By Xiangtian Zheng,
Dongqi Wu,
Liam Watts,
Efstratios N. Pistikopoulos,
and Le Xie

Analyzing Extreme Events in Power Systems

An Open, Cross-Domain Data-Driven Approach

OVER THE PAST SEVERAL YEARS THE ELECTRIC POWER SECtor has been challenged by a number of extreme events around the globe. Significant societal and economic shocks were due to the rapid spread of COVID-19 around the world. In addition to the pandemic, there have been several extreme weather and societal disruptions to the electricity sector, such as the February 2021 Texas power outage and the 9 p.m. nine-minute blackout event in India.

These major societal-level shocks interrupted the operation of the electric power system in significant manners that would have otherwise been difficult to predict. While their full impact on the electricity sector have yet to be fully realized, this article attempts to provide a summary of ongoing activities that aim at better understanding and analyzing the short-run

Digital Object Identifier 10.1109/MPE.2022.3199895
Date of current version: 18 October 2022

impact of such extreme events. In particular, this article documents several open, cross-domain, data-driven approaches to modeling and analyzing the changes in electricity consumption and related electrical variables due to these shocks. These projects collectively serve as examples of what could be done to better prepare for extreme events in future power systems.

This article summarizes several ongoing projects that are aimed at understanding the short-term impact of COVID-19 as well as other extreme events on the electric power sector. An open, cross-domain data-driven approach is shown to be effective in providing science-based decision support for operators and planners in the electricity sector. Key challenges in data collection, processing, and interpretation are illustrated through case studies. With more extreme events coming, the electric power sector would benefit from having a more systematic cross-domain data-hub approach to analyzing and predicting the impact of extreme events.

A Cross-Domain Data-Driven Approach to Analyzing the Short-Term Impact of COVID-19 on the U.S. Electricity Sector

The rapid spread of COVID-19 across the United States caused unprecedented significant impacts on the electricity sector in 2020. It is imperative to understand such an extreme event comprehensively in a scientific manner. Figure 1 shows the architecture of this cross-domain study, including steps of

- ✔ integration of cross-domain data
- ✔ backcast model training
- ✔ quantification of the impact of COVID 19 on the US power sector
- ✔ comprehension of the change in electricity consumption.

Aiming to provide insights through a data-driven approach, we first develop a cross-domain open-access data hub by integrating heterogeneous data, including region-wide electricity consumption, weather, mobile device location record, and satellite imaging data. Leveraging the data hub, we analyze the intensity and dynamics of impacts of COVID-19 on the electricity sector. Quantified changes in electricity consumption reveal that the electricity sector in the United States experienced rapid changes in 2020. Note that this study only focuses on the early stages of the pandemic in 2020. Creating a cross-domain data set that includes factors, such as the number of COVID-19 cases, the degree of social distancing, and the level of commercial activity offers fresh insights which power system operators might draw upon for impact analysis.

