INVITED PAPER

Massively Digitized Power Grid: Opportunities and Challenges of Use-Inspired AI

By LE XIE, *Fellow IEEE*, XIANGTIAN ZHENG, *Graduate Student Member IEEE*, YANNAN SUN, *Senior Member IEEE*, TONG HUANG, *Member IEEE*, AND TONY BRUTON, *Member IEEE*

ABSTRACT | This article presents a use-inspired perspective of the opportunities and challenges in a massively digitized power grid. It argues that the intricate interplay of data availability, computing capability, and artificial intelligence (AI) algorithm development are the three key factors driving the adoption of digitized solutions in the power grid. The impact of these three factors on critical functions of power system operation and planning practices is reviewed and illustrated with industrial practice case studies. Open challenges and research opportunities for data, computing, and AI algorithms are articulated within the context of the power industry's tremendous decarbonization efforts.

KEYWORDS | Artificial intelligence (AI); data-driven algorithms; decarbonization; industry use cases; machine learning; power grid.

NOMENCLATURE

AC	Alternating current.
AI	Artificial intelligence.
DC	Direct current.
DER	Distributed energy resource.
EMS	Energy management system.
EV	Electric vehicle.
PV	Photovoltaic.
SCADA	Supervisory control and data acquisition.
AGC	Automatic generation control.
SE	State estimation.
RTU	Remote terminal unit.
DMS	Distribution management system.
SSA	Static security analysis.
DSA	Dynamic security analysis.
FTR	Financial transmission right.
UC	Unit commitment.
ED	Economic dispatch.
OPF	Optimal power flow.
LMP	Local marginal price.
IBR	Inverter-based resource.
IoT	Internet of Things.
PMU	Phasor measurement unit.
DFR	Digital fault recorder.
SOE	Sequence of events.
AMI	Advanced metering infrastructure.
FDR	Frequency disturbance recorder.
CEII	Critical energy/electric infrastructure information.
HIL	Hardware-in-loop.
GPU	Graphics processing unit.
ASIC	Application-specific integrated circuit.
BLAS	Basic linear algebra subroutine.
MIO	Mixed integer optimization.
ARX	Autoregressive with exogenous input.
PCC	Point of common coupling.
NN	Neural network.
SMT	Satisfiability modulo theory.

Manuscript received December 15, 2021; revised March 13, 2022; accepted May 9, 2022. The work of Le Xie, Xiangtian Zheng, and Tong Huang was supported in part by the U.S. Department of Energy's Office of Energy Efficiency and Renewable Energy (EERE) through the Solar Energy Technologies Office (SETO) under Grant DE-EE0009031; and in part by the National Science Foundation under Grant OAC-1934675, Grant ECCS-2035688, and Grant ECCS-1611301. *(Corresponding author: Le Xie.)*

Le Xie and **Xiangtian Zheng** are with the Department of Electrical and Computer Engineering, Texas A&M University, College Station, TX 77843 USA (e-mail: le.xie@tamu.edu; zxt0515@tamu.edu).
Yannan Sun and **Tony Bruton** are with Oncor Electric Delivery, Dallas, TX 75202 USA (e-mail: yannan.sun@oncor.com; tony.bruton@oncor.com).
Tong Huang is with the Laboratory for Information and Decision Systems, Massachusetts Institute of Technology, Cambridge, MA 02139 USA (e-mail: tongh@mit.edu).

Digital Object Identifier 10.1109/JPROC.2022.3175070

RPCA	Robust principal component analysis.
RL	Reinforcement learning.
MDP	Markov decision process.
HVAC	Heating, ventilation, and air conditioning.
RIC	Residential/industrial/commercial.
ELM	Extreme learning machine.
LSTM	Long short-term memory.
CNN	Convolutional neural network.
KNN	K-nearest neighbors.
PCA	Principal component analysis.
SVR	Support vector regression.
RF	Random forest.
SVM	Support vector machine.

I. INTRODUCTION

Digitization of the electric power grid, which broadly refers to the deployment of sensing, communication, and computational capabilities, has been an integral part of the electrification process over the past century and is a key enabling factor that drives power grid transformation by spreading its outreach vertically over plants, transmission grids, distribution grids, and end-use customers. As data availability and computing capacity continue to grow, large-scale power grids are built and operated with very high levels of reliability and efficiency, providing electricity services to billions of customers. The state of today's power grids in the United States (U.S.) can be summarized in three aspects: 1) for system reliability, the average duration of annual electric power interruptions in the U.S. is varied from 3 to 8 h in the period between 2013 and 2020 [1]; 2) for the cost of electricity, the average wholesale electricity price across the U.S. is varied from \$30 to \$60 per MWh in the period between 2016 and 2021 [2]; and 3) for carbon footprint, electricity generation in the U.S. produced an average of about 0.4 kilograms of carbon dioxide emissions per kWh in 2020 [3].

In response to climate change, which has emerged as a global concern, rapid decarbonization is imperative to reduce carbon emissions, a quarter of which are contributed by the electricity sector. It is foreseeable that numerous decarbonization measures will cause profound changes in the electricity sector in the next few decades [4]. Such changes have two major drivers: 1) the energy portfolio transition from high-carbon to low-/zero-carbon generation sources, such as hydrogen, nuclear, wind, and solar-based commercial generation units and DERs and 2) electrification in other sectors, including construction, transportation, and other infrastructure systems. Deepening penetration of intermittent resources, such as wind farms and solar PV, is introducing more variability and uncertainty. The proliferation of power electronics-based inverters is changing system dynamic characteristics. Increasing numbers of DERs at the grid edge are strengthening the interaction between transmission and distribution systems. The rapid expansion of EVs will lead to substantial changes in electricity demand patterns. Therefore, it is imperative for the grid operators to adopt a more flexible and risk-aware approach. Given the massive data availability and computing capacity provided by digitized power grids, data-driven AI methods are feasible solutions for complementing traditional model-based approaches to address these complex emerging challenges.

From a broader economic perspective, AI has transformed a variety of domains over the past decade [5], including language processing [6], speech recognition [7], facial recognition [8], real-time object detection [9], multiplayer game [10]–[12], recommendation system [13], intelligent robotics [14]–[16], driving assistant system [17], disease diagnosis [18], drug discovery [19], finance [20], and others. We attribute such unprecedented success of AI as an intricate interplay between three factors, namely, *massive data acquisition*, *high computing performance*, and *advanced AI algorithms* [21]–[23]. The availability of data from heterogeneous resources has been increasing at an unprecedented rate [24]–[26] and provides fuel for developing AI-based, data-driven applications for valuable knowledge extraction in wide-range domains. In addition, remarkable improvements in computing performance have enabled a variety of practical large-scale AI models, credited to the collective advances in hardware, software, and computing architecture [27]. Alongside rapidly growing AI infrastructure that provides massive data and computing capacity, numerous advanced AI algorithms have been developed in the past decade. State-of-the-art performance on benchmark datasets for tasks in multiple research fields has been improved by pretrained models [28]–[31] and novel AI model architectures [32]–[35].

Given the widespread success of AI applications, the development and deployment of interpretable, robust, and scalable AI may help to accommodate the emerging changes brought by decarbonization, aiming to reduce carbon emission and meanwhile "keep the lights on" in a reliable and economic way (see Fig. 1). However, to facilitate the process toward decarbonization, many open questions persist in implementing practical AI approaches in digitized power grids, including domain-agnostic computing and AI advances, use-inspired AI algorithm development, and cyber–physical security and privacy in a massively digitized power grid. To this end, this article aims to provide a comprehensive review of the state-of-the-art practice of power grid digitization transformation, which focuses on three backbone factors: *data*, *computing*, and *algorithms*. Specifically, this article provides a review of the recent progress in data acquisition, computing capability, and AI algorithms that are applicable to power systems. Successful industry use cases are introduced to illustrate applications of AI algorithms on large real-world datasets.

The rest of this article is organized as follows. Section II provides an overview of power grid operation and planning practices, as well as the challenges posed by decarbonization. Sections III–V provide a comprehensive review of data, computing, and algorithmic advances in power systems. Section VI provides an industry perspective on

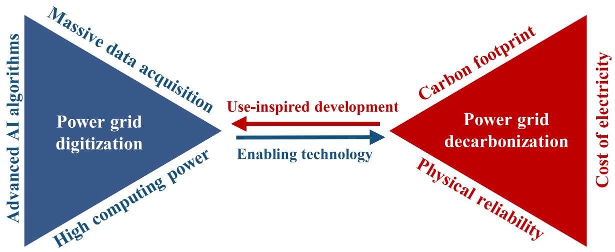


Fig. 1. ***Trifactors of digitization are enabling technologies that facilitate the process toward power grid decarbonization while simultaneously meeting requirements in the aspects of reliability, cost of electricity, and carbon emission, while power grid decarbonization steers use-inspired development of power grid digitization.***

AI adoption. Finally, Section VII concludes this article with remarks on future directions for power grid modernization.

II. PHYSICAL AND MARKET OPERATIONS OF POWER SYSTEMS

Modern power grids are being driven by the strong momentum of decarbonization [36] with decentralization and transportation electrification. Fig. 2 shows the brief conceptual diagram of a modern power grid, which can be separated into transmission and distribution systems. Transmission systems refer to bulk systems that have voltages higher than 66 kV and consist of generation, substation, and transmission lines, which are usually operated by state-wide or cross-state system operators. Distribution systems refer to close-to-users systems that have voltages lower than 33 kV and connect to residential, commercial, and industrial loads, which are usually operated by local utility companies. Power grid decarbonization is changing the energy portfolio in terms of generation resources, such as increasing commercial-size solar PV and wind farms in transmission systems, and DERs, such as rooftop solar PV in distribution systems. Power electronics-based inverters are, thus, being deployed to convert electricity by renewables from dc to ac. Transportation electrification introduces a rapidly expanding number of EVs into distribution systems.

The modern power system operations in high-voltage transmission systems can be broken down into two categories [37]. The first category is physical operations that are responsible for the grid's physical security[1] and resource adequacy[2]; the second concerns market operation. Both physical and market operations are summarized in Fig. 3.

[1]Physical security in power systems refers to the ability to resist contingency disturbances, such as a transmission line short circuit and loss of system components.

[2]Resource adequacy in power systems refers to the ability to supply electricity that accommodates load variation, renewable uncertainty, and system component outages.

A. Functions of Physical Operation and Planning

Power system operation and planning fulfill the reliability of power systems via multiple functions including real-time monitoring, control, protection, and system reliability analysis. A system-wide monitoring system collects and processes measurements, and presents intuitive information to system operators via visualization and alarming. A control system performs control actions either manually or by automated procedures. A protection system executes prescribed corrective measures upon detection of anomalies within targeted system components, which is achieved mainly by local sensors and actuators. Reliability analysis provides instructions on decision-making of multiple time horizons to guarantee the system within adequacy and security criteria.

Load and renewable forecasting provides input for both system and market operation by estimating uncertain net load and renewable generation of various projection horizons. Load forecasting covers various prediction horizons spanning hours, days, weeks, months, and years ahead, whereas renewable forecasting provides only hours and days-ahead predictions. In real-world power grids, short-term load forecasting typically has high accuracy, and renewable forecasting also has acceptable errors that can be mitigated by the real-time operation of dispatchable resources.

Real-time monitoring and control are implemented mostly by EMS in the control center, the primary functional modules of which mainly include SCADA, SE, and AGC. The SCADA system fulfills measurement acquisition and control telemetry through communication channels between the control center and RTUs, at the respective electrical station or device. Typically, the data acquisition function collects measurements every 2–10 s, of which the data stream is a key enabling factor for realizing other functionalities, such as SE, real-time control, UC, and ED. For accurate situational awareness of the system's current operation, function SE provides the steady-SE of system variables that are not directly observed in streaming SCADA data. As one of the major real-time control, primary and secondary generation controls are implemented to: 1) regulate load frequency and 2) balance power generation, load demand, and cross-area interchange in real time. Droop-based generator governors that are responsible for primary control perform instantaneous power quality corrections before triggering protection relays. AGC, considered a secondary control, mitigates unavoidable errors of primary control by sending commands from the control center to participating generation units every 2–4 s [38]. Real-time protection is mainly implemented by protective relays that are equipped with critical assets, such as generation units and substations. In high-voltage transmission systems, protective relays should clear faults within several cycles[3] to avoid further system deterioration. Similarly, a DMS enables real-time monitoring in the distribution

[3]One cycle of a 60-Hz electric power system is about 16 ms.

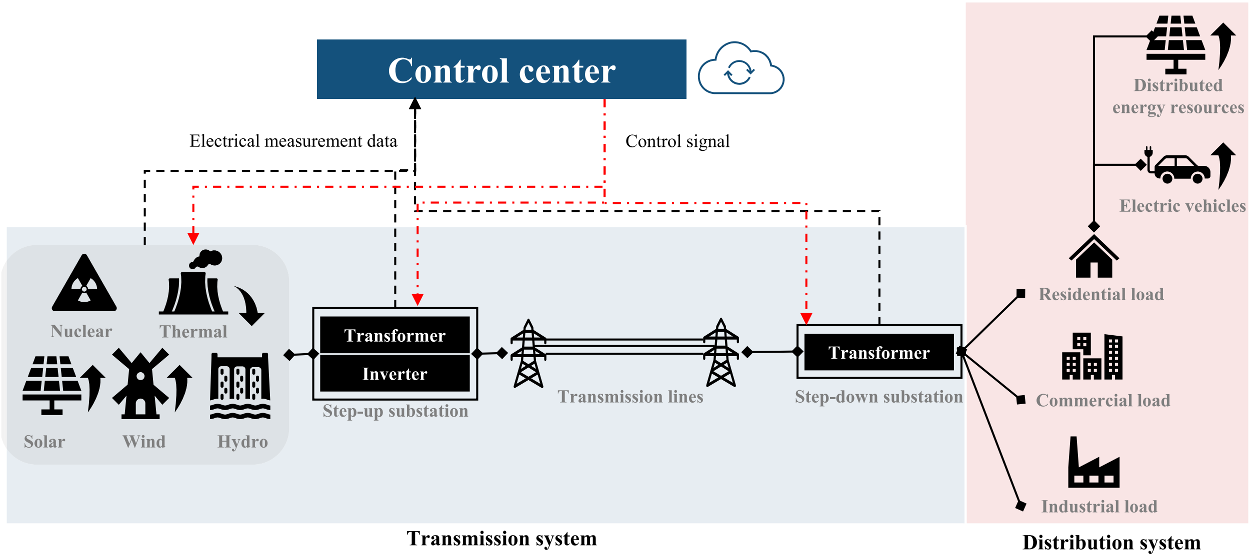


Fig. 2. *Conceptual diagram of a modern power grid, consisting of transmission and distribution systems. The high-voltage transmission system consists mainly of generation, substation, and transmission lines. The low-voltage distribution system supplies electricity to residential, commercial, and industrial loads. Decarbonization has promoted utility-scale renewable generation, DERs, and EVs while reducing investment in thermal generation. Digitization has contributed to the reform and upgrade of control centers through the development of cloud data storage and computing, and the deployment of massive digitized sensors across the grid.*

system, with a few similar functions to EMS, such as SCADA and event analysis [39]. It is worth noting that most field devices in the distribution systems are manually operated rather than remotely controlled, indicating a lower level of automation compared to the transmission system.

