Improving Load Time Series of Electric Power Systems based on the Temperatures

Farnaz Safdarian, John Penaranda, Seri Kang, Jonathan Snodgrass, Adam Birchfield, Thomas J. Overbye
Department of Electrical and Computer Engineering
Texas A&M University
College Station, TX
{fsafdarian, jpenaranda, serikang, snodgrass, abirchfield, overbye}@tamu.edu

***Abstract*—This paper finds the relationship between electrical load and temperatures to improve the creation of synthetic ("realistic but not realistic") load. The bus-level impact of temperatures is demonstrated on a synthetic Texas grid with 6717 buses. The proposed strategy can be used to scale each load entity to match the calculated load in each area. The load and temperature correlations are calculated and the curves based on recent data are presented. The curves are used to create temperature-aware load profiles. The results show that if the load curves are partitioned based on the temperatures, there is a higher correlation between load demand and temperatures in areas with higher populations. This is most likely due to higher percentages of residential and commercial load in urban areas, and illustrates the importance of updating synthetic load curves based on the changes in weather.**

***Index Terms*—Power Systems, synthetic grids, synthetic load, weather impact, linear regression**

Fig. 1. Weather zones in ERCOT [4].

I. Introduction

Weather conditions not only affect the available capacity of power generation, transmission and distribution systems, but they also change the demand as air conditioning constitutes a significant portion of the residential and commercial load. Utility companies and grid operators such as ERCOT (Figure 1) monitor weather conditions closely and adjust their operations as needed to ensure a reliable supply of electricity to customers. Planning for future weather conditions can also be considered by analyzing past events or developing hypothetical models. In our previous work [1] we proposed a strategy for direct inclusion of weather measurements such as wind speed, cloud coverage, and temperatures in AC optimal power flow (OPF) mainly to consider the availability of renewable energy resources and their available capacities at each hour. In [2] a strategy is proposed to identify wind and solar resource droughts in the United States and the renewable generation dispatch is validated in [3]. In this paper, we focus on the relationship between the weather and demand.

A variety of strategies are proposed in the literature to find the correlation between weather and load as it is essential for accurate load forecasting. Reference [5] proposes a strategy to predict the impact of weather on energy demand considering the probability distribution of power demand. A technical report from National Renewable Energy Laboratory [6] analyzes the impact of climate change on electric loads in the United States by predicting the future electricity load based on hisctorical data and climate change scenarios. Reference [7] examines the applications of weather data, including modeling outages and simulating Photovoltaic fleets. Reference [8] presents a short-term load forecasting of an urban area using an artificial neural network (ANN) and Bagged Regression Trees.

For studying the impact of weather on the demand, realistic data is required but the actual grid data are considered to be critical energy infrastructure information (CEII) with restricted access for research. We have used U.S. Energy Information Association (EIA) generation data [9], and census data to approximate the load in our previous work [10]–[12] and created realistic synthetic grids, validated using the methodology in [13]. Also, in our previous work [14], [15], we developed a strategy for creating synthetic load time series in synthetic power systems based on high-level system information, such as load types and their statistical properties. The proposed strategy includes a three-step approach, which involves selecting representative load profiles, creating synthetic loads using statistical methods, and validating the synthetic loads using real-world data and we demonstrated that the proposed approach is capable of creating realistic load profiles that represent the characteristics of real-world loads on bus level over the course of a year. However, the impact of weather changes on the load is neglected and it is highly required to adjust the load based on weather changes in different years.

In this paper, we propose a strategy to find the relationship between temperatures and load to improve the synthetic load creation in [14], [15] based on the impact of temperatures on the load at the bus level of synthetic grids. The studied case is a synthetic grid created over Texas state footprint in the U.S. with 6717 buses but the proposed strategy can be also implemented in other cases. The proposed strategy can be used to scale each load entity to match the determined load in each area. The load and temperature correlations are calculated and the curves based on the last 12 years data are presented. The curves are used to create temperature-aware load profiles. It should be noted that the proposed strategy is not limited to the synthetic grids and can be applied to industry cases as well.