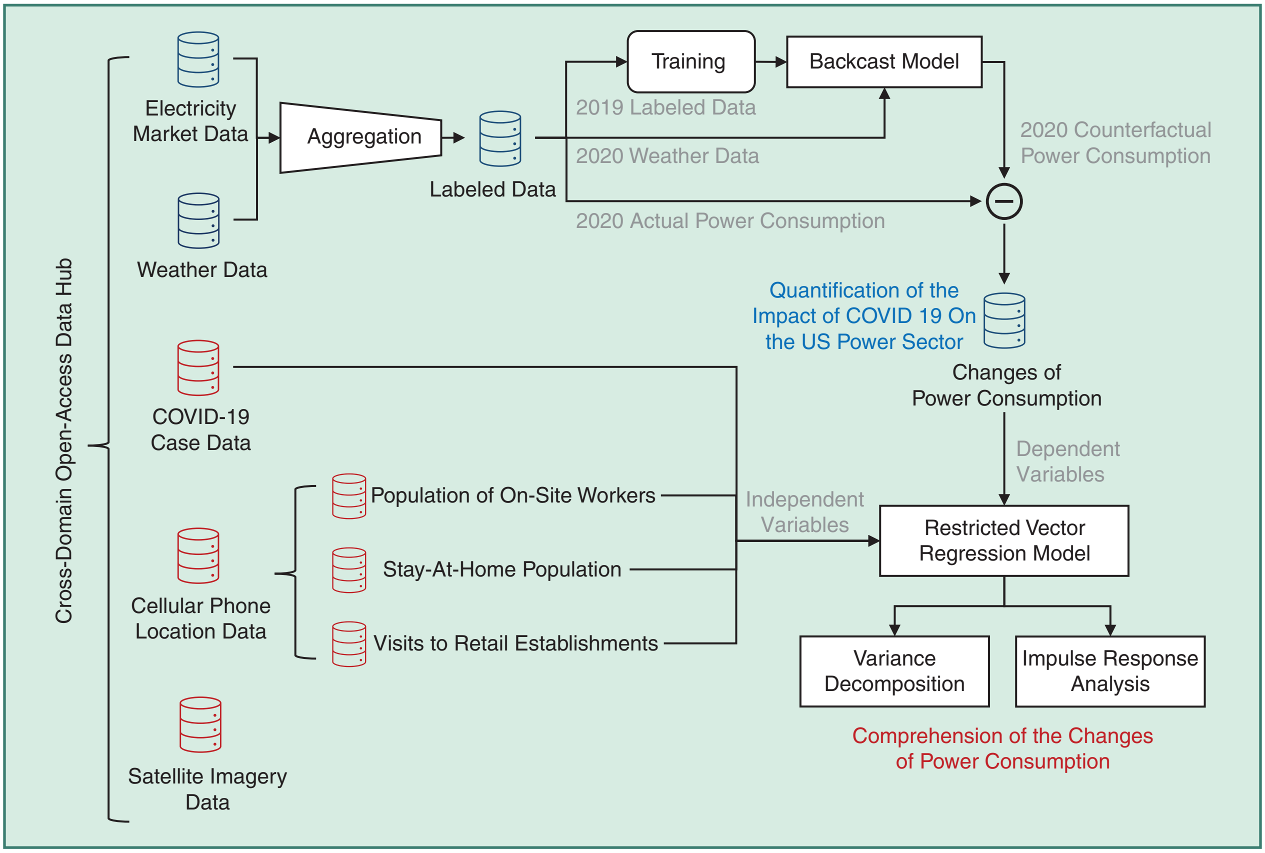

figure 1. The architecture of developing a cross-domain open-access data hub and analyzing the impact of COVID 19 on the U.S. power sector.

Developing a Cross-Domain Open-Source Data Hub

We aggregate electricity load and generation data from seven electricity markets, including California (CAISO), Midcontinent (MISO), New England (ISO–NE), New York (NYISO), Pennsylvania–New–Jersey–Maryland (PJM), Southwest Power Pool (SPP), and Texas (ERCOT). To get cross-domain insights to understand how such an extreme event affected electricity consumption, we integrate weather data, satellite imaging data, public health data, and mobile device location data in the territory of these markets. Particularly, we define several indicators by processing the mobile device location data, including "stay-at-home population," "numbers of on-site workers," and "mobility in the retail sector," which are considered as the key influencing factors in the analysis.

Here, to provide an intuitive representation of the signification reduction in electricity consumption, we visualize the night-time light intensity in Figure 2 using night-time light data from satellite imagery. It provides a preview of one subsequent analysis result that the shut-down rate of commercial activity is a key factor for the change of electricity consumption.

Quantifying Changes in Electricity Consumption in the United States

Using the aggregated historical data, we quantify the impacts of COVID-19 on the electricity sector by designing a backcast model to estimate counterfactual electricity consumption. Specifically, the goal of developing a backcast model is to provide a statistically robust baseline of electricity consumption in the absence of COVID-19 against from which the reduction in electricity consumption can be quantified. The backcast model is an ensemble of multiple neural networks, each of which is a function mapping several potential factors, including weather variables, date, and economy, to the estimated electricity consumption. The final output of the backcast model is calculated by averaging over the outputs of multiple neural networks.

We first analyze the reduction in electricity consumption in New York State. We present the comparison between the estimated counterfactual baseline for 2020, actual daily electricity consumption profiles in 2020, and historical data in 2019 in Figure 3. The strong match between the counterfactual and actual curves in 2020 before COVID-19, along with similar patterns of the counterfactual curves during COVID-19 and historical curves in 2019, demonstrates that the backcast model can reliably estimate electricity consumption in the absence of COVID-19. Furthermore, Table 1 shows the estimated change in electricity consumption in seven electricity markets and four illustrated cities to compare the impact of COVID-19 across markets and cities.

All markets experienced a reduction in electricity consumption in both April and May 2020, but with diverse magnitudes of the reductions: MISO and NYISO suffered the most significant reduction, while SPP and ERCOT experienced the least. Finally, all markets exhibited a rebound in June 2020 that may be credited to the economy reopening. The impact of COVID-19 on demand reduction is more obvious in urban areas with higher population density and larger share from the business and commercial sector, such as New York City and Boston, where the electricity consumption reduced 14.1% and 11.3%, respectively, in April 2020. On the other hand, cities like Houston, where the population and commercial activities are more dispersed, did not show a significant demand reduction.