System reliability analysis entails adequacy, static security, and DSA [23], [40], [41]. Security analysis focuses on the process of system state transitions initiated by reasonable disturbances, such as short circuits and loss of system components. SSA evaluates the viability of postevent equilibrium by calculating power flow or OPF to check whether a power or voltage violation happens after an $N-1$ contingency.[4] DSA evaluates the ability of the system to transition from one equilibrium to another postevent equilibrium within security criteria [23] by simulating on system dynamic models. Adequacy analysis quantifies the system's capacity for sustainable supply that accommodates load variation, renewable uncertainty, and system component outages by several manually defined indices. A typical method for adequacy and security analysis is numerical simulation. Due to time intensity, these reliability analysis methods tend to be impractical for real-time security control during contingencies. SSA and DSA are used in short-term scheduling, such as generation scheduling, which is performed daily or every few hours. Adequacy analysis and SSA are typically used for mid-term planning, such as facility maintenance, which is performed every several months to one year. Also, both adequacy and security analysis are used for long-term planning, which occurs annually or every few years.

[4]The $N-1$ contingency refers to the loss of a single system component, such as generation outage and transmission line tripping.

B. Functions of Market Operation

Market operation in wholesale electricity markets aims to maximize social welfare while obeying physical constraints. Wholesale markets comprise day-ahead and real-time energy markets, capacity markets, FTR markets, and ancillary service markets. Both day-ahead and real-time energy markets determine clearing prices based on bids from market participants, incorporating physical constraints and potential restrictions. Capacity markets ensure long-term system reliability. FTR markets entitle market participants to offset potential losses (hedge) related to the price risk of delivering energy to the grid. Ancillary service markets provide regulation and reserve. UC and ED are two major security-constrained, bid-based mechanisms to handle the scheduling of generation and the management of system congestion. Both UC and ED are typically formulated as large-scale nonlinear/linear programming problems, known as OPF. Providing forecast load and renewable as input, the UC function determines when and which generation units startup and shut down in day-ahead markets. The ED function calculates the power output of each committed generation unit and associated LMPs. ED is performed to meet the day-ahead hourly forecast load in day-ahead energy markets and to meet the minute-ahead forecast load every 5–10 min in real-time energy markets [42].

In today's distribution grids, the retail market contains few centralized operations or scheduling functions, such as

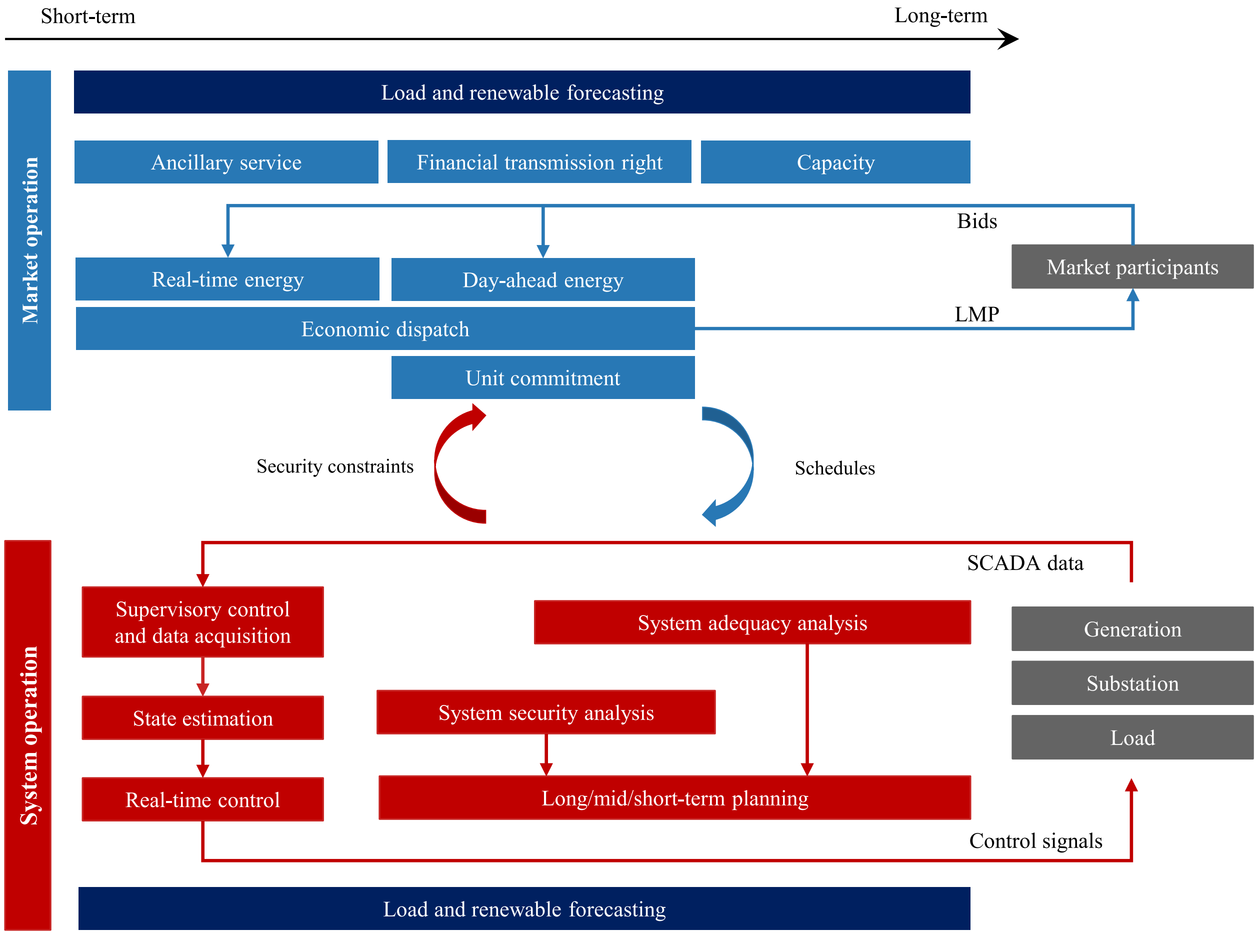


Fig. 3. ***Structure of physical and market operation in transmission systems.***

UC and ED, in the retail market. Given the proliferation of DERs in distribution grids, such as distributed generation, interruptible load, and electricity storage, the retail market will involve system upgrades and reforms in the future to accommodate DER market participation and establish an appropriate mechanism of scheduling and compensation [43].

C. Challenges of Decarbonizing Power Grid

Renewable integration and transportation electrification at scale impose challenges on the paradigm of protection and control. The emergence of massive grid-following and grid-forming inverted-based resources (IBRs) may challenge the effectiveness and efficiency of the current central control frame due to the unknown impacts of electromagnetic dynamics and low inertia. DERs at the grid edge may create bidirectional power flows that potentially incur malfunctions of the protective relays in distribution grids. Besides, typical methods for adequacy and security analysis are numerical simulations that highly rely on grid models of multiple time scales, including electromagnetic dynamic (very fast), electromechanical dynamic (fast), and steady state (slow). However, system characteristics are being changed due to the proliferation of inverter-interfaced renewable resources and EVs in modern power grids, such as low inertia and deeper integration of transmission and distribution systems. These emerging system characteristics create a need for new requirements on the existing models to determine whether the system is within critical security criteria. For example, there is an urgent need for the study of several topics in order to handle the growing system complexity, including: 1) electromagnetic transient models to reveal the fast dynamics by power electronic-based system components; 2) system-level joint simulation between transmission and distribution models to reveal the increasing cross-system interaction; and 3) cross-domain electricity-transportation models to incorporate the impacts of transportation networks on EVs.

The market operation also faces the challenge of managing potential market risks resulting from the variability and stochasticity of renewable generation [44]. Strong uncertainty is a key obstacle to ED to: 1) maintain system stability as tertiary frequency control and 2) avoid unexpected renewable curtailment to the greatest possible extent to achieve decarbonization. Current wholesale markets may not be sufficiently prepared to accommodate increasingly frequent extreme weather events, such as the 2021 Texas power outage event [45], to prevent spiking prices and mitigate energy scarcity. Specifically, strong uncertainty regarding system net load and intermittent renewables generation in future grids will raise severe challenges for the accuracy and robustness of short-term load and renewable prediction. Deepening transportation electrification may also undermine the existing end-use and econometric models for medium and long-term load forecasting [46].

The distribution system also faces a growing number of facility challenges. Aging power lines may limit the maximum use of renewable energy sources, such as wind farms and utility-scale solar, especially in less populated areas where large renewable energy installations are located. The utilization and availability of DERs installed in densely populated areas can be affected by frequent localized outages intermittently that may be recognized by the control center. Given stronger integration and correlation between transmission and distribution grids, facility outages, such as transformer failures, may cause wider impacts. Furthermore, in aiming to establish a competitive retail market in the distribution system, there multiple critical problems remain unsolved, such as LMP calculation and demand response modeling; however, these are beyond the scope of this article.

Overall, the profound changes by decarbonization are posing and will continue to pose numerous challenges to all aspects of physical reliability and economics. Given massive data acquisition as the "fuel" and high computing power as the "engine," applying advanced data-driven AI-based approaches as an "autopilot" has the potential to steer the vehicle forward in a flexible and risk-aware manner.

III. DATA ACQUISITION IN DIGITIZED POWER GRIDS

In broad industry sectors, large-volume and heterogeneously structured data have been generated at an unprecedented rate by diverse resources since 2010 [24]–[26], such as IoT records, social media, smart devices, and healthcare systems. The availability of such tremendous volumes of data has facilitated numerous applications of valuable knowledge extraction in sectors [47], such as spanning manufacturing [48], healthcare [49], government [50], retail [51], and infrastructure [52]–[54]. In particular, numerous high-quality open-source training datasets [55] have been created to boost AI research in the aspects of model training, testing, calibration, and benchmarking.

Moving with the tide of digitizing power systems, the explosive growth of data resources has also created massive volumes of data in heterogeneous formats, including electrical measurements that span across grids vertically, such as sensors installed on grid-level components, smart meters, and smart appliances, as well as nonelectrical measurements, such as weather, social media, traffic, and geographic information [56]. These data have proven very valuable in many use cases, such as asset assessment, operation planning, real-time monitoring, and protection [57]. It is worth noting that these basic functionalities have distinct requirements for data quality in perspectives of data accuracy, latency, and sampling rate [58]. This section will review data acquisition approaches for electrical measurements in the power grids.

A. Real-World Measurements in Power Systems

1) Sensors in Transmission Systems: SCADA systems, which have played an important role in transmission system operation, are capable of collecting facility information and sending control signals, which are implemented by the critical component (i.e., RTUs). SCADA systems collect asynchronous data on bus voltage magnitude, and active and reactive power flows; the typical reporting rate is merely one sample per 2–6 s. The wide-range acquisition of SCADA data has facilitated remote monitoring and system operation automation. For example, the EMS at the control center is capable of estimating physical state variables that are not directly observable based on SCADA data alone. However, due to increasing system complexity and uncertainty, even this successful SCADA-based application is becoming inadequate.

PMUs have been deployed in the bulk transmission grid at an accelerated rate after the 2003 U.S. blackout [59]. PMUs are able to measure the voltage phasors[5] at the installed bus (typically substations) and current phasors of the lines connected, along with synchronized time stamps, for which the typical reporting rate is 30 or 60 samples per second. Compared to SCADA, PMUs' high accuracy of time stamps and sensing, low latency, and a high sampling rate of PMU benefit basic functionalities to different degrees [60]: 1) more real-time control and protection applications become potentially implementable due to all of these advantages, such as remedial action schemes, including grid islanding and short-term stability control; 2) online system security analysis, such as disturbance detection and situational awareness, can be significantly improved due to low latency; and 3) system adequacy analysis for long-term planning, such as model calibration, can be improved due to high accuracy. However, it is worth noting that, due to several factors, such as high costs and time consumption of installation, only around 2500 production-grade PMUs have been installed across the North America transmission power grid [61], [62].

DFRs capture and store transient data and SOE data that can be used for various purposes, such as protection scheme monitoring and fault diagnosis, which tend to be implemented offline. DFRs have three typical recording mechanisms: steady-state, low-speed, and high-speed

[5]Phasors contain magnitude A and phase angle ϕ of sinusoidal waveforms that can be expressed as $A\sin(\omega t + \phi)$, where ω is $2\pi \times 60$ rad/s in a 60-Hz system.

disturbance recording modes. The disturbance recording modes are usually triggered by signals from protection relays. The steady-state recording mode captures the min, max, and mean values of phasors at a low sampling rate of one sample per 10 s to 1 h. The low-speed disturbance mode aims to provide phasor-domain information of long- and short-term disturbances at a sampling rate of one sample per one to ten cycles. The high-speed disturbance mode aims to record instantaneous time-domain voltage and current measurements of transient faults at a sampling rate of hundreds of samples per cycle.

2) Sensors in Distribution Systems: The rapid expansion of AMI meters at the grid edge has created massive amounts of residential electricity consumption data, typically at a rate of one sample every 1–5 min. For example, the Pacific Gas and Electric Company collects more than 3 TB of power data from nine million smart meters across the grid in the territory, and the State Grid Corporation of China collects 200 TB of data per year [63].

SCADA in distribution systems has facilitated remote monitoring and automated operation in multiple aspects, such as substation, feeder, and end-user load control. In substation systems, SCADA gathers data, including voltage magnitude, current magnitude, and binary status of facilities, such as switches, breakers, and transformers. In typical feeder systems, SCADA facilitates the collection of historical data from the feeder status of devices such as controlled load break switches and reclosers. In end-user load, SCADA collects all meter data from the end-users.

The FDR, one of the representative PMU applications in distribution systems, is a GPS-synchronized single-phase PMU at ordinary 120-V wall outlets. FDRs have the advantages of low cost and high deployability; they can be deployed even at residential households and campuses [64]. Using hundreds of FDRs that have been strategically placed across the U.S., the frequency monitoring network FNET/GridEye [65] is able to provide visualized nationwide frequency monitoring.

B. Artificially Generated Power System Data

Artificially generated data are commonly used for power system research for two major reasons: 1) most real-world operational data are protected by policies such as CEII owing to confidentiality and 2) real-world measurement datasets of high-impact events are usually insufficient for data-driven model training due to the reliability of real-world power grids, which ensures that high-impact events are rare. Alternatively, artificial data generation methods facilitate the gathering of arbitrary numbers of data samples under varying scenarios and conditions, including voltage, current, frequency, and even machine inner state measurements across grid models.

1) Model-Based Simulation: Model-based simulation is one of the most common data acquisition approaches for research and education purposes. Simulation models of transmission and distribution systems can be categorized into two major types: 1) small-scale standard systems and 2) large-scale synthetic systems, which are available at [66]. The IEEE standard test systems are typically used for investigations such as algorithm assessment and power system analysis. Researchers have recently contributed to the creation of large-scale synthetic grid models [67] that possess realistic system characteristics. These large-scale synthetic grids have been used for analysis such as macroscope energy portfolio transition [68], [69] and quantitative assessment of measures against extreme events [70]. For intuitive impression, we show the "popularity" of simulation models in Fig. 4 by counting the number of corresponding IEEE Transaction papers,[6] which are used for both machine learning model training and testing. It is clear that the most commonly used models for AI algorithm training, testing, and calibration are the IEEE 39-bus and 118-bus systems, whereas the large-scale models are rarely adopted. Please refer to Tables 2–5 for other simulation models that are not included in Fig. 4.

