IEEE Access
Multidisciplinary | Rapid Review | Open Access Journal

Received February 10, 2019, accepted February 22, 2019, date of publication March 1, 2019, date of current version March 20, 2019.

Digital Object Identifier 10.1109/ACCESS.2019.2902558

Predicting Storm Outages Through New Representations of Weather and Vegetation

DIEGO CERRAI[1], DAVID W. WANIK[2], MD ABUL EHSAN BHUIYAN[3], XINXUAN ZHANG[4], JAEMO YANG[5], MARIA E. B. FREDIANI[6], AND EMMANOUIL N. ANAGNOSTOU[1]

[1]Department of Civil and Environmental Engineering, University of Connecticut, Storrs, CT 06269, USA
[2]Eversource Energy Center, University of Connecticut, Storrs, CT 06269, USA
[3]Department of Natural Resources and the Environment, University of Connecticut, Storrs, CT 06269, USA
[4]Department of Civil, Environmental, and Infrastructure Engineering, George Mason University, Fairfax, VA 22030, USA
[5]National Renewable Energy Laboratory, Golden, CO 80401, USA
[6]National Center for Atmospheric Research, Boulder, CO 80305, USA

Corresponding author: Diego Cerrai (diego.cerrai@uconn.edu)

This work was supported by the Eversource Energy.

ABSTRACT This paper introduces new developments in an outage prediction model (OPM) for an electric distribution network in the Northeastern United States and assesses their significance to the OPM performance. The OPM uses regression tree models fed by numerical weather prediction outputs, spatially distributed information on soil, vegetation, electric utility assets, and historical power outage data to forecast the number and spatial distribution of outages across the power distribution grid. New modules introduced hereby consist in 1) a storm classifier based on weather variables; 2) a multimodel optimization of regression tree output; and 3) a post-processing routine for more accurately describing tree-leaf conditions. Model implementations are tested through leave-one-storm-out cross-validations performed on 120 storms of varying intensity and characteristics. The results show that the median absolute percentage error of the new OPM version decreased from 130% to 59% for outage predictions at the service territory level, and the OPM skills for operational forecasts are consistent with the skills based on historical storm analyses.

INDEX TERMS Power distribution, extreme events, machine learning, numerical weather predictions, power outage prediction.

I. INTRODUCTION

Electricity is a foundation of modern society [1], [2], and unreliable electric power delivery has direct and long-term socio-economic effects [3], [4]. Predicting the variability of electricity usage, in the form of forecasting the balance between energy production [5]–[7], demand [8]–[11], and outages [12], [13], is essential for addressing the reliability of power delivery.

An important reliability component is weather-based outages. Storms can cause severe disruption in the electric grid, affecting human activities [14], security and life [15], [16]. Advancing the predictive understanding of the relationships between weather and power outages is a key step for creating a resilient electric grid, capable of withstanding the current increase of weather-caused power outages [17] and the expected increase in severe weather events [18].

The associate editor coordinating the review of this manuscript and approving it for publication was Berdakh Abibullaev.

In the Northeastern United States, major power outages can be caused by singular failures on the transmission system that serves hundreds of thousands of customers [19]. In contrast, most daily outages are produced by tree branches falling on overhead distribution lines [20], that typically serve tens to hundreds of customers. During storms, power outages occurrence is determined by a complex interaction among atmospheric phenomena, vegetation cover, and infrastructure. Accurate storm power outage predictions and early communication of predicted impacts to the utilities in the face of this complexity are necessary for efficient emergency preparedness, support and response [21].

The first attempt to develop an outage prediction model can be found in [12], where a negative binomial regression model [22] was used to evaluate the relative importance of transformers, wind speed, and random effects during three hurricanes. In the decade that followed, several studies on outage prediction modeling were conducted. Liu *et al.* [23] used Generalized Linear Models (GLMs [22]) to investigate

2169-3536 © 2019 IEEE. Translations and content mining are permitted for academic research only.
Personal use is also permitted, but republication/redistribution requires IEEE permission.
See http://www.ieee.org/publications_standards/publications/rights/index.html for more information.

the importance of hurricane and ice storm variables for outage prediction, and to improve the previous formulation of the [12]'s hurricane outage prediction model. Advances were obtained with a better variable selection [24], with the use of Generalized Additive Models (GAMs [25]) for hurricane outage predictions [26], and with the implementation of a random forest (RF [27]) model [28]. The usefulness of a hybrid classification tree/regression method for handling the zero-inflation was discussed in [29], while Quiring *et al.* [30] used classification and regression trees (CARTs [31]) for demonstrating that some land cover variables are proxies for the power system, hence useful for predicting outages. Building on [28], the Spatially Generalized Hurricane Outage Prediction Model (SGHOPM [32]) combined elevation, land cover, soil and vegetation with the wind characteristics included in the first version [33], to improve prediction skill. More recently, Wanik *et al.* [34] used also vegetation management and tree height data derived from LiDAR for enhancing outage prediction accuracy by identifying the vegetation at risk for striking overhead lines.

By exploiting the knowledge derived from previous studies, this paper builds upon a comprehensive Outage Prediction Model (OPM), presented in [35] (hereafter, W15) and [36] (hereafter, H17). The model was developed for predicting outages associated with the main types of weather events affecting the Northeastern United States, including thunderstorms, snow storms, extratropical and tropical cyclones, by using state-of-the-art meteorological, statistical, and machine learning (ML) techniques. The OPM presented in W15 and H17 and its improvements have been used since 2015 to issue forecasts for Eversource Energy-Connecticut and since early 2017 for Massachusetts and New Hampshire service territories.

Previous studies on the OPM have highlighted different aspects of outage prediction. In particular, from the evaluation of eight ML techniques conducted in W15, an ensemble (ENS) model consisting of predictions from a decision tree (DT [31]), a RF and a boosted tree (BT [37]) model, that used weather hindcasts (dynamical downscaling of weather analyses), infrastructure and land cover data as inputs, emerged as spatially better performing than each individual model. In H17, a comparison between a Bayesian additive regression trees (BART [38]) model and a quantile regression forest (QRF [27], [39]) model yielded ''inconsistent performances of both models for varying season categories (tree-leaf condition) implying difficulties in predicting storm outages with leaves on trees''. This study also determined that the BART model surpassed QRF in predicting the total number of outages over the entire service territory.

Although both studies used variables extracted from weather hindcasts as inputs to the outage models, neither evaluated operational or pseudo-operational performances computed on actual weather forecasts, after the outage models were built. It is well known, however, that even within a short range of days (one to two), weather forecasts can be associated with significant error [40], due to uncertainties in initial conditions - see [41] for a theoretical approach and [42] for a case study over the East Coast - and limitations in physical parameterizations of unresolved physics, such as the cumulus [43] or planetary boundary layer schemes [44]. Therefore, performance evaluation based on weather hindcasts, despite its wide use in outage prediction literature, is not sufficient for describing the overall performance of an outage forecasting model.

In this study we aim at addressing the main challenges identified in previous works, by introducing and describing the implementation of three new modules into the OPM: 1) a *storm classifier* to allow different ML model calibrations for meteorologically distinct storm types and quantify uncertainties for each type; 2) a *leaf area index (LAI)* parameter determined from satellite data to account for dynamic leaf conditions that affects the relationship among various weather parameters (wind gust, snow, ice) and power outages; and 3) an *optimization* that combines ML model outputs based on training error characteristics, to reduce outage modeling error. We conduct cross-validations in different model configurations based on 12 years of historical events (120 storms) with the intent to demonstrate improvements brought about by the new modules. To evaluate actual operational performance, we also compare OPM skills for one year of analyzed events (25 storms) with the skills of OPM driven by weather forecasts initialized 24 hours before the selected events.

II. THE OUTAGE PREDICTION FRAMEWORK

The OPM is an operational forecasting framework that integrates weather predictions with infrastructure, land cover and vegetation characteristics to predict, through the use of machine learning models, distributed storm power outages across utility service territories in the Northeastern United States (Figure 1). This section aims at describing the OPM structure.

A. INPUT DATA

The OPM inputs are listed in Table 1, and derive from different sources: weather data (groups *b,c,d,e*) generated from numerical weather prediction (NWP) model; land cover (coniferous, deciduous and developed area) and vegetation (LAI) variables retrieved from satellites; and infrastructure data describing assets in the overhead distribution network provided by the distribution company. These input datasets have been chosen because forest and land cover type, along with the amount of leaves on trees in a given period of the year strongly influence how weather phenomena interact with overhead lines, causing power outages.

Such data are rescaled and matched on a common grid, to constitute the ''storm analysis'', a summary of measured or predicted conditions for each storm. A detailed description of the sources of each input variable, of the data preparation, and of the process that allows to rescale and match the data on a common grid is provided in the next paragraphs. The reasons behind the classification of the variables into six different groups are described in Section III.

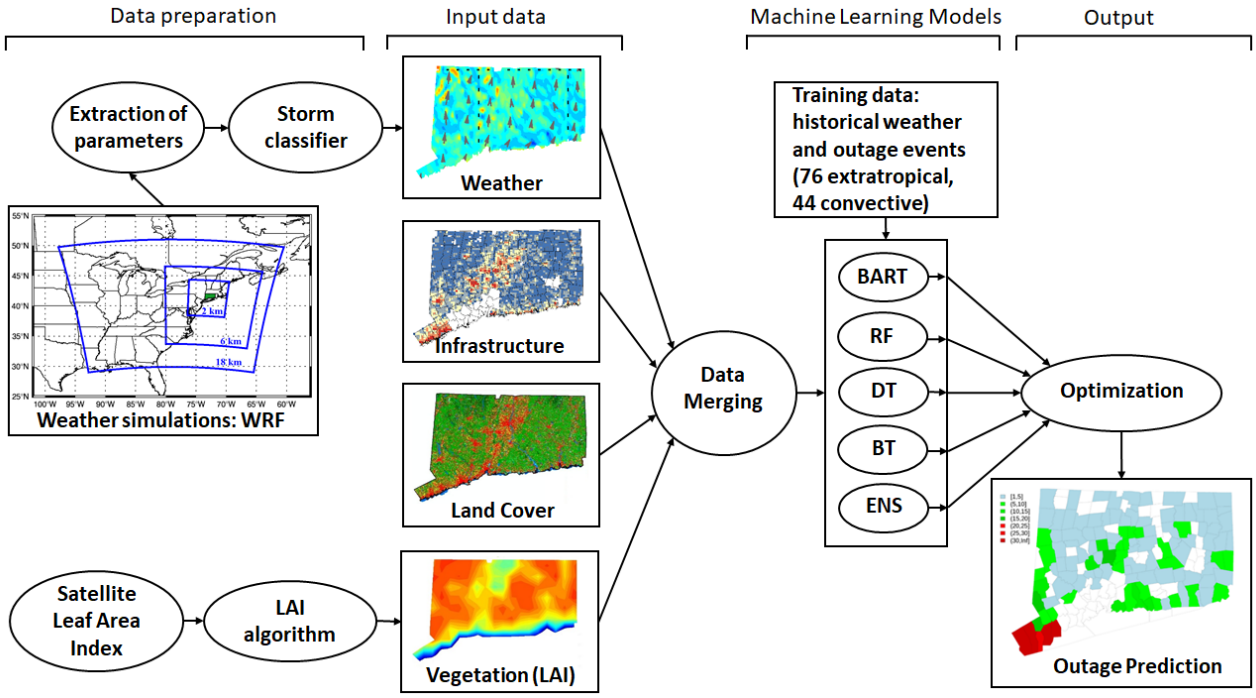


FIGURE 1. The Outage Prediction Model architecture.