TABLE I
TEMPERATURE RANGES WHERE LINEARITY WAS OBSERVED

Temperature Range (°F)
60 and below
60-70
70-80
80 and above

II. METHODOLOGY

A. Preparing Input Data of Weather Measurements

The first steps for preparing data include: 1) gathering adequate weather data for the electric grid areas of interest, 2) mapping weather information to relevant grid components, and 3) quantifying and analyzing how weather affects the generation and load. Historical weather information is gathered with hourly data from around the world dating back to the 1940s. The International Civil Aviation Organization (ICAO) and the World Meteorological Organization (WMO) provide data for thousands of weather stations worldwide, and electric utilities can supplement this information. For example, real-time weather information for about 5000 stations using the ICAO identifiers is available at [16] with several weather measurements including temperature, wind speed, and cloud cover percentage. Geographic information for grid components is becoming more readily available through geographic information systems, visualization techniques, and requirements from regulatory organizations. The weather stations are mapped to the electrical buses based on the closest geographical distance using their latitudes and longitudes. After the weather information is assigned to each load bus, the relationship between load and temperatures can be either studied at the bus level or area level. To find a more general relationship, we calculated the average hourly temperatures for each area.

B. Preparing Input Load Data

Several utilities such as Electric Reliability Council of Texas (ERCOT) disclose hourly load data for each area. [17] However, the load data at the bus level is required for several power system studies. References [14], [15] provide a detailed explanation of creating hourly synthetic load time series based on the publicly available data with realistic commercial, residential, and industrial demand ratios from different building types and their load templates for one year. Several assumptions are made for creating synthetic load data in [14], [15] including that the statistical properties in data are not changing over the years; the loads follow a seasonal pattern dependent on the time of year; there is homogeneity in the data, meaning that individual loads behave similarly and the variance is considered to be constant over time.

However, to adjust the load time series based on the weather measurements, it is required to find the relationship between the load and temperatures and consider the predicted load growth over the years. ERCOT load data is used to find this relationship and improve the synthetic load time series accordingly.

C. Correlations between hourly load and temperatures

Two techniques were used to determine the relationship of weather and demand: R^2 and correlation. Correlation explains the strength of the relationship between two variables whereas R^2 explains to what extent the variance of one variable explains the variance of the second variable. Pearson-r (or correlation) is the ratio between the covariance of two variables and the product of their standard deviations—this results in a normalized measurement of the covariance with correlation values ranging between -1 and 1. Regardless of sign, should there be a correlation between the two variables, the result will be between 0.3–0.7 for moderate correlation and above 0.7 if there is a high correlation. R^2 is a measure representing the proportion of the variance for a dependent variable that is predictable from the independent variable. In other words, this value represents how close the data are to the fitted regression line. These statistical measures are used to identify how closely related are the two variables.

D. Relationship between hourly load and temperatures

Temperatures were partitioned to the ranges where a linear relationship between the load and temperatures was observed. Once temperature ranges were identified, a linear regression model was applied using the sci-kit toolkit in python and Pearson correlation in the Panda's library. The relationship is calculated by fitting a line that is determined by finding the linear equation with the smallest sum of squared differences between observed and predicted values so that the error between the fitted curve and data was optimized without any overfitting. The calculated regression line is used to find the proportional increase or decrease in electric consumption subject to a change in temperature. Load magnitudes were normalized by dividing by the average peak demand for each zone to allow comparisons of the correlations across the various zones. Since the sign of the correlation coefficient between temperature and load changes between seasons, the load data were divided into groups based on the temperature ranges that resulted in a larger correlation between the two variables. Each zone has multiple regression lines that represent the relationship found for each temperature range. Table I shows temperature ranges where linearity was observed.

TABLE II
TEXAS SYNTHETIC GRID STATISTICS

Parameter	Numerical Value
Number of buses	6,717
Number of generators	731
Number of loads	5,095
Number of Switched Shunts	634
Number of substations	4,894
Number of transmission lines	7,173
Maximum load (MW)	74,667
Maximum generation (MW)	104,914

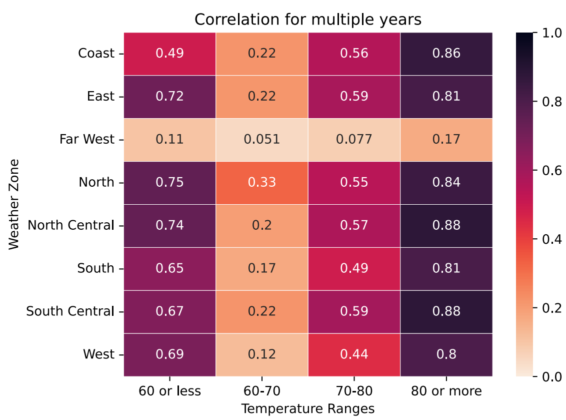


Fig. 2. Correlation for past 10 years (combined)

III. CASE STUDY

The electric grid used in this study is a synthetic network geographically sited in Texas, U.S., with around 7000 buses created based on actual generator data [18] representing transmission lines with the same voltage levels that are used in Texas grid as 345 kV, 138 kV, and 69 kV. Table II shows the main characteristics of this case.