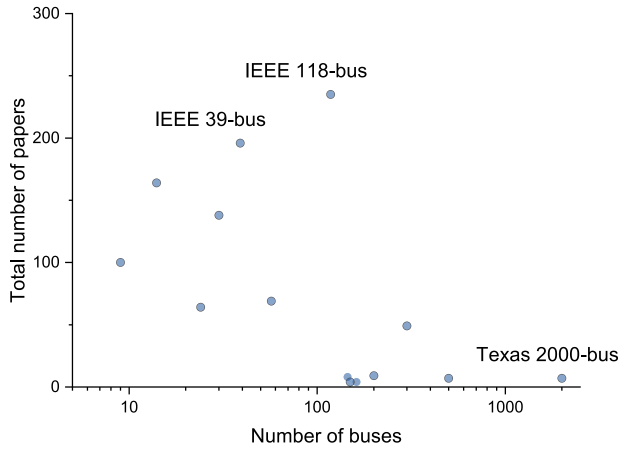


Fig. 4. ***Number of AI/ML/DL papers per simulation model of various system scales. Note that we only count typical open-source simulation models, including IEEE standard test cases and large-scale synthetic grids using Google Scholar advanced search among IEEE Transaction papers from 2016 to 2021.***

2) Hardware Test Bed: The development of HIL simulators has been used to support various types of research, including event detection, situational awareness, wide-area monitoring and control, and cybersecurity [71]. HIL leverages the interface between a real-time software simulator and a hardware system to enable closed-loop control [72]. HIL may play an important role in electromagnetic transient simulation of electronics-rich power grids because of its ability to represent realistic very-fast dynamics.

[6]We use the Google Scholar advanced search to count IEEE Transaction papers from 2016 to 2021, using several keywords, such as "IEEE Transactions on," "machine learning," "#-bus system," "power flow," and "transient."

This section gives an overview of data acquisition approaches in today's electric power grids. The rapid expansion of advanced sensors across systems and the development of simulation have facilitated massive data acquisition spanning multiple spatial and temporal scales, and have further accelerated practical data-driven applications. Efforts to explore data-driven innovation, such as big data hubs [73], [74], have also promoted data-intensive research in the power system industry, as well as academia and education. Despite these advances, there are several key challenges regarding the data for AI algorithms. First, in contrast to numerous datasets that have benefited broad AI communities, the lack of publicly accessible high-quality power datasets may be impeding the advancement of AI research in power systems. For example, insufficient data representativeness is one of the decisive factors for data-hungry AI methods. Real-world measurements cannot provide a sufficient volume of publicly available data due to confidentiality rules and strong grid reliability. Randomly sampled scenarios in simulation can generate massive amounts of data, but they do not necessarily guarantee representativeness; therefore they likely lead to unexpected training biases, which was demonstrated by the example of ACOPF scenario generation [75]. Second, the feasibility of the proposed AI algorithms may be constrained by the current data acquisition system, as indicated by the data quality requirements of major power system applications [58]. For example, limited and inappropriate placement of high-sampling sensors that determine situational awareness for a specific task may confine advanced analysis and control, including, but not limited to, practical applications of AI methods. Third, although AI methods may offer unique creativity given cross-domain datasets, they require deep interdisciplinary knowledge and collaboration to identify useful combinations of heterogeneous datasets, which has been demonstrated by few AI-based canonical studies, such as automatic classification of distribution grid phases by camera imaging [76] and comprehension COVID impacts on power sectors by mobile phone location data [77].

IV. COMPUTING IN DIGITIZED POWER GRIDS

Given sufficient available data resources, the implementation of data-driven applications in modern power grids faces computational burdens derived from large-volume, heterogeneous data. Such implementation is critical to handle the associated challenges that include data streaming storage, querying, and processing. This section will give an overview of state-of-the-art computing that has facilitated general AI and will then introduce data streaming management systems and data processing platforms [63], [78] in power systems.

A. Overview of State-of-the-Art Computing for AI

The remarkable improvement of computing performance is the key factor in the proliferation of AI, which is attributable to advances in hardware, software, and generic algorithms [27]. Quantum leaps in computing performance have yielded a variety of practical large-scale AI models, among which the amount of computation for model training has been increasing exponentially with a 3.4-month doubling period [22], [79]. The rapid progress of hardware computing resources has been the main driver behind the development of AI models. Of particular note, the emergence of general-purpose GPUs [80] and AI accelerator ASICs, such as [81]–[84], is capable of dramatically accelerating AI model training. In addition, AI-tailored software has been developed to exploit hardware computing resources [85]. For instance, BLAS libraries, which were created decades ago [86]–[88], have been used to optimize common linear algebra operations that are recursively executed in deep NNs [89]–[92]. In particular, Nvidia GPUs, which are widely supported by mainstream deep learning framework [93]–[95], have a highly optimized library cuDNN [96] enabling high-performance GPU acceleration. The progress of generic algorithms has also improved computing performance, exhibiting enormous heterogeneity on problems of different types and sizes [97]. It is worth noting that some large-size problems benefit just as much or even more from algorithmic improvement than from Moore's law. For instance, the total speedup of solving MIOs was 2.2 trillion times during the 25 years between 1991 and 2016 [21], of which a factor of 1.6 million is due to hardware speedup from 59.7 GFlop/s in 1993 to 93.0 PFlop/s in 2016; another factor of 1.4 million is due to software and algorithmic speedup from CPLEX 1.2 in 1991 to Gurobi 6.5 in 2015.

B. Data Management Platforms in Power Grids

Because power system security highly relies on real-time system operation and control, it is challenging to store and process real-time data streaming effectively and efficiently. Therefore, the building of real-time data streaming systems that mainly influence data latency is critical for the subsequent online data-driven applications, including, but not limited to, AI-based methods. In contrast to traditional database management systems that use statistical data storage, data stream management systems usually store synopsis data (instead of the entire dataset) via processing in order to handle frequent queries and data updates. We illustrate several of the most popular data stream management systems summarized in [63]: Aurora [98] has a good balance of accuracy, response time, and resource utilization; TelegraphCQ [99] is mainly used for sensor networks, which involves a front end, sharing storage, and a back end; and STREAM [100] has the advantage in situations of limited resources in that it can execute queries with high efficiency.

In particular, big data management platforms are being developed to accommodate multimodal data storage and processing of unstructured heterogeneous data. Hadoop [101] and Spark [102] are two representative open-source designs for distributed data management. Hadoop is able to process massive heterogeneous data efficiently and economically by taking advantage of a programming model [103], a distributed file system [104], and a distributed data storage system [105]. Spark, on the other hand, leverages the technology of resilient distributed datasets [106], which is more suitable for recursive computational operations in machine learning-based applications. In terms of data management platforms that are suitable for power systems, several cases of solutions have been successful in facilitating energy efficiency. For example, CenterPoint Energy has handle streaming messages from intelligent grid devices and smart meters using an IBM-developed platform to improve system reliability [107]. For its part, Oncor Energy Delivery has developed AMI data-based predictive maintenance to reduce outages and guarantee sustainable supply enabled data platforms [108].

Because of power grid digitization, computing tasks in today's power grids have been shifted and evolved to centralized clouds. Advanced computing power, along with massive data acquisition, has enabled many time-sensitive operations, such as real-time monitoring and security analysis. However, with the increasing complexity of power grids, such a computing paradigm may face several challenges, such as privacy concerns and communication bandwidth limits. In contrast, edge computing that leverages computing resources at the edge has the potential to improve computation efficiency and protect data privacy by performing data analytics close to customers [109]. Particularly, machine learning approaches that can preserve privacy, such as federated learning [110], have drawn increasing attention.

V. AI SOLUTIONS TO POWER GRID DECISION-MAKING

This section surveys recent AI solutions to the core decision-making processes in power grid operations. We report 85 papers, most of which were published in the IEEE TRANSACTIONS of the Power and Energy Society (e.g., IEEE TRANSACTIONS ON POWER SYSTEMS and IEEE TRANSACTIONS ON SMART GRID) from 2019 to 2021. For earlier works about AI algorithms for grid operations, we refer readers to previous survey papers [23], [111]–[113]. Table 1 classifies the approaches used in these 85 papers according to the category to which these approaches belong (i.e., supervised, unsupervised, and RL). In addition, for each decision-making process, we provide not only an overview of the state-of-the-art, AI-powered grid solutions but also illustrative examples that give readers a sense of how specific AI techniques can be leveraged to solve grid challenges. We use an independent notation system in each subsection.

Table 1 Classification of Grid Solutions Based on AI Methods

Key Decision-making Module	Supervised Learning	Unsupervised Learning	Reinforcement Learning
Renewable/Load Modeling	[114]–[132]	[118], [121] [123], [124]	[133]
Grid Economic Operation	[134]–[150]	-	-
Grid Security & Resource Adequacy	[151]–[163]	[159], [164]	-
Grid Monitoring, Control, & Protection	[165]–[193]	[194], [195]	[196] [197], [198]

A. Renewable/Load Modeling

Renewables and load introduce many uncertainties to the operation of low-carbon power grids. One way to address such uncertainties in grid operation is to develop an accurate forecast algorithm for renewables and load. The topic areas in renewable/load modeling include renewable (e.g., wind and solar) generation forecasting, load forecasting, and load clustering. Table 2[7] lists the most recent works in these topic areas. Table 2 also summarizes the data source, AI method, and computation resource used in the references provided.

Next, we provide an example to elaborate on how AI can be leveraged to solve PV forecasting tasks in the grid. The technical details are reported in [199]. Fig. 5 shows the geographic locations of a target solar site C6 and its neighboring N solar sites. Let us suppose that we want to predict the solar irradiance of the target solar site C6 at time step $(k+1)$. Yang *et al.* [199] formulate the forecasting problem into one of estimating the parameters of the following ARX model [199]:

$$\begin{aligned} x[k+1] &= f(x[k], \ldots, x[k-n+1], \\ &\quad w_1[k-d_1+1], \ldots, w_1[k-m_1-d_1+2], \\ &\quad \ldots w_i[k-d_i+1], \ldots, w_i[k-m_i-d_i+2], \\ &\quad \ldots w_N[k-d_N+1], \ldots, w_N[k-m_N-d_N+2]) \end{aligned} \tag{1}$$

where $x[k]$ is the solar irradiance at the target solar site at time step k; w_i is the solar irradiance at the neighboring solar site i; $f(\cdot)$ is an ARX-structured function; and positive integers n, d_i, and m_i are user-defined parameters that can be determined at training stages [199]. The intuition of the formulation (1) is that the next-step solar irradiance $x[k+1]$ at the target solar site depends not only on the local solar irradiance but also on the solar irradiance at its neighboring solar sites. The case studies based on real-world renewable data from California and Colorado suggest that such an algorithm is suitable for 1- and 2-h ahead PV forecasting [199]. However, the algorithm proposed in [199] does not provide a probability description for the forecast quality. One potential avenue for future work is to investigate such a description [199].

[7]In the table, "—" indicates that no computation resource is reported in reference.

Table 2 Trifactors of AI-Based Renewable/Load Forecasting

Task [Ref.]	Data Sources (Availability)	Computation Resources	AI Method
PV prediction on shipboards [114]	Solar irradiance and weather data (N)	i7-7700 CPU	Neural network (NN), Extreme learning machine (ELM)
Wind power prediction [115]	Actual wind power data from Glens of Foudland wind farm (N)	i7-7700 CPU, 16GB RAM	ELM
Wind power prediction [116]	Wind generation data from Irish transmission system operator (N)	i7-7700 CPU; 16GB RAM	ELM
Socio-demographic information identification [117]	Residential smart meter data (N)	i7-4770MQ CPU	Convolutional neural network (CNN)
Feature selection for PV forecasting [118]	Three PV arrays' measured datasets in Australia (N)	-	Principal component analysis (PCA); K-nearest neighbors (KNN)
Wind power ramp forecasting [119]	Wind Integration National Dataset (WIND) Toolkit [119] (Y)	Intel-e5-2603 32 GB RAM	Gaussian mixture model
Solar forecasting [120]	Desert Knowledge Australia Solar Centre (DKASC) (N)	Nvidia GTX 1080 GPU	Spectral graph convolution
PV forecasting [121]	DKASC (N)	-	ELM, Auto-encoder
Short-term load forecasting [122]	Actual data of Tianjin power Grid in China (N)	i5-6700 CPU, 16GB RAM	Deep belief network
Load forecasting [133]	Hourly load data of the University of Texas at Dallas (UTD) (Y)	i7-4870HQ CPU, 16GB RAM	Reinforcement learning
Customer segmentation [123]	Real-world utility data (N)	-	Spectral clustering; regression
Load clustering, prediction, and inference [124]	18-node real utility feeder (N)	-	Spectral clustering; Recursive Baysian Learning
Wind power prediction [125]	Data from three wind farm data in China (N)	-	Multi-source and temporal attention network
Demand flexibility estimation [126]	Datasets from ERCOT, PJM, CAISO (Y)	i7-8550U CPU, 16GB RAM	Long short term memory (LSTM) NNs
Wind power prediction [127]	Wind farm data from NREL (Y)	i7-7700 CPU, 16GB RAM	ELM
PV power prediction [128]	5-MW PV power plant (Y)	i7-2600 CPU	Deep belief network
Cold load pick-up demand assessment [129]	Real-world smart meter data (N)	-	Regression; Gaussian mixture model
Dynamic load modeling [130]	CIGRE benchmark low voltage network (N)	-	Decision trees; Ant colony optimization
PV power forecasting [131]	PV plant dataset (N)	-	Graph neural network
Load forecasting [132]	Real data set from residential Irish customers (N)	-	Random forest

B. Grid Economic Operation

The large-scale deployment of renewables poses unprecedented challenges to the electricity market operation. Conventional deterministic tools may not be able to support the electricity market operation of the electricity infrastructure with a significant amount of uncertain renewables. Modarresi *et al.* [200] propose a scenario-based approach that unlocks the potential of data in order to incorporate renewables' uncertainties into the dispatch of grid resources. Let us suppose that there are N historical scenarios $\Delta_N = \{\delta_1, \delta_2, \ldots, \delta_n, \ldots, \delta_N\}$ that is a subset of all possible scenarios Δ. In each historic scenario δ_n, the net-load forecasting errors at each bus are recorded. Modarresi *et al.* [200] formulate the ED problem as follows [200]:

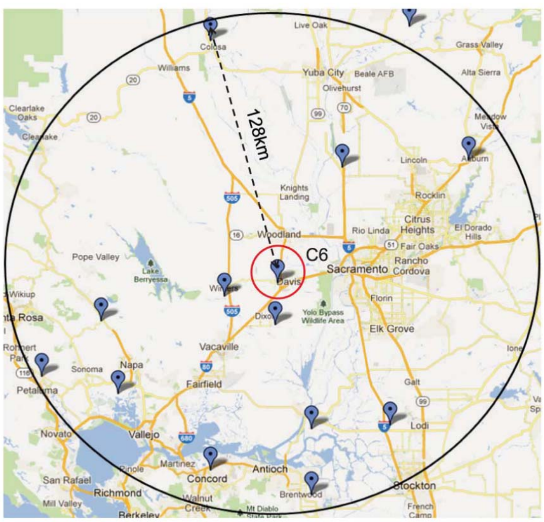


Fig. 5. *Target solar farm site C6 (in the red circle) and its neighboring solar farms. (Source: Fig. 1 of [199] © IEEE 2015.)*

$$\min_{p} \; c^{\top} p \tag{2a}$$

$$\text{s.t. } g_1(p) \le 0 \tag{2b}$$

$$g_2(p, \delta_n) \le 0 \quad \forall \delta_n \in \Delta_N \tag{2c}$$

where vector p concerns the power generation of all generators in all intervals during a planning horizon; vector c collects cost coefficients associated with generators; (2b) represents the scenario-independent constraints [200], such as ramp and capacity constraints of generators; and (2c) represents the scenario-dependent constraints [200], such as generation-load balance constraints. Suppose that p_N^* is the solution to the optimization (2) given N historical samples. Because Δ_N is a subset of all possible scenarios Δ, it is possible that there exists a scenario δ that

