Patterns

CellPress
OPEN ACCESS

Perspective

Energy system digitization in the era of AI: A three-layered approach toward carbon neutrality

Le Xie,[1,*] Tong Huang,[2] Xiangtian Zheng,[1] Yan Liu,[3] Mengdi Wang,[4,5,8] Vijay Vittal,[6] P.R. Kumar,[1] Srinivas Shakkottai,[1] and Yi Cui[7]

[1]Department of Electrical and Computer Engineering, Texas A&M University, College Station, TX 77843, USA
[2]Department of Electrical and Computer Engineering, San Diego State University, San Diego, CA 92182, USA
[3]Computer Science Department, University of Southern California, Los Angeles, CA 90007, USA
[4]Department of Electrical Engineering, Princeton University, Princeton, NJ 08540, USA
[5]DeepMind, Mountain View, CA 94043, USA
[6]School of Electrical, Computer and Energy Engineering, Arizona State University, Tempe, AZ 85281, USA
[7]Department of Materials Science and Engineering, Stanford University, Stanford, CA 94305, USA
[8]Center for Statistics and Machine Learning, Princeton University, Princeton, NJ 08540, USA
*Correspondence: le.xie@tamu.edu
https://doi.org/10.1016/j.patter.2022.100640

THE BIGGER PICTURE As one of the biggest game changers in addressing climate change, the transition to a carbon-neutral electric grid poses significant challenges to conventional paradigms of modern grid planning and operation. Artificial intelligence (AI) has the potential to address the challenges posed by large decision-making scale and increased uncertainty, as many key decision-making mechanisms of grid planning and operation can be formulated as representative AI problems. However, AI needs to be tailored for power system applications in three layers of technology, market, and policy to meet the needs of safety criticality, time sensitivity, and interpretability.

Proof-of-Concept: Data science output has been formulated, implemented, and tested for one domain/problem

SUMMARY

The transition toward carbon-neutral electricity is one of the biggest game changers in addressing climate change since it addresses the dual challenges of removing carbon emissions from the two largest sectors of emitters: electricity and transportation. The transition to a carbon-neutral electric grid poses significant challenges to conventional paradigms of modern grid planning and operation. Much of the challenge arises from the scale of the decision-making and the uncertainty associated with the energy supply and demand. Artificial intelligence (AI) could potentially have a transformative impact on accelerating the speed and scale of carbon-neutral transition, as many decision-making processes in the power grid can be cast as classic, though challenging, machine-learning tasks. We point out that to amplify AI's impact on carbon-neutral transition of the electric energy systems, the AI algorithms originally developed for other applications should be tailored in three layers of technology, markets, and policy.

INTRODUCTION

To grapple with climate change, many countries are striving to achieve carbon neutrality of their electricity sectors. As an example, the US aims for 100% electricity generation from zero-carbon resources by 2035. However, today's decarbonization rate in the US electricity sector may not be able to realize such an aggressive agenda sector.[1]

Speeding up the carbon-neutral transition of the electricity sector requires massive integration of renewable generation at an unprecedented rate. Such large-scale renewable integration poses significant challenges to the operational paradigm that today's grid employs. Reliable operation depends on both offline planning as well as real-time decision-making, including monitoring, control, and protection. In the emerging electricity landscape with deep renewables, due to the temporally variable nature of renewable generation, the number of scenarios needing to be considered in planning studies is extremely large, which creates significant difficulties for grid planners. Moreover, impactful anomalies in a low-carbon grid, e.g., oscillations and large voltage deviations, may appear more often than in the conventional grid. Consequently, there