The Texas synthetic grid models are divided into the same areas as the actual Texas grid so that the study can be facilitated. Historical publicly available load data is used for each area. Figure 1 shows the eight weather zones; regions in light grey fall outside of ERCOT and are serviced by other system operators.

IV. RESULTS

Table III shows the relationship between the average temperatures and average load in different areas of Texas and Figure 2 shows the absolute correlation of load and temperatures in Texas zones based on the last 12 years. As expected, there is a higher correlation between the load and temperatures in higher and lower temperatures where air conditioning is required. Also, areas with dominant residential and commercial loads have a higher correlation but areas such as the Far West with dominant industrial loads have a lower correlation. Also as expected, the temperature ranges between 60–70°F and 70–80°F which are referred to as “nice weather” and reduce the need for air conditioners, have the lowest correlation for

TABLE III
THE PARAMETERS OF FITTED LINE IN DIFFERENT TEMPERATURE RANGES AND AREAS, CORRELATIONS AND R-SQUARED CRITERIA

Temperature F	Coefficient	Intercept	R^2	Correlation
below 60	-0.00878	1.284209	0.242725	-0.49267
60-70	0.007216	0.345529	0.04897	0.221292
70-80	0.022631	-0.73588	0.308775	0.555676
80 and above	0.032822	-1.50514	0.66224	0.813781

(a) Coast

Temperature F	Coefficient	Intercept	R^2	Correlation
below 60	-0.00878	1.284209	0.242725	-0.49267
60-70	0.007216	0.345529	0.04897	0.221292
70-80	0.022631	-0.73588	0.308775	0.555676
80 and above	0.032822	-1.50514	0.66224	0.813781

(b) East

Temperature F	Coefficient	Intercept	R^2	Correlation
below 60	-0.00352	1.118819	0.011083	-0.10528
60-70	0.00513	0.600463	0.002554	0.050536
70-80	0.008527	0.35701	0.005908	0.076863
80 and above	0.010601	0.215498	0.027725	0.166508

(c) Far West

Temperature F	Coefficient	Intercept	R^2	Correlation
below 60	-0.01172	1.449032	0.56044	-0.74863
60-70	0.008399	0.294721	0.107646	0.328094
70-80	0.020315	-0.53602	0.298574	0.54642
80 and above	0.027657	-1.10742	0.705793	0.840115

(d) North

Temperature F	Coefficient	Intercept	R^2	Correlation
below 60	-0.01607	1.651761	0.548497	-0.74061
60-70	0.006726	0.341361	0.039138	0.197833
70-80	0.024513	-0.89971	0.329607	0.574114
80 and above	0.035559	-1.76425	0.778759	0.882473

(e) North Central

Temperature F	Coefficient	Intercept	R^2	Correlation
below 60	-0.01966	1.869804	0.418358	-0.64681
60-70	0.00622	0.367771	0.029079	0.170526
70-80	0.02048	-0.62244	0.241543	0.491471
80 and above	0.031014	-1.42672	0.650781	0.80671

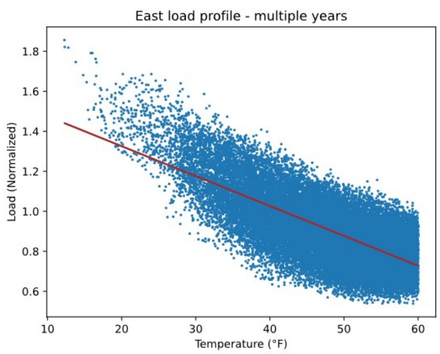
(f) South

Temperature F	Coefficient	Intercept	R^2	Correlation
below 60	-0.01546	1.623227	0.453906	-0.67373
60-70	0.007477	0.294607	0.047924	0.218916
70-80	0.025879	-0.99479	0.351395	0.592786
80 and above	0.034299	-1.63408	0.773004	0.879206