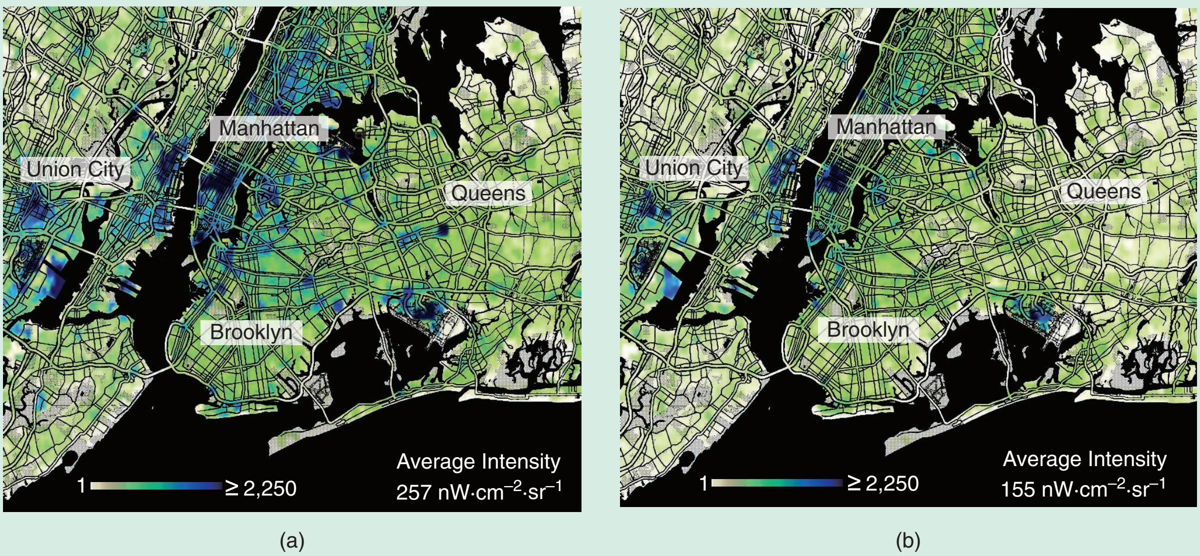

figure 2. A comparison of night-time electricity consumption before and during the COVID-19 pandemic in New York City by satellite image data. (a) Night-time light intensity at 1 a.m. on 8 February 2020. (b) Night-time light intensity at 1 a.m. on 25 April 2020.

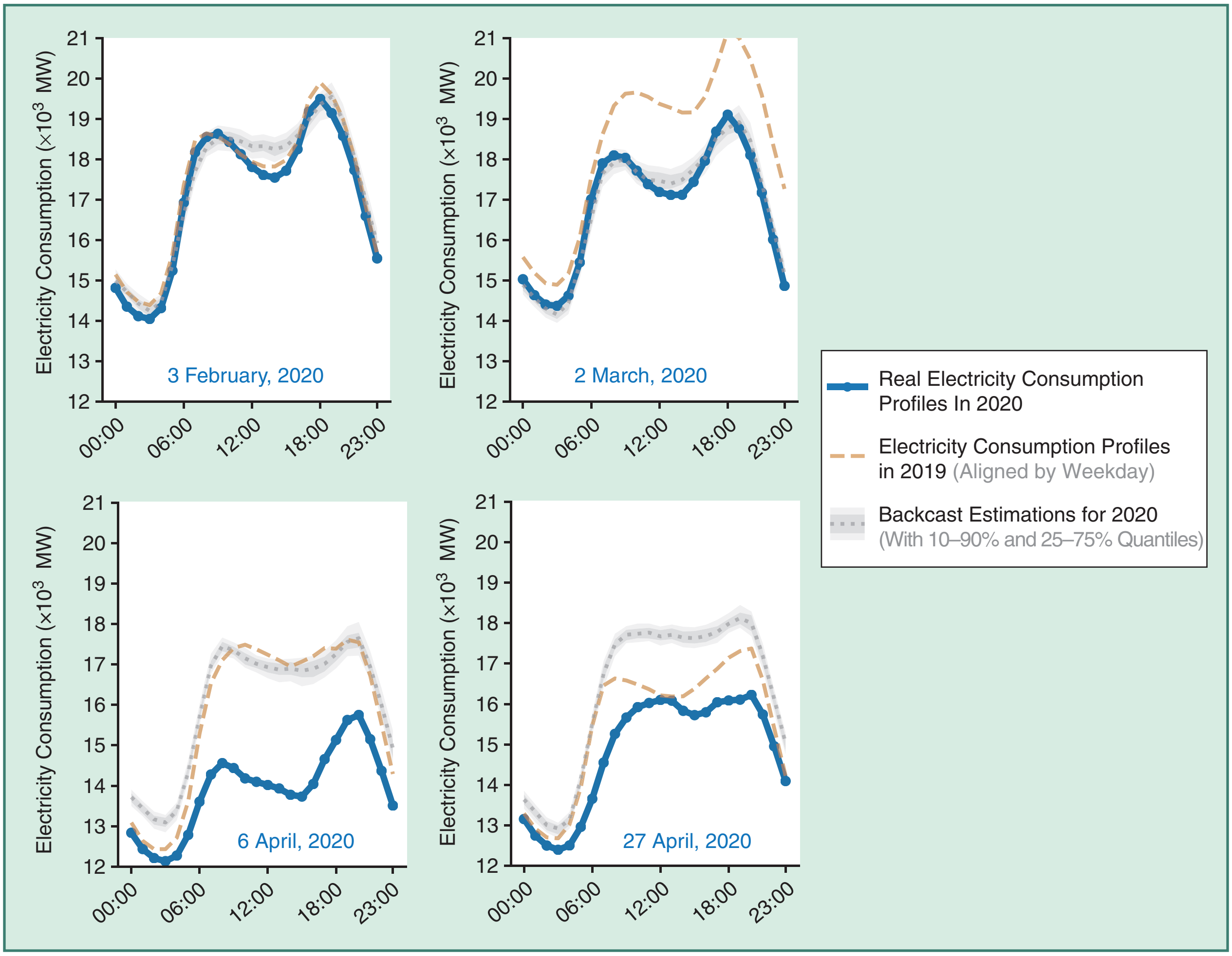

figure 3. A comparison between the estimated counterfactual baseline for 2020, actual daily electricity consumption in 2020, and historical daily electricity consumption profiles in 2019 in NYISO. (Source: Ruan et al., 2020; used with permission.)

table 1. A comparison of changes in electricity consumption across major electricity markets in the United States