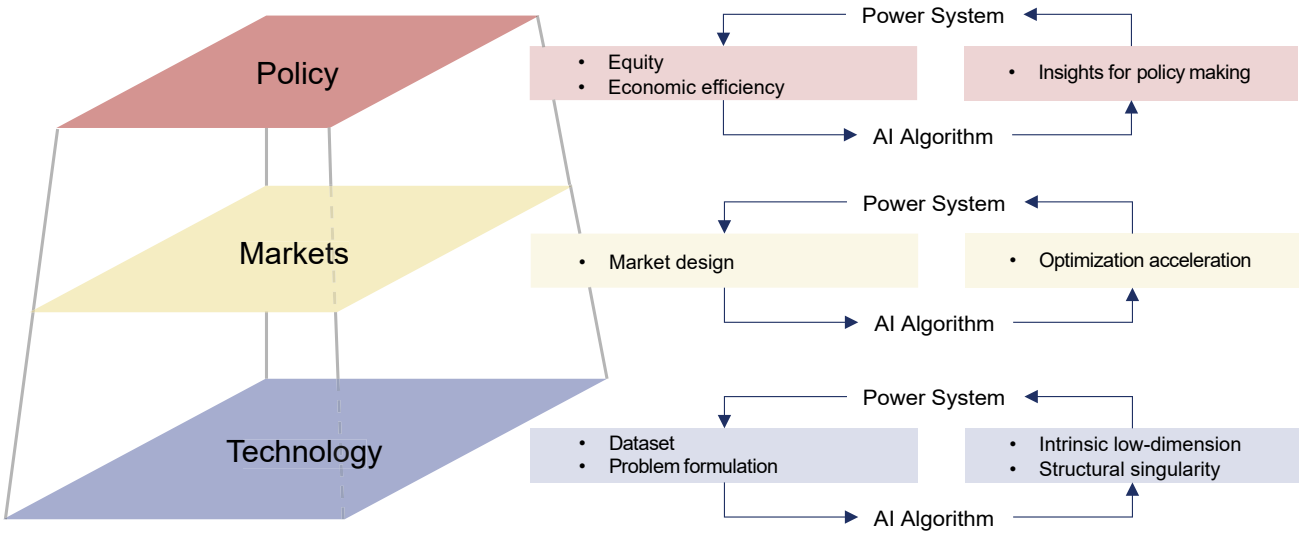

Figure 1. A layered vision of energy system digitization

is a stringent need to monitor and correct these anomalies in a timely manner. Such anomaly monitoring and correction require closed-loop decision-making tools that can convert high-dimensional streaming data into reliable decisions and apply the decisions to the physical infrastructure in a timely fashion. However, there is generally a lack of such tools for most such anomalous scenarios.

The above operational challenges are likely to become a major bottleneck to accelerating the carbon-neutral transition of the electricity sector. These challenges may benefit from artificial intelligence (AI)-based solutions. In recent years, AI-based applications are transforming all aspects of human society and endeavors.[2] In many successful applications, AI algorithms make decisions based on data without requiring detailed models. This is a highly desirable feature for future power grid operation since accurate physical models of phenomena based on weather or consumer behavior are likely either unknown or complex. On the other hand, as will be seen in what follows, some grid operational needs can be translated into classic AI problems. However, this will require the translation of the potential of AI into solutions at full scale that can empower the ambitious carbon-neutral transition in the electricity grid[3,4] under the additional requirements of safety, time sensitivity, and interpretability.

While several survey papers[2,5,6] review the applications of AI in electric energy systems and point out the research directions of this field from a technological viewpoint, how to unleash the power of AI in decarbonizing the electricity sector is a complex problem in a broader social-economic-technological space. This perspective argues that general-purpose AI needs to be carefully tailored in three layers before it can be applied in safety-critical grid infrastructure applications. These three layers include technology, markets, and policy, as summarized in Figure 1.

The rest of this perspective is organized as follows. First, we briefly introduce key decision-making mechanisms associated with the operation of modern power grids. Then, we highlight the connection between grid operation and AI. Finally, we elaborate on the tri-layer tailoring of AI for the carbon-neutral transition of the electric grids.

THE ROLE OF AI IN CARBON-NEUTRAL TRANSITION OF POWER GRID

This section elaborates on AI's potential for addressing challenges in the carbon-neutral transition of the electricity energy systems. We provide a brief overview of power grid operation (see a detailed overview of power system basic functionalities in the review paper by Xie et al.[6]) and highlight the decision-making processes (Figure 2). Then, we present the connection between these decision-marking processes and representative problems in the AI field.

Overview of modern power grid operation

The stringent balance requirement between electricity supply and demand in the modern grid is accomplished by a two-stage strategy, which consists of offline planning and real-time operation (the grid operation described here is for transmission systems; the decision-making processes in current distribution system operation generally are a subset of those in transmission systems). Offline planning includes generation and transmission planning and several forward financial markets. Generation planning leverages reliability studies to decide when and how much additional generation capacity should be added over a planning horizon of several years.[7] The objective of transmission planning is to ensure that the transmission lines deliver electricity from the generators to the loads. In the transmission planning stage, system planners scrutinize the grid behavior at multiple timescales by comprehensively conducting physical model-based simulations for future possible scenarios based on past operational experience. Such a procedure is called system security analysis. Once the power grid is physically established, system operators can leverage financial markets, such as day-ahead, reserve,

