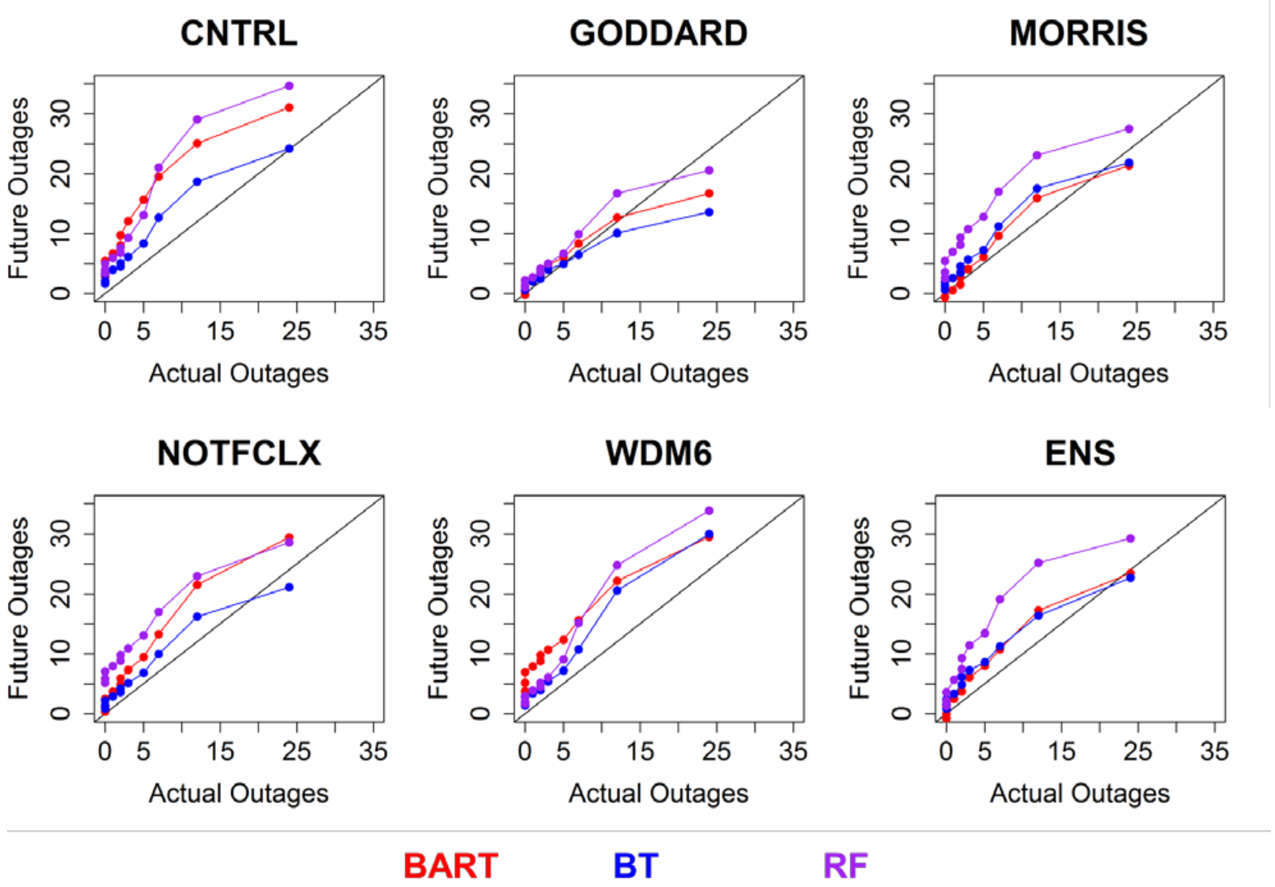

Figure 17: Quantile-quantile plot showing the increase in predicted outages per grid cells for Future Sandy (Y axis) compared to actual Current Sandy outages per grid cell (X axis) for BART, BT and RF models with the full data input. Quantiles represent the 5, 10, 20, 30, 40, 50, 60, 70, 80, 90, and 95th percentiles.