Electricity Consumption Reduction (%)	February	March	April	May	June
CAISO	–1.3	–2.7	9.2	6.5	0.3
MISO	–0.1	1.8	10.2	10.7	3.5
ISO-NE	2.2	5.2	9.5	10.4	1.8
NYISO	0.8	4.5	10.2	10.5	5.7
PJM	0.5	2.7	9.4	7.4	0.1
SPP	–0.9	2.5	7.7	9.2	2.7
ERCOT	–1.5	–1.3	6.4	4.4	2.4
Boston	0.4	7.1	11.3	9.4	0.4
New York City	0.4	5.3	14.1	14.8	11.1
Houston	–0.6	–0.5	5.3	3.6	4.4
Kansas City	0.1	0.2	9.0	7.0	0.2

Analyzing Influencing Factors of the Change in Electricity Consumption

To comprehend the changes in electricity consumption across the United States during COVID-19, we investigated three influencing factors, including public health (the number of COVID-19 cases), social distancing (the size of the stay-at-home population and the population of on-site workers), and commercial activity (the frequency of visits to retail establishments). These influencing factors need to be considered because they reflect social activities closely related with electricity consumption from different perspectives. We leverage a restricted vector autoregression (VAR) model to analyze the complex multidimensional corelationship between multiple variables, including the number of COVID-19 cases, population of on-site workers, stay-at-home population, mobility in the retail sector, and electricity consumption. A restricted VAR model is essentially a linear regression model, which is a function that maps the historical values of multiple variables to their current values. The model parameters can be used

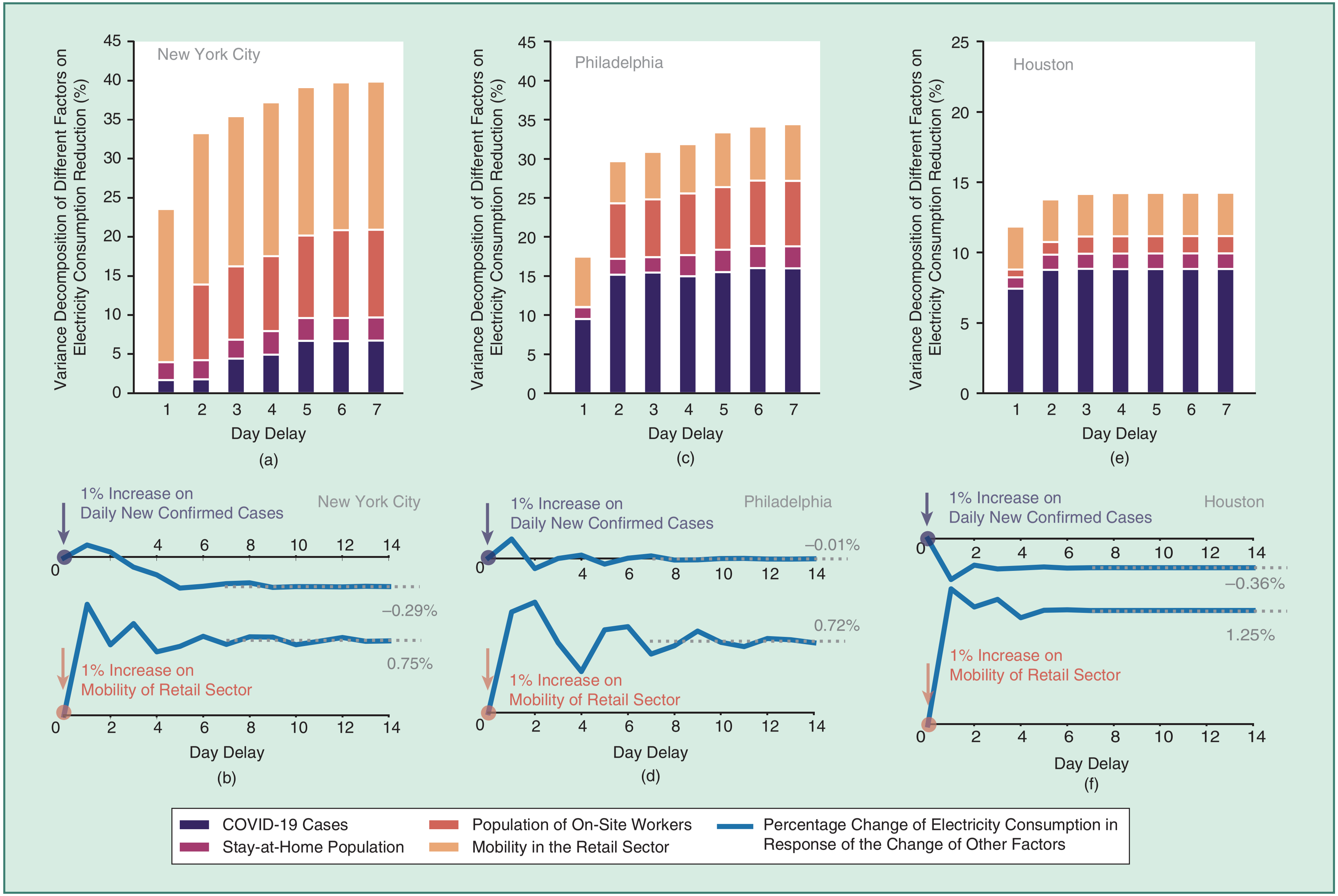

figure 4. The results of variance decomposition and impulse response analysis for New York City, Philadelphia, and Houston. (a) Variance decomposition for New York City. (b) Impulse response analysis for New York City. (c) Variance decomposition for Philadelphia. (d) Impulse response analysis for Philadelphia. (e) Variance decomposition for Houston. (f) Impulse response analysis for Houston. (Ruan et al., 2020; used with permission.)

