Applied Energy 205 (2017) 1408–1418

Contents lists available at ScienceDirect

Applied Energy

journal homepage: www.elsevier.com/locate/apenergy

Evaluating penalized logistic regression models to predict Heat-Related Electric grid stress days

L.M. Bramer[a,*], J. Rounds[a], C.D. Burleyson[a], D. Fortin[a], J. Hathaway[a], J. Rice[b], I. Kraucunas[a]

[a] Pacific Northwest National Laboratory, 902 Battelle Blvd, Richland, WA 99352, USA
[b] Smarter Decisions, LLC, 7057N Omaha Ave, Portland, OR 97217, USA

HIGHLIGHTS

- Statistical models for predicting grid stress using weather data are developed.
- The relative importance of weather variables and observed time scale are evaluated.
- Models fit to specific operation zones provide benefits over a globally-fitted model.
- Temperature, absolute humidity, precipitation, and previous days' precipitation are key predictive variables for all zones.

ARTICLE INFO

Keywords:
Electrical grid
Grid stress
Heatwave
Statistical modeling

ABSTRACT

Understanding the conditions associated with stress on the electricity grid is important in the development of contingency plans for maintaining reliability during periods when the grid is stressed. In this paper, heat-related grid stress and the relationship with weather conditions were examined using data from the eastern United States. Penalized logistic regression models were developed and applied to predict stress on the electric grid using weather data. The inclusion of other weather variables, such as precipitation, in addition to temperature improved model performance. Several candidate models and combinations of predictive variables were examined. A penalized logistic regression model which was fit at the operation-zone level was found to provide predictive value and interpretability. Additionally, the importance of different weather variables observed at various time scales were examined. Maximum temperature and precipitation were identified as important across all zones while the importance of other weather variables was zone specific. The methods presented in this work are extensible to other regions and can be used to aid in planning and development of the electrical grid.

1. Introduction

Extreme events, such as heat waves and drought, have historically posed reliability challenges for the electric grid by simultaneously increasing electricity demand, affecting natural resource availability for power generation and power plant cooling, and reducing the capacity and efficiency of power plants and transmission lines [1]. For example, during the summer 2003 heat wave in Europe, which was the hottest since 1500 in Switzerland and led to upwards of 15,000 deaths in France alone [2], electricity demand for cooling peaked coincidentally with the shutdown of multiple nuclear power plants in France due to restricted access to cooling water. Similar conditions occurred in the southeastern U.S. during the summer of 2007 when a combination of a heat wave and drought led to reduced generation and higher electricity prices across the Tennessee Valley Authority [3] and in Connecticut in 2012 when high water temperatures reduced cooling efficiencies and led to the shutdown of the Millstone Nuclear Power Station [4]. Perhaps the best-known incident occurred in 2003, when an extended period of heat over the eastern U.S. contributed to a series of events leading to a failure of the electricity grid that at its peak impacted over 50 million people in the U.S. and Canada [5].

The electricity sector has contingency planning and reserves in place to maintain reliability during periods when the grid is stressed, but these plans are based on the historical frequency and severity of extreme events. The future effectiveness of these contingency plans may be in question due to climate non-stationarity: the expectation of changes in the frequency, duration, or intensity of climate extremes such as heat waves due to climate change [5–9]. Similarly, periods of extreme drought in the future will have significant impacts on electricity production by reducing the amount and temperature of water

* Corresponding author.
E-mail address: lisa.bramer@pnnl.gov (L.M. Bramer).

http://dx.doi.org/10.1016/j.apenergy.2017.09.087
Received 1 June 2017; Received in revised form 12 September 2017; Accepted 13 September 2017
Available online 22 September 2017
0306-2619/

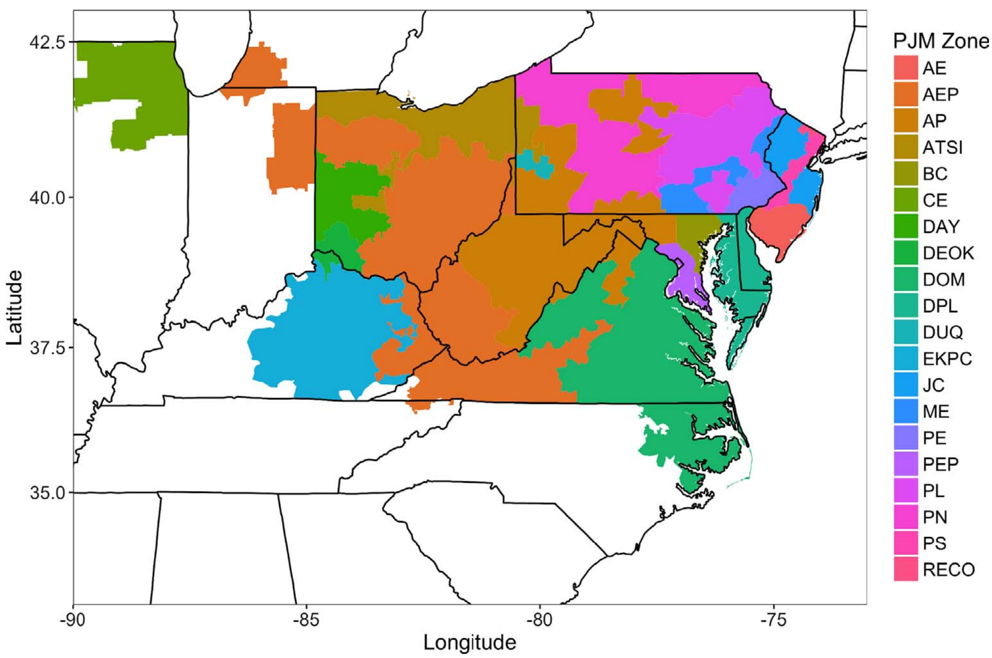

Fig. 1. Map of the 20 PJM control zones.

available for cooling power plants [10]. Planning and operations teams within the electricity sector need to understand the risks presented by a non-stationary climate to perform capacity planning and revise operational protocols as needed (e.g., [4,11,12]). Understanding the "new normals" for climate and weather and their implications for energy sector vulnerability is the subject of ongoing research at the U.S. Department of Energy (DOE) and its national laboratories [4], including Pacific Northwest National Laboratory (PNNL) [13]. Because of the complexity of the U.S. electric grid – the Eastern Interconnection (EIC) [14] has been called "the most complex machine in the world" [15] – understanding the interactions among weather, natural resources, electricity supply and demand, and transmission and distribution systems is a significant science challenge.

The first step in understanding the time-evolving vulnerability of the electric grid to climate change is to understand the environmental conditions that generate stress on the current configuration of the grid. This paper focuses on the development and application of a robust statistical model designed to identify the weather conditions associated with stress on the electric grid in a specific service area. We focus on weather-induced grid stress in this work, which will be referred to generically as grid stress throughout the rest of this manuscript. We define grid stress as a function of the load, in megawatts (MW), and the locational marginal price (LMP), in dollars per megawatt hour ($/MWh). LMP indicates the marginal cost of supplying the next increment of energy in a particular zone of the electric grid. As load increases (for example, during a heat wave), LMP rises as system dispatchers call on increasingly more expensive resources from within the zone or purchase power from the transmission system—whichever is cheaper. We define grid stress events as those days in which both the load and LMP reach unusually high levels.

A significant amount of work has been done to forecast electrical load based on seasonality, weather, and socioeconomic conditions. Our approach is novel because we develop a predictive model of grid stress as opposed to electrical load and because we model a wide range of weather variables, including time lags and variable interactions, whereas most other studies focus on temperature [16–18] or temperature-derived variables.

A literature review indicates that electrical load modeling efforts use two primary classes of models: neural network models and statistical or probabilistic models. Numerous studies detail methods for using neural networks or fuzzy logic to forecast load (e.g., [19–22]). These approaches are often lacking in interpretability as there is inherently very little transparency in these types of models, thus understanding what drives variability in electrical load is not straightforward.