Table 3 Trifactors of AI Solutions to Market Operation

Task [Ref.]	Data Resources (Availability)	Computation Resources	AI Methods
UC & ED [134]	A real-world 1881-bus system in China (N)	i7-1065G7 CPU, 16GB RAM	NN and oblique decision trees
OPF [135]	Polish 2383-bus system (N)	i7-8700K CPU, 32GB RAM	Stacked ELM
OPF [136]	129-node feeder (N)	-	Regression
Probabilistic power flow [137]	A 33 bus distribution system from MATPOWER	i7-7820HQ CPU, 31.9GB RAM	Copula function
ACOPF [138]	A 6700-bus French system (N); a 9000-bus system (N)	i7 CPU 16GB RAM; Nvidia Tesla V100 GPUs, 16GB RAM	NN
Binding constrain prediction in UC [139]	500-bus synthetic South Carolina system (N)	24-core CPU, 128GB RAM	Bagged trees
SCOPF [140]	118-IEEE, 1354-PEG; 1888-RTE system (N)	Nvidia Tesla V100 GPUs	Deep learning
OPF [141]	30-, 57-, 118-, 300-IEEE (N)	4-core i7-3770 CPU, 16GB RAM	Deep learning
Financial return maximization [142]	261,489-consumer real dataset from a Brazilian utility (N)	-	Rotation forest; XGBoost
Energy pricing [143]	13-IEEE (N)	-	Regression
Energy bidding [144]	Real-world price data (Y)	-	Risk-averse learning
Constrain screening of UC [145]	IEEE RTS-96; 2000-bus Texas synthetic grid (N)	2.6 GHz CPU, 2GB RAM	KNN XGBoost
Reserve capacity estimation [146]	821-day data (N)	-	ELM
ED and UC [147]	24-, 118- IEEE (N)	Intel Xeon CPU, 16GB RAM	XGBoost
Energy bidding [148]	Market data from PJM, CAISO, and ISO-NE (N)	-	XGBoost
Electricity price scenario generation [149]	Dutch market prices (N)	Nvidia Tesla V100, 61GB RAM	LSTM; recurrent neural network
Price forecasting [150]	Historic price data (N)	-	ELM

causes the scenario-dependent constraints to be violated, i.e., $g_2(p_N^*, \delta) > 0$. The probability that such an event may occur is termed the "risk" in [200]. Formally, the risk $v(p_N^*)$ for the solution p_N^* is defined by

$$v(p_N^*) = \text{Prob.}(\delta \in \Delta : g_2(p_N^*, \delta) > 0) \quad (3)$$

where $\text{Prob.}(\cdot)$ denotes the probability that event "·" occurs. We expect that the probability that the risk $v(p_N^*)$ of the solution p_N^* exceeds a small number ϵ will be small, i.e.,

$$\text{Prob.}(v(p_N^*) > \epsilon) < \gamma \quad (4)$$

where $0 < \epsilon, \gamma \ll 1$. With the risk preference parameters ϵ and γ, a natural question is how to determine the size of Δ_N, i.e., N, to achieve the risk preference (4). Modarresi *et al.* [200] provide a lower bound of N that depends solely on the look-ahead intervals and risk preference parameters [200]. Such a lower bound can help system operators determine how many scenarios must be drawn from the historical observations based on their risk preference. For example, in an open-source, 2000-bus synthetic Texas grid, if we suppose that the risk preference parameters of the system operators are $\gamma = 10^{-6}$ and $\epsilon = 0.0083$, then 2000 historical scenarios are needed to be embedded into the ED formulation (2) [200]. A rigorous investigation of the relationship between the number of support constraints and the design parameters (γ and ϵ) is still needed to further refine the algorithm in [200].

Other recent AI solutions to the problems of UC, ED, and OPF are summarized in Table 3. The AI methods associated with the data sources and computation resources in the references are listed in Table 3.

C. Grid Security and Resource Adequacy

To decarbonize the power grids, fossil-fueled generators are being replaced by IBRs, e.g., wind/solar farms and energy storage. To assess grid security and resource adequacy, it is necessary to develop new planning tools that explicitly consider these new elements. The grid security and resource adequacy analysis include steady-state, DSA, and reliability analyses. Table 4 summarizes the state-of-the-art AI adoption in these analyses. Next, we will present a learning-based approach to networked microgrid security analysis [164], in order to show how an AI technique can be adopted in this specific topic area.

Fig. 6 shows the physical architecture of n networked microgrids, where the n microgrids interact with one another via distribution lines. The dynamics of the networked microgrids can be described by $\dot{\mathbf{x}} = f(\mathbf{x})$, where the state vector $\mathbf{x}$ is related to voltage magnitudes and phase angles at the PCCs. In the networked microgrids, large disturbances may come from: 1) the microgrid operating mode change, e.g., one microgrid enters an islanded mode and 2) the distribution network, e.g., distribution line tripping. The security analysis attempts to quantify the disturbance magnitude that the networked microgrids can tolerate [164]. The result of this analysis is critical for both distribution system planners and operators.

Huang *et al.* [164] formulate the security analysis problem as one of searching for a legitimate Lyapunov function, i.e., a system-behavior summary function for a dynamic

Table 4 Trifactors of AI Solutions to Grid Security and Resource Adequacy Analysis

Task [Ref.]	Data Sources (Availability)	Computation Resources	AI Method
Reliability study for power-gas systems [151]	73-bus power system w/ 40-node gas system (N)	-	Random Forest; XGBoost
Reliability study w/ rich PE [152]	IEEE RTS-24 bus w/ 40-node gas system (N)	Intel CPU, 16GB RAM	Support vector regression (SVR); random forests (RF)
Energy Loss estimation [153]	33-bus distribution system w/ 40-node gas system (N)	PC, 8GB RAM	Regression trees
Distribution system phase identification [154]	25-bus, 123-bus, 450-bus systems (N)	-	Modified k-means
Outage scheduling [155]	IEEE-RTS79; IEEE- RTS96 (N)	300-core Xeon CPUs, 2GB RAM for each	k-nearest neighbor classification
Energy Disaggregation at Substations [156]	Ind-Solar dataset; EnerNOC GreenButton Data; Solar generation from National Renewable Energy Laboratory (N)	i9 CPU, 32 GB RAM	Bayesian dictionary learning
Distribution power flow [157]	8-, 123- IEEE; 362-node utility distribution network (N)	-	Support Matrix Regression
Microgrid Scheduling [158]	33-node system (Y)	-	Deep learning
Stability Assessment of Networked Microgrids [164]	123-IEEE (N)	i5 CPU, 8GB RAM	Neural Lyapunov method
Residential PV localization [159]	Umass Smart data (Y)	Xeon CPU E5-2687W v4, 64 GB RAM	Autoencoder; semi-supervised learning
Phase Identification [160]	Systems from SCE, PGEC, and FortisBC (Y)	-	Information theoretic machine learning
Reliability study [161]	ISO-NE area data (Y)	-	Regression
Extreme outage prediction [162]	24-IEEE (N)	-	Bayes Decision Theory
Network management [163]	4,000 circuits from U.K. utilities (N)	-	Decision tree

system. A Lyapunov function $V(\mathbf{x})$ satisfies two conditions: 1) $V(\mathbf{x})$ is a positive-definite function in a region $\mathcal{R}$ around the system equilibrium point and 2) the time derivative $\dot{V}$ is a negative-definite function in $\mathcal{R}$. In [164], the Lyapunov function is assumed to possess an NN structure with parameter vector $\boldsymbol{\theta}$. To make the NN-structured function satisfy the two conditions of a Lyapunov function, a cost function $c(\boldsymbol{\theta})$ is designed. The cost function incurs a positive penalty if the NN with $\boldsymbol{\theta}$ violates one or both of the two Lyapunov function conditions. Vector $\boldsymbol{\theta}$ is tuned by the following procedure.

1) Create a sample pool by randomly drawing a large number of states $\mathbf{x}$ within the region $\mathcal{R}$.
2) Update $\boldsymbol{\theta}$ n times based on the cost function $c(\boldsymbol{\theta})$ and the gradient descent algorithm [164].
3) For the NN with the latest $\boldsymbol{\theta}$, search for samples that violated one or both of the two Lyapunov conditions via the SMT tool. If no sample is found, claim that the NN is a Lyapunov function; otherwise, add the samples to the sample pool in step 1) and repeat step 2).

Fig. 7 visualizes a Lyapunov function learned from a state space for a grid-tied microgrid [164]. The parameters of the system are reported in [164]. It takes 32.18 s to learn the Lyapunov function [164]. Having learned the Lyapunov function shown in Fig. 8, a security region (SR) can be estimated, which is visualized in Fig. 9. If a disturbance leads the state vector to deviate from the equilibrium (the origin of Fig. 9) while also remaining within the solid red circle in Fig. 9, one can conclude immediately that the system trajectory will converge to the equilibrium without conducting any simulations. The

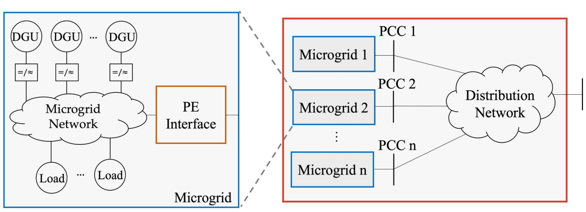


Fig. 6. ***Physical architecture of networked microgrids with power electronics interfaces. (Source: Fig. 1 of [164] © IEEE 2021.)***

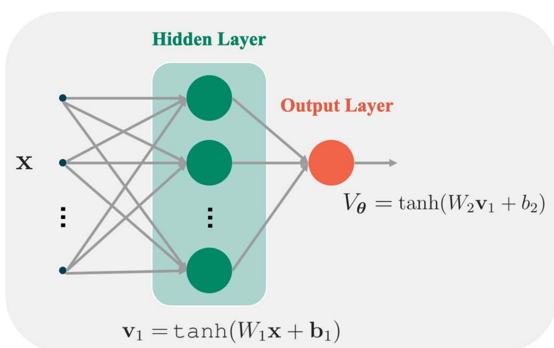


Fig. 7. ***NN-structured Lyapunov function: the tunable parameter vector θ is related to weights W_1 and W_2, and biases b_1 and b_2 in the hidden and output layers. (Source: Fig. 3 of [164] © IEEE 2021.)***

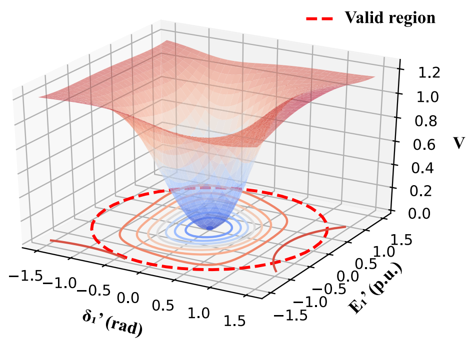


Fig. 8. *Lyapunov function learned for a grid-tied microgrid. (Source: Fig. 6(a) of [164] © IEEE 2021.)*

region in the solid blue circle is the SR estimated by a conventional approach. It can be observed in Fig. 9 that the learning-based approach is much less conservative than the conventional approach since the red-solid circle is larger than the blue circle. Although the approach in [164] can address heterogeneous interface dynamics and can provide less conservative results than the conventional approach, it incurs large computational costs when analyzing large-scale systems.

D. Grid Monitoring, Control, and Protection

Deep penetration of clean energy resources is changing power grid behavior (for example, clean-energy resources may lack physical inertia). As a result, the power grids are becoming increasingly sensitive to disturbances, and impact anomalies may become more frequently observable. Effectively monitoring and correcting these anomalies in real time define a key challenge facing system operators. A large body of literature in the last three years has argued in favor of leveraging streaming data to make

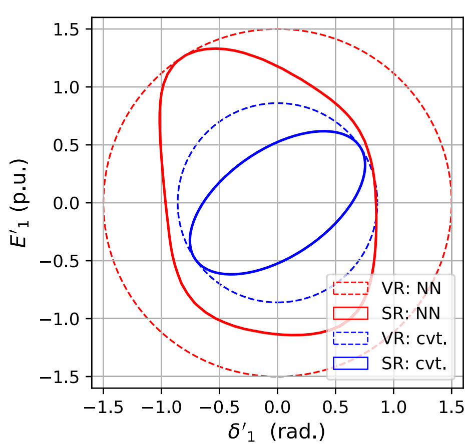


Fig. 9. *SRs and valid regions (VRs): the NN approach and the conventional (cvt.) approach. (Source: Fig. 7(a) of [164] © IEEE 2021.)*

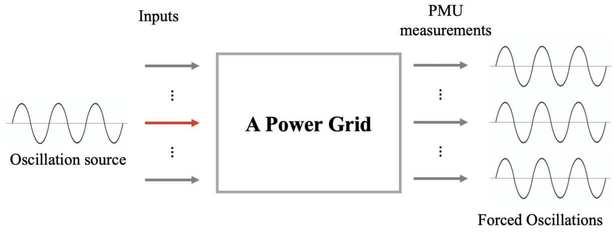


Fig. 10. *Forced oscillation mechanism.*

operational decisions in real time. Table 5 summarizes these recent works from the perspectives of data sources, methods, and computation resources. The following are two specific examples that address online operational challenges in the grid.

1) Forced Oscillation Localization Based on Robust Principal Component Analysis: Forced oscillations are one type of critical phenomenon that concerns system operators because these oscillations may cause large-scale blackouts and decrease the lifespans of power grid components [201]. Fig. 10 illustrates the mechanism of forced oscillations. Let us consider a power grid as a blackbox with some inputs and outputs, as shown in Fig. 10. The inputs can be thought of as setpoints of generators, while the outputs are PMU measurements. If one of the inputs varies periodically, oscillations can be observed in the PMU measurements. These oscillations are termed "the forced oscillations," and the periodic input is called the source of the forced oscillations. Different PMU measurements have different geographical distances from the oscillation source. The objective of the forced oscillation localization is to pinpoint which PMU measurements are close to the oscillation source based only on the PMU data without information on the inputs and the power grid models.

Locating the oscillation source is a challenging task because the measurement closest to the source may not exhibit the largest oscillations. Fig. 11 shows such a counterintuitive case in which the measurement (the red curve) closest to the oscillation source does not exhibit the largest oscillation magnitude. Nezam Sarmadi *et al.* [202] report a real-world, counterintuitive case in which the distance between the source and the measurement exhibiting large oscillations is more than 1100 mi [201]. Huang *et al.* [201] formulate the forced oscillation localization as decomposing the measurement matrix Y_t into a low-rank matrix L_t and a sparse matrix S_t, namely, $Y_t = L_t + S_t$. This matrix decomposition problem can be solved by RPCA as follows:

$$\min_{S_t} \quad \|Y_t - S_t\|_* + \epsilon\|S_t\|_1 \tag{5}$$

where $\|\cdot\|_*$ and $\|\cdot\|_1$ denote the nuclear norm and the l_1 norm, respectively, Y_t represents a measurement matrix up to time t where each row of the matrix represents a time series from one PMU, and S_t is the corresponding approx-

This article has been accepted for inclusion in a future issue of this journal. Content is final as presented, with the exception of pagination.