TABLE 1. Description of the predictor variables used in the OPM.

Variable	Description	Group	Units
PercConif	Percentage of coniferous forest	a	%
PercDecid	Percentage of deciduous forest	a	%
PercDevel	Percentage of developed area	a	%
Wgt5	Duration of wind at 10 m above 5 m/s	b	hr
Wgt9	Duration of wind at 10 m above 9 m/s	b	hr
Wgt13	Duration of wind at 10 m above 13 m/s	b	hr
MaxW10m	Maximum wind at 10 m	b	$m\ s^{-1}$
Cowgt5	Continuous duration of wind at 10 m above 5 m/s	b	hr
Cowgt9	Continuous duration of wind at 10 m above 9 m/s	b	hr
Cowgt13	Continuous duration of wind at 10 m above 13 m/s	b	hr
MeanW10m	Mean wind at 10 m	b	$m\ s^{-1}$
Ggt13	Duration of wind gusts at 10 m above 13 m/s	c	hr
Ggt17	Duration of wind gusts at 10 m above 17 m/s	c	hr
MaxGust	Maximum wind gusts at 10 m	c	hr
MeanGust	Mean wind gusts at 10 m	c	$m\ s^{-1}$
MaxTotPrec	Total precipitation	d	mm
MaxPreRate	Maximum precipitation rate	d	$mm\ h^{-1}$
MeanPreRate	Mean precipitation rate	d	$mm\ h^{-1}$
MaxSoilMst	Maximum Soil Moisture	e	$m^3\ m^{-3}$
MeanSoilMst	Mean Soil Moisture	e	$m^3\ m^{-3}$
MaxSpHum	Maximum Specific Humidity	e	$g\ g^{-1}$
MaxTemp	Maximum Temperature	e	K
MeanTemp	Mean Temperature	e	K
LAI	Leaf area index	f	$m^2\ m^{-2}$

1) WEATHER

For each storm, weather data are computed at 2 km grid spacing through the Weather Research and Forecasting (Advanced Research WRF-ARW, v.3.7.1 [45], [46]) NWP model. The WRF produces high-resolution weather forecasts (future conditions) and hindcasts (past conditions) by dynamically downscaling Global Forecast System (GFS) data, used as inital and boundary conditions. GFS data are produced by the National Center for Environmental Prediction (NCEP) and are available at six-hourly intervals on a 0.25 degree grid. Weather forecasts and hindcasts are created by applying a two-way nesting technique on three nested domains. This technique allows to remove numerical noise at the boundaries between different domains, through interactive inter-domain communication [47]–[49].

We fixed an outer domain covering most of the Eastern United States with 18 km grid spacing; a 6 km intermediate domain capturing the Northeastern United States, and an inner domain at 2 km grid spacing, centered over Connecticut, providing the final high-resolution forecast for the study area. The WRF model is set up with the configuration described by Table 2. WRF operational forecasts are updated daily at http://cee-wrf.engr.uconn.edu/.

TABLE 2. Parametrization schemes used in the WRF model.

Parametrization	Scheme
Microphysics	New Thompson et al. [50]
Cumulus	Grell 3D Ensemble Scheme [51] [52]
Planetary Boundary Layer	Yonsei University [53]
Surface Layer	MM5 Similarity [54], [55], [56], [57] [58]
Land Surface Model	Unified Noah [59]
Longwave Radiation	RRTM [60]
Shortwave Radiation	Goddard RRTMG [61] [62]

Weather simulations are processed by extracting new parameters summarizing wind, precipitation and soil data over the entire storm life (48 hours for extratropical and 36 hours for convective storms) for each grid cell:

- Maximum value (MaxW10m, MaxGust, MaxSoilMst, MaxSpHum, MaxTemp, MaxTotPrec, MaxPreRate): maximum value of any selected variable;
- Mean value (MeanW10m, MeanGust, MeanPreRate, MeanSoilMst, MeanTemp): mean value of any selected variable computed in the 4-hour window of highest mean 10m wind;
- Occurrence value (Wgt5, Wgt9, Wgt13, Ggt9, Ggt17): number of hours of wind speed or wind gusts above given thresholds over the specified storm duration;
- Continuous duration value (Cowgt5, Cowgt9, Cowgt13): maximum continuous duration (in hours) of sustained winds above a certain threshold.

Wind- and gust-related variables are used because the primary stress on trees during most of the storms is caused by winds. Precipitation changes the resistance of the wood, adds weight to tree (particularly in the form of snow or ice), and increases soil moisture, contributing to loosen the bonds between roots and soil, which may cause uprooting.Specific humidity and temperature depict additional stress on trees and infrastructure. Moreover, given that the majority of the outages occur during the storm peak, a characterization of both the peak intensity and of the duration of adverse meteorological conditions during an event are used in our model.

Storms are subsequently classified (through the classification introduced in Section II-D) according to their dominant meteorological characteristics. This allows to select different relevant input variables for different storm types, for facilitating the learning tasks of the machine learning models.

2) INFRASTRUCTURE

Utility infrastructure data were provided by Eversource Energy and reported as geolocated protective devices. Such data are proprietary and owned by the utility company. We aggregated infrastructure data at the 2 km grid cells used by the WRF model, in order to create a variable, *sum of assets*, representing the total number of electric transformers, fuses, reclosers, and switches in each grid cell.

3) LAND COVER

The land cover dataset was created by the University of Connecticut Center for Land Use Education and Research (CLEAR) at 30 m resolution [63]. Data were aggregated into 2 km grid cells by considering the land cover in the immediate proximity (60 m) of overhead lines, following W15.

4) VEGETATION

Vegetation characteristics are present in the model through the LAI, a dimensionless quantity that measures the green leaf area per unit ground surface area [64]. The original dataset is available from the NASA Earth Observations (NEO) project [65], which is part of the NASA Earth Observing System project [66]. It is derived from the Moderate Resolution Imaging Spectroradiometer (MODIS) aboard the Terra and Aqua satellites, which acquires images at 500 meters resolution, and is scaled and resampled at 0.1° spatial and eight-day temporal resolutions. The dataset is available for the 2000-2016 time period, and an algorithm that we developed for this study (Section II-D1) allows to process, quality control and rescale the entire dataset at the 2 km grid.

B. RESPONSE VARIABLE: HISTORICAL OUTAGE DATA

The number of electric power outages per 2 km grid cell per storm represents the response variable of the OPM. Outages are defined as locations where at minimum a two-man restoration crew is needed for manual intervention to restore power (W15, H17). Historical, geolocated power outages recorded by the Eversource Outage Management System (OMS) were aggregated on the 2 km grid by storm. We identified the storm dataset using an outage intensity and a weather severity approach: (i) we computed the 95^{th} percentile of the daily outage distribution on the Eversource service territory, and, in agreement with utility managers, defined storm days as the days that exceeded this percentile; (ii) we applied the storm classifier (introduced in Section II-D2) to the Automated Surface Observing System (ASOS) airport station data, to include the storms with significant weather, but whose outages were not in the top 5^{th} percentile. Based on these approaches we identified 76 extratropical storms and 44 convective storms, which were simulated using the WRF model, and we aggregated such data with the other input data for the OPM to create a storm analysis dataset. The total service territory outages per storm in the database range between 24 and 3,590 for extratropical storms, and between 92 and 1,042 for thunderstorms or convective events, following a log-normal distribution of outages for the territory and a general negative binomial distribution at town and grid cell level totals (not shown). Most outages occurred over densely populated Southwestern Connecticut (Figure 2a and 2b).

For having an overview of the most vulnerable areas, we removed the dependence of outages on infrastructure (Figure 2c and 2d). Specifically, we normalized the outages occurring in each town by the sum of the assets in that territory, and we found differences in the impact between extratropical and convective storms. Since many extratropical storms lie over the ocean, major outages were concentrated across the coastal area, mainly in the southwestern coastal region (with more than two outages per thousand assets in an average extratropical storm) and in high-elevation areas in northeastern Connecticut (Figure 2c). The impact of convective storms, in contrast, was inland (Figure 2d), especially in the highly vegetated areas of northeastern and northwestern Connecticut. Despite this statistical information on the average storm impact is not used by the OPM, it provides a descriptive overview of the spatial characteristics of the interactions between infrastructure and different storm types.