> Texas, among all affected states, was hit particularly hard mainly from a state-wide power outage that lasted more than three days, which affected more than 4.5 million customers.

to understand the linear correlation between time series of these variables. Note that while the restricted VAR model can explain how multiple variables interact, we only focus on the part of the model parameters related to how other influencing factors impact the dynamic process of electricity consumption. We use variance decomposition analysis [Figure 4(a), (c), and (e)] that leverages the corresponding parameters of current/historical influencing factors to quantify the contribution of influencing factors to the changes in electricity consumption. We also use impulse response analysis [Figure 4(b), (d), and (f)] to reflect how electricity consumption evolves in response to a unit change of one influencing factor.

Based on the analysis results, we have several key findings:

- ✔ The mobility in the retail sector is the most significant factor that accounts for a major proportion of the change in electricity consumption, which is supported by a consistently leading proportion across different cities in both the variance decomposition analysis [Figure 4(a), (c), and (e)] and impulse response analysis [Figure 4(b), (d), and (f)].
- ✔ The number of COVID-19 cases is not a key factor, which is supported by a low sensitivity in the impulse response analysis [Figure 4(b), (d), and (f)].
- ✔ Electricity consumption in cities with a mild reduction, such as Houston, may be highly sensitive to some influencing factors such as the level of commercial activity [Figure 4(f)].

Cross-Domain Data-Driven Approach to Analyzing the Impact of the Extreme 2021 Winter Storm in Texas

The severe Winter Storm Uri in February 2021 hit the southern states of the United States and caused record levels of low temperature, snow, and freezing, which caused widespread disruption of many public services including electricity supply. Texas, among all affected states, was hit particularly hard mainly from a state-wide power outage that lasted more than three days, which affected more than 4.5 million customers, caused supply shortage of food, water, heating, and medical care, and in turn led to the death of more than 246 people.

Assessing Possible Corrective Measures for the 2021 Texas Power Outage

Figure 5 presents how load shedding and generation outage evolved in the Texas grid from 14–18 February 2021. On the night of 14 February 2021, the temperature throughout Texas dropped below subfreezing temperatures and caused a spike in residential energy consumption, because most Texas residences had poor heat insulation and used resistive electric heaters with relatively low efficiency. The total demand in Texas surged to a historical 69,692 MW that was roughly 3,200 MW higher than the previous winter record in 2018. The sudden drop in load capacity around 2 a.m. on 15 February marked the beginning of an about 70-h forced load shedding that once reached a maximum of about 20,000 MW. The total electricity cost for ERCOT in a single day (16 February) alone exceeded

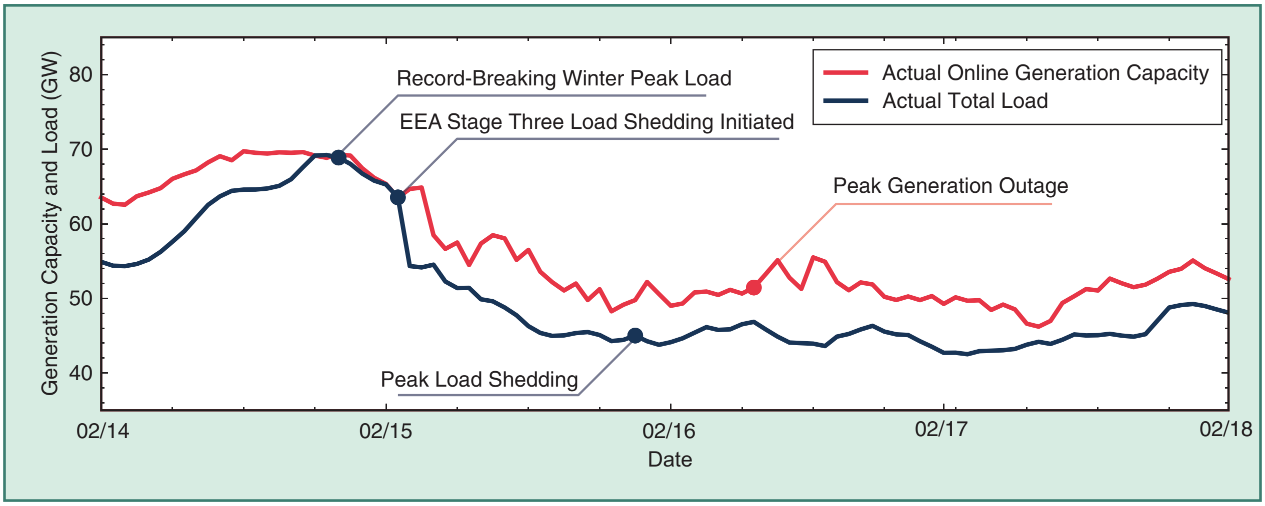

figure 5. The actual online generation capacity and load in Texas interconnection between 14–18 February 2021. EEA: Energy Emergency Alert.