Table 5 Trifactors of AI Solutions to Grid Monitoring, Control, and Protection

Task [Ref.]	Data Sources (Availability)	Computation Resource	AI Method
Cyberattack detection [165]	13-, 123- IEEE (N)	-	Bayes classifier
Actuator placement [194]	118-, 123- IEEE (N)	2-core Xeon CPU, 32GB RAM	K-means clustering
Frequency prediction assessment and control [166]	140-NPCC (N)	i5-5200U CPU, 8GB RAM	ELM
Emergency control [196]	39-IEEE (N)	AMD Opteron CPU, 64GB RAM	Deep reinforcement learning
Under-voltage load shedding [167]	77-Nordic (N)	8-core i7-7700 CPU; Nvidia GTX-1080 GPU	Deep feedback learning machine
Transient stability prediction [168]	39-IEEE (N)	i7-7700 CPU, 32GB RAM; Nvidia GTX-1080 GPU	CNN
Transient stability prediction and control [169]	39-IEEE (N)	-	AdaBoost
Generator dynamic behavior prediction [170]	39-, 68-, 140-, 145- IEEE (N)	Intel CPU, 16GB RAM	Ensemble decision trees
Voltage stability margin prediction [171]	118-NREL; Taiwan Power Systems (N)	8-core CPU, 16GB RAM	ELM
Training data preparation for transient stability prediction [172]	39-IEEE; 2417-bus GD Power Grid in South China (N)	8-core i7 CPU, 8GB RAM	Semi-supervised ensemble learning
Representative state selection for security prediction [173]	118-IEEE (N)	8-core Intel Xeon PC	C-Vine pair-copula decomposition; PCA
Hydrostatic tidal turbine (HTT) control [174]	HTT simscape model (N)	-	Extreme learning machine
Cyber-physical anomaly detection [175]	HiL test of Kundur two area system (N)	-	Decision tree; KNN
Inverter control [176]	123-IEEE (N)	i5 2.4 GHz, 8GB RAM	Support vector machine (SVM)
Transient stability prediction [177]	39-IEEE; 2417-bus GD Power Grid in South China (N)	8-core i7 CPU, 8GB RAM	Time series shapelet learning
Line outage detection [178]	30-, 118-, 300- IEEE (N)	i7 CPU, 8GB RAM	Learning-to-infer
Dynamic security prediction [179]	39-IEEE (N)	-	CNN, LSTM network
Distribution system state estimation [180]	37-IEEE (N)	-	NN
Local control design for active distribution grids [181]	Typical European radial LV grid (N)	Intel Core i7-2600 CPU 16GB RAM	Regression; SVM
Anomaly detection, localization and classification [195]	14-, 39- IEEE (N)	Multiple CPUs	Autoencoder
Volt-VAR optimization [197]	13-, 123- IEEE (N)	i5 CPU, 8GB RAM	Reinforcement learning
Cyber anomaly detection [182]	39-IEEE (N)	-	SVM, Decision tree
Faulted line localization [183]	39-, 68- IEEE (N)	i7 CPU, 32GB RAM	CNN
Feature extraction for Security assessment [184]	118-IEEE (N)	8-core Intel Xeon PC	Deep learning
Building Energy Optimization [198]	Real-world data from Pecan Street Inc. (N)	-	Reinforcement learning
PV frequency control [185]	6,102-bus system; HiL tests (N)	-	Multivariate random forest regression
Wind turbine control [186]	HiL tests (N)	-	Cascade-forward neural network
Non-intrusive load monitoring [187]	Reference Energy Disaggregation Dataset (Y)	i7-4770 CPU, 12GB RAM	Semi-supervised multi-label algorithms
Building occupancy detection [188]	Electricity consumption and occupancy (ECO) dataset; Building-level fully labeled Electricity disaggregation dataset; UMass Smart Home Dataset Almanac of Minutely Power Dataset (Y)	Intel Xeon E5-2603 CPU; Nvidia TITAN V GPU	CNN, LSTM network
Outage duration prediction [189]	E15-year outage records (N)	-	Natural language processing
Event detection and classification [190]	HiL tests (N)	i7-9700K CPU; Nvidia GTX 2080Ti GPU	CNN
Cyber attack detection [191]	57-, 118-, 300- IEEE (N)	i9-8950 HK CPU 2.90GHz Nvidia GeForce RTX 2070 GPU	Graphic Neural Network
Distribution system topology identification [193]	33-IEEE; 135-bus, 874-bus systems (N)	-	Split expectation maximization
Grid restoration [192]	70-bus 4-feeder system (N)	i5 CPU, 8GB RAM	Regression

imate sparse matrix. Fig. 12 visualizes matrices Y_t, L_t, and S_t, respectively. The computation complexity analysis of RPCA is reported in [203]. The measurement near the source can be located by identifying the largest absolute element in the sparse matrix. Huang *et al.* [201] also provide a possible interpretation to justify the effectiveness of the RPCA-based source localization algorithm. Huang *et al.* [201] create 44 counterintuitive cases in an open-source, benchmark system. The RPCA-based algorithm can pinpoint the sources in 43 cases, and in the wrong case, the algorithm can narrow the searching scope [201]. However, when the RPCA can exactly locate the true source remains an open-ended question.

2) Reinforcement Learning-Based Protection Scheme for Renewable-Rich Distribution Systems: The conventional protection paradigm in distribution systems has been challenged by the increasing amount of DERs. Fig. 13 presents the overcurrent protection scheme that is widely deployed in power distribution systems. Such a protection scheme will trip the line once the line current exceeds a threshold value, e.g., five times the current I_0 under normal conditions. However, if a DER is installed nearby, it may decrease the fault current by injecting reverse power flow. As a consequence, the current under the faulty condition might be much less than the relay threshold.

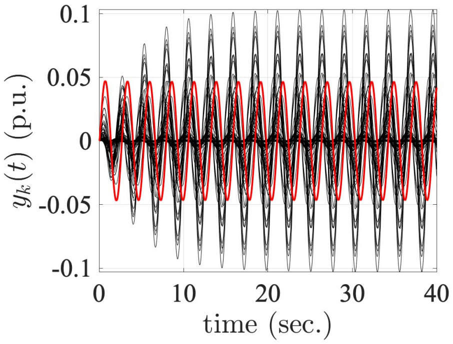


Fig. 11. ***Voltage deviations in a counterintuitive case: the red curve is the voltage deviation closest to the oscillation source; the black curves are other voltage deviation measurements. (Source: Fig. 1 of [201] © IEEE 2021.)***

In order to address the protection challenges in a renewable-rich distribution system, Wu *et al.* [204] place the protection problem into an RL framework (see Fig. 14) in which the protection scheme is learned by interacting with a distribution system simulator. In the RL framework, the distribution system is modeled by an MDP described by states $s \in \mathcal{S}$, actions $a \in \mathcal{A}$, a reward function $r(s,a)$, transition probability P, and a user-defined discount factor $\beta \in (0,1]$. The implication of the states, action, and reward function in the protection problem is annotated in Fig. 13. In particular, the state $s_{i,t}$ and action $a_{i,t}$ of relay i at time t are defined by

$$s_{i,t} = \{s^c_{i,t}, s^b_{i,t}, s^d_{i,t}\} \tag{6a}$$
$$a_{i,t} = \{a^{\text{set}}_{i,t}, a^d_{i,t}, a^{\text{reset}}_{i,t}\} \tag{6b}$$

where $s^c_{i,t}$ represents local current measurements, $s^b_{i,t}$ represents the status of the local breaker, $s^c_{i,t}$ represents the value of the countdown timer, $a^{\text{set}}_{i,t}$ represents the action of triggering the countdown timer, $a^d_{i,t}$ represents the action of decreasing the value of the counter by one, and $a^{\text{reset}}_{i,t}$ represents the action of resetting the counter. The reward function gives deterministic positive rewards to the tripping action under fault conditions and stay-in-silence action under normal condition, and it gives negative rewards to malfunctions. The transition probability is determined by the distribution system; in practice, it is unknown. The optimal action $a^*(s)$ at state s is obtained by

$$Q(s,a) = \mathbb{E}\left(r(s,a) + \beta \max_{a' \in \mathcal{A}} Q(s',a')\right) \tag{7a}$$
$$a^*(s) = \arg\max_{a' \in \mathcal{A}} Q(s,a') \tag{7b}$$

where $\mathbb{E}(\cdot)$ is the expectation operator, a' is the possible next-step action, and s' is the next-step state given the current state and action; it is determined by the distribution system. In [204], the Q function in (7) is approximated by an NN. The NN's parameters are learned by a sequence of $\{s,a,r,s'\}$ observations from the framework shown in Fig. 14. The dataset reported in [205] can be used for training the algorithm. The simulation results in [204] suggest that the failure rate of the RL-based relay is only 0.32% in a distribution system with 30% DER penetration, whereas the conventional overcurrent relay has a much higher failure rate, i.e., 15.46%, under the same condition. One future direction of this work is to investigate a rigorous convergence guarantee for the sequential RL algorithm [204].

To summarize this section, we provide twofold guidance on applying use-inspired AI methods in power systems. First, it is critical to find appropriate application scenarios that take precedence over proposing innovative methodology. With deep NNs as representatives, current AI techniques that are essentially model-agnostic function approximators usually present outperforming performance in application scenarios where there is only heuristic experience with no clear first-principle physical model, such as in load and renewable prediction. The illustrated NN-based Lyapunov function [164] is another example. Although a Lyapunov function itself has a rigorous definition, there is no traditional cost-effective analytical or numerical way to construct such a function for a large-scale real-world dynamical system, in which NNs can provide an alternative effective solution. Second, it is desirable to intelligently and insightfully formulate critical challenges in traditional power systems into AI-friendly formats. Consider illustrated forced oscillation source localization [201] as one example. Intuitively, it can be formulated as a typical classification problem by taking system global states as inputs and discrete location labels as outputs. However, formulated as a matrix decomposition problem, this problem can be solved by RPCA that is commonly used for image processing, which has both outperforming accuracy and explainability.

VI. USE CASES OF INDUSTRY ADOPTION

As more measurement data and data-driven algorithms become available, the power industry continues to adapt and improve operations by leveraging new technology and systems that enable it to meet and exceed customer expectations. This section presents some industry use cases to illustrate the continuing adoption of machine learning techniques by Oncor, a regulated utility that operates the largest distribution and transmission system in Texas. The following use cases were selected to show instances of AI adoption with relatively high maturity. In addition, we illustrate use cases (e.g., asset management) that are not considered in power-systems research but are essential for business operations with physical devices spread over large distances. All use case developments are based on business needs, and the value of the investment must be justified before a use case is developed, even if data are

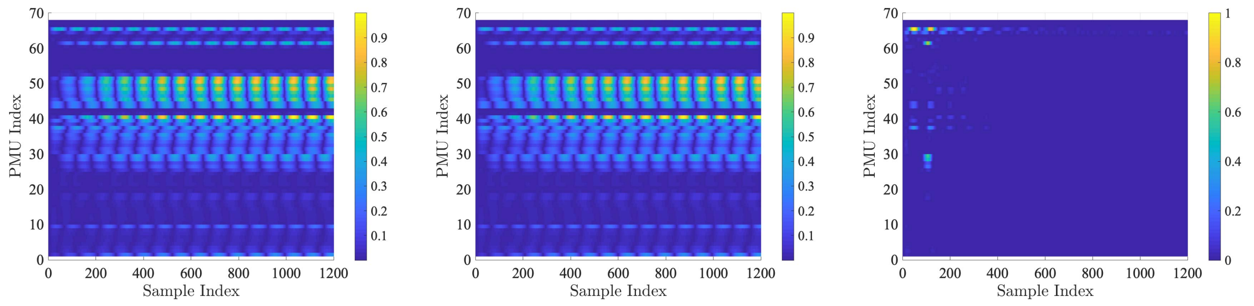


Fig. 12. ***Illustration of the RPCA-based source localization algorithm: (a) measurement matrix, (b) low-rank matrix, and (c) sparse matrix. The (normalized) magnitudes of matrix entries are color-coded. The measurement closest to the source can be tracked by identifying the largest absolute entry in the sparse matrix, i.e., the entry with the brightest color. (Source: Fig. 2 of [201] © IEEE 2021.)***

readily available. Moreover, the value-add of some high-performance algorithms in many cases may not offset the maintenance cost required to keep such models operating properly (e.g., due to model drift). Table 6 provides a brief introduction to the industry use cases that will be described in detail. Because some use cases involve proprietary information, details about preprocessing and postprocessing steps, and model accuracy level will not be disclosed.

In many industry use-cases, the methods currently used may appear simplistic compared to the latest research; however, these use-cases are of high value, and large amounts of data are readily available. Utilities usually have multiple databases for various systems, such as outage management, advanced metering, work orders, geographical and meteorological data, and financial info. An essential challenge for conducting any big data analysis is to unify this data and enforce consistent formats for each data type. At Oncor, a datalake was created to consolidate the data needed for analytics. The datalake replicates data from all of Oncor's operational databases. In addition to supporting uniformity, this approach also minimizes stress on operational databases because they are accessed only during each scheduled copy rather than whenever an analyst makes a query.

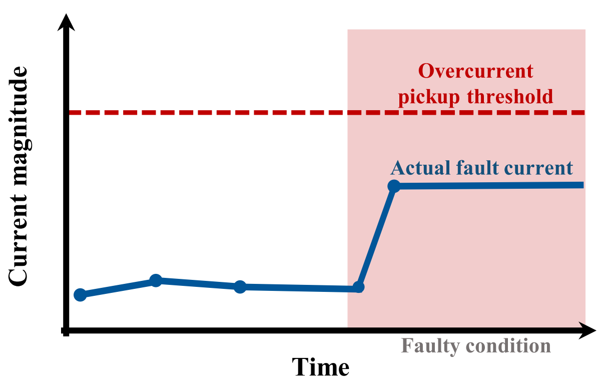


Fig. 13. ***Conventional threshold-based protection scheme may fail due to low fault current. (Modified from source: Fig. 1 of [204].)***

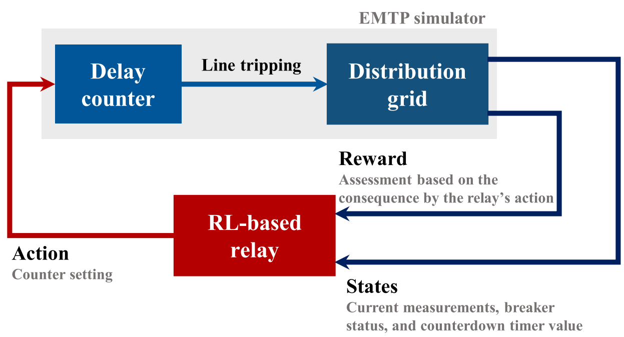


Fig. 14. ***Obtaining the Q function for the RL-based relay: the optimal policy embedded in the RL-based relay is obtained by interacting with a distribution system simulator.***

As the industry continues to adopt machine learning continues, and available platforms become more mature, advanced techniques will be more feasible at a lower cost; these will be necessary to address more complex problems in power systems. Most importantly, the collaboration between practitioners and researchers must intensify to achieve efficient and continuous adoption.