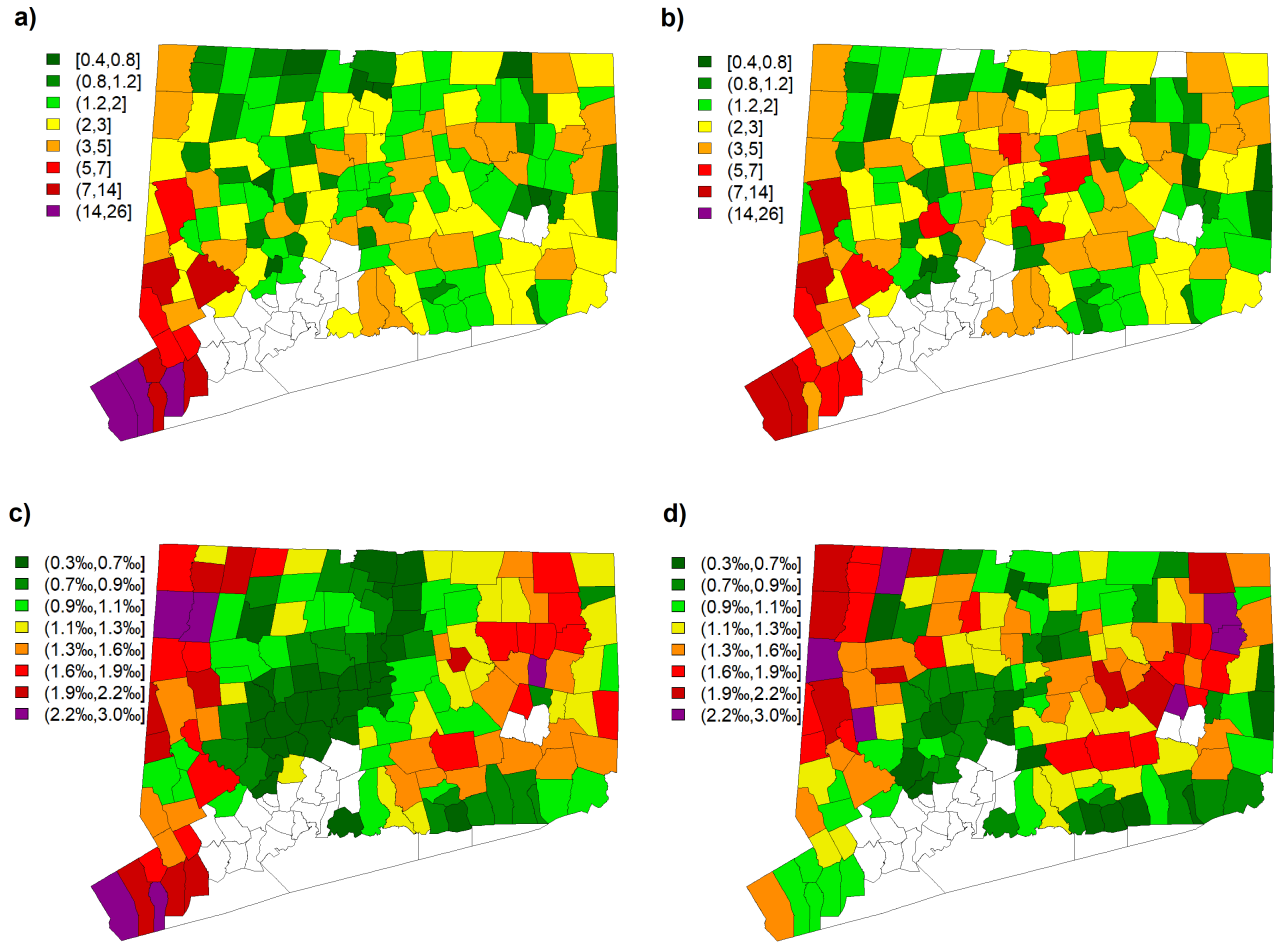


FIGURE 2. **Mean storm outages by town for: a) extratropical, b) convective storms. Mean outages per storm per assets by town for c) extratropical, d) convective storms.**

C. MACHINE LEARNING MODELS

For training the nonparametric models used in the OPM, we used historical storm analyses, composed of a summary of weather hindcasts, land cover, LAI, electric utility assets and historical outages for a range of storms with substantial impact within Connecticut. The models allow to predict the number of future (using weather forecasts) or past (using weather hindcasts) power outages during a storm, at the same 2 km grid used for the input variables.

To limit the effect of correlated predictors, input data were preprocessed by performing a principal component analysis (PCA [67], [68]) using a varimax rotation [69], and by keeping the first 10 components, that explained between 85% and 95% of the model variability in our datasets. The machine learning models used in the OPM are the following:

- Decision Tree (DT): a series of decision nodes ('if-then' statements) that, starting from a 'root node', allows to recursively split the training data into subsets of similar values for the response variable (outages) [70]. In this work we use the CART [31] method, that allows a generation of only two "branches" from each decision node. Dataset partitioning occurs through the minimization of the sum of square error (SSE).
- Gradient Boosted Tree (BT) [37]: a model that fits small decision trees "weak lerners" on the residuals of decision trees of fixed size "base learners". The residuals are the deviations of the observed values from the mean value of each partition. Since the trees fitted on the residuals may lead to overfitting, a learning rate is used to reduce the corrections obtained by such trees. Similarly to W15, one thousand trees were fitted, with a learning rate of 0.1.
- Random Forest (RF) [27]: a collection of decision trees trained on a random subset of training data, using a random subset of predictors. Similarly to W15, two hundred trees were created for our random forest model.
- Ensemble regression (ENS): a decision tree model trained on the DT, BT and RF outputs to refine their predictions. The ENS attempts to find new patterns in the ML outputs, which may yield better results than any single model. The DT, BT and RF outputs are the only three variables that are inputted into a final decision

tree model (ENS). This model was previously found to have better spatial accuracy than individual ML models in W15.

- Bayesian Additive Regression Tree (BART) [38]: it is a sum of regression tree models:

$$Y = \sum_{j=1}^{m} g_j(x, T_j, M_j) + \epsilon. \quad (1)$$

where $g_j(x; T_j, M_j)$ is the contribution provided to the model by the j^{th} regression tree; x is the vector of predictors; T_j and M_j are respectively the set of tree nodes and of terminal nodes for the j^{th} tree; and ϵ is the variance component, assumed to be $N(0, \sigma^2)$. We set a number of trees $m = 30$, and starting from prior specifications for T_j, M_j and σ, we used ten thousand iterations in the Markov Chain Monte Carlo (MCMC [71]) algorithm to reach convergence. Once convergence was reached, we used four thousand iterations to obtain the predictions. A similar configuration was used in H17.

Model outputs, in the form of outage predictions, are either aggregated at the town level and presented as power outage maps, or aggregated at the distribution company level, and presented as total outages on the territory. Since discrepancies in the predictions were found among the different ML models, in the new OPM version, the predictions of the five models are post-processed through a statistical optimization (Section II-D). After the final outage predictions are generated, they are communicated to emergency response personnel, who use this information to develop preparedness measures for facing the incoming storm and to increase the resilience of the affected territories through more accurate planning and an estimation of the time to restoration [72], [73].

D. NEW MODULES

Three new modules have been added to the OPM framework and are presented in this work: the LAI was added to take into account of the interactions between power lines and vegetation; the storm classifier was introduced to allow the ML models to be calibrated with different variables for different storm types; and an optimization technique was implemented for improving OPM results based on past performance of the different ML models.

1) LEAF AREA INDEX

Following the recommendations of H17, we created an algorithm for implementing a quality controlled LAI information in the OPM (Figure 3). The LAI variable allows to consider the variability of power outage occurrence created by the varying amount of leaves that are on the trees. The algorithm was necessary because the original NEO dataset contained a substantial amount of missing data, primarily due to clouds and interpolation effects, and its use was indicated for basic analysis or trend detection.

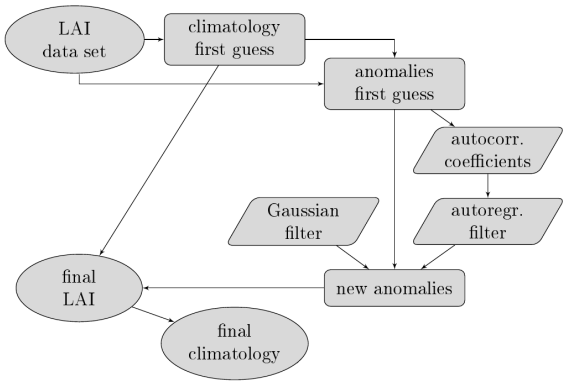


FIGURE 3. Algorithm for LAI data processing and correction.

The first step of the LAI data processing algorithm was the formulation, for each eight-day period, of a first guess at the climatological LAI values and anomalies for the region of interest. The missing data present in the original dataset were not considered in the computation of climatological values. The second step consisted in the estimation of the missing values by using the temporal autocorrelation function of the first guess climate anomalies (departures from the first guess climatology) of the dataset. We found the LAI values, at any point, to be (weakly) correlated at a 0.05 significance level with the LAI values measured in the previous eight-day period; but no significant correlation occurred for a distance of two or more time steps. We estimated the anomaly for each missing point, by using a combination of an autoregressive model of order 1 -AR(1), by considering both previous and successive values - and a Gaussian filter on the valid data. The new anomalies were added to the first guess of the climatology to compute the corrected LAI values, whose missing data were compensated for because in locations for which the LAI could not be retrieved using neighboring values, the value of the first guess at the climatology was used. From this new LAI dataset we computed the new eight-day climatological values of LAI that are used in the OPM, as well as the new anomalies.

2) STORM CLASSIFIER

Weather prediction accuracy as well as the interactions among weather parameters, vegetation and infrastructure, vary for distinct storm types. In order to consider the different errors occurring in both meteorological analysis and forecasts for different types of weather, and to allow the models to use only the variables that successfully describe the processes involved for each type, we introduced a storm classification, grouping events to extratropical and thunderstorms. In future model developments we will also introduce hurricanes/tropical system and snow/ice events.

Following [74] and [75], the storms hitting the considered territory were classified into four different categories, according to the dominant meteorological conditions and to the scale of the processes involved:

- Convective storms at the mesoscale: events of short duration (minutes to a few hours) characterized

by lightning, high rain rate, strong wind gusts, and, in some cases, hail. These thunderstorms may either be embedded in a frontal system (in any season) or isolated and created by diurnal heating of the surface (usually in summer and late spring).
- Extratropical storms at the synoptic scale: extratropical storms exhibiting heavy rain and winds and lasting from several hours to a few days. Such events in our database are most frequent during the fall.
- Snow and ice storms: events characterized by wintry precipitation, such as snow, sleet, and freezing rain, and often associated with strong winds.
- Hurricanes: tropical cyclones characterized by long duration of strong sustained winds and wind gusts, and high precipitation rates.

We created an algorithm for retrospectively classifying all the storms in the database using surface meteorological observations and forecasts valid 24 hours prior to the storm start time. The algorithm used the following criteria:
- If the storm produced one or several tornadoes (identified by the National Weather Service), the event was not included in any data set, because the short-range weather hindcasts and the forecasts analyzed lacked tornado predictability.
- If the storm was classified as a ''hurricane'' or ''tropical storm'' by the National Hurricane Center, we applied the same classification in the OPM. Two hurricanes have hit New England in the last 10 years: Irene (August 2011) and Sandy (October 2012). They were extensively studied in W15 and in [13].
- If the storm produced a water equivalent of at least 5 mm of wintry precipitation over at least 25 percent of the territory, we classified it as a snow or ice storm. Wintry precipitation was not included in this study.
- If wind gusts of at least 13 m/s lasting less than five consecutive hours associated with precipitation of at least 10 mm in one hour were measured or predicted, or if the event was officially reported as ''thunderstorm'' at at least one airport station, and if gusts exceeding 13 m/s for more than five hours were not measured or predicted at all the other airport stations, we classified the storm as convective at the mesoscale. We classified 44 thunderstorms with these characteristics into this category.
- If the above conditions were not satisfied, and wind gust exceeding 17 m/s were measured or predicted in any location, we classified the storm as extratropical at the synoptic scale. We selected 76 extratropical storms for this study.