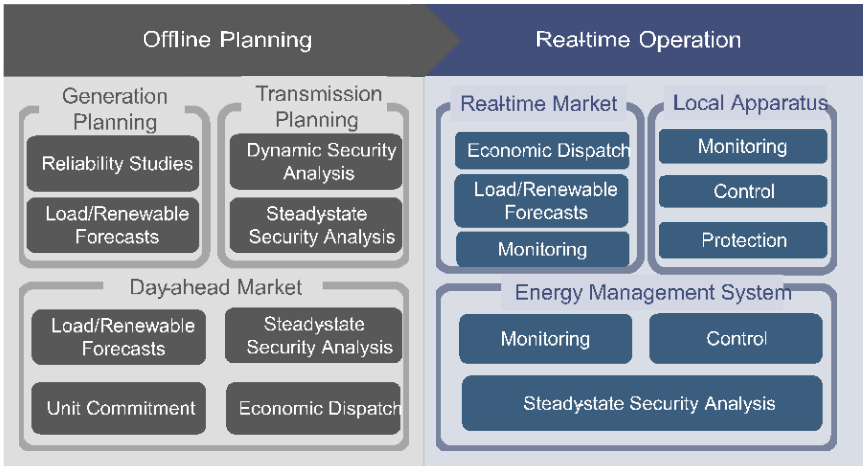

Figure 2. Decision-making modules of power (transmission) system operation

and capacity markets, for enhancing the grid's economical efficiency. Unit commitment (UC) and economic dispatch (ED) are key modules that support energy trading in a day-ahead market.

The decision-making in real-time operation leverages real-time measurements,[8] e.g., those obtained over supervisory control and data acquisition (SCADA), and phasor measurement units (PMUs), in a hierarchical manner. At the global level, the energy management system (EMS) at a control center pre-processes raw measurements from the grid and provides system operators with monitoring and control functionalities. System security analysis is also conducted based on the most recent measurements. At the local level, grid components are regulated and protected by local control and protection apparatus. The economic efficiency of the grid is ensured by a real-time (spot) market. The spot electricity prices are determined by solving ED problems based on the most recent measurements and short-term load forecasts.

Background of grid operation with AI support

This subsection establishes connections between the decision-making modules of grid operation and the three representative formulations of AI problems, i.e., supervised, unsupervised, and reinforcement learning.

Supervised learning for prediction and detection

Given a set of vectors with labels, a supervised-learning method is used for constructing the relationship between the vectors and their labels. Formally, suppose that a procedure takes vector X as input and generates a label vector Y, where the label can be either numerical or categorical. The elements of X are called "features." The supervised-learning method aims to construct a function f that maps X to Y by learning from N historical records of input-output pairs generated from the procedure, i.e., $\{(x_1,y_1),(x_2,y_2), .,(x_N,y_N)\}$. When a new input, say, x_{N+1}, appears, the resulting y_{N+1} is expected to be approximated by $f(x_{N+1})$ with reasonable accuracy. If the number of options for Y is finite, the problem is called one of classification; if continuous variables constitute the output vector Y, the problem is called regression.[9]

Many decision-making modules for grid operation involve classification and regression with clear definitions of the input vector X and the response vector Y. For example, it can be shown that battery cycle life can be predicted by linear regression with properly selected input features.[10] Another example concerns grid online stability assessment that aims to determine if the system is stable given real-time measurements. This is a standard classification problem, with the measurements as inputs and a binary variable (1: stable or 0: unstable) as output. Conventional stability assessment relies heavily on time-consuming simulations that cannot be carried out in real time. It has been reported that a convolution neural network (CNN) can quickly achieve high accuracy.[11] Table 1 shows other AI adoption examples with inputs and outputs clearly defined in the context of power grids. Researchers are striving to build state-of-the-art machine-learning benchmarks for several challenging use cases with open datasets.[12]

Unsupervised learning for modeling high-dimensional data

Unsupervised learning addresses samples without labels. Suppose that only N samples $\{x_1,x_2, .,x_N\}$ are available, without any labels. Unsupervised learning aims to achieve one of the following objectives: (1) clustering aims to find samples that share similarities; (2) density estimation aims to determine the probability distribution governing the given samples (explicitly or implicitly) and to generate new samples with the same probability distribution as the given samples; and (3) dimensionality reduction attempts to project the high-dimensional samples into a low-dimensional space that allows for visualization or easily discovering irregularities in the samples.