A. Asset Health

For all utility companies, monitoring and maintaining their assets are critical to realizing system reliability and providing the highest quality service to their customers. Some assets, such as distribution class transformers, can be monitored by utilizing AMI meter data, such as voltage and kWh readings. For assets where digital measurements are not available, health monitoring may be possible by analyzing asset images using advanced image processing techniques. Several Oncor use cases are presented below to illustrate how asset health can be monitored by utilizing machine learning methods.

As the largest utility company in the state of Texas, Oncor provides power to nearly four million customers

Table 6 Industry Data-Driven Use Cases

Application	Algorithm			Data				Computing Platform
	Main method	Type	Accuracy	Source	Resolution	# Sample	Dimension	
Transformer voltage anomaly classification	Change point detection	Offline	94%	AMI	15min	$3.7 * 10^6$	672×1	Spark*
Asset defect detection	YOLOv3	Offline	$\geq 87\%$	Aerial images	-	500	608×608	Tesla V100 GPU
Short-/Mid-term load forecast	Regression trees	Offline	-	AMI	15min	$3.7 * 10^6$	$4.1 * 10^4 \times 1$	Spark*
Cold load characterization	Linear regression	Offline	-	SCADA	1min	1300	Various	-
RIC Categorization	Cluster analysis	Offline	-	AMI	15min	4.9×10^5	$8,064 \times 1$	Spark*

*Spark: 2 namenodes, dual Xeon-4208 (8-core); 8 datanodes, dual Xeon-5218 (16-core); 768 GB RAM per node.

through more than one million distribution class transformers, which can fail from damaged coils or overload degradation. Reactive replacement of a failed transformer can take more than 4 h, but proactive replacements often take less than 1 h. Thus, detecting failure precursors can significantly reduce both labor costs and outage time.

Fig. 15 shows a plot of the voltage and load measurements from a single-phase 240-V AMI meter. Both voltage "V1" (in Volts) and load "LOAD" (in kWh) time series, in red and gray, respectively, have a 15-min resolution. The two horizontal lines are the upper and lower limits of the operating voltage ratings defined by the American National Standards Institute (ANSI C84.1-2020), which are ±5% of the nominal voltage. On June 24, 2018, the voltage suddenly rose above the upper limit due to a damaged coil on the primary side of the transformer. The sudden drop in voltage on July 18, 2018, denotes the time of the replacement. Typically, a transformer will not fail immediately after a coil is damaged. Therefore, proactive replacement is realistic and valuable if a change in voltage can be detected soon enough.

After examining the preoutage voltage profiles of all transformers replaced in Oncor's system during an 18-month period, a change point detection algorithm was designed to detect over/under voltage issues. A change in mean and/or variance of a meter's voltage was detected by a PySpark implementation of the functions provided in [206]. Several postprocessing steps were implemented to remove change points due to outages or temporary voltage changes. The thresholds for these steps were selected from the ground-truth data. Based on the number of issues

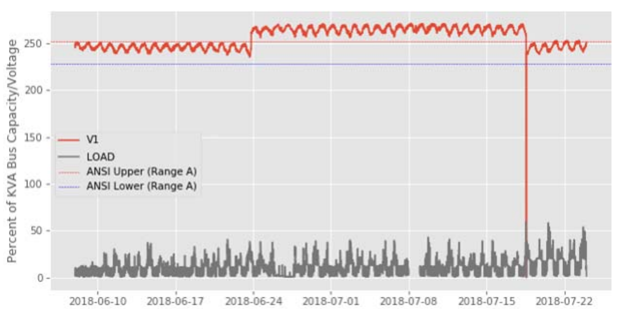


Fig. 15. ***Voltage profile of a transformer with a coil damage and subsequent replacement. The voltage profile is in red, and the load profile is in gray.***

seen on the same feeder, the detected issues were then categorized into various types, such as meter, transformer, or regulation issues, to enhance the troubleshooting process of the distribution operations organization. The algorithm and thresholds were tuned and improved using feedback received from the field. Currently, the voltage monitoring process runs every weekday on data from 3.7 million AMI meters. The weekly average accuracy for June–November 2021 is 94%.

Oncor began to monitor distribution transformer health in 2016. As of November 2021, 3834 issues have been resolved proactively using transformer health analysis. These issues include damaged transformers or meters, as well as installation, regulation, and secondary issues that affect voltage measurements. Proactive transformer maintenance has saved Oncor approximately $3.25 million in equipment, labor, and expenses, as well as 5.5 million customer interruption minutes.

Another asset health use case is defective insulator detection, due to, for example, lightning strikes, forceful impacts, or aging. Defective insulators are hazardous to the operation of power lines and pose a risk to system reliability. Oncor has more than 18 000 circuit miles of transmission lines with over 500 000 transmission insulators. Rapid identification of damaged insulators, especially after a storm, is, therefore, a critical task in asset management. Due to the scale of Oncor's transmission system, manual inspection is infeasible. An automated inspection method was developed, which uses aerial/drone images of transmission lines and convolutional NNs.

The insulator defect detection method employs You Only Look Once, Version 3 (YOLOv3 [9]), which is a real-time object detection model that uses Darknet-53 [207] as the backbone feature extractor in a deep convolutional NN. The model was initialized with YOLO's pretrained weights using the Microsoft Common Objects in Context (COCO) dataset [208], and insulator images, provided by the Electric Power Research Institute (EPRI), were used for transfer learning and validation (confidential data).

The defect detector successfully recognized the insulators in an image, pinpointed those issues of each damaged insulator, and classified the issues as either "broken" or "flashed." For the 50 testing images, each containing multiple flashed/broken locations, 100% of the broken points were detected correctly, and 90% of the flashed points were detected. There were no misclassified issues.

The recent Texas House Bill 4150, also known as the "William Thomas Heath Power Line Safety Act," which was passed through the Legislature in May 2019, requires all utilities to make regular inspections of their power lines to ensure that they comply with state and federal safety regulations. Although Oncor completes routine inspections of all transmission power lines, detailed manual inspections of all structures are time-consuming and impactful to land owners and costly. In an effort to reduce resources, such as on right-of-way truck traffic, another deep model was trained and applied to aerial images of the power lines. This model is being developed in stages to ultimately identify reliability risks due to structures damaged by impacts or aging. The first stage of this model requires Oncor to verify all structure asset information in the Oncor Transmission Information System. Because many transmission lines are 40+ years old, information in historical records may be inaccurate for structures where components were replaced or added after the initial installation. Additional stages include identifying attributes that can indicate structural issues that may cause outages and affect reliability performance. These attributes include the following.

1) *Composition:* Wood, steel, and concrete.
2) *Design:* H-frame, A-frame, lattice tower, multipole, and single-pole.
3) *Cross arm:* Beam and double-plank.
4) *Brace:* V, X, and knee.

The effort to classify transmission line attributes made use of YOLOv3; the initial results were promising, with accuracy rates of 89% for braces and 87% for cross arms. Figs. 16 and 17 show several examples of successful classification results. As more images are labeled to augment training data, the model's performance is expected to improve; furthermore, by including images with defective structures, the system can be used to inventory components and their degradation levels.

B. Load Forecasting

Load forecasting is an essential building block in operating and planning tasks in both the power industry [209] and commercial building energy [210]. It is needed in many decision-making processes for electric energy generation, DERs management, transmission, distribution, markets, and demand–response. The pursuit of models

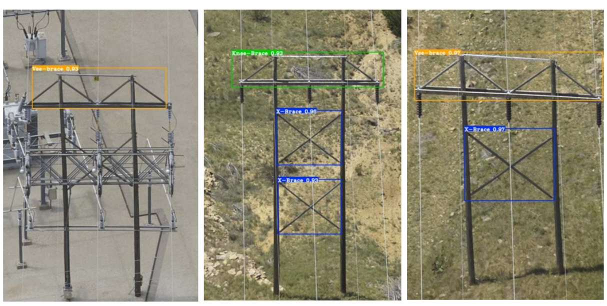


Fig. 16. ***Examples for brace type classification. Yellow boxes: V brace; green box: knee brace; and blue boxes: X brace.***

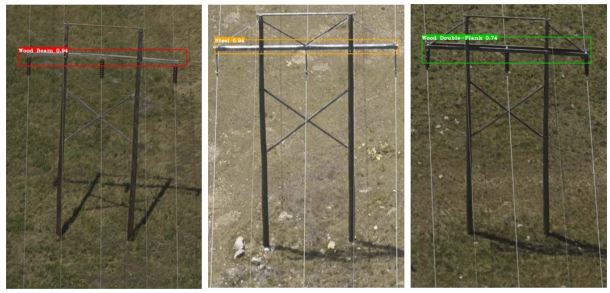


Fig. 17. ***Examples for beam type classification. Red box: wood beam; yellow box: steel beam; and green box: wood double-plank.***

that can achieve accurate load forecasts for short-, mid-, and/or long-term purposes is a long-standing research area with a large body of the literature [211], [212].

For utility companies, short- and mid-term load forecasts are used to plan switching operations in control centers. Moreover, load forecasts contribute to network reconfiguration and infrastructure development/improvement decisions. For example, to better prepare for high-power-demand seasons, Oncor conducts load analyses to forecast summer and winter feeder load peaks. In some cases, a contingency plan will be made ahead of these peak seasons for feeders that are at risk of overload based on historical load data leveraged by analytics.

These efforts have significantly improved Oncor's reliability performance; there has not been a feeder lockout event due to overload since 2018. Switching operations, however, are a major challenge for feeder load forecasting because a feeder's load can change significantly due to a load switching event (e.g., feeder reconfiguration due to an outage or planned maintenance). A robust model is needed to respond to these events quickly and adjust the forecasts correspondingly. Oncor currently is developing deep learning methods to surpass the performance of the current approach.

Besides feeder load forecasts, load forecasting at any device is needed for making operational decisions in the control rooms. One approach is to forecast the load at each distribution transformer using AMI meter data and then aggregate it at each device as needed. With a large number of distribution transformers (e.g., more than one million in Oncor's system), if computational power is limited, cluster analysis can be used to group transformers with similar load behaviors. Normalization (rescaling each load profile to range $[0, 1]$) is needed before clustering so that the clustering results are affected mainly by the shape of the load profiles. After the transformers have been assigned into clusters, load forecasts for each cluster center (the representative of all transformers in that cluster) can be obtained; they are then scaled back to each transformer's load level by undoing the normalization steps. If distributed computing platforms are available, transformer load forecasting can be conducted by directly

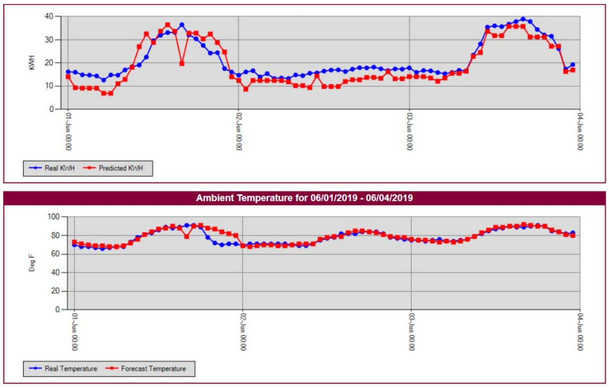


Fig. 18. ***Example of transformer load forecast results. Blue dots: actual measurements; red dots: predicted values. Top: predicted and actual loads. Bottom: predicted and actual temperatures.***

training individual models for every transformer, which will introduce fewer errors.

Oncor implemented a regression tree model [213] on Spark that serves both short- and mid-term needs. The load of a transformer is affected by both numerical and categorical factors. The most important numerical factors include temperature, wind speed, humidity, and solar radiation, whereas categorical factors include the time of day, day of week, month, and so on. To avoid overfitting, the maximum numbers of layers and leaves were tuned based on model performance.

Fig. 18 shows an example of the hourly load forecasting results for one distribution transformer over the course of three days. The blue and red curves on the top plot give the actual and predicted load based on the predicted temperatures in the bottom plot (blue curve) using a regression tree model trained for a particular transformer. There is a tradeoff between model performance (error level) and computing time, which can be calibrated to suit shifting business needs at any given time. This approach is able to capture nonperiodic activity that sometimes deviates from the temperature, as seen on Day 2 in Fig. 18. The accuracy of load forecasts is highly dependent on the accuracy of weather forecasts, which utility companies usually obtain from vendors. The uncertainty in the exogenous factors must be accounted for in the final forecast, and because several of those factors are forecasts themselves, errors can be large. In this case, the model's performance is sufficient to add value to business operations at normal operating levels and in typical seasonal weather. The accuracy will be reduced during the time of extreme cold or heat due to the lack of historical meter data.

A special case in load forecasting is cold load characterization. During a steady state, the heating or cooling load on a feeder is typically a smaller percentage of the total heating or cooling load. This reduced load results from the diversity of HVAC units simultaneously running due to normal cycling between on and off. After an extended outage, the temperature in the residence will likely fall outside the setpoint range. Once the power is restored to the feeder, the diversity of the heating or cooling load would be lost due to all the units turning on at the same time. This increase in load is referred to as "cold load." After some time period passes, the diversity will be restored because the unit run times will vary depending on factors such as HVAC rating, home size, and temperature setpoints. Cold load peak values are affected by preoutage load behavior, season (winter/summer), time of day, ambient temperature, and load composition (customer types). Predicting these values at feeder breakers or other downstream protective devices enables optimal sequencing of operations to restore power quickly while minimizing the likelihood of damaging equipment. In addition, EMS typically has a load shed/restoration tool that can automatically conduct outage rotations among all feeders in the system during a short supply situation, such as the recent Texas power crisis [70]. With predictions of each feeder's postoutage load peaks, the EMS can automatically and accurately follow ISO's load-shed requirements to protect the entire power grid.

Oncor is currently testing a linear regression model to predict the ratio of the peak cold load (postoutage) and preoutage load of a feeder. The data used are outage duration, preoutage and postoutage temperatures, and the fraction of residential customers on the feeder. The residential load fraction is a good proxy for feeder load diversity (i.e., the independently controlled cyclic loads, such as HVAC systems that may be energized at any given time during normal operating conditions). Since feeder breaker level outages are relatively rare, feeders are grouped by their residential fractions, and a model is learned for each feeder group. A total of 1127 breakers were evaluated, and training data were collected for fitting the regression model. To accurately capture the cold load behavior, switch operation logs and fuse level events were reviewed to ensure that the cold load peaks were neither overestimated due to switching operations nor underestimated due to fuse level events behind the breakers. During an emergency situation, this model will take the preselected outage durations for feeder rotations and postoutage temperatures as inputs. The model will output a predicted load ratio for each (phase) feeder and the power ratio, and then, the cold-load peaks can be estimated. These four predictions are useful for unbalanced feeders; in balanced feeders, a single estimate of the power ratio is sufficient.

Fig. 19 shows an example of the cold load peak prediction for one feeder using the trained regression model. The two highlighted points in the figure mark the preoutage current and predicted postoutage current for one phase of a feeder. The predicted value is marked at the same location as the postoutage load peak only for better visualization and easier comparison.