It is widely recognized that temperature plays a key role in driving energy demand [14,23] and many statistical studies have utilized temperature or metrics derived from temperature (e.g., heating degree days) to predict future electrical loads using statistical models (e.g., [16–18,24]). A small subset of these models attempt to leverage other weather information such as humidity [25–27] and wind speed [26]. Our review did not discover any previous studies focused on a multivariate prediction of grid stress.

We investigated the following research questions: (1) Can a general framework for predicting grid stress, which might be applied to other study regions, be developed? (2) How predictable is a grid stress event given weather observations as explanatory variables in a statistical model (i.e., how well can the model perform)? and (3) What are the key weather-related variables, other than temperature, that can aid in predicting grid stress events?

The remainder of this work is presented as follows: In Section 2 we discuss the available data, steps taken in data processing, and methods for identifying grid stress events in the observational record. We describe our statistical modeling framework and methods for model fitting and evaluation in Section 3. In Section 4 we present our results from fitting statistical models to predict grid stress. Finally, we close with a discussion of our results and a summary of future research needs in Section 5.

2. Data

2.1. Study region

We selected the PJM Interconnection within the EIC[1] for our study because its historical load and LMP data were publicly available for a large number of years.[2] PJM operates a wholesale electric power market that spans a large portion of the EIC.[3] PJM's market extends into

[1] http://www.pjm.com/library/~/media/about-pjm/pjm-zones.ashx.
[2] https://esuite.pjm.com/mui/index.htm.
[3] https://energy.gov/oe/services/electricity-policy-coordination-and-implementation/transmission-planning/recovery-act-0.

a total of 13 states and Washington D.C., with a population of more than 61 million people served and a net generating capacity of 177,683 MW as of December 31, 2015 [28]. The PJM Interconnection is broken down into 20 control zones, which are the fundamental spatial scale of analysis in this work (Fig. 1).

2.2. Identifying grid stress events

In this work we were interested only in heat-related grid stress (as opposed to cold-related stress), thus we considered non-holiday weekdays (when peak loads tend to be higher due to the use of commercial buildings) during the months of June, July, August, and September. The time frame used was 2005–2015. Grid stress days over these date ranges were defined in terms of the measured load and LMP using the following criteria: (1) The maximum daily load exceeded the 90th percentile of maximum daily load values and (2) the maximum daily LMP exceeded the 90th percentile of maximum daily LMP values. The first criterion establishes that the peak load was unusually high for the day in question while the second criterion establishes that the LMP was unusually high in response to the demand. The maximum daily load values within each zone were population-adjusted, by dividing load values by the annual population[4] and then scaling to the 2015 population estimate, to allow load values to be comparable over multiple years. Four of the PJM zones (ATSI, DEOK, EKPC, RECO) were not examined because data were not available for the entire time period due to the constantly evolving configuration of zones. We focused our evaluation on the remaining 16 zones whose boundaries and configuration remained stable for the entire study period.

The choice of the 90th percentile in maximum daily load and LMP values to define a grid stress day was determined by an exploratory analysis of change points (the point at which the distribution of values changes significantly) in the relationship between the two variables. We found that the 90th percentile approximates the change point in most PJM zones. Fig. 2 illustrates this for two zones. In Fig. 2, the rate of change of LMP with respect to maximum daily load increases when both values exceed their 90th percentiles compared to data below the 90th percentile. Analogous plots for all PJM zones are included in Figure S1 of the Supplementary Material. In addition to helping focus our analysis on the extreme conditions that are most worrisome to grid operators, considering both load and LMP helps us avoid selecting days when a non-weather-related event, such as a forced outage, might have been the primary reason for high LMPs. Table 1 gives the number of grid stress days for each zone that were identified over the June-September 2005–2015 time period.

PJM emergency message data have potentially useful information for understanding how stress on the power grid manifests at the operational level; however, challenges exist in using these data to identify grid stress days. The majority of the PJM emergency messages are assigned to the regional level (PJM-RTO) as opposed to specific zones, making it a challenge to use these data to characterize zone-specific events. Moreover, challenges exist in understanding the cause-and-effect influences of specific messages. For example, the presence of a hot weather alert may be indicative of a heat-related grid stress event, or it may set in motion the appropriate actions to mitigate stress on the grid. Due to these challenges, it is not feasible to utilize specific messages to define a grid stress day. However, we cross-checked the grid stress days determined by the two criteria above against PJM emergency messages to confirm emergency warnings were issued and/or actions were taken to relieve the load pressure on or before the grid stress day. By not including the emergency messages in our definition of a grid stress day, our working definition of grid stress can be applied in energy markets other than PJM where detailed emergency messages may not be readily available. A detailed description of the challenges associated with the emergency message data and their association with grid stress days is given in Section 2 of the Supplementary Material.

2.3. Weather data

Hourly surface meteorology observations from weather stations across the U.S. are available in the NOAA ISD-lite data set.[5] Our work used weather data from 2005–2015, the same period for which we obtained PJM load and LMP data. The meteorological data from 2005–2015 served as an input data set from which training data sets for our statistical model were selected. There were 420 stations within the PJM territory that had a sufficiently complete data record from 2005–2015. We used four base hourly meteorological variables (Table 2): Air temperature (Temp), wind speed (WindSpeed), sea-level pressure (Pressure), and precipitation (Precipitation). Precipitation in the ISD-lite data set is reported as an hourly-mean rain rate, which we converted to a binary variable so that a day with any measurable precipitation was recorded as 1 and days with no precipitation were recorded as 0. From the ISD-lite data we derived absolute humidity (AbsHum) and relative humidity (RelHum). Additionally, to capture weather conditions that may be persistent over consecutive days, such as would occur during a heat wave event, four-day lagged weather variable values were calculated by taking the average of the previous 4 days for each of the six weather variables (Table 2). We considered and evaluated many different potential lag periods to include in model. In order to limit the collinearity of variables, we evaluated the correlation of different lagged variables to the value of the variable for the current day. Four days was the minimum amount of time required for all weather variables to reach minimal correlation with the day zero value (analogous to a decorrelation time scale). Additionally, a sensitivity study showed that for lags longer than four days, no significant gains in model performance were observed. Results from this analysis are included in Section 3 of the Supplementary Material. The inclusion of lagged variables resulted in a total of 12 weather variables (6 weather variables + 6 lagged variables) used as input to the statistical models.

Because our analysis was done on the spatial scale of PJM zones, the daily weather variables from each station within a given zone were population weighted so that combined they represented the weather that was impacting the largest amount of people within each zone. To accomplish this, we subsetted a 1/8° population dataset [29] for all the grid cells within each PJM zone. These grid cells were assigned to the nearest weather station using Haversine distances, which give the shortest distance between two points on a sphere. The total population assigned to each weather station was divided by the total population within the zone to create a weighting for that station. All variables from each station were multiplied by this population weight, reducing the daily weather variables per station to daily weather variables per zone. Because the available weather stations and the location and spatial distribution of the population both varied from year-to-year, the population weights were recalculated each year using interpolated county-level census data.[6] From each of the hourly variables, with the exception of temperature, we computed the daily mean population weighted values for each zone. Instead of a daily mean temperature, population weighted daily maximum temperature was used to capture the occurrence of extreme heat. Daily mean values were calculated for other variables because several of the variables have minimal variability within a day and their means are often more representative of the true conditions compared to their maximum values. For example, wind speed is highly variable and the maximum value may only occur for a few minutes within a given day.

[4] https://factfinder.census.gov/faces/nav/jsf/pages/index.xhtml.

[5] https://www.ncdc.noaa.gov/isd/data-access.

[6] https://factfinder.census.gov/faces/nav/jsf/pages/index.xhtml.