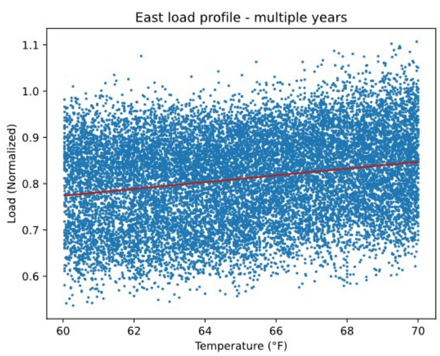
(g) South Central

Temperature F	Coefficient	Intercept	R^2	Correlation
below 60	-0.01383	1.602931	0.475928	-0.68988
60-70	0.004206	0.556583	0.015264	0.123548
70-80	0.017777	-0.39074	0.19486	0.44143
80 and above	0.02721	-1.12819	0.636748	0.797965

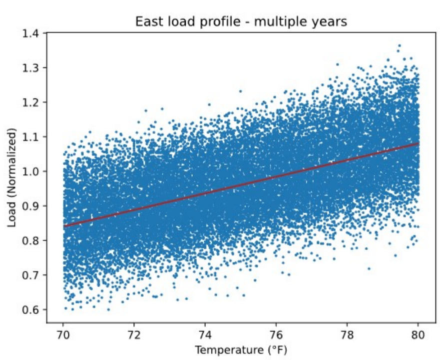
(h) West



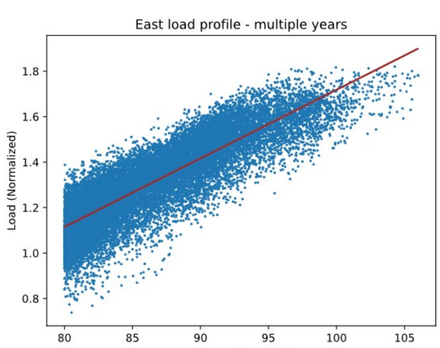
(a) 10-60 °F



(b) 60-70 °F



(c) 70-80 °F



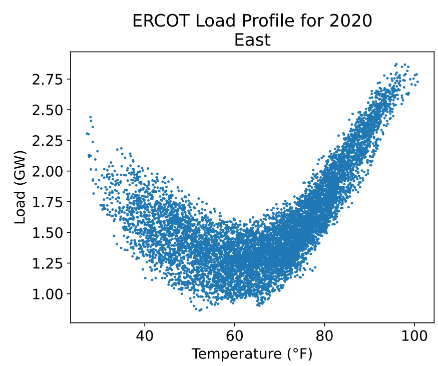
(d) 80-105 °F

Fig. 3. Normalized load based on temperatures in East area over years 2010-2021.

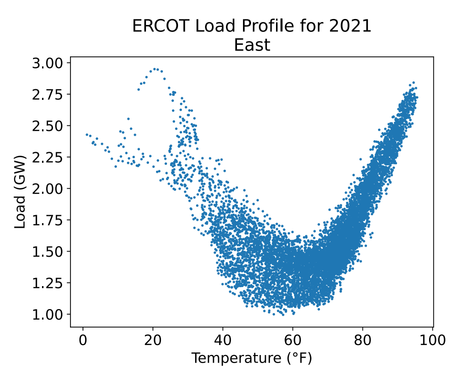
all weather zones and also contained the most inconsistent characteristics due to the differences with end-use behavior. Lower correlations can also be due to the lower number of weather-dependent loads and dominant industrial load in areas such as the Far West weather zone, where correlation coefficients were lower than in other regions. In general, areas with higher populations and residential or commercial loads have a higher impact from temperatures whereas areas with dominant industrial loads are not influenced by temperature as much.

For a closer look into an area with a dominant residential and commercial load, the East area is selected. Figure 3 shows the linear relationship of the hourly load and hourly temperatures while the average values over this area are considered. The blue dots show the normalized demand based on temperatures over the recent 12 years (from 2010 to 2021). The normalized load is the hourly load divided by the average load for the whole year. The red line is the fitted curve that shows the linear relationship between the load and temperatures in this area.

Also, for a more detailed study, we select the years 2020 and 2021. Both of these years are interesting for study as in 2020 COVID became prevalent and changed the overall load profile and in February 2021 the winter storm Uri in Texas created unusually low temperatures and high demand. Figure Figure 4 shows a comparison of the overall load and temperature curves in these two years.



(a) 2020



(b) 2021

Fig. 4. Electric consumption profiles for East weather zone.