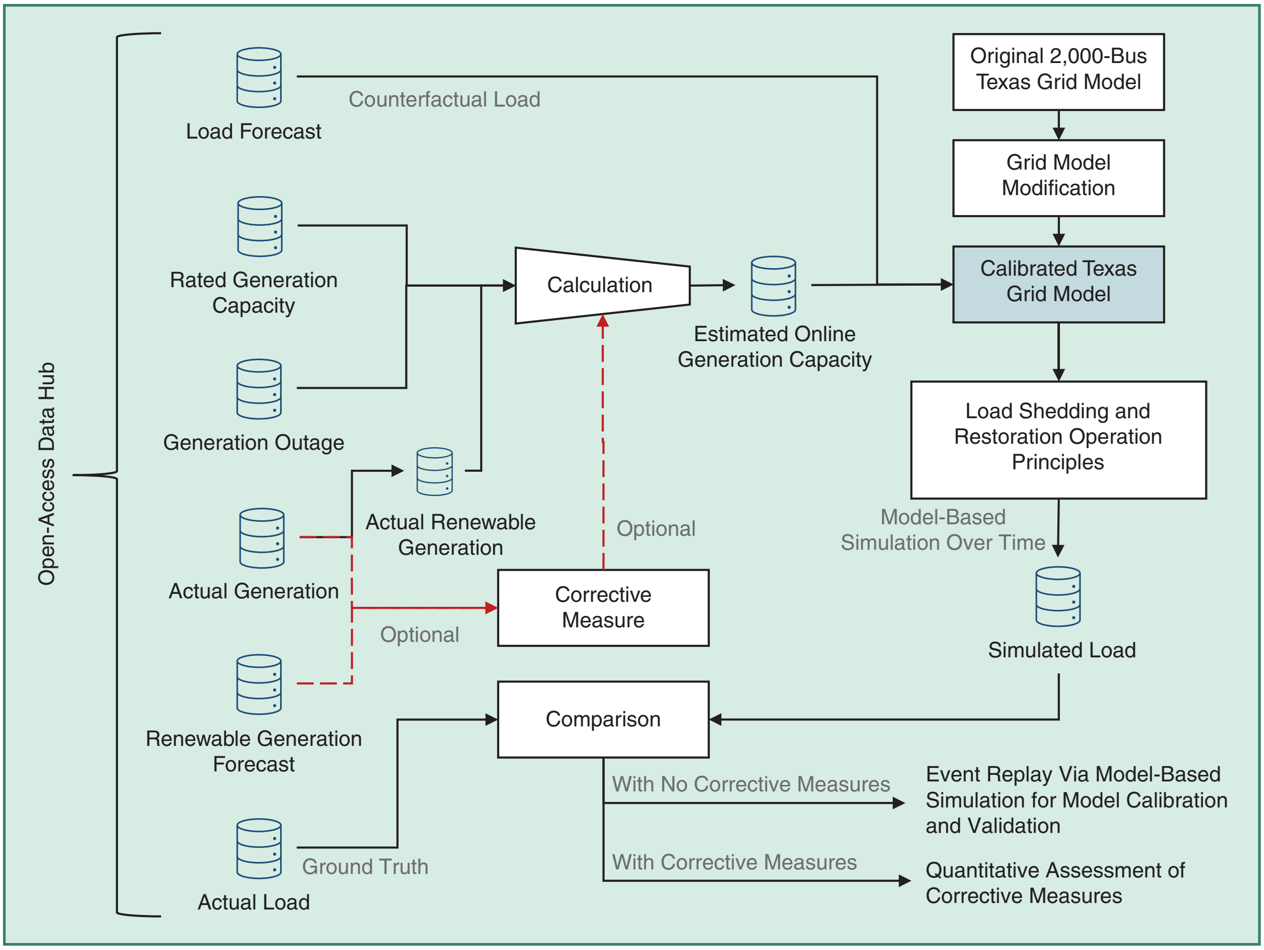

figure 6. The architecture of data collection, Texas grid model calibration, power outage event replay, and assessment of possible corrective measures.

10 billion dollars, which was more than the total ERCOT electricity cost of the entire year 2020 at 9.8 billion dollars, due to the exorbitant wholesale market price. The total economic damage caused by this power outage was estimated to be US$200 billion dollars.

In the aftermath of this widespread blackout, it is imperative for the broader energy community to investigate the following:

- ✔ why and how such a disastrous blackout occurred in the Texas power grid
- ✔ what could potentially be done to reduce and eventually eliminate the extent of the outage.

Therefore, as shown in Figure 6, we developed a data-driven approach for replaying the outage event and assessing possible corrective measures.

To assist researchers and policy makers with different backgrounds of expertise, we developed an open-access data hub using only nonrestrictive publicly available data about the Texas power grid and the 2021 power outage event. This data hub consists of a calibrated 2,000-bus synthetic grid model and a collection of cross-domain data related to the power outage from various sources that are prepared to be used with the synthetic model directly. Specifically, the topology, generation capacity, and loads of the grid model are calibrated so that the simulation results through direct current optimal power flow closely

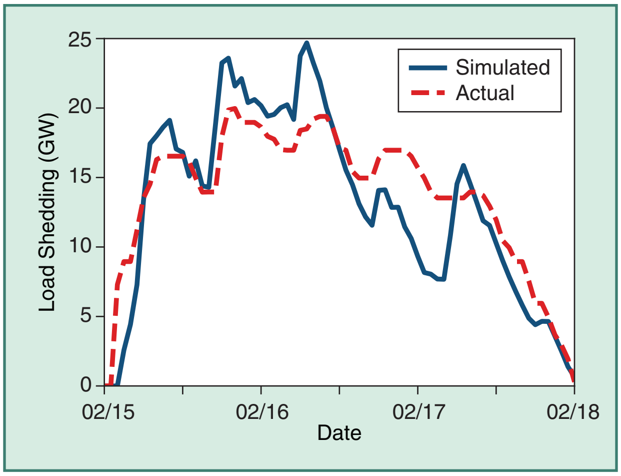

figure 7. The loading shedding curves of blackout event reproduction via simulation on the synthetic grid.

representing the state of the Texas grid around February 2021. In our study, we first used this synthetic model and data set to perform a simulated reproduction of the 2021 Texas power outage event, and then quantitatively evaluated the effectiveness of possible technical solutions that could potentially mitigate the electricity scarcity under similar extreme weather conditions.