		Reduced Data Input						Full Data Input						Δ(Full - Reduced)					
		Fairfield County			Eversource-Connecticut Territory			Fairfield County			Eversource-Connecticut Territory			Fairfield County			Eversource-Connecticut Territory		
		MAX Gust	MAX Wind 10m	Tot Prec	MAX Gust	MAX Wind 10m	Tot Prec	MAX Gust	MAX Wind 10m	Tot Prec	MAX Gust	MAX Wind 10m	Tot Prec	MAX Gust	MAX Wind 10m	Tot Prec	MAX Gust	MAX Wind 10m	Tot Prec
BART	CNTRL	0.41	0.38	-0.50	0.14	0.13	-0.19	0.30	0.28	-0.37	0.04	0.02	-0.13	-0.11	-0.10	0.13	-0.10	-0.11	0.06
	ENS	0.30	0.52	-0.57	0.13	0.16	-0.15	0.27	0.43	-0.49	0.23	0.17	-0.27	-0.03	-0.09	0.08	0.10	0.01	-0.12
	GODDARD	0.36	0.53	0.33	0.01	0.14	0.13	0.45	0.63	0.42	0.09	0.26	0.17	0.09	0.10	0.09	0.08	0.12	0.04
	MORRIS	0.51	0.33	0.17	0.12	0.02	0.12	0.51	0.33	0.11	0.11	0.03	0.09	0.00	0.00	-0.06	-0.01	0.01	-0.03
	NOTCFLX	0.53	0.44	-0.46	0.18	0.12	-0.19	0.55	0.45	-0.46	0.16	0.09	-0.12	0.02	0.01	0.00	-0.02	-0.03	0.07
	WDM6	0.58	-0.12	0.29	0.25	0.06	0.02	0.53	-0.14	0.28	0.28	0.05	0.03	-0.05	-0.02	-0.01	0.03	-0.01	0.01
BT	CNTRL	0.34	0.30	-0.44	0.17	0.13	-0.20	0.31	0.27	-0.37	0.24	0.23	-0.23	-0.03	-0.03	0.07	0.07	0.10	-0.03
	ENS	0.26	0.51	-0.48	0.02	0.12	0.01	0.26	0.51	-0.53	0.00	0.11	0.00	0.00	0.00	-0.05	-0.02	-0.01	-0.01
	GODDARD	0.40	0.58	0.38	0.13	0.25	0.14	0.39	0.53	0.36	0.15	0.28	0.16	-0.01	-0.05	-0.02	0.02	0.03	0.02
	MORRIS	0.63	0.44	0.20	0.17	0.06	0.11	0.63	0.44	0.17	0.15	0.11	0.10	0.00	0.00	-0.03	-0.02	0.05	-0.01
	NOTCFLX	0.55	0.46	-0.48	0.12	0.10	-0.11	0.55	0.43	-0.43	0.09	0.02	-0.01	0.00	-0.03	0.05	-0.03	-0.08	0.10
	WDM6	0.22	-0.02	-0.04	0.31	0.09	0.03	-0.04	0.03	-0.20	0.38	0.14	0.01	-0.26	0.05	-0.16	0.07	0.05	-0.02
RF	CNTRL	-0.23	-0.18	0.36	-0.11	-0.16	0.01	-0.32	-0.27	0.45	-0.20	-0.24	0.09	-0.09	-0.09	0.09	-0.09	-0.08	0.08
	ENS	-0.24	-0.46	0.58	-0.22	-0.09	0.25	-0.24	-0.43	0.53	-0.25	-0.10	0.28	0.00	0.03	-0.05	-0.03	-0.01	0.03
	GODDARD	-0.11	-0.07	-0.11	0.10	0.33	0.11	-0.13	-0.08	-0.14	0.10	0.29	0.10	-0.02	-0.01	-0.03	0.00	-0.04	-0.01
	MORRIS	-0.08	0.01	-0.02	-0.06	-0.05	0.09	0.15	0.19	0.03	-0.03	-0.02	0.14	0.23	0.18	0.05	0.03	0.03	0.05
	NOTCFLX	-0.56	-0.48	0.52	-0.15	-0.28	0.27	-0.60	-0.49	0.52	-0.15	-0.30	0.26	-0.04	-0.01	0.00	0.00	-0.02	-0.01
	WDM6	0.61	-0.11	0.37	0.29	0.04	0.03	0.60	-0.11	0.35	0.35	0.05	0.04	-0.01	0.00	-0.02	0.06	0.01	0.01