This algorithm was not only used for classifying historical storms, but is also operationally used for triggering the OPM. The algorithm replaces the storm categorization algorithm based on the time of the year, introduced in W15 and H17. In that algorithm the absence of information on the condition of tree leaves was partially compensated for by the categorization into ''cold season'' (having little or no leaf coverage), ''warm season'' (having high leaf coverage) and ''transition period'' (having intermediate leaf coverage). However, despite the occurrence of predominant processes across different seasons climatologically reflects the storm type categorization, different storm types can occur in each season, which limits the ability of seasonal categorization to capture the unique storm characteristics.

The foundation of the storm classification was based on the fact that different variables are important for different storm types for physical (predominant processes) and model structure (weather model inaccuracy) reasons. While it could be argued that the inclusion of all weather model output variables in the OPM could lead to better performance, empirical evidence (Section IV-A) demonstrates that spurious interactions may lead to overfitting.

3) OPTIMIZATION

We introduced a module for optimally combining the ML model predictions and providing the best estimate of the outage prediction, by taking into account historical model performances. The outage prediction F_k for the k^{th} storm was obtained as a linear combination of the five non-parametric models predictions ($p_{i,k}$):

$$F_k = \sum_{i=1}^{5} c_{i,k} \cdot p_{i,k} \tag{2}$$

The coefficients $c_{i,k}$'s were computed by maximizing an objective function with the characteristics of robustness for different orders of magnitude and of absence of overweightness of some statistics with respect to others. The oversensitivity to very damaging storms rendered inadequate the results produced by the least squares minimization, hence we introduced the ratio U between the mean absolute percentage error (MAPE) and the Nash-Sutcliffe coefficient (NSE [76]):

$$U = \frac{NSE}{MAPE} \tag{3}$$

where the NSE (defined in Appendix) is a nondimensional measure of efficiency, ranging between $-\infty$ and 1, that determines the magnitude of the residual variance of a regression analysis relative to the initial, observed variance; and the MAPE is defined in Appendix. The estimation of the coefficients $c_{i,k}$ for each storm is performed by considering only the remaining storms. Such coefficients are used for predicting outages related to the excluded storm.

The function U allowed us to simultaneously optimize two of the statistics (NSE, MAPE) indicated as most important by Legates and McCabe [77], while taking into account the main characteristic of the dataset: that the storm total outages follow a strongly right-skewed log-normal distribution (not shown). For this reason the overfitting toward strong events obtained by the maximization of the NSE was balanced by the high weight given to the many low impact events, which was obtained with the denominator of the function.

III. METHODOLOGY

This section describes the methodology used for demonstrating outage prediction skills improvements brought about by the new modules implemented into the OPM for the Eversource Connecticut service territory.

A. CROSS-VALIDATION

We used the historical storm dataset described in Section II-B to conduct leave-one-(storm)-out-cross-validations (LOOCV [78]–[80]) experiments in presence and absence of the implemented modules.

The LOOCV at the storm level consists in predicting outages of any storm by calibrating a model using all the storms present in the historical dataset except for the considered storm. We chose this method because both the k-fold cross-validation and the repeated random holdout validation - e.g. [81] for a complete overview - used in W15 and H17, respectively, do not reflect the scope needed in an outage prediction model - that is, the ability to forecast damages from a storm whose outages are originally unknown at any location. Both k-fold and repeated randomized holdout may use strongly correlated neighboring data for the storm of interest in the model calibration. In a LOOCV framework, on the other hand, the knowledge of weather-related outages in neighboring areas from the same storm cannot be used. Moreover, results obtained using a LOOCV are directly comparable with operational model settings, with the unique difference that a storm forecast is used in place of the analysis.

B. EXPERIMENTAL SETUP

The LOOCV experiments on the storm dataset were organized as follows:

i) We tested the skills of the five machine learning models (BART, ENS, RF, DT, and BT) in the same configuration used in W15 and H17 (with a seasonal categorization).

ii) We implemented the optimization module and created the *OPT* model. We quantified the improvements brought about by the optimization by comparing the *OPT* skills with the skills of each ML model.

iii) We additionally implemented the the LAI module and created the *OPT_LAI* model. We compared the *OPT* model skills with the *OPT_LAI* ones.

iv) For testing the importance of the implementation of the storm classifier, we designed 20 LOOCV experiments (Table 3), that allowed us to evaluate the sensitivity of the model to different categories of input variables (Table 1). In category *(a), land cover and assets*, we included static variables related to infrastructure and land cover. In groups *(b), wind, (c), gust, and (d), precipitation*, we included all the variables outputted from the weather model or computed during a post-processing that were, respectively wind-, gust-, or precipitation-related. In group *(e), near-surface*, we included the remaining important weather variables computed below the ground or at the surface. Finally, for demonstrating the improvements from adding the vegetation annual variability, we created category *(f), leaves*, with LAI as its unique variable. The only group common to all sensitivity experiments was *(a), land cover*, since W15 demonstrated that the variables in this category were significant for the OPM. Experiments 11-20 corresponded to 1-10, with the addition of *(f), leaves*. Each odd-numbered experiment differed from the even-numbered one after it in that *(c), gusts*, were added to the latter. Experiments 1-8 and 11-18 inclued *(b), wind*, while 9, 10, 19, and 20 did not. Experiment 18 was the most complex, containing all the predictors listed in Table 1.

TABLE 3. Sensitivity tests.

Experiment	Group	Experiment	Group
1	a, b	11	a, b, f
2	a, b, c	12	a, b, c, f
3	a, b, d	13	a, b, d, f
4	a, b, c, d	14	a, b, c, d, f
5	a, b, e	15	a, b, e, f
6	a, b, c, e	16	a, b, c, e, f
7	a, b, d, e	17	a, b, d, e, f
8	a, b, c, d, e	18	a, b, c, d, e, f
9	a, d, e	19	a, d, e, f
10	a, c, d, e	20	a, c, d, e, f

v) We then selected the two best performing models (one with and another without LAI) for each storm class, and compared their skills with the *OPT* model skills, that used a seasonal categorization.

vi) We also computed the aggregated (extratropical and convective) skills of the new model that has all the implementations (*OPT_LAI_CLA*) and compared with the skill of both the *OPT_LAI* model and of the *OPT_CLA* model, selected from experiments 1-10.

C. VERIFICATION METRICS

We computed the error metrics described in Appendix for evaluating model skills, and the *percent improvement* (PI) for comparing changes in an error metric M between different model versions, namely 1 and 2:

$$PI = \begin{cases} 100 \cdot \dfrac{M_2 - M_1}{M_1} & \text{if higher values of M correspond to better performances} \\ 100 \cdot \dfrac{M_1 - M_2}{M_1} & \text{if lower values of M correspond to better performances} \end{cases} \quad (4)$$

We used the Taylor [82] diagram for evaluating model performances. The Taylor diagram uses the relationship among NCRMSE, NSD, and r to plot these three statistics in a standardized quadrant.

Since outage data are spread over different orders of magnitude, however, "correlation-based measures are

inappropriate and should not be used to evaluate the goodness-of-fit of model simulation'' [77], because these measures are likely to be sensitive to extreme values. Legates and McCabe [77] also recommend reporting ''the observed and modeled means and standard deviations, as well as MAE or RMSE (and probably both)'', despite indicating the coefficient of efficiency as the most appropriate measure. For this reason, we chose to present and compare several statistics to find the best overall model.

In evaluating the significance of the results, it was not possible to apply common statistical techniques for error evaluation. In fact, the strong non-linearities in the model did not allow the error to be Gaussian. Hence, we considered five different pairs of cross-validations performed with the same input data and model setting. In comparisons within the pairs, each cross-validation was found not to differ more than 3% from each other on any error metric. For this reason, despite the limits imposed by the long computational time on the number of experiments we could perform, we assumed a difference of less than 3 percent between two statistics as representative of similar results while considering a greater difference as different.

IV. RESULTS AND DISCUSSION

In this section we present the results obtained by progressively implementing and evaluating each new module into the OPM. Results were rigorously validated using sensitivity tests in a LOOCV framework on the historical storm events. The new OPM version, obtained after implementation of all the new modules, was further evaluated in an operational framework, using the weather forecasts available the day before each storm. This last step was necessary because many decisions on the crews needed for restoration are taken one day before each storm, since this is the tradeoff between weather forecast accuracy and adequate preparedness. However, the possible presence of errors in the short-term forecasts makes performance evaluation using the weather hindcasts alone inadequate for describing the operational performance.

A. SENSITIVITY TESTS

The original, operational OPM version, described in W15 and H17, was composed of five different ML models that generated separate outage forecasts. We evaluated these models on the 120 storms that are the object of this study. Results of the model evaluation showed MdAPEs between 57% (BART) and 88% (RF), MAPEs between 117% (BT) and 160% (RF), and NSE between 0.17 (DT) and 0.45 (BART) (Table 4, first box). The typical feature of some of these models was a remarkable overestimation of low impact events, as shown in W15 and H17, which led to very high values of MAPE.

The information produced by each ML model was optimized using the techinque described in Section II-D3 to create the *OPT* model version, that was operational between June 2016 and January 2017 (Table 4, second box). The combination between the different ML models in the optimization allowed the reduction of most of the very high errors (up to 10 times) for the low-impact storms, and to correct overall biases. This optimization led to a drastic decrease of the MAPE, which halved to 65%. Analyzing the performance by storm type it is possible to notice that most of its skills came from the extratropical model, since the NSE coefficient of the convective model was not significantly different than zero, despite having lower absolute percentage errors (because of the smaller range).

The addition of the LAI module to the *OPT* model allowed us to create the *OPT_LAI* model, operational between January and June 2017. Such model was characterized by a further performance improvement, estimated between 2% and 9%, with respect to the *OPT* model. However, since we consider two statistics as different when their values differ each other by more than 3% (section III-C), it is not possible to state that the improvement brought about by LAI lead to subsantially different results for all the considered metrics.

The validity of a storm classification for outage prediction purposes was demonstrated through 20 sensitivity tests, described in Table 3. The results of the sensitivity experiments 11-20 (with LAI) are presented in the central part of Table 4, from which, for brevity, we have omitted experiments 1-10 (their aggregated results are summarized at the bottom of the table in the *OPT_CLA* model). Since the table is not easily interpreted, we follow [77] in giving primary importance to the NSE. In cases of similar results, the MdAPE and MAPE complete the evaluation of the results.