Load-temperature profile for 2020 remains similar to previous years although the residential load was increased in this year due to the lock-down. In 2021, unusually low temperatures are observed in Figure Figure 4 but the load shedding was unavoidable as enough generation was not available and this changed the linear behavior of the load in lower temperatures [19].

However, overall load-temperature profiles were consistent in the same temperature ranges even considering the effects of winter storm in 2021. However, one should note the very low temperatures in 2021 are outliers in Texas which resulted in the unavailability of several generators and multiple load shedding. [19] If the unusually low temperature values were removed, this resulted in a load profile that was nearly identical to the previous year. The fitted lines for the years 2020 and 2021 are shown in Figure 5 for comparisons. These preliminary results show a similar linear fit curve for ERCOT's East region with minimal differences between the two years which remain consistent for older years with a shift in the load as mentioned in [5].

According to the results, there are significant increases in electric consumption for extremely high and low temperatures. The fitted lines are used to adjust the hourly synthetic load in the bus level of synthetic grids based on the temperatures from the closest weather station to each bus.

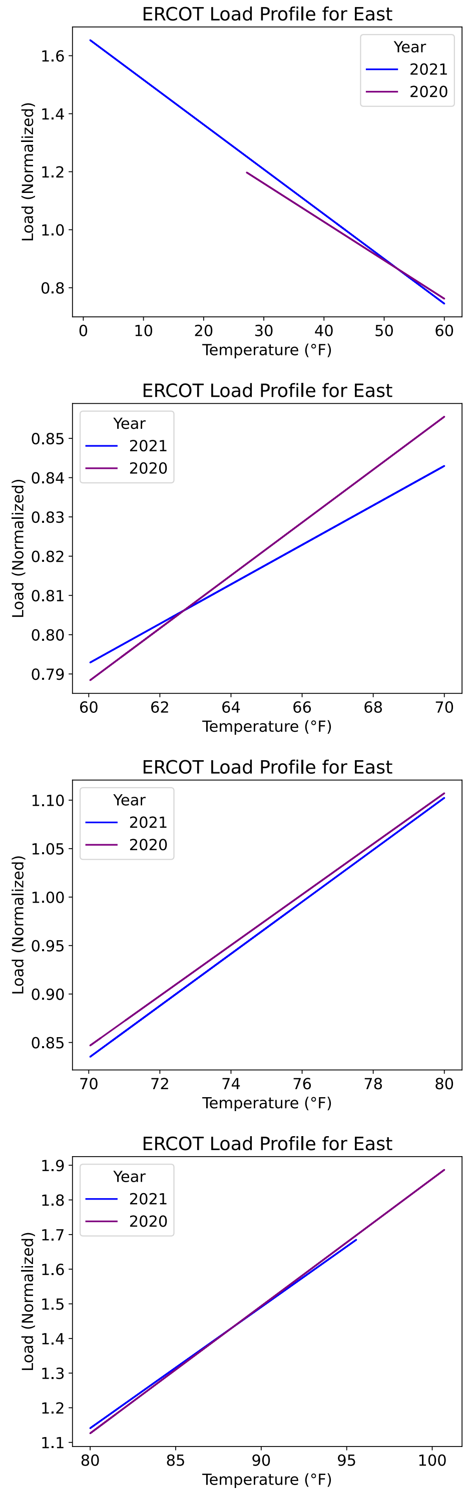


Fig. 5. Linear fit curves for East weather zone.

V. Conclusion

Weather conditions (mainly temperature) will either increase or decrease residential and commercial demand because of the large HVAC (heating, ventilation and air condition) load. To quantify this impact, we calculated the correlation between the average of hourly load in each ERCOT weather area and the averages of hourly temperatures in the same area to improve the creation of synthetic but realistic load. This improved the strategy from our previous work by including the impact of temperatures on the load at the bus level of large electrical grids. The proposed strategy is used to scale each load entity to match the determined load in each area.

The results show that in areas with higher populations and dominating residential and commercial load, there is a higher correlation between load demand and temperatures. We quantify this impact using linear regression, and use it to create future load scenarios for synthetic grids.

Future work includes using multivariate regression to model the impact of other weather factors on load, such cloud coverage, time of the day, humidity, wind speed, and wind direction. These factors, especially cloud coverage, could be used to model the variation in electricity load for temperature ranges between 60–70°F and 70–80°F. As demonstrated this paper, in this temperature range, the temperate has a low correlation to electrical load.