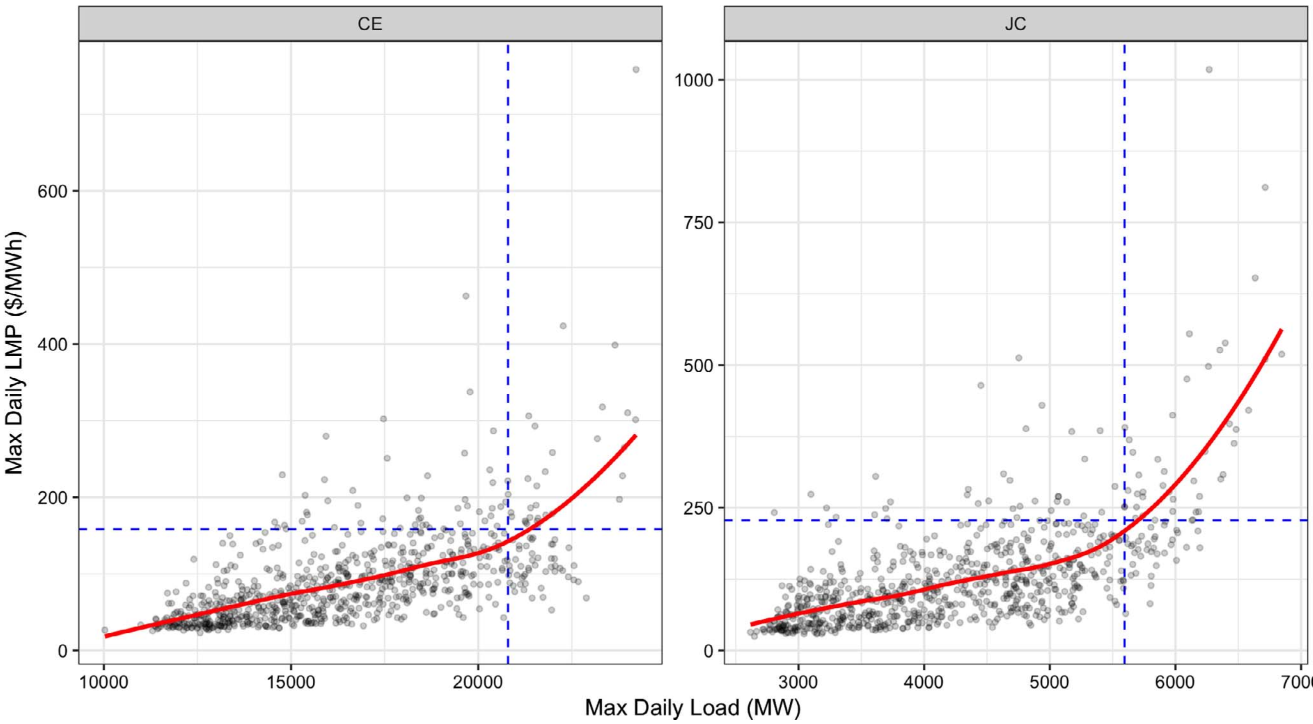

Fig. 2. Scatterplot of maximum daily load (x-axes; MW) versus maximum daily LMP (y-axes; $/MWh) for the CE (left) and JC (right) PJM zones derived using weekday data from June–September 2005–2015. The dashed blue vertical and horizontal lines show the 90th percentiles of both distributions and the solid red line is a nonparametric smoothing spline fit to the data. We defined grid stress days as those points in the upper-right quadrant of the joint distributions. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

Table 1
The name of each of the 20 zones within the PJM Interconnection, the abbreviation we assigned to each zone, postal service abbreviations representing the states that each zone services, and the total number of grid stress day identified over the period 2005–2015.

Zone name	Zone ID	States of service	Grid stress days
Atlantic City Electric Company	AE	NJ	49
American Electric Power Company, Inc.	AEP	IN, OH, KY, VA, WV	42
Allegheny Power	AP	PA, WV, MD, VA	43
American Transmission Systems, Inc.	ATSI	OH, PA	–
Baltimore Gas and Electric Company	BC	MD	42
Commonwealth Edison Company	CE	IL	34
Dayton Power and Light Company	DAY	OH	36
Duke Energy – Ohio and Kentucky	DEOK	OH, KY	–
Virginia Electric and Power Company	DOM	VA, NC	34
Delmarva Power and Light Company	DPL	MD, DE	40
Duquesne Light Company	DUQ	PA	42
Eastern Kentucky Power Cooperative	EKPC	KY	–
Jersey Central Power and Light Company	JC	NJ	48
Metropolitan Edison Company	ME	PA	51
PECO (formerly Philadelphia Electric Company) Energy Company	PE	PA	48
Potomac Electric Utilities Corporation	PEP	MD	37
PPL (formerly Pennsylvania Power and Light) Electric Utilities Corporation	PL	PA	44
Pennsylvania Electric Company	PN	PA	47
Public Service Electric and Gas Company	PS	NJ	48
Rockland Electric Company	RECO	NJ	–

3. Modeling framework

3.1. Statistical models

A large number of models are available to researchers for classification. For this application, we considered models that have a binary response variable (1 = grid stress day and 0 = not a grid stress day) for a set of training data and used a given set of weather values as potential explanatory variables. A range of different classification models and approaches were considered and evaluated. We found that penalized logistic regression (PLR) [30] performed well and provided an interpretable model that allowed us to investigate our research questions. PLR provides a closed-form model for classification and shrinks the coefficients of variables with little predictive information to 0 so only useful explanatory variables are retained. We then focused on different variable sets and formulations of the model to address the research questions defined in the Introduction. The specific equations for the PLR models used in this paper are given in Section 4 of the Supplementary Material.

Table 2
Observed and derived meteorological variables obtained from the NOAA ISD-Lite data set.

Observation	Name	Units
Raw Variables		
Temperature	Temp	°C
Wind speed	WindSpeed	m s^{-1}
Sea-level pressure	Pressure	hPa
Precipitation	Precipitation	Binary (1 = yes, 0 = no)
Derived Variables		
Relative humidity	RelHum	%
Absolute humidity	AbsHum	g m^{-3}

3.2. Explanatory variable sets

Temperature information is easily available to grid operation teams and can be used to assess the potential risk of grid stress. It is the primary and often only weather variable considered by load forecasters, as detailed in the Introduction. Thus, we considered a model using only maximum temperature as an explanatory variable as a baseline model for comparison purposes.

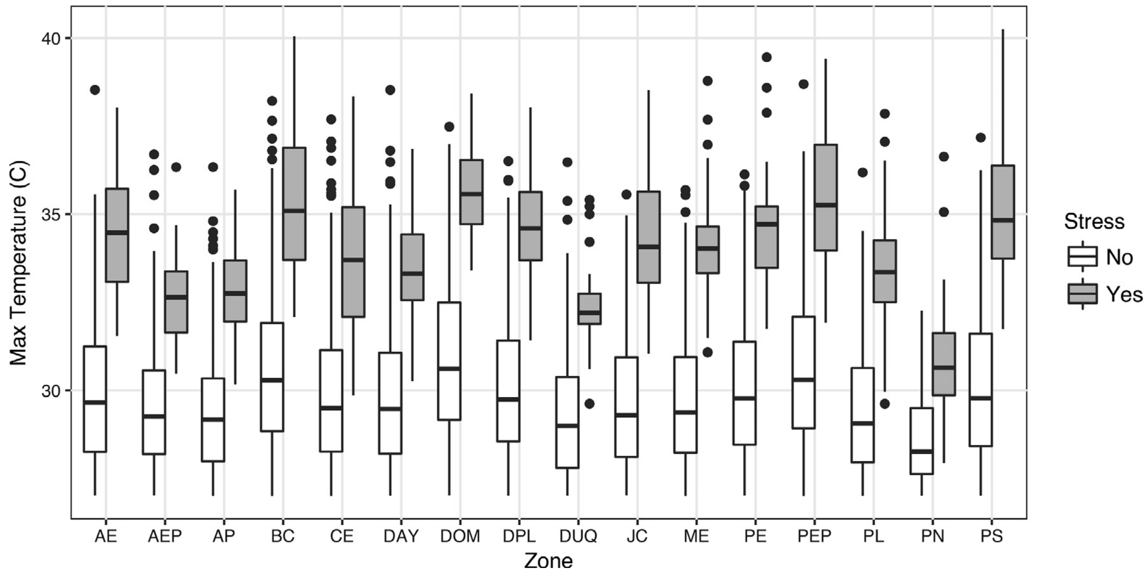

Fig. 3. Observed maximum temperatures on grid stress days (shaded bars) and non-grid stress days (white bars) within the training data set. This data was taken as all summer days in which the maximum temperature was above 27.9 °C.