The three objectives above are relevant in addressing some challenges associated with the grid. For example, the offline planning problem requires realistic scenarios, including load data conditioned on weather and the season. Generating the scenarios can be formulated as a density estimation problem, which can then generate new scenarios with the same probability density as historic scenarios. The density estimation problem can be solved by a generative adversarial network (GAN).[21] In a distribution system, a utility company divides its customers into several representative clusters in order to understand customers' behaviors.[22] This problem can be solved by various clustering algorithms, such as k-means.[22] In addition, decision-making in the control center often involves high-dimensional data collected from a wide area of the grid. It can be shown that the high-dimensional grid data can be projected into a low-dimensional space by principal-component analysis (PCA), allowing for efficient data storage and early event detection.[23] Table 2 maps some power system challenges to the three goals of the unsupervised learning.

Reinforcement learning

The goal of reinforcement learning (RL) is to make decisions by interacting with an environment under study.[9] Generally, an RL algorithm includes the following elements: states, actions, a

Table 1. Power grid applications with supervised-learning formulation

Challenge	Application	Input: *X*	Output: *Y*
Load/renewable forecasts	short-term forecast[13]	data on energy and/or weather[13]	photovoltaic (PV) forecasts[13]
Reliability studies	fast state evaluation,[14] state enumeration[15]	bus generation info,[14] weather[15]	success/failure status of buses,[14] renewable generation[15]
Security analysis	external grid modeling,[16] contingency screening[17]	interface voltage,[16] bus flow injection[17]	interface power flow,[16] bus voltage and security index[17]
Unit commitment and economic dispatch	binding constraint prediction,[18] constrain convex relaxation[19]	scenarios,[18] 3-phase voltage[19]	binary variable for each constrain,[18] power flow[19]
Monitoring	anomaly classification,[20] stability assessment,[11] battery cycle life prediction[10]	streaming data,[11,20] info. from discharge voltage curves and capacity fade curves[10]	binary indicator,[11] event type,[20] battery cycle life[10]

policy, and a reward function. A state s_k at time k can be measured from the environment, and an action a_k can be applied to the environment at time k, resulting in a new state s_{k+1}. The state transition from s_k to s_{k+1} is determined by the environment, whose governing laws are either unknown or complex. The action and the resulting state transition lead to a real-value reward assigned by a reward function $R(s_k,a_k,s_{k+1})$. A policy p maps a state to an action. RL searches for an optimal policy that maximizes the total reward over a planning horizon. Possible applications of RL in power systems include model calibration,[26] demand modeling,[27] energy trading,[28] reactive power control,[29] tap changer control,[30] and relay timer setting in distribution systems.[31] Table 3 identifies the standard RL ingredients for each of the algorithms.

A THREE-LAYERED APPROACH TO TAILOR DESIGN AI FOR CARBON-NEUTRAL ELECTRIC GRIDS

Given the safety criticality, time sensitivity, and desire for interpretability in the electric power system operation, we postulate a three-layered approach to tailor AI for power system applications, namely technology, markets, and policy. At each layer, the domain-specific constraints require AI development to be suitably designed; innovations in AI also shed light on how each layer can be further evolved by taking full advantage of future developments.

At the technology layer

Power grid operation entails complicated interactions of millions of physical components. The decision-making processes over the electricity infrastructure typically employ the mathematical descriptions of these interactions, which in turn are derived from first principles, such as Newtonian mechanics and electrodynamics. The mathematical descriptions of these processes may possess some properties that may not be straightforward to discern, even for power engineers. For example, high-dimensional data concerning electrical variables, such as measurements of voltage and current over the power grid, may possess a low-rank structure. This is because electrical variables at different locations are correlated by transmission/distribution lines. The hidden properties of the nature of the power grid phenomena can guide the selection/design of AI algorithms to address the operational challenges facing power grids. One can consider the localization of the source of the forced oscillation as an example, which consists of the determination the location of anomalous sources given the measurements over the grid. At first glance, the source localization problem seems to be closely related to a typical supervised classification problem in that we aim to classify the locations into two categories, i.e., locations close to the anomaly source versus locations far away from the source. The performance of the supervised-learning algorithm generally depends heavily on the size and data quality of the training sets. However, it is generally challenging to generate a large, high-quality training set for large-scale power systems. By recognizing the low-rank property of the sensor measurements over the grid, and the sparsity property of anomalous sources, one can choose an unsupervised learning algorithm called robust PCA (RPCA) to pinpoint the source. Such an algorithm does not require a large amount of training data and exhibits promising performance.[25] While it is well accepted that domain knowledge can help identify application scenarios of general-purpose AI algorithms, as well as to select useful features feeding the AI algorithms, the structural properties that are hidden under the complicated mathematical descriptions of the electricity infrastructure should be exploited at all stages of development of the AI algorithms for grid applications in order to obtain robust, interpretable AI-powered tools for the grid applications.