VI. Acknowledgements

This work was partially supported through funding provided by the Power Systems Engineering Research Center (PSERC) through project S-99.

References

[1] T. J. Overbye, F. Safdarian, W. Trinh, Z. Mao, J. Snodgrass, and J. Yeo, "An approach for the direct inclusion of weather information in the power flow," in *Proc. 56th Hawaii International Conference on System Sciences (HICSS)*, 2023.

[2] J. L. Wert, F. Safdarian, A. Gonce, T. Chen, D. Cyr, and T. J. Overbye, “Wind resource drought identification methodology for improving electric grid resiliency,” 2023.

[3] J. L. Wert, T. Chen, F. Safdarian, J. Snodgrass, and T. J. Overbye, “Calculation and validation of weather-informed renewable generator capacities in the identification of renewable resource droughts,” 2023.

[4] ERCOT, “Maps,” Ercot.com, 2022. [Online]. Available: https://www.ercot.com/news/mediakit/maps

[5] J. Lee and A. E. Dessler, “The impact of neglecting climate change and variability on ercot’s forecasts of electricity demand in texas,” *Weather, Climate, and Society*, vol. 14, no. 2, pp. 499–505, 2022.

[6] E. K. Patrick Sullivan, Jesse Colman, “Predicting the response of electricity load to climate change,” Tech. Rep., 2015. [Online]. Available: https://www.nrel.gov/docs/fy15osti/64297.pdf

[7] J. Black, T. H. A. Hoffman, J. Roberts, and P. Wang, “Weather data for energy analytics: From modeling outages and reliability indices to simulating distributed photovoltaic fleets,” *IEEE Power and Energy Magazine*, vol. 16, no. 3, 2018.

[8] V. Dehalwar, A. Kalam, M. L. Kolhe, and A. Zayegh, “Electricity load forecasting for urban area using weather forecast information,” pp. 355–359, 2016.

[9] (EIA), US Energy Information Administration, “Form eia-860,” Tech. Rep., 2021. [Online]. Available: https://www.eia.gov/electricity/data/eia860/

[10] A. B. Birchfield, K. M. Gegner, T. Xu, K. S. Shetye, and T. J. Overbye, “Statistical considerations in the creation of realistic synthetic power grids for geomagnetic disturbance studies,” *IEEE Transactions on Power Systems*, vol. 32, no. 2, pp. 1502–1510, 2017.

[11] K. M. Gegner, A. B. Birchfield, T. Xu, K. S. Shetye, and T. J. Overbye, “A methodology for the creation of geographically realistic synthetic power flow models,” in *2016 IEEE Power and Energy Conference at Illinois (PECI)*. IEEE, 2016, pp. 1–6.

[12] T. Xu, A. B. Birchfield, K. M. Gegner, K. S. Shetye, and T. J. Overbye, “Application of large-scale synthetic power system models for energy economic studies,” in *Proceedings of the 50th Hawaii International Conference on System Sciences*, 2017.

[13] A. B. Birchfield, E. Schweitzer, M. H. Athari, T. Xu, T. J. Overbye, A. Scaglione, and Z. Wang, “A metric-based validation process to assess the realism of synthetic power grids,” *Energies*, vol. 10, no. 8, 2017. [Online]. Available: https://www.mdpi.com/1996-1073/10/8/1233

[14] H. Li, A. L. Bornsheuer, T. Xu, A. B. Birchfield, and T. J. Overbye, “Load modeling in synthetic electric grids,” in *2018 IEEE Texas Power and Energy Conference (TPEC)*. IEEE, 2018, pp. 1–6.

[15] H. Li, J. H. Yeo, A. L. Bornsheuer, and T. J. Overbye, “The creation and validation of load time series for synthetic electric power systems,” *IEEE Transactions on Power Systems*, vol. 36, no. 2, pp. 961–969, 2021.

[16] “Current Weather and Wind Station Data”. [Online]. Available: https://aviationweather.gov/adds/dataserver_current/current/metars.cache.csv

[17] ERCOT, “Hourly load data archives,” Ercot.com, 2021. [Online]. Available: https://www.ercot.com/gridinfo/load/load_hist

[18] [Online]. Available: https://electricgrids.engr.tamu.edu/

[19] “FERC—NERC—The Report in Region”. [Online]. Available: https://www.naesb.org/pdf4/ferc_nerc_regional_entity_staff_report_Feb2021_cold_weather_outages_111621.pdf