In the simulations for extratropical storms, results from experiment 18 (with the inclusion of all the predictor variables) showed the best performance with a NSE of 0.50, a median absolute percentage error of 47 percent, and a mean percentage error of 64 percent. The most important variables for this experiment were: sum of assets, total precipitation, maximum and mean temperature, and mean wind gusts. It is noteworthy that model errors around or above 50 percent derive from an outage distribution spanning more than two orders of magnitude. Similar results for all the metrics were found also in experiment 14 (without near-surface variables), which did not produce any improvement to or setback for the model. Gusts were also important, as strong improvements in the NSE could be seen between each odd experiment and the following even one. Two other relevant variables were wind and precipitation; when wind was removed, the NSE decreased by 20 percent (experiment 20), and the absence of precipitation as predictor variable (experiment 16) implied a 28 percent decrease.

For thunderstorms, the results showed a different pattern. The most skillful experiment was number 15, which used neither gusts nor precipitation and whose NSE was 0.42. For this experiment the most important variables were: sum of assets, maximum specific humidity, maximum and mean temperature, and maximum soil moisture. In all the other thunderstorm simulations, the NSE assumed lower values; simulations 11-14 in particular had no skills (NSE $\leq$ 0.15).

TABLE 4. Summary table for the cross validations who led to model improvements. Bolded values indicate the statistics, unbolded values indicate the percent improvement (PI) with respect to the respective, bolded statistic in the same box, and values in bracket indicate the PI with respect to the previous, bolded model version.

Model	MdAPE	MAPE	NSE			
BART	**57%**	**130%**	**0.45**			
DT	-24%	-21%	-63%			
BT	-9%	+10%	-39%			
RF	-37%	-23%	-29%			
ENS	-19%	-12%	-20%			
OPT	**47%** (+18%)	**65%** (+50%)	**0.47** (+4%)			
OPT_LAI	+2%	+9%	+8%			
	Extratropical			Convective		
Experiment	**MdAPE**	**MAPE**	**NSE**	**MdAPE**	**MAPE**	**NSE**
OPT	**48%**	**80%**	**0.49**	**45%**	**49%**	**0.01**
11	-6%	+2%	-62%	-35%	-43%	-100%
12	+6%	-27%	-14%	-38%	-37%	-100%
13	+4%	+9%	-34%	-45%	-35%	-74%
14	-2%	-3%	+2%	-28%	-32%	-83%
15	+11%	+5%	-48%	**29%**	**46%**	**0.42**
16	+0%	-13%	-28%	+3%	-7%	-17%
17	+21%	+0%	-24%	-3%	-9%	-24%
18	**47%**	**64%**	**0.50**	+14%	-4%	-17%
19	-28%	-41%	-36%	-17%	-9%	-49%
20	-4%	-18%	-20%	-10%	-7%	-17%
Model	**MdAPE**	**MAPE**	**NSE**			
OPT_LAI_CLA	**43%** (+9%)	**59%** (+9%)	**0.53** (+13%)			
OPT_CLA	-2%	-2%	-9%			

This means any model that did not use near-surface parameters was unable to correctly predict outages related to thunderstorms, hence only experiments 15-20 were worth investigating. In experiment 16, the addition of gusts to the best performing model produced a significant decrease in performance in the NSE and MAPE. This could be explained by overfitting in the model (due to particular, hidden features in the limited number of storms) that emerged in these results. The overfitting could be avoided by removing some predictor variables and creating a simpler model. For the same reasons, the implementation of precipitation in the model (in experiments 17 and 18) also failed to improve the performance relative to experiments 15 and 16.

These results highlight the complexity of predicting outages when thunderstorms occur. In fact, while thunderstorms are characterized by heavy precipitation and strong gusts, these variables were not important in the models due to the difficulty of predicting the exact location of the phenomena by the WRF model - hence, the *double-penalty effect* predominated. More specifically, errors were introduced by the misplacing of the typical thunderstorm features and this penalized the model twice - once for not predicting the event in the location where it occurred (miss) and once for forecasting it at the wrong location (false alarm) [83]. The double-penalty effect, derived from important errors in thunderstorm forecasts and analyses, is reflected in the creation of models overfitted on some meteorological features that have been artificially created by the weather simulations, but which did not manifest. Hence, when the model was validated in the LOOCV, the overfitting emerged as a decrease in performance. Despite these limitations, however, the thunderstorm model had lower MAPE and MdAPE with respect to the model for extratropical storms, because of the lower range of thunderstorm-related outages.

Gaining an understanding of the strengths and weaknesses of the models for different storm types allowed us to create the new OPM version (*OPT_LAI_CLA*), operational since June 2017. This model allows us to use the most suitable set of variables for each storm type. The combined introduction of the classification and LAI modules led to a 13% improvements in the NSE and 9% in both MdAPE and MAPE, over the *OPT* version. When comparing this new OPM version with the BART model, performance improvements are drastic, varying between 17% and 63%, depending on the statistics. As the Taylor diagram (Figure 4c) shows, the NCRMSE decreased from 0.79 of the BART to 0.61 of the new OPM, and the correlation increased from 0.63 to 0.79.

Consequently, the new model allowed us to capture the order of magnitude of outages (the dashed lines in Figure 4a and 4b) for 95 percent of the storms and to predict 65 percent of the storms within a 50 percent error (thick red lines in Figure 4b). This constituted a big step toward the milestone

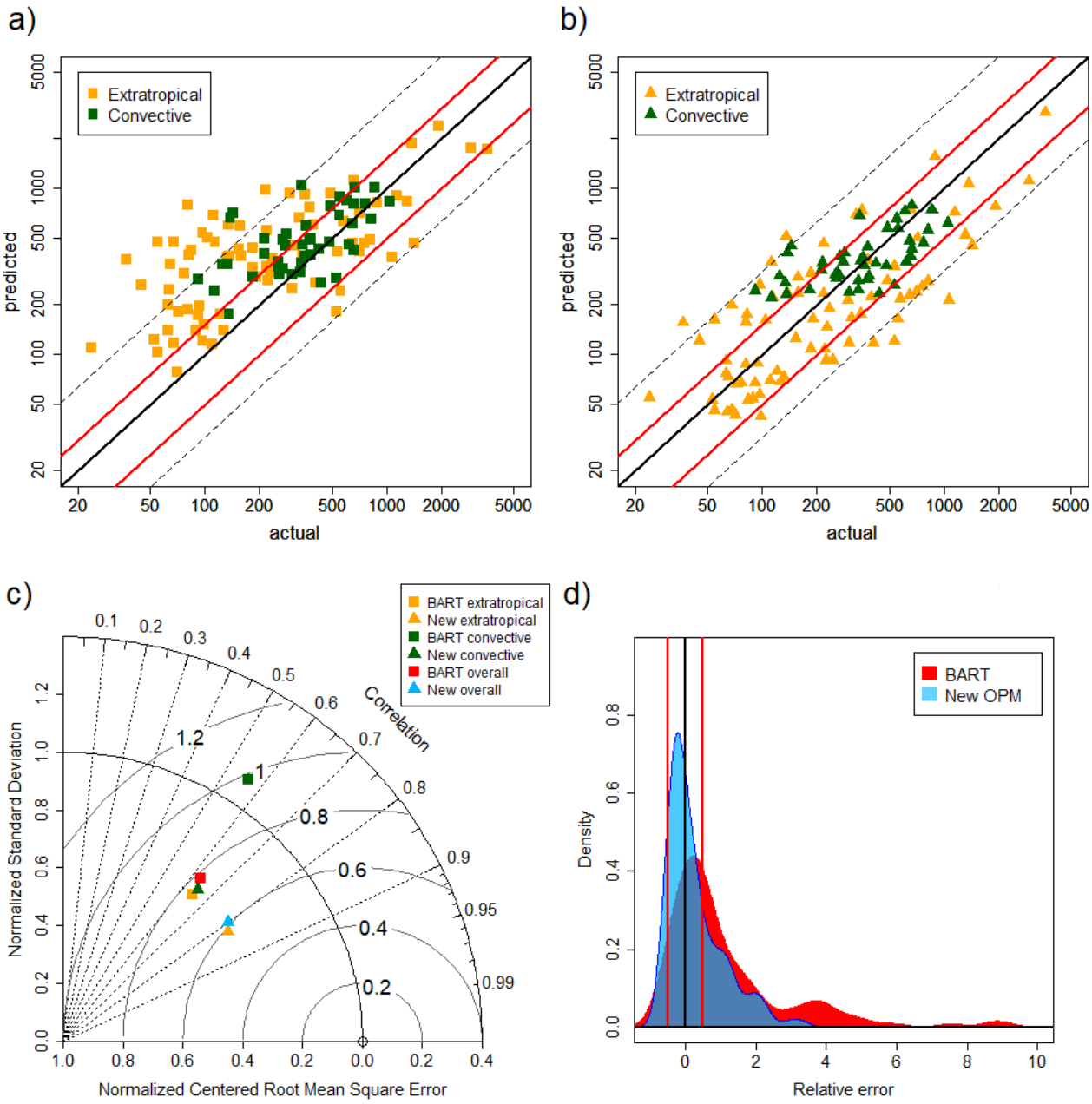


FIGURE 4. Scatter plots of predicted versus actual outages for a) the BART and b) the new OPM versions. c) Taylor diagram comparing the performances for Extratropical and Convective storms for the BART and new OPM. d) Density plot of storm total relative errors for the BART (red) and the new OPM (blue) versions. Red lines represent a 50% error, while dashed lines delimit an error of half order of magnitude.

of predicting 90 percent of the events within that error and an important improvement over the 51 percent of the BART model (Figure 4a). From the analysis of the relative errors, it emerged that the new model has a much narrower and more centered distribution than the BART (Figure 4d). The main feature of the BART was the overestimation of medium and low impact events, that manifested in false alarms. The increased predictability of such events is noteworthy, since the decision-making process involves determination of the number of in- and out-of-state crews, crews travel time and associated costs, expected duration of the restoration process, and prevention from eventual penalities that utilities may face. False alarms may trigger costly prevention measures such as crews and resources allocation, and a high number of false alarms threatens the credibility of the model.