The fidelity of our synthetic model, data, and the associated simulation method is validated by reproducing the event timeline in simulation and comparing with real ERCOT and Energy Information Administration records. We mimic the forced load shedding and restoration process based on the operation principles from ERCOT operation protocols and use our model to compute the minimally necessary load shedding that is necessary for safe operation of the grid. We use energy-not-served (ENS), a widely used power system reliability index, as a metric to quantitatively assess the extent of forced load shedding in the event reproduction and hypothetic scenarios. The simulated blackout reproduction in Figure 7 shows the following:

- ✔ the total ENS of the simulated load shedding is 999 GWh, having a 4.3% difference in comparison with the actual 956 GWh
- ✔ the correlation coefficient between the simulated and actual load shedding curves is 0.88
- ✔ the largest gap between the simulated and actual load shedding is 6.7 GW

The unavoidable mismatch can be attributed to errors in the modeling of synthetic network and uncertainty in system operation under such emergency conditions. Further research could be devoted to estimate the error range in a more systematic manner.

We then used the data hub as a platform to model, simulate, and evaluate various potential corrective methods that could strengthen the Texas grid under extreme weather. Specifically, the simulation results of generator winterization are illustrated in Figure 8. These simulation results captured some key characteristics of the outage event that would otherwise be difficult to obtain directly from data, such as the regional and fuel-type disparity of generator winterization effectiveness and the interdependency of performance across different corrective measures. Such open-source simulation allows scholars to develop and propose optimal investment allotment on Texas grid enhancement and efficiently inform policy makers about their importance and impact.

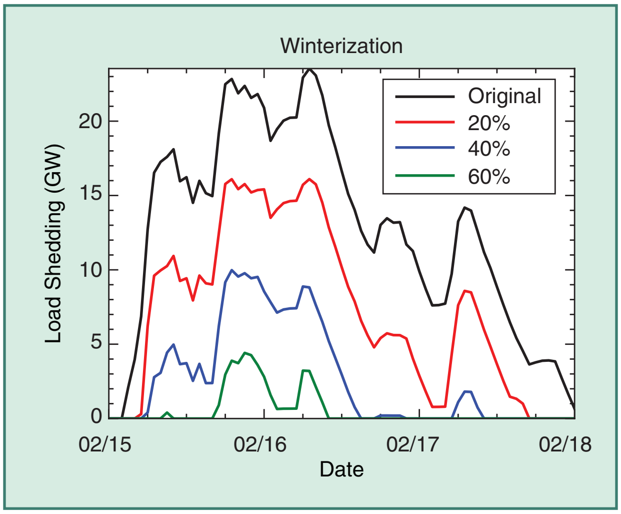

figure 8. The load shedding curves for different degrees of winterization.

Analyzing the Interdependency Between the Natural Gas and Electricity Sectors

We have also created another simulation using the existing work and data from the Texas 2021 power outage event and adapting it to visually recreate the natural gas and electricity generation infrastructure present in the ERCOT grid. This simulation seeks to understand and visualize the interdependencies of the natural gas and electricity generation sectors during Winter Storm Uri using available natural gas data over 14–19 February 2021.

The scope of the simulation is the ERCOT service area inside Texas. The project focused strictly on the natural gas and electricity generation infrastructure, so only that side of the energy mix is plotted as represented in the simulation. Geographic data for all sites was collected from the U.S. Energy Information Administration, which contained shapefile layers for many components of the U.S. energy mapping system. Geographic information for Texas and ERCOT shapefiles came from the ERCOT website.

Data surrounding the 2021 Texas power outages were derived from several different sources. Natural gas electricity generator derate data initially came from a ERCOT report, but the data frame was downloaded from the "2021 Texas Power Outage" project. Data regarding specific natural gas storage units was unobtainable, but general trends of this resource allocation were made available through Wood Mackenzie reports as seen in the University of Texas' paper. Likewise, data trends for natural gas processing plants were obtained the same way and extrapolated to fit the scope of this simulation.

Next, these two types of data were combined to create a geospatial representation of capacity and flowrate values for individual plants. In total, three types of polygon elements were added: Texas and county borders (black), ERCOT service area (blue), and natural gas shale plays (pink). Three node elements were added: natural gas processing facilities (red), natural gas powerplants (light blue), and natural gas storage facilities (green). Finally, one polyline or line-string element was added: natural gas intrastate pipelines (pink) to indicate flows between the nodes.