Table 2. Power grid applications with unsupervised-learning formulation

Challenge	Application	Samples	Goal
Load/renewable Forecasts	scenario generation[21]	data of energy and weather[21]	density estimation[21]
Security analysis	load modelling,[22] generator coherency identification[24]	customer energy data,[22] synchrophasor data[24]	clustering[22,24]
Monitoring	anomaly detection,[23] localization[25]	streaming data[23,25]	dimensionality reduction[23,25]

Table 3. Power grid applications with reinforcement-learning formulation

Challenge	Application	State s_k	Action a_k	Policy Goal R
Security analysis	model calibration,[26] demand modeling[27]	model parameters,[26] load fraction[27]	model parameter modification,[26] load fraction modification[27]	response mismatch minimization[26,27]
Unit commitment and economic dispatch	energy trading[28]	past wholesale price-quantity pairs and retail demand-price pairs[28]	retail price[28]	profit maximization of a load serving entity[28]
Control	reactive power control,[29] tap changer control[30]	binary security status,[29] voltage and tap ratio[30]	tap ratio,[29,30] reactive power compensation[29]	flow/voltage tracking[29,30]
Protection	relay timer setting in distribution systems[31]	line current, breaker status, and timer value[31]	timer setting[31]	minimizing misoperation rate[31]

At the markets layer

One of major AI applications in electricity markets is to accelerate large-scale optimization. For instance, UC and ED, which are fundamental problems in electricity markets, face critical challenges due to increasing uncertainty caused by deepening penetration of renewables. Several data-driven, scenario-based optimization approaches have been proposed to efficiently obtain optimal solutions that explicitly provide the probability that the solution is feasible.[32–35] Another machine-learning-based approach has been proposed to identify active sets of safety constraints to obtain optimality more efficiently.[36] Additionally, the advances in AI can potentially help to optimize the allocation of market investment resources. As an example, machine-learning-based demand responses from just a few targeted locations have been found to be effective in mitigating price volatility. Therefore, resources for demand response programs can be strategically concentrated in targeted areas, rather than one size fits all, to achieve the most effective social welfare improvements.[37]

Although there have been many studies showing the potential of AI technology in improving the reliability and economy of electric grids, the lack of appropriate market design is holding it back. For example, an RL-based framework has been proposed to provide voltage regulation via reactive power support in distribution grids with deep solar photovoltaic penetration.[38] However, the lack of auxiliary markets in distribution grids hinders the realization of economic revenue from providing reactive power support, thus weakening the economic impacts of AI techniques. Therefore, changes to the design of the electricity market are imperative to accommodate technological innovation and translate advances in AI into reality.

At the policy layer

Electric grids are heavily regulation and policy driven. With the increasing complexity of electric grids, AI ushers in unique opportunities to integrate interdisciplinary knowledge and leverage heterogeneous data to provide effective insights for policy making. For example, a machine-learning-based approach has been proposed to comprehend the correlation between electricity consumption, number of COVID-19 cases, level of social distancing, and degree of commercial activity during the COVID-19 pandemic using cross-domain datasets. It can be used as an indicator for predicting changes caused by such an unprecedented event.[39] As another example, the design of green energy policies, such as encouraging the adoption of household rooftop solar panels, will require insights from AI techniques. An example is a machine-learning-based solar deployment database in the US[40] to improve energy justice and equity for diverse populations and analyze social impacts on job creation. An AI-informed policy design would be a prudent approach to driving the carbon-neutral transition in the electricity sector. On the other hand, new regulations must be developed to regulate AI applications in such a critical infrastructure system to ensure reliability and privacy. For example, an AI application must possess some key properties, such as interpretability, to facilitate further inspection and investigation by human operators. The increasing demand for data acquisition in power grids also requires new regulations on data privacy and data availability and regulating data acquisition processes to protect the privacy of data owners while maximizing the utility of data.

CONCLUDING REMARKS

Energy system decarbonization is one of the most challenging and exciting areas of research and innovation for the 21st century. This perspective presents a view on how digitization and AI could play a crucial role in the carbon-neutral transition of the energy sector. We argue that higher impact of digitization and AI in the electric energy industry could be achieved through a "three-layered" integrated approach that encompasses technology, markets, and policy layers. Domain-tailored digitization and AI will draw upon unique specifications in all three layers in the energy sector while providing fertile ground for use-inspired innovations in methodology and algorithms. There are several actionable recommendations that would potentially bring the energy and AI communities closer together.