Finally, we again evaluated the importance of the LAI in the new OPM by analyzing experiments 1-10 (not shown) and summarizing the performance difference between the new OPM with LAI and without (last row of Table 4). After the removal of this variable, the performance decrease (between 2 percent and 9 percent) was consistent with that found for the previous model version, consolidating, but not proving,

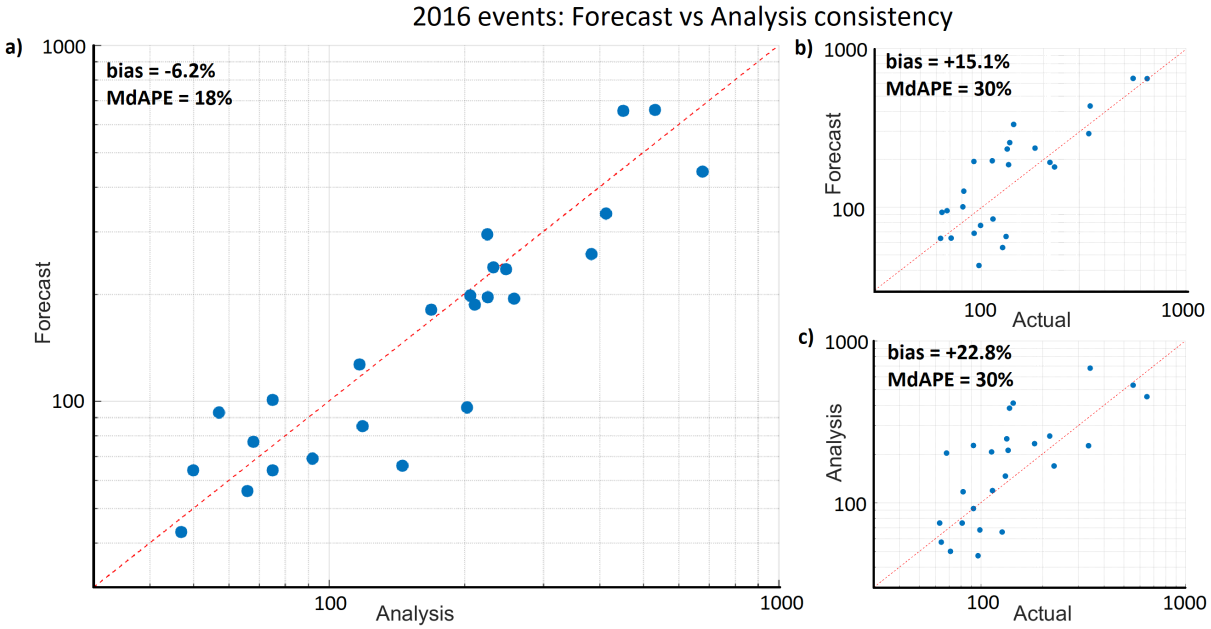


FIGURE 5. Plot of storm total outage predictions using storm forecasts versus outage predictions using analyses for storms that occurred in Connecticut during 2016 (a); plots of outage predictions using storm forecasts (b) and analyses (c) versus actual outages, for the same storms of (a).

the hypothesis that the LAI is an important variable for outage prediction.

Hence, the analysis of several statistical measures of error proved that the new modules implemented in the OPM allowed drastic outage prediction improvements over each ML model introduced in W15 and H17.

B. OPERATIONAL PERFORMANCES

The outage prediction framework developed with the storm analyses would be a mere numerical exercise if a consistency between weather hindcasts and forecasts (that is, the absence of a significant bias) did not hold. To evaluate the hypothesis of consistency between outages predicted using storm forecasts and those predicted using storm analysis, we considered the events that occurred in 2016. From the comparison of the outage predictions using forecasts and analyses, shown in Figure 5a, it emerged that the median absolute percentage difference among each couple of forecast-analysis (18 percent) was much smaller than the MdAPE of the forecasts and of the analyses versus the actual outages (30% in both cases, Figure 5b and 5c).

Hence, the errors deriving from the weather model uncertainties between forecasts and analyses were a small part of the total OPM error. The major contribution to the error came from consistent inaccuracies of the weather model and from the statistical models. The biases for the 2016 storms for forecasts and analyses with respect to the actual data are 15 and 23 percent, respectively. The fact that these values are much higher than the bias between forecasts and analyses (-6.2 percent) strengthens the finding of consistency. Furthermore, the positive values can be explained with the outage model underdispersion, which produces overestimation for medium-to-low impact events.

V. CONCLUDING REMARKS

This study presented three new modules that were implemented in an existing outage prediction model and evaluated the new model predictions and improvements based on 76 extratropical and 44 convective storms. The optimization module introduced hereby led to a version that significantly outperformed each of the five machine learning models comprising the OPM. Further improvements - the addition of the leaf area index and the weather classification with a consequent variable selection - allowed the new OPM to perform with a Nash-Sutcliffe efficiency coefficient of 0.53 and to predict most of the events with a service-territory-total-outages maximum error of 43 percent. Extratropical storms, associated with a more extended historical dataset, exhibited a higher Nash-Sutcliffe coefficient than convective storms, but also a higher median absolute percentage error, because of the much wider range of outages.

We also presented operational OPM performance using the weather forecasts for a subset of the analyzed events and compared it with performance using weather hindcasts. Comparison at the service territory level showed that the high correlation and low bias between the predicted outages in the two configurations exhibited the validity of the current operational framework, which consists of calibrating the model using weather hindcasts and predicting outages using consistent weather forecasts. Errors in both forecasts and analyses can be attributed to a random component, as the misplacing of weather phenomena or to a systematic component dependet

on orography, infrastructure resilience measureas and other local conditions.

We suggest further improvements for addresing the double-penalty effect, which played an important role in the random component for the thunderstorm model: these are the application of a spatial translation and matching of the forecast and observed fields [84]; the neighborhood method for upscaling the fields [85]–[87], or a combination of the two. The upscaling would imply an evaluation of the outage prediction obtained by aggregating the predictors at different resolutions to find the best resolution for each storm type.

Future steps also include the exploration of probabilistic outage forecasts, based on ensemble numerical weather predictions as input. Since the machine learning models use weather data as inputs and weather data have their own errors, providing a probabilistic forecast based on machine learning model uncertainty only would not be sufficient for capturing the actual variability of the predictions. A probabilistic framework based on weather ensemble forecasts would be the best candidate to capture the actual predictive uncertainty.

Finally, we would like to mention that, since the OPM is operational and emergency response personnel partially base their decisions on this outage prediction model, both the reduction of the model errors demonstrated in this paper, and future improvements have immediate social and economic benefits. Accurate outage predictions allow emergency managers to allocate the correct number of crews and equipment for preparing for upcoming storms. This results in shorter outage times, and reduces restoration costs. When the impact of a weather event is underestimated, higher restoration costs are associated with the need of calling out-of-state crews after the event. Moreover, crews travel time would delay the restoration process leading to longer outages. Higher costs are also associated with overestimation of a storm's impact, due to excessive crew allocation, therefore unnecessary restoration cost.

APPENDIX
ERROR METRICS

The error metrics used in the paper are listed below:

- The *median absolute percentage error* (MdAPE): maximum percentage error that can be committed 50 percent of the times.
- The *mean absolute percentage error* (MAPE):

$$MAPE = \frac{1}{n}\sum_{k=1}^{n}\left|\frac{F_k - O_k}{O_k}\right| \tag{5}$$

- The *Pearson product-moment correlation coefficient* (r): linear correlation between predictions and observations.
- The *normalized standard deviation* (NSD): ratio between the standard deviation of the model and the standard deviation of the observations.
- The *normalized centered root mean squared error* (NCRMSE) measures the random component of the error:

$$NCRMSE = \sqrt{\frac{\sum_{k=1}^{n}\left[(F_k - \bar{F}) - (O_k - \bar{O})\right]^2}{\sum_{k=1}^{n}\left[(O_k - \bar{O})\right]^2}} \tag{6}$$

- the *Nash-Sutcliffe Efficiency* (NSE [76]), a nondimensional measure of efficiency, ranging between $-\infty$ and 1, that determines the magnitude of the residual variance of a regression analysis relative to the initial, observed variance:

$$NSE = 1 - \left[\frac{\sum_{k=1}^{n}(F_k - O_k)^2}{\sum_{k=1}^{n}(O_k - \bar{O})^2}\right] \tag{7}$$

where, for all the metrics, n is the total number of storms, F_k and O_k are respectively the k^{th} prediction and observation, and $\bar{O}$ the mean observed data.

ACKNOWLEDGEMENT

This work was entirely performed at the Eversource Energy Center at the University of Connecticut. This publication partially uses classified datasets of the electric grid. The authors hold stock in ACW Analytics. They have full access to all of the data in this study and take complete responsibility for the integrity of the data and the accuracy of the data analysis.

REFERENCES

[1] D. A. Jones, "Electrical engineering: The backbone of society," *IEE Proc. A Sci., Meas. Technol.*, vol. 138, no. 1, pp. 1–10, Jan. 1991.
[2] P. Schavemaker and L. V. der Sluis, *Electrical Power System Essentials*. Hoboken, NJ, USA: Wiley, 2008.
[3] M. Munasinghe and A. Sanghvi, "Reliability of electricity supply, outage costs and value of service: An overview," *Energy J.*, vol. 9, pp. 1–18, May 1988.
[4] M. Schmidthaler and J. Reichl, "Assessing the socio-economic effects of power outages ad hoc," *Comput. Sci.-Res. Develop.*, vol. 31, no. 3, pp. 157–161, 2016.
[5] I. Sánchez, "Short-term prediction of wind energy production," *Int. J. Forecasting*, vol. 22, no. 1, pp. 43–56, 2006.
[6] M. Zamo, O. Mestre, P. Arbogast, and O. Pannekoucke, "A benchmark of statistical regression methods for short-term forecasting of photovoltaic electricity production, part I: Deterministic forecast of hourly production," *Sol. Energy*, vol. 105, pp. 792–803, Jul. 2014.
[7] C. Wan, Z. Xu, P. Pinson, Z. Y. Dong, and K. P. Wong, "Probabilistic forecasting of wind power generation using extreme learning machine," *IEEE Trans. Power Syst.*, vol. 29, no. 3, pp. 1033–1044, May 2014.
[8] J. W. Taylor, "Short-term electricity demand forecasting using double seasonal exponential smoothing," *J. Oper. Res. Soc.*, vol. 54, no. 8, pp. 799–805, 2003.
[9] J. W. Taylor, L. M. de Menezes, and P. E. McSharry, "A comparison of univariate methods for forecasting electricity demand up to a day ahead," *Int. J. Forecasting*, vol. 22, no. 1, pp. 1–16, 2006.
[10] T. Hong, P. Wang, and L. White, "Weather station selection for electric load forecasting," *Int. J. Forecasting*, vol. 31, no. 2, pp. 286–295, 2015.
[11] T. Hong and S. Fan, "Probabilistic electric load forecasting: A tutorial review," *Int. J. Forecasting*, vol. 32, no. 3, pp. 914–938, 2016.
[12] H. Liu, R. A. Davidson, D. V. Rosowsky, and J. R. Stedinger, "Negative binomial regression of electric power outages in hurricanes," *J. Infrastruct. Syst.*, vol. 11, no. 4, pp. 258–267, 2005.
[13] D. W. Wanik *et al.*, "A case study on power outage impacts from future hurricane sandy scenarios," *J. Appl. Meteorol. Climatol.*, vol. 57, no. 1, pp. 51–79, 2018.
[14] D. Chapman, "The cost of poor power quality," in *Leonardo Power Quality Application Guide*. Washington DC, USA: Copper Development Association, Mar. 2001. [Online]. Available: http://copperalliance.org.uk/uploads/2018/03/21-the-cost-of-poor-power-quality.pdf