Two additional sets of explanatory variables were considered. One set was comprised of the 12 values from the aforementioned weather variables and their corresponding lags (Section 2.2). We refer to this set as the “in-zone” variable set. To evaluate the importance of weather conditions occurring in other parts of the PJM Interconnection, we created a second set of explanatory variables made up of the “in-zone” variable set combined with the 6 basic weather variables from each of the other 15 PJM zones. We refer to this set as the “across-zone” variable set. These two variable sets can be thought of as capturing locally-occurring grid stress events (“in-zone”) as well as grid stress events that form as a result of hot temperatures occurring further away and being transmitted across the Interconnection (“across-zone”).

3.3. Temperature thresholds for grid stress

Grid stress was never observed with population weighted max-temperatures less than 27.9 °C (Fig. 3). Because of this there was little use in training the models on days where that temperature is never reached, so we removed any day in which the population weighted max-temperature was below 27.9 °C from the training or testing data. We classified all such days as non-grid stress and the probability of grid stress was assigned to be zero. These days were also not included in the evaluation of model performance. The remaining days with max-temperatures over 27.9 °C contained a mixture of grid stress and non-grid stress days, with no single temperature threshold uniquely separating the two. A secondary effect of the decision to remove days with max-temperatures colder than 27.9 °C was that it helped balance the sample sizes of the two data classes (grid stress and non-grid stress) we sought to distinguish. Data with major class imbalances can cause serious issues in fitting a binary classifier [31]. By eliminating days with a temperature under 27.9 °C, the ratio of grid stress days to non-grid stress days decreased from approximately 1:25 to 1:10 within each zone.

3.4. Model evaluation

We used cross-validation [32] to evaluate the performance of the statistical models. The practice of cross-validation, dividing data into training and testing data, is used to estimate how accurately a predictive/classification model will perform in practice. Random sampling of days is one possible way of dividing data into training and testing sets. However, this method could have confounded results, since lagged variable values from previous days were included in some of the sets of explanatory variables. We evaluated model performance by leaving out observations from each year from 2005 to 2015 in turn, one at a time. This is a stratified leave-one-out cross-validation method. For example, we left out data from 2015 and trained the model using data from 2005 to 2014, then evaluated model performance based on the 2015 data. We repeated this until all years were evaluated as a testing data set. Cross-validation was also used to determine a reasonable range for the shrinkage parameter, λ, in the penalized logistic regression model. In these types of models, the shrinkage parameter controls how much regression coefficient estimates are shrunk to prevent overfitting the model to the training data. The shrinkage parameter was chosen in a grid search of values between 0.0002 and 0.01 by increments of 0.0002.

We calculated several metrics of model performance for each testing data set. Table 3 gives a list of these metrics and their definitions. We focused on positive predictive value (PPV), balanced accuracy (BACC), and the F1 measure as they are relevant and more informative when class imbalance is present and correctly predicting positive events is of particular interest [33]. PPV gives the proportion of positive predictions that are made which are correct. BACC averages the proportion of correctly predicted positives and negatives. The F1 measure takes into account both false negatives and false positives. BACC, PPV, and the F1 measure all range from 0 to 1, with larger values indicating better model performance.

Table 3
Metrics for evaluating model performance and their definitions.

Metric	Definition
True Positives (TP)	Number of grid stress days correctly predicted
True Negatives (TN)	Number of non-grid stress days correctly predicted
False Positives (FP)	Number of non-grid stress days incorrectly predicted as grid stress days
False Negatives (FN)	Number of grid stress days incorrectly predicted as non-grid stress days
Balanced Accuracy (BACC)	0.5 * [TP/(FN + TP) + TN/(FP + TN)]
Positive Predictive Value (PPV)	TP/(TP + FP)
F1 measure	2 * TP/(FP + 2*TP + FN)

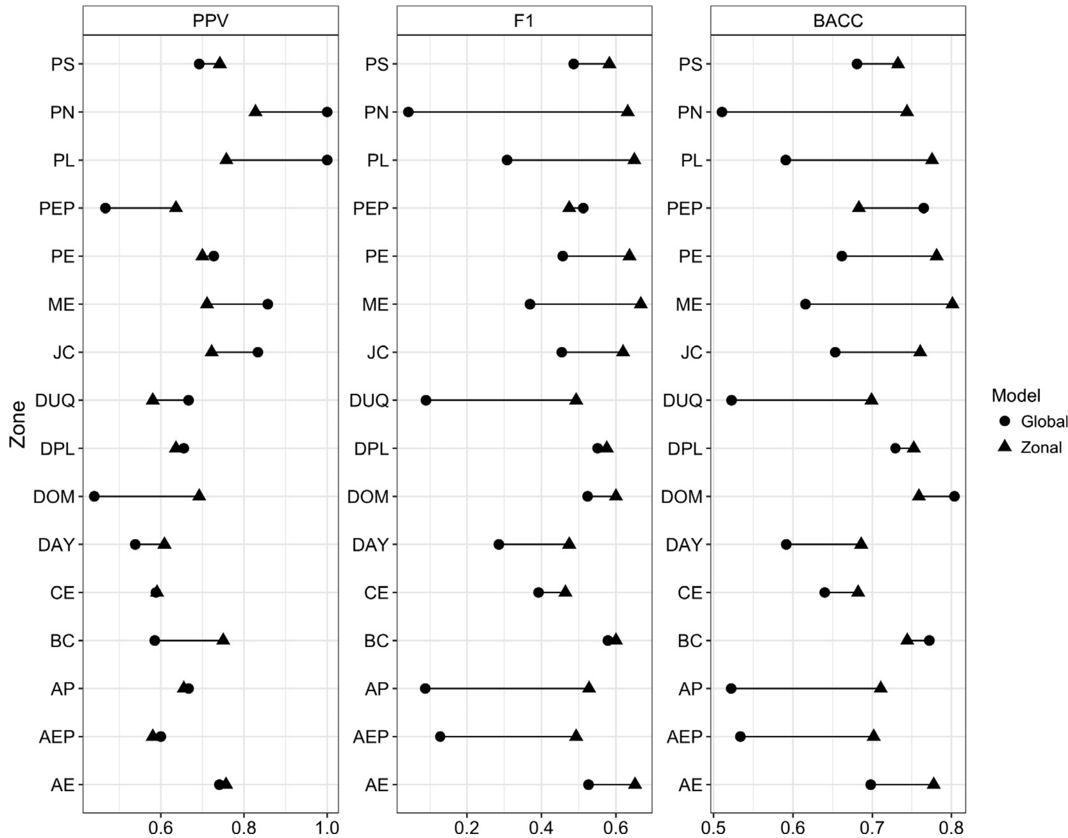

Fig. 4. Model performance metrics for PLR models fit to data from all zones at once (Global) and separate models fit to each zone (Zonal).