- To involve more power domain-agnostic researchers from the broader AI community, the energy and power community should develop a suite of problem formulations that are accessible to general AI researchers. These problems should be motivated by real-world needs and have the potential to engage the AI community. As an example, the workshop on "Learning to run a power network (L2RPN)" provides a platform for such problem definition and solution.[41]
- A suite of open, cross-domain datasets that are well benchmarked and labeled for representative power system operating conditions should be developed and shared with the broader AI community. A possible example is an

open-source, cross-domain dataset[39] that includes electricity consumption, public health, and mobility data, which was released for broader communities to understand the short-run impact of COVID on the US electricity sector in a data-driven manner.

- On the educational front, both the power/energy community and the AI/digitization community should provide use-inspired cases and tools that will be accessible to undergraduates in both areas. Helping students become "bilingual" in both energy and AI terms would be important. Examples include data science and machine-learning courses for power systems in universities, such as Texas A&M University,[42] University of Texas, Austin,[43] and University of Washington, Seattle.[44]

MATERIALS AVAILABILITY

This study did not generate new unique reagents.

DATA AND CODE AVAILABILITY

This study does not involve data or code.

ACKNOWLEDGMENT

This material is based upon work partially supported by the US Department of Energy's Office of Energy Efficiency and Renewable Energy (EERE) under the Solar Energy Technologies Office award number DE-EE0009031 and by the US National Science Foundation under the grants CMMI-2130945 and ECCS-2038963. The views expressed herein and the conclusions contained in this document are those of the authors and should not be interpreted as representing the views or official policies, either expressed or implied, of the Department of Energy or the United States government.

AUTHOR CONTRIBUTIONS

Conceptualization, L.X.; methodology, L.X. and T.H.; investigation, L.X. and T.H.; writing – original draft, L.X., T.H., and X.Z.; writing – review & editing, L.X., T.H., X.Z., Y.L., M.W., V.V., P.R.K., S.S., and Y.C.; funding acquisition, L.X.; supervision, L.X.

DECLARATION OF INTERESTS

The authors declare no competing interests.