[15] C. Klinger, V. Murray, and V. Landeg, "Power outages, extreme events and health: A systematic review of the literature from 2011–2012," *PLoS Currents*, vol. 6, Jan. 2014. doi: 10.1371/currents.dis.04eb1dc5e73dd1377e05a10e9edde673.
[16] G. B. Anderson and M. L. Bell, "Lights out: Impact of the August 2003 power outage on mortality in New York, NY," *Epidemiology*, vol. 23, no. 2, pp. 189–193, 2012.
[17] D. T. Ton and W.-T. P. Wang, "A more resilient grid: The U.S. department of energy joins with stakeholders in an R&D plan," *IEEE Power Energy Mag.*, vol. 13, no. 3, pp. 26–34, May/Jun. 2015.
[18] R. K. Pachauri *et al.*, *Climate Change 2014: Synthesis Report. Contribution of Working Groups I, II and III to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change*. Geneva, Switzerland: IPCC, 2014.
[19] R. J. Campbell, "Weather-related power outages and electric system resiliency," Congressional Res. Service, Washington, DC, USA, Tech. Rep., 2012.
[20] C. Light and Power, "Transmission and distribution reliability performance report," Connecticut Light Power, Berlin, CT, USA, Tech. Rep., 2014.
[21] D. Lubkeman and D. E. Julian, "Large scale storm outage management," in *Proc. IEEE Power Eng. Soc. Gen. Meeting*, Jun. 2004, pp. 16–22.
[22] A. C. Cameron and P. K. Trivedi, *Regression Analysis of Count Data*, vol. 30. Cambridge, U.K.: Cambridge Univ. Press, 1998.
[23] H. Liu, R. A. Davidson, and T. V. Apanasovich, "Spatial generalized linear mixed models of electric power outages due to hurricanes and ice storms," *Rel. Eng. Syst. Saf.*, vol. 93, no. 6, pp. 897–912, 2008.
[24] S.-R. Han, S. D. Guikema, S. M. Quiring, K.-H. Lee, D. Rosowsky, and R. A. Davidson, "Estimating the spatial distribution of power outages during hurricanes in the Gulf coast region," *Rel. Eng. Syst. Saf.*, vol. 94, no. 2, pp. 199–210, 2009.
[25] T. J. Hastie and R. J. Tibshirani, *Generalized Additive Models*. Hoboken, NJ, USA: Wiley, 1990.
[26] S.-R. Han, S. D. Guikema, and S. M. Quiring, "Improving the predictive accuracy of hurricane power outage forecasts using generalized additive models," *Risk Anal.*, vol. 29, no. 10, pp. 1443–1453, 2009.
[27] L. Breiman, "Random forests," *Mach. Learn.*, vol. 45, no. 1, pp. 5–32, 2001.
[28] R. Nateghi, S. Guikema, and S. M. Quiring, "Power outage estimation for tropical cyclones: Improved accuracy with simpler models," *Risk Anal.*, vol. 34, no. 6, pp. 1069–1078, 2014.
[29] S. D. Guikema and S. M. Quiring, "Hybrid data mining-regression for infrastructure risk assessment based on zero-inflated data," *Rel. Eng. Syst. Saf.*, vol. 99, pp. 178–182, Mar. 2012.
[30] S. M. Quiring, A. B. Schumacher, and S. D. Guikema, "Incorporating hurricane forecast uncertainty into a decision-support application for power outage modeling," *Bull. Amer. Meteorol. Soc.*, vol. 95, no. 1, pp. 47–58, 2014.
[31] L. Breiman, J. Friedman, C. J. Stone, and R. A. Olshen, *Classification and Regression Trees*. Boca Raton, FL, USA: CRC Press, 1984.
[32] D. B. McRoberts, S. M. Quiring, and S. D. Guikema, "Improving hurricane power outage prediction models through the inclusion of local environmental factors," *Risk Anal.*, vol. 38, pp. 2722–2737, Dec. 2018.
[33] S. D. Guikema, R. Nateghi, S. M. Quiring, A. Staid, A. C. Reilly, and M. Gao, "Predicting hurricane power outages to support storm response planning," *IEEE Access*, vol. 2, pp. 1364–1373, 2014.
[34] D. W. Wanik, J. R. Parent, E. N. Anagnostou, and B. M. Hartman, "Using vegetation management and LiDAR-derived tree height data to improve outage predictions for electric utilities," *Electr. Power Syst. Res.*, vol. 146, pp. 236–245, 2017.
[35] D. W. Wanik, E. N. Anagnostou, B. M. Hartman, M. E. B. Frediani, and M. Astitha, "Storm outage modeling for an electric distribution network in northeastern USA," *Natural Hazards*, vol. 79, no. 2, pp. 1359–1384, 2015.
[36] J. He, D. W. Wanik, B. M. Hartman, E. N. Anagnostou, M. Astitha, and M. E. B. Frediani, "Nonparametric tree-based predictive modeling of storm outages on an electric distribution network," *Risk Anal.*, vol. 37, no. 3, pp. 441–458, 2017.
[37] J. H. Friedman, "Greedy function approximation: A gradient boosting machine," *Ann. Statist.*, vol. 29, no. 5, pp. 1189–1232, 2001.
[38] H. A. Chipman, E. I. George, and R. E. McCulloch, "BART: Bayesian additive regression trees," *Appl. Statist.*, vol. 4, no. 1, pp. 266–298, 2010.
[39] N. Meinshausen, "Quantile regression forests," *J. Mach. Learn. Res.*, vol. 7, pp. 983–999, Jun. 2006.
[40] M. S. Jones, B. A. Colle, and J. S. Tongue, "Evaluation of a mesoscale short-range ensemble forecast system over the northeast United States," *Weather Forecasting*, vol. 22, no. 1, pp. 36–55, 2007.
[41] E. N. Lorenz, "Deterministic nonperiodic flow," *J. Atmos. Sci.*, vol. 20, no. 2, pp. 130–141, 1963.
[42] R. H. Langland, M. A. Shapiro, and R. Gelaro, "Initial condition sensitivity and error growth in forecasts of the 25 January 2000 east coast snowstorm," *Monthly Weather Rev.*, vol. 130, no. 4, pp. 957–974, 2002.
[43] G. A. Grell, Y.-H. Kuo, and R. J. Pasch, "Semiprognostic tests of cumulus parameterization schemes in the middle latitudes," *Monthly Weather Rev.*, vol. 119, no. 1, pp. 5–31, 1991.
[44] M. E. B. Frediani, J. P. Hacker, E. N. Anagnostou, and T. Hopson, "Evaluation of PBL parameterizations for modeling surface wind speed during storms in the northeast united states," *Weather Forecasting*, vol. 31, no. 5, pp. 1511–1528, 2016.
[45] W. C. Skamarock *et al.*, "A description of the advanced research WRF version 3," NCAR, Boulder, CO, USA, NCAR Tech. Note NCAR/TN-475+STR, 2008.
[46] W. Wang, C. Bruyère, M. Duda, J. Dudhia, D. Gill, and H.-C. Lin, "ARW version 3 modelling system user's guide," Mesosc. Microsc. Meteorol. Division, Nat. Center Atmos. Res., Boulder, CO, USA, 2010.
[47] N. A. Phillips and J. Shukla, "On the strategy of combining coarse and fine grid meshes in numerical weather prediction," *J. Appl. Meteorol.*, vol. 12, no. 5, pp. 763–770, 1973.
[48] T. L. Clark and R. D. Farley, "Severe downslope windstorm calculations in two and three spatial dimensions using anelastic interactive grid nesting: A possible mechanism for gustiness," *J. Atmos. Sci.*, vol. 41, no. 3, pp. 329–350, 1984.
[49] D.-L. Zhang, H.-R. Chang, N. L. Seaman, T. T. Warner, and J. M. Fritsch, "A two-way interactive nesting procedure with variable terrain resolution," *Monthly Weather Rev.*, vol. 114, no. 7, pp. 1330–1339, 1986.
[50] G. Thompson, P. R. Field, R. M. Rasmussen, and W. D. Hall, "Explicit forecasts of winter precipitation using an improved bulk microphysics scheme. Part II: Implementation of a new snow parameterization," *Monthly Weather Rev.*, vol. 136, no. 12, pp. 5095–5115, 2008.
[51] G. A. Grell, "Prognostic evaluation of assumptions used by cumulus parameterizations," *Monthly Weather Rev.*, vol. 121, no. 3, pp. 764–787, 1993.
[52] G. A. Grell and D. Dévényi, "A generalized approach to parameterizing convection combining ensemble and data assimilation techniques," *Geophys. Res. Lett.*, vol. 29, no. 14, pp. 38-1–38-4, 2002.
[53] S.-Y. Hong, Y. Noh, and J. Dudhia, "A new vertical diffusion package with an explicit treatment of entrainment processes," *Monthly Weather Rev.*, vol. 134, pp. 2318–2341, Sep. 2006.
[54] C. A. Paulson, "The mathematical representation of wind speed and temperature profiles in the unstable atmospheric surface layer," *J. Appl. Meteorol.*, vol. 9, no. 6, pp. 857–861, 1970.
[55] A. J. Dyer and B. B. Hicks, "Flux-gradient relationships in the constant flux layer," *Quart. J. Roy. Meteorol. Soc.*, vol. 96, no. 410, pp. 715–721, 1970.
[56] E. K. Webb, "Profile relationships: The log-linear range, and extension to strong stability," *Quart. J. Roy. Meteorol. Soc.*, vol. 96, no. 407, pp. 67–90, 1970.
[57] A. C. Beljaars, "The parametrization of surface fluxes in large scale models under free convection," *Quart. J. Roy. Meteorol. Soc.*, vol. 121, no. 522, pp. 255–270, 1995.
[58] D. Zhang and R. A. Anthes, "A high-resolution model of the planetary boundary layer—Sensitivity tests and comparisons with SESAME-79 data," *J. Appl. Meteorol.*, vol. 21, no. 11, pp. 1594–1609, 1982.
[59] M. Tewari *et al.*, "Implementation and verification of the unified Noah land-surface model in the WRF model [presentation]," in *Proc. 20th Conf. Weather Anal. Forecasting, 16th Conf. Numer. Weather Predict.*, vol. 1115. Seattle, WA, USA: American Meteorological Society, 2004.
[60] E. J. Mlawer, S. J. Taubman, P. D. Brown, M. J. Iacono, and S. A. Clough, "Radiative transfer for inhomogeneous atmospheres: RRTM, a validated correlated-k model for the longwave," *J. Geophys. Res.*, vol. 102, no. D14, pp. 16663–16682, 1997.
[61] M.-D. Chou, M. J. Suarez, X.-Z. Liang, M. M.-H. Yan, and C. Cote, "A thermal infrared radiation parameterization for atmospheric studies," NASA, Washington, DC, USA, Tech. Rep. NASA/TM-2001-104606/VOL19, 2001.