4. Results

4.1. Global vs zone-specific models

We first consider whether it is sufficient to fit a single model for all zones or if there are benefits to fitting individual models for each zone. We fit a PLR model to training data from all zones at once, thus forcing the estimated regression coefficient of each variable in the model to be equal across zones. We refer to this as a "Global" model. We also fit a PLR model to each zone's training data separately, allowing the variable effects to vary from zone to zone. We refer to the set of these as "Zonal" models. Fig. 4 shows the BACC, PPV, and F1 measure values for each zone using the "in-zone" variable set as input to the Global and Zonal models. In all but 3 zones (BC, PEP, and DOM), the BACC values are higher for the Zonal model compared to the Global model. The F1 measure values are higher for the Zonal models compared to the Global model for all but 1 zone (PEP). However, trends in PPV values are less consistent across zones. For 9 of the 16 zones, the PPV value using the Global model is higher than the Zonal models. The PPV values are lower in most Zonal models due to an increase in the number of false positives predicted compared to the Global model. Zonal models falsely predicted grid stress 31.7% more often than the Global model. This is equivalent to an average of 2.9 more false positives per year across all of PJM. Higher F1 and BACC values in most zones for the Zonal model compared to the Global model are largely due to 41.3% more false negatives predicted by the Global model. This is equivalent to the Global model failing to predict grid stress for an average of 12.9 more days per year across all zones.

The Zonal models more often correctly classifies grid stress days, but do so at the cost of an increase in the number of false positives. There are costs associated with both falsely predicting grid stress and failing to predict grid stress that should be taken into account when choosing between a Zonal and Global modeling approach. The decreased skill of the Global model to predict grid stress indicates that the weather characteristics associated with grid stress are variable across PJM zones. We use Zonal models throughout the rest of this work because the number of false positives generated by the Zonal models are substantially fewer than the number of false negatives from the Global model.

4.2. Weather variable set comparisons

As explained in Section 3.2, we developed a baseline model for evaluation purposes that uses only in-zone maximum temperature to predict grid stress. To create this baseline model, we fit a logistic regression model for each zone with maximum temperature as the explanatory variable. The BACC values for this baseline model ranged from 0.57 to 0.78, PPV performance for each zone ranged from 0.39 to 0.74, and the F1 measures ranged from 0.13 to 0.48. The mean model performance metrics, averaged across all zones, were 0.68, 0.62, and 0.31 for BACC, PPV, and F1 values, respectively. Fig. 5 shows the F1 measure of the baseline model plotted by zone. In general, maximum temperature is more predictive of grid stress for zones along the east coast compared to zones further west.

We then fit PLR models for each zone using the "in-zone" variable set to evaluate what improvements, if any, can be made in our ability to predict grid stress compared to the temperature only baseline model. Fig. 6 gives a summary of these models' performance compared to the temperature only models. The PLR models using the "in-zone" variable set consistently outperform the temperature only model in terms of F1 measure and BACC value for all zones with the exception of JC, which is a small zone along the coast of New Jersey (Fig. 1). The PPV for the "in-zone" variable set outperforms the baseline model for all but four zones (JC, ME, DPL, and DOM), where differences between the models were less than 0.03. Overall, this analysis demonstrates that the inclusion of other meteorological variables in addition to maximum temperature

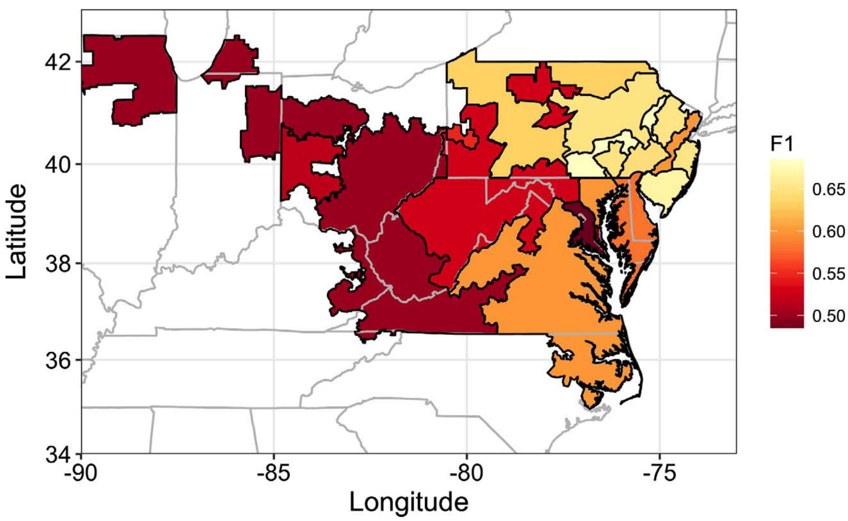

Fig. 5. The F1 measure of the temperature only model plotted by zone.

generally leads to improved model performance.

4.3. Predictive weather variables

Penalized logistic regression models shrink the coefficients of variables that are not useful in the model to zero, thus useful variables are identified by their non-zero weights. Examining the variables selected by the optimal PLR model in each zone indicates which weather variables are most useful for predicting grid stress days. For the Global PLR model, maximum temperature, mean relative humidity, and mean relative humidity as well as the four-day lagged derivations of each of these variables were found to be useful in predicting grid stress days. Additionally, while mean day 0 precipitation across all PJM zones was found to be useful, its four-day lag was not retained. This suggests that rain in the days leading up to a grid stress event is not a robust predictor of grid stress. Further examination of the sign of the estimated coefficients for the retained variables in this model found 3 of the useful variables (maximum temperature plus mean absolute and relative humidity) had positive coefficients. In other words, an increase in any one of these variables (holding all other variables constant) results in an increased probability of grid stress for a given day. In contrast, mean precipitation had a negative coefficient, indicating that precipitation is negatively correlated with the predicted probability of grid stress. Intuitively this makes sense because large mean precipitation values in the Global model are likely associated with widespread precipitation and extensive cloud cover across the PJM Interconnection – both of which would act to suppress maximum temperatures.

We also examined the Zonal PLR models and the variables selected for each zone using the "in-zone" variable set. Fig. 7 shows the number of times each weather variable (broken down by specific time lags) was selected in the optimal model for each zone. The combination of variables selected in each zone is unique, but some variables are selected very often across zones. Maximum temperature and mean absolute humidity were included in the optimal model for each of the 16 zones. This verifies that maximum temperature has demonstrable value for predicting grid stress, confirming what we saw in the performance of the temperature only model in Section 4.2. Precipitation and the four-day lagged precipitation were selected in 15 of the 16 zonal models. Further investigation showed that the effect (estimated coefficient) of the day 4 lagged precipitation variable was negative in some zones and

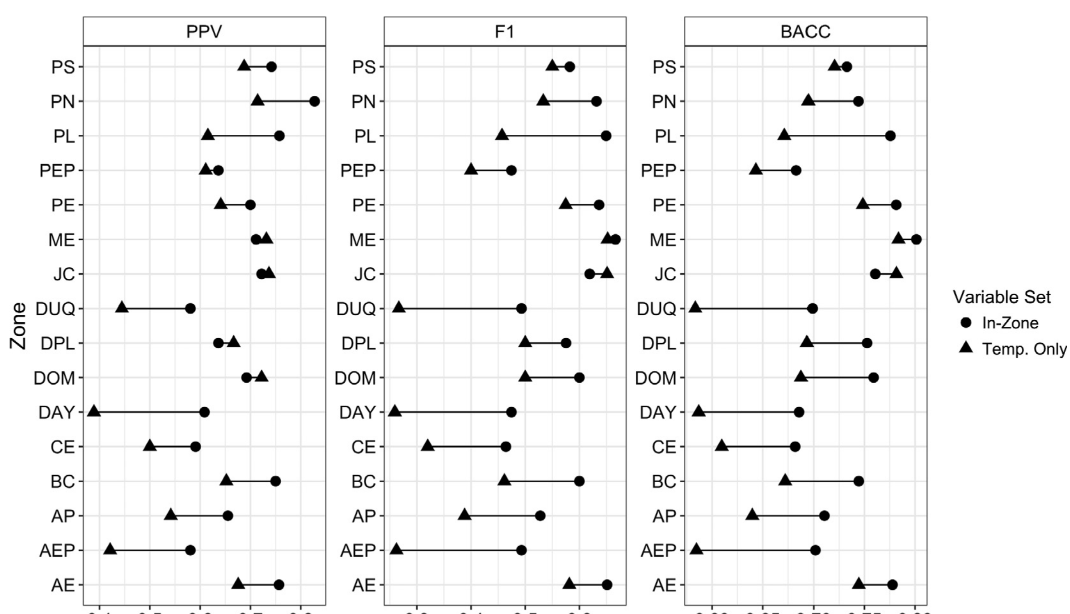

Fig. 6. Model performance metrics for Zonal PLR models with maximum temperature as the only explanatory variable (Temp. Only) and models using the full "in-zone" variable set (In-Zone).