REFERENCES

1. Monthly Energy Review, Tech. Rep. (2021) (Energy Information Administration).
2. Rolnick, D., Donti, P.L., Kaack, L.H., Kochanski, K., Lacoste, A., Sankaran, K., Ross, A.S., Milojevic-Dupont, N., Jaques, N., Waldman-Brown, A., et al. (2019). Tackling climate change with machine learning. Preprint at arXiv. https://doi.org/10.48550/arXiv.1906.05433.
3. Duchesne, L., Karangelos, E., and Wehenkel, L. (2020). Recent developments in machine learning for energy systems reliability management. Proc. IEEE *108*, 1656–1676. https://doi.org/10.1109/JPROC.2020.2988715.
4. Chatterjee, J., and Dethlefs, N. (2022). Facilitating a smoother transition to renewable energy with ai. Patterns *3*, 100528. https://doi.org/10.1016/j.patter.2022.100528.
5. Dalal, G., Gilboa, E., Mannor, S., and Wehenkel, L. (2019). Chance-constrained outage scheduling using a machine learning proxy. IEEE Trans. Power Syst. *34*, 2528–2540. https://doi.org/10.1109/TPWRS.2018.2889237.
6. Xie, L., Zheng, X., Sun, Y., Huang, T., and Bruton, T. (2022). Massively digitized power grid: opportunities and challenges of use-inspired AI. Proc. IEEE, 1–26. https://doi.org/10.1109/JPROC.2022.3175070.
7. Bebic, J. (2008). Power System Planning: Emerging Practices Suitable for Evaluating the Impact of High-Penetration Photovoltaics (National Renewable Energy Lab.(NREL)). Tech. rep.. https://doi.org/10.2172/924647
8. Sztipanovits, J., Hemingway, G., Bose, A., and Srivastava, A. (2012). Model-based integration technology for next generation electric grid simulations. In 2012 IEEE Power and Energy Society General Meeting, IEEE, pp. 1–8.
9. Bishop, C.M., and Nasrabadi, N.M. (2006). Pattern Recognition and Machine Learning, *4* (Springer).
10. Severson, K.A., Attia, P.M., Jin, N., Perkins, N., Jiang, B., Yang, Z., Chen, M.H., Aykol, M., Herring, P.K., Fraggedakis, D., et al. (2019). Data-driven prediction of battery cycle life before capacity degradation. Nat. Energy *4*, 383–391. https://doi.org/10.1038/s41560-019-0356-8.
11. Zhu, L., Hill, D.J., and Lu, C. (2020). Hierarchical deep learning machine for power system online transient stability prediction. IEEE Trans. Power Syst. *35*, 2399–2411. https://doi.org/10.1109/TPWRS.2019.2957377.
12. Zheng, X., Xu, N., Trinh, L., Wu, D., Huang, T., Sivaranjani, S., Liu, Y., and Xie, L. (2022). A multi-scale time-series dataset with benchmark for machine learning in decarbonized energy grids. Sci. Data *9*, 359. https://doi.org/10.1038/s41597-022-01455-7.
13. Yang, C., Thatte, A.A., and Xie, L. (2015). Multitime-scale data-driven spatio-temporal forecast of photovoltaic generation. IEEE Trans. Sustain. Energy *6*, 104–112. https://doi.org/10.1109/TSTE.2014.2359974.
14. Urgun, D., and Singh, C. (2019). A hybrid Monte Carlo simulation and multi label classification method for composite system reliability evaluation. IEEE Trans. Power Syst. *34*, 908–917. https://doi.org/10.1109/TPWRS.2018.2878535.
15. Lin, T., and Shang, C. (2021). Reliability evaluation on a joint machine learning and optimization framework. IEEE Trans. Power Syst. *36*, 49–57. https://doi.org/10.1109/TPWRS.2020.3007618.
16. Ma, F., and Vittal, V. (2012). A hybrid dynamic equivalent using ann-based boundary matching technique. IEEE Trans. Power Syst. *27*, 1494–1502. https://doi.org/10.1109/TPWRS.2012.2182783.
17. Du, Y., Li, F., Li, J., and Zheng, T. (2019). Achieving 100x acceleration for n-1 contingency screening with uncertain scenarios using deep convolutional neural network. IEEE Trans. Power Syst. *34*, 3303–3305. https://doi.org/10.1109/TPWRS.2019.2914860.
18. Mohammadi, F., Sahraei-Ardakani, M., Trakas, D., and Hatziargyriou, N. (2021). Machine learning assisted stochastic unit commitment during hurricanes with predictable line outages. IEEE Trans. Power Syst. *36*, 5131–5142. https://doi.org/10.1109/TPWRS.2021.3069443.
19. Hu, R., Li, Q., and Qiu, F. (2021). Ensemble learning based convex approximation of three-phase power flow. IEEE Trans. Power Syst. *36*, 4042–4051. https://doi.org/10.1109/TPWRS.2021.3055481.
20. Li, W., Wang, M., and Chow, J.H. (2018). Real-time event identification through low-dimensional subspace characterization of high-dimensional synchrophasor data. IEEE Trans. Power Syst. *33*, 4937–4947. https://doi.org/10.1109/TPWRS.2018.2796442.
21. Chen, Y., Wang, Y., Kirschen, D., and Zhang, B. (2018). Model-free renewable scenario generation using generative adversarial networks. IEEE Trans. Power Syst. *33*, 3265–3275. https://doi.org/10.1109/TPWRS.2018.2794541.
22. Tsekouras, G.J., Hatziargyriou, N.D., and Dialynas, E.N. (2007). Two-stage pattern recognition of load curves for classification of electricity customers. IEEE Trans. Power Syst. *22*, 1120–1128. https://doi.org/10.1109/TPWRS.2007.901287.
23. Xie, L., Chen, Y., and Kumar, P.R. (2014). Dimensionality reduction of synchrophasor data for early event detection: linearized analysis. IEEE Trans. Power Syst. *29*, 2784–2794. https://doi.org/10.1109/TPWRS.2014.2316476.

24. Raak, F., Susuki, Y., and Hikihara, T. (2016). Data-driven partitioning of power networks via koopman mode analysis. IEEE Trans. Power Syst. *31*, 2799–2808. https://doi.org/10.1109/TPWRS.2015.2464779.

25. Huang, T., Freris, N.M., Kumar, P.R., and Xie, L. (2020). A synchrophasor data-driven method for forced oscillation localization under resonance conditions. IEEE Trans. Power Syst. *35*, 3927–3939. https://doi.org/10.1109/TPWRS.2020.2982267.

26. Wang, S., Diao, R., Xu, C., Shi, D., and Wang, Z. (2021). On multi-event co-calibration of dynamic model parameters using soft actor-critic. IEEE Trans. Power Syst. *36*, 521–524. https://doi.org/10.1109/TPWRS.2020.3030164.