Figure 18: Correlation between increased outages and weather magnitude using Spearman correlation for Fairfield County and the Eversource Connecticut service territory for the full and reduced data input. Red cells indicate positive correlation, blues cells indicate negative correlation, and white cells indicate a lack of correlation. Difference between Spearman correlations for full and reduced data input are also tabulated in right third (positive values indicate an improvement in the full model).

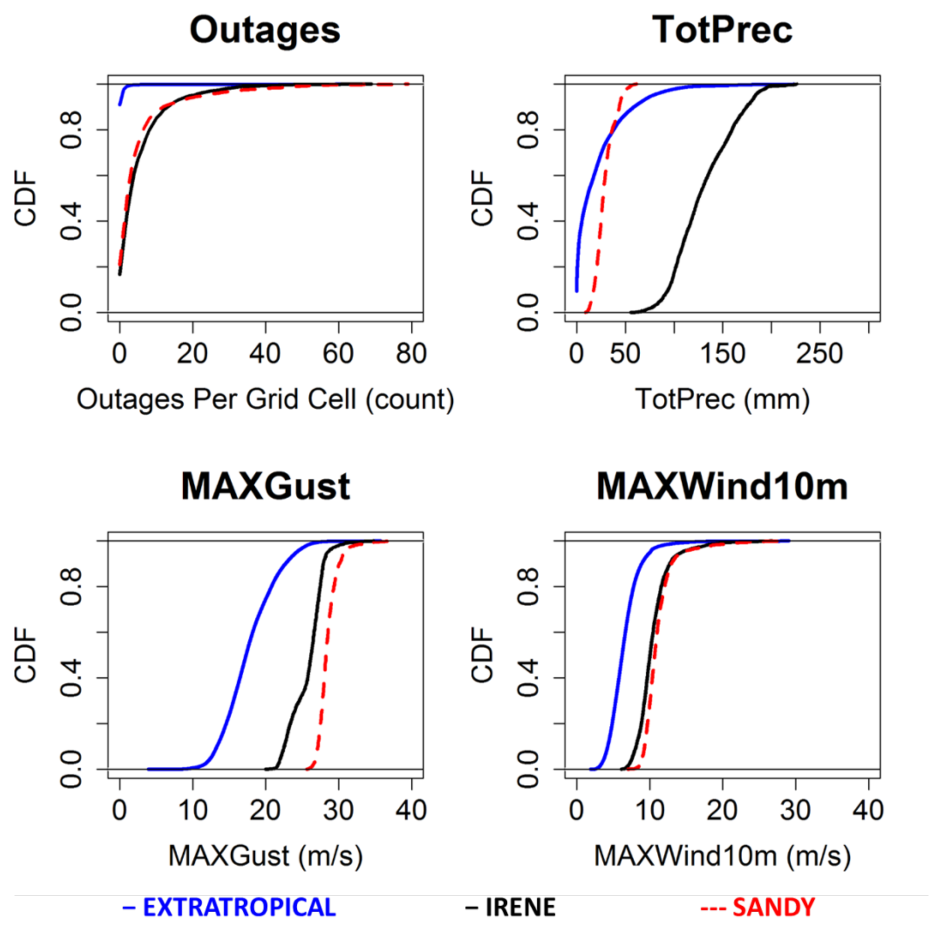

Figure A.1: Comparison of CDF plots for select weather variables for 76 extratropical storms (occurred between 2005 and 2017), Hurricane Irene (2011), and Hurricane Sandy (2012).