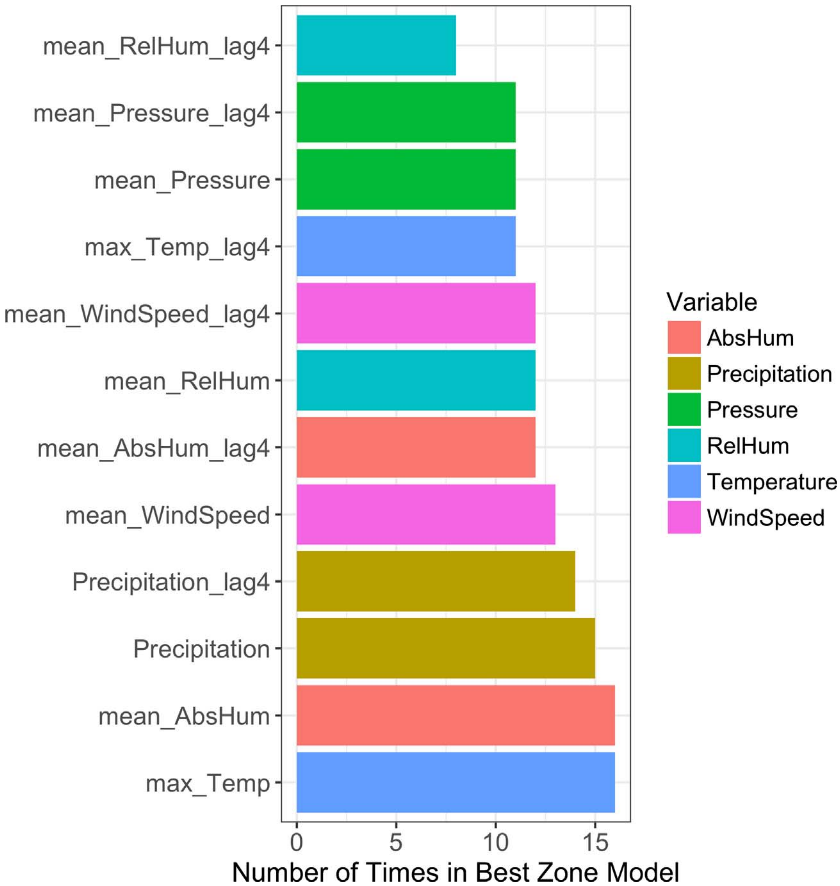

Fig. 7. Number of zones in which a variable was included in the best PLR model fit using the "in-zone" variable set.

positive in other zones. This inconsistent effect is likely why the variable was not selected in the Global model. In most zones, the mean wind speed tended to be included in the model as was the mean relative humidity and the four-day lagged variables for mean wind speed, maximum temperature, and mean absolute humidity. This suggests that extended periods of warm temperatures (increases the likelihood) or widespread precipitation (decreases the likelihood) are predictive of grid stress.

One benefit of the PLR models is that a closed form of the equation used to predict grid stress can be obtained and the impact of specific variables on the probability of grid stress can be evaluated (while holding all other variables constant). The estimated coefficients reflect the change in the predicted log odds ratio, the log of the probability of grid stress divided by the probability of non-grid stress, of grid stress for each one unit change in the variable of interest, while holding all variables constant. For example, Table 4 gives the selected variables and estimated coefficients (rounded to 3 decimal places) for the optimal PLR model fit using the "in-zone" variable set in the PN zone. For every 1 °C increase in the maximum temperature in PN, the log odds ratio of grid stress increases by 0.923. In other words, the odds of a grid stress day becomes 2.517 ($\exp\{0.923\} \approx 2.517$) times more likely on average. Increases in the mean absolute humidity also lead to an increase in the average log odds ratio. This is also true for the four-day lagged mean absolute humidity, but it has smaller overall impact on the probability of grid stress compared to the current mean absolute humidity. The mean wind speed and precipitation are negatively correlated with grid stress. Overall, maximum temperature and mean absolute humidity had positive estimated coefficients in all 16 zones and precipitation had a negative estimated coefficient for each model in all 15 zones where precipitation was included in the model. We found that mean wind speed had a negative estimated coefficient for all zones where mean wind speed was included in the model. The coefficients for mean absolute humidity and mean pressure had mixed signs depending on the zone (6 were positive and 6 were negative for mean absolute humidity and 6 were positive and 5 were negative for mean pressure).

Table 4
Estimated coefficients for best PLR model using the "in-zone" variable set in the PN zone.

Variable	Estimated coefficient
(Intercept)	−33.660
max_Temp	0.923
mean_AbsHum	0.569
Precipitation	−1.342
mean_AbsHum_lag4	0.093
Precipitation_lag4	0.342
mean_Pressure_lag4	−0.004
mean_WindSpeed_lag4	−0.099

4.4. Comparison of IN-Zone and Across-Zone models

Electrical grids are inherently connected across regions or zones of service. Thus, it is natural to consider whether weather conditions (and thus energy demand) in other zones might provide information that allows for improved prediction of grid stress in a single zone of interest. We investigate this concept by fitting an additional set of PLR models in each zone using the "across-zone" variable set, which contains weather variables for each of the 15 other PJM zones as potential explanatory variables. Fig. 8 shows the optimal PLR models' F1 measure, PPV, and BACC for each zone using both the "in-zone" and "across-zone" variable sets as potential explanatory variables.

Including the information from other zones can improve model performance in some of the zones. The models using the "in-zone" variable set outperformed those using the "across-zone" variable set