C. Residential/Industrial/Commercial Categorization

For many transmission and distribution planning models, RIC percentages at each substation transformer bank

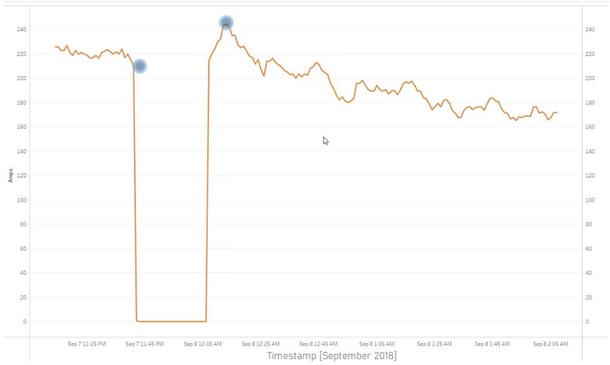


Fig. 19. ***Example of the cold load peek forecasting result for one feeder. Orange curve: SCADA current time series data before and after a feeder level outage. Gray points: preoutage current reading and predicted postoutage current.***

are used to allocate load in the base-case models. These percentages are also used to derive the number of various motor types for dynamics models and simulations. Likewise, distribution planners must sometimes perform weather corrections for load projections. In these cases, industrial and other nonweather-sensitive loads (such as water pumping and/or oil field pumping loads) are not weather-corrected because these load types are rarely weather-sensitive or weather-dependent. Traditionally, the customer category of a premise is established at the creation of the premise and may not get updated when the customer type changes. For example, a commercial building can be leased to a new business that has a completely different load profile from that of the previous business, but the utility may not be aware of the change.

Before the system-wide installation of advanced meters, the RIC process used typical summer and winter hourly load profiles for each category of the building distribution feeder models. With the availability of AMI interval data and distributed computing, the process can be improved by directly analyzing the load profile of each premise. Because residential meters can usually be identified using the information provided by ISOs, it is more valuable to focus on nonresidential meters.

As an initial approach, cluster analysis was conducted on data from approximately 490 000 nonresidential meters. Domain experts selected 12 weeks (noncontiguous) over a one-year period that adequately covered different seasonal and holiday effects (e.g., extended hours during holidays). The 15-min interval load data were collected from each week, and the time series for each meter was stacked into 8064-D vectors ($12 \times 7 \times 24 \times 4$). K-means clustering was applied to the data with, initially, $k = 100$. The initial parameter values were chosen as subject matter experts' estimations. Subsequently, large clusters were checked by comparison of random samples within the cluster to the cluster center (i.e., comparing the average load profile with the other profiles within the cluster). If a large deviation was found, then the cluster was split. A less heuristic approach would be to use the V-measure or silhouette-coefficient to determine an optimal number of clusters [214], [215]; however, cluster splitting was found to be effective for this use case.

This cluster analysis was conducted using Spark; 3–4 h were needed for a cluster with two name nodes (dual Xeon-4208, 768-GB RAM per node) and eight data nodes (dual Xeon-5218, 768-GB RAM per node), as shown in Table 6. The analysis will be repeated annually to capture any premises with changes in load type.

D. Other Use Cases

Many companies in the power industry have been developing data-driven methods for their business needs. Exelon Utility and ComEd applied classification methods to aerial/satellite images and light detection and ranging (LiDAR) data for vegetation management to better understand the system's tree trimming workload in the system seeking to cut rimming costs while reducing the number of tree-related outages [216], [217]. ISO New England proposed a prediction method based on a decision tree to instruct interface limit values for different operating conditions [218]. Researchers in Hitachi proposed a three-layer wind power prediction model based on the data from historical power measurements and numerical weather prediction tools [219]. In addition, Bhattarai *et al.* [56] reviewed related literature on big data analytics from the perspectives of electric utilities and industry.

VII. CONCLUSION AND OUTLOOK

In this article, we have briefly reviewed the structure of power system physical and market operation, today's AI infrastructure of data acquisition and computation in power systems, state-of-the-art AI-based approaches for multiple critical functions, and industrial use cases of AI methods. In the following, we propose several research directions from the aspects of data, computing, and AI algorithms.

A. High-Quality Open-Source Datasets

Despite the advances in data acquisition, in contrast to numerous datasets that have benefited broad AI communities, the lack of publicly accessible high-quality power datasets may be impeding the advancement of AI research in power systems. There are several reasons for the limited public access to power datasets. First, most real-world operational data are protected by policies such as CEII in the interest of confidentiality. Second, due to the reliability of real-world power grids, the rarity of opportunities to observe high-impact events may produce an insufficiently robust real-world measurement dataset. Third, the value of creating comprehensive and trustworthy benchmark power datasets has been overlooked by the power system community. There have been few open-source datasets [220], [221] and online contests dedicated to topics such as forced oscillation localization [222] and power system operation [223], [224]. However, far more will be needed to build a standard library of open-source benchmark datasets along with critical tasks in a clear mathematical formulation that can be used to train, calibrate, test, and benchmark data-driven models. One critical challenge is that commonly used random sampling and data generation methods do not guarantee representativeness [44] and may introduce unexpected biases into subsequent data-drive methods. Therefore, it is critical to investigate data generation methods that guarantee comprehensiveness and representativeness; these may be dataset-inspired or task-tailored. In the meantime, it is also necessary to propose algorithm-agnostic metrics to consistently assess the property of representativeness.

B. Advanced Computing

As mentioned in Section II, complex control algorithms are too time-consuming for real-time security control, especially in contingency scenarios. The rapid expansion of sensors has enabled massive data acquisition; however, although these data are necessary for realizing a digitized power grid, using all of it is beyond the current computing capacity for centralized methods. Therefore, to explore and exploit advanced algorithms and massive streaming data, hybrid edge and cloud computing are necessary to dynamically balance the computational load and escalate computing power as needed. For example, edge devices can compute partial results across several hundred sensors (e.g., half of NN's layers) and forward the results to the control center for final computations, effectively distributing the computational load. Furthermore, new ASIC devices, dedicated to power system computations, could be used in edge devices for real-time data processing and to accelerate simulations. In addition, communications between edge and cloud may contain sensitive information, requiring privacy-preserving methods, such as federated learning [110].

Besides accelerating computation, platforms are needed to manage the complexity introduced by digitization. The software development industry uses a set of (automation) practices called "DevOps" to manage the development, integration, testing, deployment, and monitoring of distributed software systems. In sectors where data-driven and machine learning algorithms are used, another layer is added to DevOps [225], [226] that encompasses automated training, testing, deployment, and monitoring of models—this is called "MLOps" [227], [228]. Both DevOps and MLOps lower the maintenance cost of complex software systems through automation, but the initial investment is high. For efficient digitization of the power grid, both DevOps and MLOps will be necessary; however, there are unique aspects of power systems that require investigation. Because the grid is primarily hardware, it would be highly imprudent to blindly adopt methods developed for pure software environments.

The instrumentation and sensors being deployed into modern grids also bring cybersecurity challenges. If the data and controls are transmitted over the internet (e.g., cloud computing), the grid is vulnerable to the same cyberattacks as a website, except the stakes are much higher: outages, energy theft, and loss of private data. Monitoring and detecting cyberthreats to the grid are important areas for cross-disciplinary research combining power systems, cybersecurity, and AI.

C. Use-Inspired AI Methods for Practical Applications

Because power grids are large-scale critical infrastructure systems for human society, future research efforts ought to use-inspired AI algorithms that possess three key properties, namely, interpretability, robustness, and scalability, aiming to facilitate practical applications. First, AI algorithms ought to be explainable by first-principle-based physical models because only interpretable algorithms are acceptable for participation in the human-in-the-loop decision-making process. In particular, interpretable AI approaches should provide clear causal inference for the purposes of real-time monitoring, control, and diagnosis, such as identifying the root cause of complex observations. Preliminary efforts have been devoted to physics-informed ML, as summarized in [229]. The principle is to steer the learning process toward identifying physically consistent solutions, of which instructive guidance contains three aspects, namely, data processing, loss function modification, and model architecture design. For example, incorporating ordinary different equation (ODE) formats into loss function as regularization terms can improve the performance of system identification algorithms based on transient data or improve the fidelity of transient data generation methods. Second, AI algorithms must have performance guarantees extending beyond the basic, unperturbed scenarios. Particularly, the robustness to perturbation is critically important for RL-based algorithms for decision-making. Meta-RL [230], [231] and transfer learning can potentially accommodate the gap between reality and simulation environment, thereby rendering the decision-making adaptive to varying conditions and scenarios. Third, another highly desirable feature of AI algorithms is scalability, which refers to adequate effectiveness and efficiency in large-scale real-world systems. The concern regarding scalability arises from the aforementioned observation that the performance of existing AI algorithms in the power system domain is mostly demonstrated by small-scale grids without validation in large-scale cases. As high-dimensional measurements in power systems empirically have properties such as approximate low-rankness and sparsity, they may be potentially efficacious to discover intrinsic low-dimensional manifolds and linear coordinates in data structure [232].

In summary, digitization of the power grid will play a major role in transforming the electricity sector into a decarbonized system while simultaneously improving grid reliability. The synergy of high-dimensional dynamic data, increased computing power, and use-inspired AI algorithms will enable improvements to the reliability and operational efficiency of the power grid at multiple scales. Challenges remain in the integration of heterogenous datasets, cyber–physical security, and the development of robust, interpretable AI algorithms. Strong collaboration between industry and academia will be crucial for the successful adoption of use-inspired AI methods in a decarbonized power system. ■

Acknowledgment

The authors sincerely thank Jimmy Liu, Steven Dennis, and Thomas Wilson for their help on the Oncor use cases presented in this article.

REFERENCES

[1] U. S. Energy Information Administration. (2021). *U.S. Electricity Customers Experienced Eight Hours of Power Interruptions in 2020*. [Online]. Available: https://www.eia.gov/todayinenergy/detail.php?id=50316

[2] L. Thomas. (2021). *U.S. vs U.K. Wholesale Electricity Market Outlook*. [Online]. Available: https://www.ussolarfund.co.uk/content/us-vs-uk-wholesale-electricity-market-outlook

[3] U.S. Energy Information Administration. (2021). *How Much Carbon Dioxide is Produced Per Kilowatthour of U.S. Electricity Generation?* [Online]. Available: https://www.eia.gov/tools/faqs/faq.php?id=74&t=11#:~:text=In%202020%2C%20total%20U.S.%20electricity,CO2%20emissions%20per%20kWh

[4] J. Rogelj *et al.*, "Mitigation pathways compatible with 1.5°C in the context of sustainable development," in *Global warming of 1.5°C*. Intergovernmental Panel on Climate Change, 2018, pp. 93–174.

[5] M. L. Littman *et al.*, "Gathering strength, gathering storms: The one hundred year study on artificial intelligence (AI100) 2021 study panel report," 2021. [Online]. Available: http://ai100.stanford.edu/2021-report

[6] X. Qiu, T. Sun, Y. Xu, Y. Shao, N. Dai, and X. Huang, "Pre-trained models for natural language processing: A survey," *Sci. China Technol. Sci.*, vol. 63, no. 10, pp. 1872–1897, 2020.

[7] A. B. Nassif, I. Shahin, I. Attili, M. Azzeh, and K. Shaalan, "Speech recognition using deep neural networks: A systematic review," *IEEE Access*, vol. 7, pp. 19143–19165, 2019.

[8] Rec Faces. (2020). *What is Facial Recognition Used for?* [Online]. Available: https://recfaces.com/articles/what-is-facial-recognition-used-for

[9] J. Redmon and A. Farhadi, "YOLOv3: An incremental improvement," 2018, *arXiv:1804.02767*.

[10] O. Vinyals *et al.*, "Grandmaster level in starcraft ii using multi-agent reinforcement learning," *Nature*, vol. 575, no. 7782, pp. 350–354, 2019.

[11] D. Silver *et al.*, "Mastering the game of go without human knowledge," *Nature*, vol. 550, pp. 354–359, Oct. 2017.

[12] D. Silver *et al.*, "A general reinforcement learning algorithm that masters chess, shogi, and go through self-play," *Science*, vol. 362, no. 6419, pp. 1140–1144, 2018.

[13] Q. Zhang, J. Lu, and Y. Jin, "Artificial intelligence in recommender systems," *Complex Intell. Syst.*, vol. 7, no. 1, pp. 439–457, Feb. 2021.

[14] Boston Dynamics. (2021). *Spot*. [Online]. Available: https://www.bostondynamics.com/products/spot

[15] Boston Dynamics. (2021). *Atlas*. [Online]. Available: https://www.bostondynamics.com/atlas

[16] Agility Robotics. (2021). *Cassie*. [Online]. Available: https://www.agilityrobotics.com/robots#digit

[17] Tesla. (2021). *AutoPilot*. [Online]. Available: https://www.tesla.com/autopilot

[18] P. Rajpurkar *et al.*, "CheXNet: Radiologist-level pneumonia detection on chest X-rays with deep learning," 2017, *arXiv:1711.05225*.

[19] D. Paul, G. Sanap, S. Shenoy, D. Kalyane, K. Kalia, and R. K. Tekade, "Artificial intelligence in drug discovery and development," *Drug Discovery Today*, vol. 26, no. 1, pp. 80–93, Jan. 2021.

[20] F. Wu *et al.*, "Towards a new generation of artificial intelligence in China," *Nature Mach. Intell.*, vol. 2, no. 6, pp. 312–316, Jun. 2020.

[21] *Interpretable AI by Dimitris Bertsimas, MIT Plus Opening of OR62 Conference*. Accessed: Nov. 25, 2021. [Online]. Available: https://www.youtube.com/watch?v=gAZ4YRngEj0

[22] D. Amodei, D. Hernandez, G. Sastry, J. Clark, G. Brockman, and I. Sutskever. (2018). *AI and Compute*. [Online]. Available: https://openai.com/blog/ai-and-compute/

[23] L. Duchesne, E. Karangelos, and L. Wehenkel, "Recent developments in machine learning for energy systems reliability management," *Proc. IEEE*, vol. 108, no. 9, pp. 1656–1676, Sep. 2020.

[24] A. Oussous, F. Z. Benjelloun, A. A. Lahcen, and S. Belfkih, "Big data technologies: A survey," *J. King Saud Univ. Comput. Inf. Sci.*, vol. 30, no. 4, pp. 431–448, Oct. 2018.

[25] A. Z. Faroukhi, I. El Alaoui, Y. Gahi, and A. Amine, "Big data monetization throughout big data value chain: A comprehensive review," *J. Big Data*, vol. 7, no. 1, pp. 1–22, Dec. 2020.

[26] A. Agrahari and D. Rao, "A review paper on big data: Technologies, tools and trends," *Int. Res. J. Eng. Technol.*, vol. 4, no. 10, p. 10, Oct. 2017.

[27] C. E. Leiserson *et al.*, "There's plenty of room at the top: What will drive computer performance after Moore's law?" *Science*, vol. 368, no. 6495, Jun. 2020.

[28] Z. Yang, Z. Dai, Y. Yang, J. Carbonell, R. R. Salakhutdinov, and Q. V. Le, "XLNet: Generalized autoregressive pretraining for language understanding," in *Proc. Adv. Neural Inf. Process. Syst.*, vol. 32, 2019, pp. 1–11.