27. Wang, X., Wang, Y., Shi, D., Wang, J., and Wang, Z. (2020). Two-stage wecc composite load modeling: a double deep q-learning networks approach. IEEE Trans. Smart Grid *11*, 4331–4344. https://doi.org/10.1109/TSG.2020.2988171.

28. Xu, H., Sun, H., Nikovski, D., Kitamura, S., Mori, K., and Hashimoto, H. (2019). Deep reinforcement learning for joint bidding and pricing of load serving entity. IEEE Trans. Smart Grid *10*, 6366–6375. https://doi.org/10.1109/TSG.2019.2903756.

29. Vlachogiannis, J., and Hatziargyriou, N. (2004). Reinforcement learning for reactive power control. IEEE Trans. Power Syst. *19*, 1317–1325. https://doi.org/10.1109/TPWRS.2004.831259.

30. Xu, H., Domínguez-García, A.D., and Sauer, P.W. (2019). Optimal tap setting of voltage regulation transformers using batch reinforcement learning. IEEE Trans. Power Syst. *35*, 1990–2001. https://doi.org/10.1109/TPWRS.2019.2948132.

31. Wu, D., Zheng, X., Kalathil, D., and Xie, L. (2019). Nested reinforcement learning based control for protective relays in power distribution systems. In 2019 IEEE 58th Conference on Decision and Control (CDC) (IEEE), pp. 1925–1930. https://doi.org/10.1109/CDC40024.2019.9029268.

32. Geng, X., and Xie, L. (2019). Data-driven decision making in power systems with probabilistic guarantees: theory and applications of chance-constrained optimization. Annu. Rev. Control *47*, 341–363. https://doi.org/10.1016/j.arcontrol.2019.05.005.

33. Modarresi, M.S., Xie, L., Campi, M.C., Garatti, S., Care, A., Thatte, A.A., and Kumar, P.R. (2019). Scenario-based economic dispatch with tunable risk levels in high-renewable power systems. IEEE Trans. Power Syst. *34*, 5103–5114. https://doi.org/10.1109/TPWRS.2018.2874464.

34. Ming, H., Xie, L., Campi, M.C., Garatti, S., and Kumar, P.R. (2019). Scenario-based economic dispatch with uncertain demand response. IEEE Trans. Smart Grid *10*, 1858–1868. https://doi.org/10.1109/TSG.2017.2778688.

35. Xavier, Á.S., Qiu, F., and Ahmed, S. (2020). Learning to solve large-scale security-constrained unit commitment problems. Inf. J. Comput. *33*, 739–756. https://doi.org/10.1287/ijoc.2020.0976.

36. Misra, S., Roald, L., and Ng, Y. (2022). Learning for constrained optimization: identifying optimal active constraint sets. Inf. J. Comput. *34*, 463–480. https://doi.org/10.1287/ijoc.2020.1037.

37. Lee, K., Geng, X., Sivaranjani, S., Xia, B., Ming, H., Shakkottai, S., and Xie, L. (2022). Targeted demand response for mitigating price volatility and enhancing grid reliability in synthetic Texas electricity markets. iScience *25*, 103723. https://doi.org/10.1016/j.isci.2021.103723.

38. El Helou, R., Kalathil, D., and Xie, L. (2021). Fully decentralized reinforcement learning-based control of photovoltaics in distribution grids for joint provision of real and reactive power. IEEE Open J. Power Energy *8*, 175–185. https://doi.org/10.1109/OAJPE.2021.3077218.

39. Ruan, G., Wu, D., Zheng, X., Zhong, H., Kang, C., Dahleh, M.A., Sivaranjani, S., and Xie, L. (2020). A cross-domain approach to analyzing the short-run impact of covid-19 on the us electricity sector. Joule *4*, 2322–2337. https://doi.org/10.1016/j.joule.2020.08.017.

40. Yu, J., Wang, Z., Majumdar, A., and Rajagopal, R. (2018). Deepsolar: a machine learning framework to efficiently construct a solar deployment database in the United States. Joule *2*, 2605–2617. https://doi.org/10.1016/j.joule.2018.11.021.

41. ChaLearn (2020). RTE, L2RPN NeurIPS 2020. [Online]. https://competitions.codalab.org/competitions.

42. Xie, L. (2022). Data science and machine learning for modern power systems. [Online]. https://smartgridcenter.tamu.edu/index.php/data-science-and-machine-learning-for-modern-power-systems/.

43. Zhu, H. (2021). Data analytic course. [Online]. https://utexas.app.box.com/v/EE394VDataInPowerSys.

44. Zhang, B. (2022). Data science for power systems. [Online]. http://zhangbaosen.github.io/teaching/EE555.