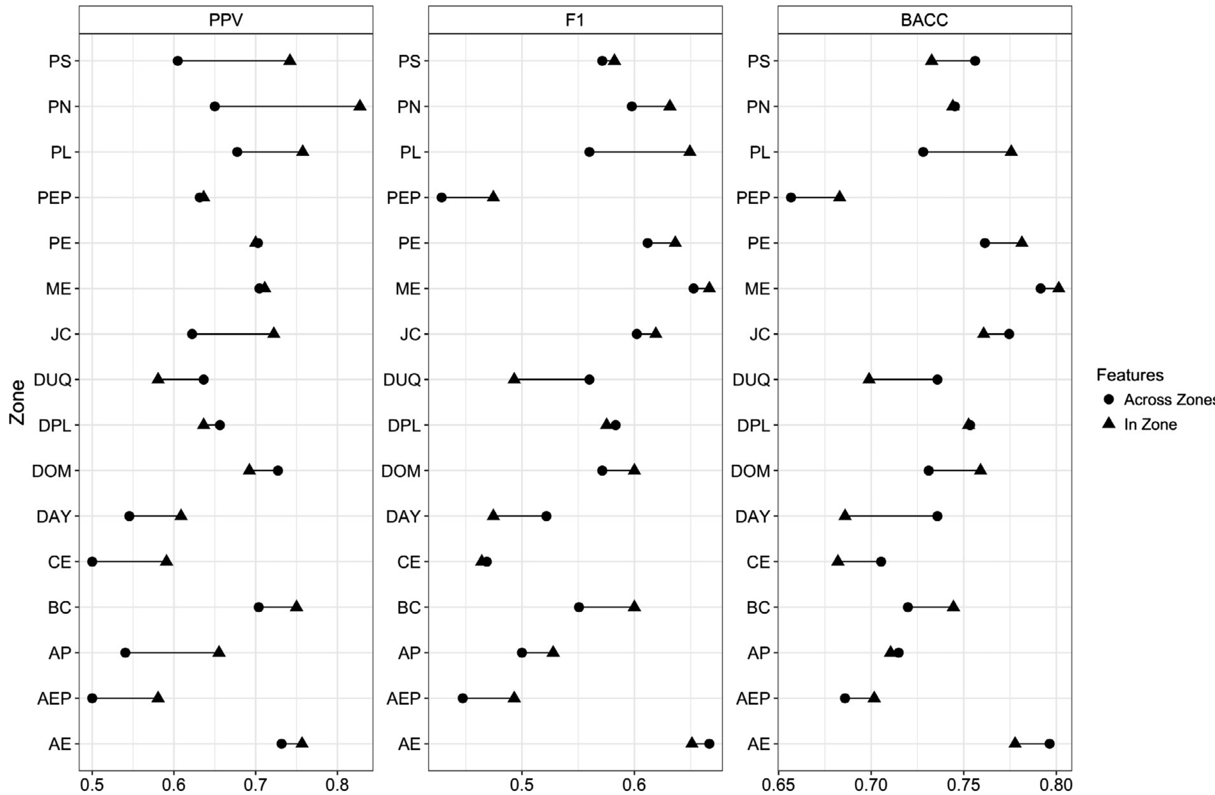

Fig. 8. Model performance metrics for PLR models using the "across-zone" variable set (Across Zones) and using the "in-zone" variable set (In Zone).

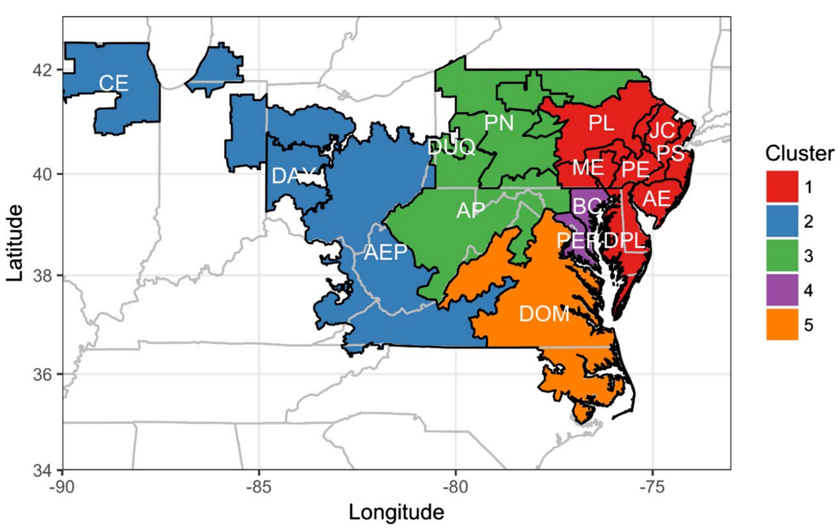

Fig. 9. Map of clusters resulting from analysis of grid stress co-occurrence hierarchical clustering analysis.

across all three metrics for half of the zones. In contrast, using the "across-zone" variable set provided some improvement in at least one of the metrics for the other half of the zones. All three metrics are improved for the models using the "across-zone" variable set for DPL and DUQ. We examined the characteristics (geographic location, population, etc.) of zones for which the "across-zone" variable set provided value compared to zones for which the "in-zone" variable set was the best, but found no clear patterns or explanations. For the zones where using the "across-zone" variable set provided value, we also investigated which other zones' weather information was retained in each model. In some cases, geographically close zones were retained in combination with other zones, but this was not true for all cases. For example, the optimal "across-zone" model for DPL retains weather information from two zones which are geographically close (AE and BC) and two zones further away (AP, CE). The model for DUQ includes weather information from six other zones (CE, DPL, JC, PE, PEP, and PN) of which only PN is geographically adjacent. A subset of zones: AE, CE, DAY, JC, and PS, showed improvement in some combination of BACC and F1 measures for the model using the "across-zone" variable set, although improvements were relatively small for AE, JC, and PS. The difference in model performance for these zones was due to an increased number of true positives for the "across-zone" model, however in each of these cases there was also an increase in the number of false positives.

In an effort to better understand potential relationships among zones, we investigated the co-occurrence of grid stress in PJM. We then examined if any of the relationships discovered were preserved in the zones for which the model using the "across-zone" variable set provided improvement over the "in-zone" variable set. For each pair of zones, we first calculated the proportion of grid stress days identified in either zone that occurred in both zones. This value can be thought of a metric of similarity between each pair of zones with respect to the co-occurrence of grid stress. Figure S4 in the Supplementary Material gives these values for all pairs of zones. We then performed hierarchical clustering with complete linkage [34] using the matrix of similarity values. The resulting dendogram, a tree diagram that illustrates results of clustering algorithms and shows the distance between clusters, revealed five clusters of zones. Fig. 9 shows the zones and resulting clusters based on the co-occurrence of grid stress events. The co-occurrence of grid stress has some clear geographic structure, as clusters are made up of zones that tend to be geographically close to one another.

For zones where the model using the "across-zone" variable set outperformed the "in-zone" model, we examined which other zones were retained in the model. For four of these zones, weather information from other zones in the same cluster were retained. For example, the PLR model using the "across-zone" variable set for CE retained weather information from DAY and AEP only. The analogous model for AE included weather information from JC, PS, ME, and PE. In contrast, some zones tended to pull weather information primarily from other clusters. For example, DPL retained information from one other zone in its cluster, AE, but also retained weather information from BC, AP, and CE - all of which are from a different cluster. Similarly, the model from DOM kept weather information from one zone in each of the other clusters (AE, AP, BC, and DAY). The net result of this analysis is that, while there is a clear co-occurrence of grid stress in geographically-contiguous zones (i.e., there are clear clusters), there are no clear trends that demonstrate that providing weather information from other zones in the same cluster will always improve the performance of the grid stress model.

5. Discussion

In this paper we described the steps of formulating a statistical model, in this case penalized logistic regression (PLR), to predict grid stress events based on observed or, potentially, predicted weather variables. Our goals were to determine how well such a model could perform and if variables other than temperature, which is commonly used across the industry, had explanatory value. We utilized a PLR model because it performed well in an initial analysis of several varying types of competing statistical models. Additionally, a PLR model has a closed-form expression and thus provides interpretable results. Future studies might do a comprehensive evaluation across a suite of possible statistical classification models. Any such study should focus on the performance as well as the interpretability of the results of each potential model.

Using data from the PJM Interconnection, daily load and locational marginal price (LMP) values were used to identify summertime weekdays where there was stress on the electrical grid. This analysis was done independently for each of the 16 PJM control zones for which there was sufficient data. We saw no clear trends in the number of grid stress days across the 11 years we analyzed. However, a potential area for future study is the application of these models for predictions of grid stress events under a range of future climate scenarios.