The radius of each node corresponds to its generating capacity, throughput, or storage capacity depending on each type of unit. These nodes grow larger and smaller as the simulation runs to mimic the derating or processing failures of powerplants or natural gas processing plants over the

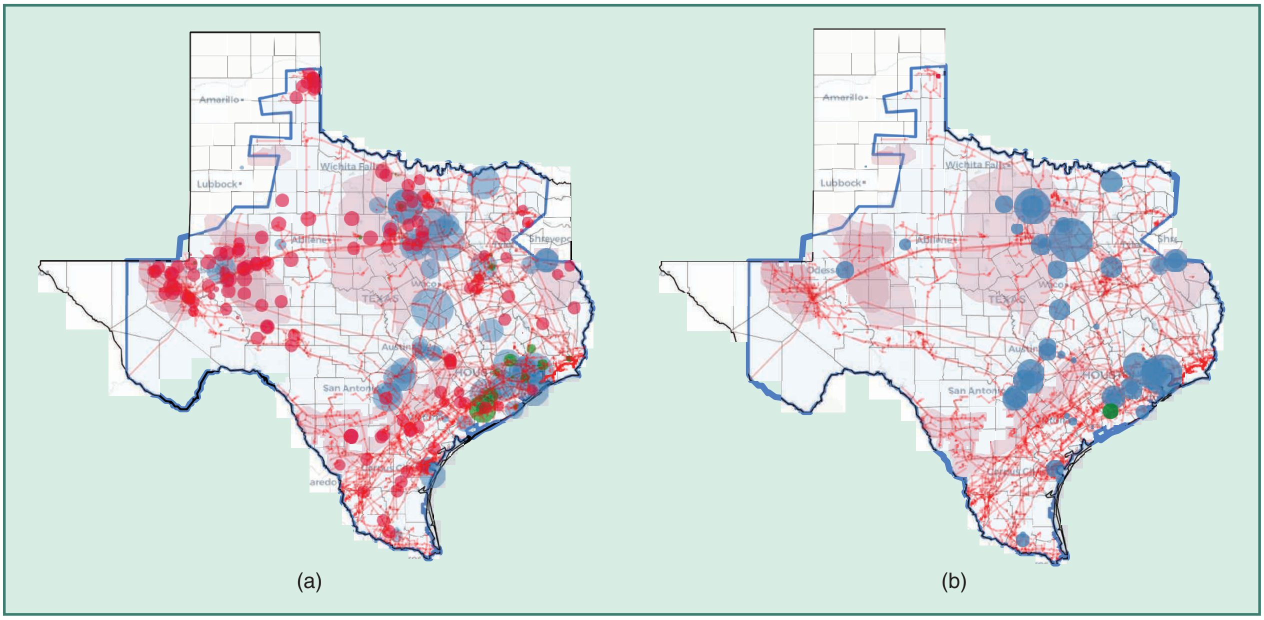

figure 9. A visualization of the synthetic ERCOT natural gas grid with the simulation. (a) Natural gas at 12:00 a.m. on 14 February 2021. (b) Natural gas at 2:00 a.m. on 16 February 2021.

five-day period. The simulation is visually scalable, so it can be viewed zoomed out, looking at holistically at ERCOT, or zoomed in to specifically analyze one node or area of interest. In addition, hovering the mouse over a node or pipeline will provide the user with descriptive data for that element of the simulation at that given time (Figure 9). Finally, the simulation can be paused, run back, and slowed down to allow for ease of analysis.

This model can be utilized for other data sets and can incorporate transmission or pipeline flows if available. All these functions allow the user to retroactively study specific sites after an extreme weather event. For example, a close look can be taken at the generators that maintained optimal functionality during the event to understand which factors contributed to their success (e.g., more interconnections to natural gas processing facilities, winterization packages, backup generation, and so on).

Concluding Remarks

This article presented several studies that aim at understanding the short-term impacts of extreme events, such as the COVID-19 pandemic and unprecedented winter storm, on the electric power sector by cross-domain data-driven approaches. The data-driven approach for COVID 19 helped to understand the change in electricity consumption in 2020 and pointed to population mobility as a key driver. The approaches for the 2021 Texas winter outage provided quantified corrective measure assessments and interdependency between infrastructure systems, which can be used as a reference for policy making. Besides, the open-source data set is expected to provide a common basis for potentially fostering transparent and efficient intra- and interdisciplinary collaboration.

Acknowledgment

This work is supported in part by NSF Grants ECCS-2035688, CMMI-2130945, ECCS-2038963, in part by Breakthrough Energy Sciences, and in part by Texas A&M Energy Institute.

For Further Reading

G. Ruan et al., "A cross-domain approach to analyzing the short-run impact of COVID-19 on the US electricity sector," *Joule*, vol. 4, no. 11, pp. 2322–2337, Nov. 2020, doi: 10.1016/j.joule.2020.08.017.

D. Wu et al., "An open-source extendable model and corrective measure assessment of the 2021 Texas power outage," *Adv. Appl. Energy*, vol. 4, p. 100,056, Nov. 2021, doi: 10.1016/j.adapen.2021.100056.

K. T. Gillingham, C. R. Knittel, J. Li, M. Ovaere, and M. Reguant, "The short-run and long-run effects of COVID-19 on energy and the environment," *Joule*, vol. 4, no. 7, pp. 1337–1341, 2020, doi: 10.1016/j.joule.2020.06.010.

Biographies

Xiangtian Zheng is with Texas A&M University, College Station, TX, 77843 USA.

Dongqi Wu is with Texas A&M University, College Station, TX, 77843 USA.

Liam Watts is with Texas A&M University, College Station, TX, 77843 USA.

Efstratios N. Pistikopoulos is with Texas A&M University, College Station, TX, 77843 USA.

Le Xie is with Texas A&M University, College Station, TX, 77843 USA.