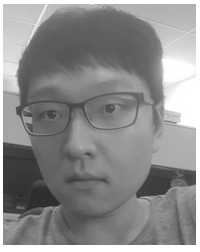
[62] M.-D. Chou and M. J. Suarez, "A solar radiation parameterization for atmospheric studies," NASA, Washington, DC, USA, Tech. Rep. NASA/TM-1999-104606/VOL15, 1999, vol. 15.
[63] University of Connecticut. (2006). *Connecticut's Changing Landscape Land Cover*. Accessed: Dec. 17, 2017. http://clear.uconn.edu/projects/landscape/
[64] D. J. Watson, "Comparative physiological studies on the growth of field crops: I. Variation in net assimilation rate and leaf area between species and varieties, and within and between years," *Ann. Botany*, vol. 11, no. 41, pp. 41–76, 1947.
[65] NASA Earth Observations. (2017). *Leaf Area Index (1 Month-Terra/MODIS)*. Accessed: Dec. 17, 2017. [Online]. Available: https://neo.sci.gsfc.nasa.gov/view.php?datasetId=MOD15A2_M_LAI
[66] NASA. (2017). *NASA's Earth Observing System*. Accessed: Dec. 17, 2017. [Online]. Available: https://eospso.nasa.gov/
[67] A.-L. Cauchy, "Sur l'équationa l'aide de laquelle on dÃlermine les inégalités séculaires des mouvements des planetes," *Exer. Math.*, vol. 4, no. 1829, pp. 174–195, 1829.
[68] F. R. S. K. Pearson, "LIII. On lines and planes of closest fit to systems of points in space," *London, Edinburgh, Dublin Philosoph. Mag. J. Sci.*, vol. 2, no. 11, pp. 559–572, 1901.
[69] H. F. Kaiser, "The varimax criterion for analytic rotation in factor analysis," *Psychometrika*, vol. 23, no. 3, pp. 187–200, 1958.
[70] D. T. Larose and C. D. Larose, *Data Mining and Predictive Analytics*. Hoboken, NJ, USA: Wiley, 2015.
[71] T. Hastie and R. Tibshirani, "Bayesian backfitting (with comments and a rejoinder by the authors," *Stat. Sci.*, vol. 15, no. 3, pp. 196–223, 2000.
[72] D. Wanik, E. Anagnostou, B. Hartman, and T. Layton, "Estimated time of restoration (ETR) guidance for electric distribution networks," *J. Homeland Secur. Emergency Manage.*, vol. 15, no. 1, 2018. doi: 10.1515/jhsem-2016-0063.
[73] T. Walsh, T. Layton, D. Wanik, and J. Mellor, "Agent based model to estimate time to restoration of storm-induced power outages," *Infrastructures*, vol. 3, no. 3, p. 33, 2018.
[74] J. M. Wallace and P. V. Hobbs, *Atmospheric Science: An Introductory Survey*, vol. 92. New York, NY, USA: Academic, 2006.
[75] National Severe Storms Laboratory. (2017). *Severe Weather 101*. Accessed: Dec. 17, 2017. [Online]. Available: http://www.nssl.noaa.gov/education/svrwx101/
[76] J. E. Nash and J. V. Sutcliffe, "River flow forecasting through conceptual models part I—A discussion of principles," *J. Hydrol.*, vol. 10, no. 3, pp. 282–290, 1970.
[77] D. R. Legates and G. J. McCabe, "Evaluating the use of 'goodness-of-fit' measures in hydrologic and hydroclimatic model validation," *Water Resour. Res.*, vol. 35, no. 1, pp. 233–241, 1999.
[78] M. Stone, "Cross-validatory choice and assessment of statistical predictions," *J. Roy. Stat. Soc. B, Methodol.*, vol. 36, no. 2, pp. 111–147, 1974.
[79] D. M. Allen, "The relationship between variable selection and data agumentation and a method for prediction," *Technometrics*, vol. 16, no. 1, pp. 125–127, 1974.
[80] S. Geisser, "The predictive sample reuse method with applications," *J. Amer. Stat. Assoc.*, vol. 70, no. 350, pp. 320–328, 1975.
[81] J. Han, J. Pei, and M. Kamber, *Data Mining: Concepts and Techniques*. Amsterdam, The Netherlands: Elsevier, 2011.
[82] K. E. Taylor, "Summarizing multiple aspects of model performance in a single diagram," *J. Geophys. Res., Atmos.*, vol. 106, no. D7, pp. 7183–7192, 2001.
[83] A. Rossa, P. Nurmi, and E. Ebert, "Overview of methods for the verification of quantitative precipitation forecasts," in *Precipitation: Advances in Measurement, Estimation and Prediction*. Berlin, Germany: Springer, 2008.
[84] E. E. Ebert and J. McBride, "Verification of precipitation in weather systems: Determination of systematic errors," *J. Hydrol.*, vol. 239, no. 1, pp. 179–202, 2000.
[85] J. Zepeda-Arce, E. Foufoula-Georgiou, and K. K. Droegemeier, "Space-time rainfall organization and its role in validating quantitative precipitation forecasts," *J. Geophys. Res., Atmos.*, vol. 105, no. D8, pp. 10129–10146, 2000.
[86] T. Cherubini, A. Ghelli, and F. Lalaurette, "Verification of precipitation forecasts over the alpine region using a high-density observing network," *Weather Forecasting*, vol. 17, no. 2, pp. 238–249, 2002.
[87] E. Yates, S. Anquetin, V. Ducrocq, J.-D. Creutin, D. Ricard, and K. Chancibault, "Point and areal validation of forecast precipitation fields," *Meteorol. Appl.*, vol. 13, no. 1, pp. 1–20, 2006.

DIEGO CERRAI received the B.S. degree in physics from the University of Pisa, Italy, in 2012, and the M.S. degree in physics of the earth system from the University of Bologna, Italy, in 2015. He is currently pursuing the Ph.D. degree with the Department of Civil and Environmental Engineering, University of Connecticut. His primary Ph.D. research focuses on the prediction of the weather-related power outages using machine learning and statistical models. He is studying the dynamical processes involved in the formation of Mediterranean hurricanes and extreme hydro-meteorological events in the Mediterranean region.

DAVID W. WANIK received the B.S. degree in environmental science and the M.S. and Ph.D. degrees in environmental engineering from the University of Connecticut, where he is currently an Adjunct Assistant Professor. He is also a Research Fellow with the Mississippi State University's Social Science Research Center. His major research interests include the areas of data mining and machine learning, natural hazards impact modeling, and business analytics.

MD ABUL EHSAN BHUIYAN received the B.S. degree in water resources engineering from the Bangladesh University of Engineering and Technology, in 2009, the M.S. degree in environmental engineering from the University of Louisiana, Lafayette, in 2013, and the Ph.D. degree in civil and environmental engineering from the University of Connecticut, in 2018, where he is currently a Postdoctoral Research Fellow with the Department of Natural Resources and the Environment. His research focuses on the application of non-parametric statistical and machine learning techniques in remote sensing, weather forecasting, and the modeling of hydro-meteorological processes.

XINXUAN ZHANG received the B.S. degree in applied meteorology from the Nanjing University of Information Science and Technology (formerly named the Nanjing Institute of Meteorology), China, in 2007, and the M.S. and Ph.D. degrees in environmental engineering from the University of Connecticut, in 2012 and 2018, respectively. She was with the China Meteorological Administration, from 2007 to 2010. She is currently a Postdoctoral Research Fellow with George Mason University. Her primary researches focus on the numerical weather model-based satellite precipitation adjustment and data assimilation in land surface modeling. She also participates in the Project of Power Outage Prediction and provides weather forecasts to the outage model.



JAEMO YANG received the B.S. degree in civil and environmental engineering and the M.S. degree in hydraulics/hydrology engineering from Korea University, in 2009 and 2011, respectively, and the Ph.D. degree from the Department of Civil and Environmental Engineering, University of Connecticut, in 2018. He is currently a Postdoctoral Researcher with the National Renewable Energy Laboratory. His Ph.D. research focuses on developing methodologies that are computationally inexpensive to improve numerical weather prediction of extreme weather events, which are defined by the occurrence of high wind speed and/or intense precipitation (i.e., thunderstorms, rain/wind events, and tropical storms).

MARIA E. B. FREDIANI received the bachelor's and master's degrees in meteorology from the University of Sao Paulo and the Ph.D. degree in civil engineering from the University of Connecticut, in 2017. Her Ph.D. research focused on the methods to improve numerical wind speed forecasts to ultimately extend the skill of the storm-induced power outage prediction. Her research career includes a Postdoctoral Fellowship with The University of British Columbia and a Marie-Curie Research Fellowship with the Hellenic Centre for Marine Research, Greece. She is currently a Postdoctoral Fellow with the National Center for Atmospheric Research. She is part of teams developing a stochastic microphysics parameterization and implementing a spot fire ignition module into the Colorado Fire Prediction System. Her main research interests include numerical weather prediction, forecast verification, forecast post-processing, cloud microphysics, and planetary boundary layer processes.

EMMANOUIL N. ANAGNOSTOU is currently a Civil and Environmental Engineering Professor and the Eversource Energy Endowed Chair in environmental engineering with the University of Connecticut. He is also the Founding Director of the Eversource Energy Center and the Applied Research Director of the Connecticut Institute for Resilience and Climate Adaptation. His research expertise is on remote-sensing applications in water resources with a focus on the improvement of predictability of hydro-meteorological extremes and associated hazards, and the synthesis of sociological and engineering methods to create solutions to the nexus problem of water, food, and energy insecurities and risks. He also coordinates the National Science Foundation-Partnership for International Research and Education Multi-Institutional Food Security Project in Ethiopia (pire.engr.uconn.edu). In his capacity as the Eversource Energy Center Director and as one of the CIRCA's Applied Research Directors, he is leading power delivery resilience and inland and coastal flood vulnerability studies.