We explored both global and zone-specific models and determined that modeling grid stress at the zone-specific level provided measurable improvements in the classification of grid stress days. The global model systematically under-predicted grid stress frequency across zones, leading to an increase in false negatives compared to models fit to individual control zones. We explored which weather variables had consistent predictive value across multiple PJM zones. A model using only the in-zone daily maximum temperature as a predictive variable was most valuable for zones located along the east coast. The performance of these temperature-only models decreased moving westward from the coast. We found that the incorporation of weather variables and lagged weather variables other than maximum temperature led to improved predictive performance across all but one zone, where the difference between the temperature-only model and a model using the full set of available weather variables was marginal. While the optimal weather variables for predicting grid stress varied from zone to zone, maximum temperature, mean absolute humidity, precipitation, and day 4 lagged precipitation were found to be important in almost all of the 16 PJM zones.

We also compared the optimal zone-specific PLR models using a variable set that included only in-zone weather information to zone-specific PLR models using a variable set that included weather information from all PJM zones. We found that adding weather information from other zones improved model performance in half of the zones. Exploration of zone characteristics (e.g., population and geographic location) revealed no clear patterns that could explain why some zones benefited from including weather information from other zones. We found that the co-occurrence of grid stress between zones was linked to the geographic proximity of zones. A clustering analysis identified five clusters of zones in which grid stress days tended to occur together. For zones where including weather data from other zones was beneficial, we observed that some zones retained weather information from within its cluster while others retained information across clusters. Overall, these results indicated that there is value in including weather information from other zones. However, given the lack of consistent patterns among zones where this was true, our approach of including weather information from all other zones and allowing the PLR model to select which are important could likely be improved upon. Future work is needed to investigate whether incorporating information about the transmission structure of an operating region would result in better model performance.

Assuming that data is available, the methods presented in this work could be extended to other regions of the electrical grid in the U.S. Additionally, the models developed could generate predictions and probabilities of grid stress for future climate scenarios (under the assumption that the configuration of the grid is the same as at the time the models were developed). These types of statistical models could thus be used to assess risk and aid in planning of capacity expansion to ensure adequate power supplies during periods of heat-related grid stress. If the current electric industry trends toward more expansive collection and sharing of higher resolution data were to continue, analytic tools and model formulations such as those presented in this work will become increasingly important and useful to grid operators and capacity expansion planners.

Acknowledgments

This research was supported by the Office of Science of the U.S. Department of Energy as part of the Integrated Assessment Research Program. The project team also thanks the National Protection and Programs Directorate, Office of Cyber and Infrastructure Analysis, at the U.S. Department of Homeland Security for support and guidance.

Appendix A. Supplementary material

Supplementary data associated with this article can be found, in the online version, at http://dx.doi.org/10.1016/j.apenergy.2017.09.087.

References

[1] Schaeffer R, et al. Energy sector vulnerability to climate change: a review. Energy 2012;38(1):1–12.
[2] Pirard P, et al. Summary of the mortality impact assessment of the 2003 heat wave in France. Euro Surveillance: Bull. Europeen sur les Maladies Transmissibles = European Communicable Disease Bulletin 2005;10(7):153–6.
[3] T. Kimmell, J. Veil, Impact of drought on US steam electric power plant cooling water intakes and related water resource management issues, Argonne National Laboratory (ANL), 2009.

[4] D.o. Energy, The Water-Energy Nexus: Challenges and Opportunities, U.S. Department of Energy, 2014.
[5] NYISO, Interim Report on the August 14, 2003 Blackout, New York Independent System Operator, 2004.
[6] Alexander L, et al. Global observed changes in daily climate extremes of temperature and precipitation. J. Geophys. Res.: Atmos. 2006;111(D5).
[7] Diffenbaugh NS, et al. Quantifying the influence of global warming on unprecedented extreme climate events. Proc. Natl. Acad. Sci. 2017;114(19):4881–6.
[8] C.B. Field, Managing the risks of extreme events and disasters to advance climate change adaptation: special report of the intergovernmental panel on climate change, Cambridge University Press, 2012.
[9] D. Jakob, Nonstationarity in extremes and engineering design, in Extremes in a Changing Climate, Springer, 2013, pp. 363–417.
[10] Bartos MD, Chester MV. Impacts of climate change on electric power supply in the Western United States. Nat. Clim. Change 2015;5(8):748–52.
[11] Beniston M. The 2003 heat wave in Europe: a shape of things to come? An analysis based on Swiss climatological data and model simulations. Geophys. Res. Lett. 2004;31(2).
[12] Van Vliet MT, et al. Vulnerability of US and European electricity supply to climate change. Nat. Clim. Change 2012;2(9):676–81.
[13] Kraucunas I, et al. Investigating the nexus of climate, energy, water, and land at decision-relevant scales: the Platform for Regional Integrated Modeling and Analysis (PRIMA). Clim. Change 2015;129(3–4):573–88.
[14] C. Crowley, F.L. Joutz, Weather effects on electricity loads: Modeling and forecasting 12 December 2005. Final report for US EPA on weather effects on electricity loads, 2005.
[15] Gellings CW. Transforming the electric infrastructure. Phys. Today 2004;57(12):45–51.
[16] Chikobvu D, Sigauke C. Modelling influence of temperature on daily peak electricity demand in South Africa. J. Energy Southern Africa 2013;24(4):63–70.
[17] Pardo A, Meneu V, Valor E. Temperature and seasonality influences on Spanish electricity load. Energy Econ. 2002;24(1):55–70.
[18] Robinson PJ. Modeling utility load and temperature relationships for use with long-lead forecasts. J. Appl. Meteorol. 1997;36(5):591–8.
[19] Ekonomou L. Greek long-term energy consumption prediction using artificial neural networks. Energy 2010;35(2):512–7.
[20] Hong W-C. Electric load forecasting by seasonal recurrent SVR (support vector regression) with chaotic artificial bee colony algorithm. Energy 2011;36(9):5568–78.
[21] Lee W-J, Hong J. A hybrid dynamic and fuzzy time series model for mid-term power load forecasting. Int. J. Electr. Power Energy Syst. 2015;64:1057–62.
[22] Nagi J, et al. A computational intelligence scheme for the prediction of the daily peak load. Applied Soft Computing 2011;11(8):4773–88.
[23] D. Paravan, et al., Influence of temperature on short-term load forecasting using the EPRI-ANNSTLF, in: Proc. 2 Balkan Power Conf, 2002.
[24] Apadula F, et al. Relationships between meteorological variables and monthly electricity demand. Appl. Energy 2012;98:346–56.
[25] Y.R. Chen, How does relative humidity affect electricity demand? The University of North Carolina at Charlotte, 2015.
[26] Griffin JS. Impacts of weather variations on energy consumption efforts at US air force installations. Air Force Inst of Tech Wright-Patterson AFB OH School of Eng. Manage. 2008.
[27] Mirasgedis S, et al. Models for mid-term electricity demand forecasting incorporating weather influences. Energy 2006;31(2):208–27.
[28] M. Analytics, State of the Market Report for PJM, Volume 2: Detailed Analysis. Eagleville, PA Monitoring Analytics, Independent Market Monitor for PJM, 2013.
[29] Zhou Y, et al. Modeling the effect of climate change on US state-level buildings energy demands in an integrated assessment framework. Appl. Energy 2014;113:1077–88.
[30] Wu TT, et al. Genome-wide association analysis by lasso penalized logistic regression. Bioinformatics 2009;25(6):714–21.
[31] Japkowicz N, Stephen S. The class imbalance problem: a systematic study. Intell. Data Anal. 2002;6(5):429–49.
[32] S. Geisser, Predictive inference. Vol. 55, CRC press, 1993.
[33] L.A. Jeni, J.F. Cohn, F. De La Torre. Facing imbalanced data–Recommendations for the use of performance metrics. in Affective Computing and Intelligent Interaction (ACII), 2013 Humaine Association Conference on, IEEE, 2013.
[34] Murtagh F, Legendre P. Ward's hierarchical agglomerative clustering method: which algorithms implement Ward's criterion? J. Classif. 2014;31(3):274–95.