[29] J. Devlin, M.-W. Chang, K. Lee, and K. Toutanova, "BERT: Pre-training of deep bidirectional transformers for language understanding," 2018, *arXiv:1810.04805*.

[30] C. Szegedy, S. Ioffe, V. Vanhoucke, and A. A. Alemi, "Inception-v4, inception-resNet and the impact of residual connections on learning," in *Proc. 21st AAAI Conf. Artif. Intell.*, Feb. 2017, pp. 4278–4284.

[31] A. Radford, K. Narasimhan, T. Salimans, and I. Sutskever, "Improving language understanding by generative pre-training," 2018. [Online]. Available: https://www.cs.ubc.ca/~amuham01/LING530/papers/radford2018improving.pdf

[32] A. Krizhevsky, I. Sutskever, and G. E. Hinton, "ImageNet classification with deep convolutional neural networks," in *Proc. Adv. Neural Inf. Process. Syst. (NIPS)*, vol. 25. Stateline, NV, USA, Dec. 2012, pp. 1097–1105.

[33] K. Simonyan and A. Zisserman, "Very deep convolutional networks for large-scale image recognition," 2014, *arXiv:1409.1556*.

[34] K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jun. 2016, pp. 770–778.

[35] A. Vaswani *et al.*, "Attention is all you need," in *Proc. Adv. Neural Inf. Process. Syst.*, 2017, pp. 5998–6008.

[36] J. D. McDonald, *The Digitized Grid*. IEEE, 2019.

[37] F. F. Wu, K. Moslehi, and A. Bose, "Power system control centers: Past, present, and future," *Proc. IEEE*, vol. 93, no. 11, pp. 1890–1908, Nov. 2005.

[38] D. Apostolopoulou, P. W. Sauer, and A. D. Dominguez-Garcia, "Automatic generation control and its implementation in real time," in *Proc. 47th Hawaii Int. Conf. Syst. Sci.*, Jan. 2014, pp. 2444–2452.

[39] W. R. Cassel, "Distribution management systems: Functions and payback," *IEEE Trans. Power Syst.*, vol. 8, no. 3, pp. 796–801, Aug. 1993.

[40] K. Morison, L. Wang, and P. Kundur, "Power system security assessment," *IEEE Power Energy Mag.*, vol. 2, no. 5, pp. 30–39, Sep./Oct. 2004.

[41] R. N. Allan *et al.*, *Reliability Evaluation of Power Systems*. Berlin, Germany, Springer, 2013.

[42] S. Stoft, *Power System Economics: Designing Markets for Electricity*, vol. 468. Hoboken, NJ, USA: Wiley, 2002.

[43] R. Haider, S. Baros, Y. Wasa, J. Romvary, K. Uchida, and A. M. Annaswamy, "Toward a retail market for distribution grids," *IEEE Trans. Smart Grid*, vol. 11, no. 6, pp. 4891–4905, Nov. 2020.

[44] P. L. Joskow, "Challenges for wholesale electricity markets with intermittent renewable generation at scale: The U.S. experience," *Oxford Rev. Econ. Policy*, vol. 35, no. 2, pp. 291–331, Apr. 2019.

[45] Federal Energy Regulatory Commission, North American Electric Reliability Corporation, and Regional Entities, "The February 2021 cold weather outages in Texas and the south central United States," 2021. [Online]. Available: https://www.ferc.gov/media/february-2021-cold-weather-outages-texas-and-south-central-united-states-ferc-nerc-and

[46] E. A. Feinberg and D. Genethliou, "Load forecasting," in *Applied Mathematics for Restructured Electric Power Systems* (Power Electronics and Power Systems). Cham, Switzerland: Springer, 2005, pp. 269–285.

[47] National Research Council *et al.*, *Frontiers in Massive Data Analysis*. Washington, DC, USA: National Academies Press, 2013.

[48] M. Hilbert, "Big data for development: A review of promises and challenges," *Develop. Policy Rev.*, vol. 34, pp. 135–174, Jan. 2016.

[49] D. V. Dimitrov, "Medical Internet of Things and big data in healthcare," *Healthcare Inform. Res.*, vol. 22, no. 3, pp. 156–163, 2016.

[50] I. Pencheva, M. Esteve, and S. J. Mikhaylov, "Big data and AI—A transformational shift for government: So, what next for research?" *Public Policy Admin.*, vol. 35, no. 1, pp. 24–44, Jan. 2020.

[51] L. Einav and J. Levin, "Economics in the age of big data," *Science*, vol. 346, no. 6210, p. 1243089, 2014. [Online]. Available: https://www.science.org/doi/abs/10.1126/science.1243089, doi: 10.1126/science.1243089.

[52] C. Chen, J. Ma, Y. Susilo, Y. Liu, and M. Wang, "The promises of big data and small data for travel behavior (aka human mobility) analysis," *Transp. Res. C, Emerg. Technol.*, vol. 68, pp. 285–299, Jul. 2016.

[53] K. Zhou, C. Fu, and S. Yang, "Big data driven smart energy management: From big data to big insights," *Renew. Sustain. Energy Rev.*, vol. 56, pp. 215–225, Apr. 2016.

[54] E. Baccarelli, N. Cordeschi, A. Mei, M. Panella, M. Shojafar, and J. Stefa, "Energy-efficient dynamic traffic offloading and reconfiguration of networked data centers for big data stream mobile computing: Review, challenges, and a case study," *IEEE Netw.*, vol. 30, no. 2, pp. 54–61, Mar. 2016.

[55] Wikipedia. (2021). *List of Datasets for Machine-Learning Research*. [Online]. Available: https://en.wikipedia.org/wiki/List_of_datasets_for_machine-learning_research

[56] B. P. Bhattarai *et al.*, "Big data analytics in smart

grids: State-of-the-art, challenges, opportunities, and future directions," *IET Smart Grid*, vol. 2, no. 2, pp. 141–154, Jun. 2019. [Online]. Available: https://digital-library.theiet.org/content/journals/10.1049/iet-stg.2018.0261

[57] M. Kezunovic, L. Xie, and S. Grijalva, "The role of big data in improving power system operation and protection," in *Proc. Symp. Bulk Power Syst. Dyn. Control-Optim., Secur. Control Emerg. Power Grid*, Aug. 2013, pp. 1–9.

[58] C. Huang *et al.*, "Data quality issues for synchrophasor applications. Part I: A review," *J. Modern Power Syst. Clean Energy*, vol. 4, no. 3, pp. 342–352, Jul. 2016.

[59] U.S.-Canada Power System Outage Task Force, "Final report on the August 14, 2003 blackout in the United States and Canada," 2004. [Online]. Available: https://www.energy.gov/sites/default/files/oeprod/DocumentsandMedia/BlackoutFinal-Web.pdf

[60] M. Hojabri, U. Dersch, A. Papaemmanouil, and P. Bosshart, "A comprehensive survey on phasor measurement unit applications in distribution systems," *Energies*, vol. 12, no. 23, p. 4552, Nov. 2019.

[61] M. U. Usman and M. O. Faruque, "Applications of synchrophasor technologies in power systems," *J. Mod. Power Syst. Clean Energy*, vol. 7, no. 2, pp. 211–226, Mar. 2019.

[62] NASPI. (2017). *MASPI PMU Map with and Without Data Connections*. [Online]. Available: https://www.naspi.org/node/749

[63] Y. Guo, Z. Yang, S. Feng, and J. Hu, "Complex power system status monitoring and evaluation using big data platform and machine learning algorithms: A review and a case study," *Complexity*, vol. 2018, pp. 1–21, Sep. 2018.

[64] L. Wang *et al.*, "Frequency disturbance recorder design and developments," in *Proc. IEEE Power Eng. Soc. Gen. Meeting*, Jun. 2007, pp. 1–7.

[65] P. I. T. Laboratory and U. O. Tennessee. (2021) *FNET/GridEye Web Display*. [Online]. Available: http://fnetpublic.utk.edu/index.html

[66] (2021). *Electric Grid Test Case Repository*. [Online]. Available: https://electricgrids.engr.tamu.edu/electric-grid-test-cases/

[67] A. B. Birchfield, T. Xu, K. M. Gegner, K. S. Shetye, and T. J. Overbye, "Grid structural characteristics as validation criteria for synthetic networks," *IEEE Trans. Power Syst.*, vol. 32, no. 4, pp. 3258–3265, Jul. 2017.

[68] Y. Xu *et al.*, "U.S. Test system with high spatial and temporal resolution for renewable integration studies," in *Proc. IEEE Power Energy Soc. Gen. Meeting (PESGM)*, Aug. 2020, pp. 1–5.

[69] Breakthrough Energy Sciences, *A 2030 United States Macro Grid: Unlocking Geographical Diversity to Accomplish Clean Energy Goals*, Breakthrough Energy Sciences, Kirkland, WA, USA, 2021.

[70] D. Wu *et al.*, "An open-source extendable model and corrective measure assessment of the 2021 Texas power outage," *Adv. Appl. Energy*, vol. 4, Nov. 2021, Art. no. 100056.

[71] U. Adhikari *et al.*, "Development of power system test bed for data mining of synchrophasors data, cyber-attack and relay testing in RTDS," in *Proc. IEEE Power Energy Soc. Gen. Meeting*, Jul. 2012, pp. 1–7.

[72] J. Montoya *et al.*, "Advanced laboratory testing methods using real-time simulation and hardware-in-the-loop techniques: A survey of smart grid international research facility network activities," *Energies*, vol. 13, no. 12, p. 3267, Jun. 2020.

[73] S. B. D. I. Hub. (2017). *Big Data Regional Innovation Hubs and Spokes Workshop*. [Online]. Available: https://www.nsf.gov/cise/bdspokes/index.jsp

[74] N. S. Foundation. (2018). *The South Big Data Innovation Hub*. [Online]. Available: https://southbigdatahub.org/

[75] T. Joswig-Jones, K. Baker, and A. S. Zamzam, "OPF-learn: An open-source framework for creating representative AC optimal power flow datasets," 2021, *arXiv:2111.01228*.

[76] M. Sheinin, Y. Y. Schechner, and K. N. Kutulakos, "Computational imaging on the electric grid," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jul. 2017, pp. 6437–6446.

[77] G. Ruan *et al.*, "A cross-domain approach to analyzing the short-run impact of COVID-19 on the U.S. Electricity sector," *Joule*, vol. 4, no. 11, pp. 2322–2337, Nov. 2020.

[78] M. F. Khan *et al.*, "A review of big data resource management: Using smart grid systems as a case study," *Wireless Commun. Mobile Comput.*, vol. 2021, pp. 1–18, Oct. 2021.

[79] N. C. Thompson, K. Greenewald, K. Lee, and G. F. Manso, "The computational limits of deep learning," 2020, *arXiv:2007.05558*.

[80] *NVIDIA TESLA V100 GPU Architecture*. Santa Clara, CA, USA: NVIDIA, 2017.

[81] Google. (2021). *Cloud TPU*. [Online]. Available: https://cloud.google.com/tpu

[82] P. A. Merolla *et al.*, "A million spiking-neuron integrated circuit with a scalable communication network and interface," *Science*, vol. 345, no. 6197, pp. 668–673, Aug. 2014.

[83] GraphCore. (2021). *Intelligence Processing Unit*. [Online]. Available: https://www.graphcore.ai/products/ipu

[84] NVIDIA. (2021). *Jetson Nano Developer Kit*. [Online]. Available: https://developer.nvidia.com/embedded/jetson-nano-developer-kit

[85] M. Capra, B. Bussolino, A. Marchisio, G. Masera, M. Martina, and M. Shafique, "Hardware and software optimizations for accelerating deep neural networks: Survey of current trends, challenges, and the road ahead," *IEEE Access*, vol. 8, pp. 225134–225180, 2020.

[86] C. L. Lawson, R. J. Hanson, F. T. Krogh, and D. R. Kincaid, "Algorithm 539: Basic linear algebra subprograms for fortran usage [F1]," *ACM Trans. Math. Softw.*, vol. 5, no. 3, pp. 324–325, Sep. 1979.

[87] J. J. Dongarra, J. D. Cruz, S. Hammarling, and I. S. Duff, "Algorithm 679: A set of level 3 basic linear algebra subprograms: Model implementation and test programs," *ACM Trans. Math. Softw.*, vol. 16, no. 1, pp. 18–28, Mar. 1990.

[88] G. H. Golub and C. F. Van Loan, *Matrix Computations*. Baltimore, MD, USA: JHU Press, vol. 3, 2012.

[89] AMD. (2021). *AMD Math Library (LibM)*. [Online]. Available: https://developer.amd.com/amd-aocl/amd-math-library-libm/

[90] Intel. (2021). *Intel Oneapi Math Kernel Library*. [Online]. Available: https://www.intel.com/content/www/us/en/developer/tools/oneapi/onemkl.html#gs.gq697a

[91] NVIDIA. (2021). *Basic Linear Algebra on NVIDIA GPUs*. [Online]. Available: https://developer.nvidia.com/cublas

[92] The Khronos Group. (2021). *Opencl Overview*. [Online]. Available: https://www.khronos.org/opencl/

[93] M. Abadi *et al.*, "TensorFlow: Large-scale machine learning on heterogeneous distributed systems," 2016, *arXiv:1603.04467*.

[94] Y. Jia *et al.*, "Caffe: Convolutional architecture for fast feature embedding," in *Proc. 22nd ACM Int. Conf. Multimedia*, Nov. 2014, pp. 675–678.

[95] A. Paszke *et al.*, "Automatic differentiation in PyTorch," 2017. [Online]. Available: https://openreview.net/pdf?id=BJJsrmfCZ

[96] NVIDIA. (2021). *NVIDIA CuDNN*. [Online]. Available: https://developer.nvidia.com/cudnn

[97] Y. Sherry and N. Thompson, "Point of view: How fast do algorithms improve?" *Proc. IEEE*, vol. 109, no. 11, pp. 0018–9219, 2021.

[98] D. Carney *et al.*, "Monitoring streams—A new class of data management applications," in *Proc. 28th Int. Conf. Very Large Databases*. Amsterdam, Netherlands: Elsevier, Jan. 2002, pp. 215–226.

[99] J. M. Hellerstein *et al.*, "Adaptive processing: Technology in evolution," *IEEE Data Eng. Bull.*, vol. 23, no. 2, pp. 7–18, Jun. 2000.

[100] B. Babcock, S. Babu, M. Datar, R. Motwani, and J. Widom, "Models and issues in data stream systems," in *Proc. 21st ACM SIGMOD-SIGACT-SIGART Symp. Princ. Database Syst.*, 2002, pp. 1–16.

[101] K. Shvachko, H. Kuang, S. Radia, and R. Chansler, "The Hadoop distributed file system," in *Proc. IEEE 26th Symp. Mass Storage Syst. Technol. (MSST)*, 2010, pp. 1–10.

[102] M. Zaharia *et al.*, "Apache spark: Aunified engine for big data processing," *Commun. ACM*, vol. 59, no. 11, pp. 56–65, Oct. 2016.

[103] J. Dean and S. Ghemawat, "MapReduce: Simplified data processing on large clusters," *Commun. ACM*, vol. 51, no. 1, pp. 107–113, Jan. 2008.

[104] S. Ghemawat, H. Gobioff, and S.-T. Leung, "The Google file system," in *Proc. 19th ACM Symp. Operating Syst. Princ. (SOSP)*, 2003, pp. 29–43.

[105] F. Chang *et al.*, "Bigtable: A distributed storage system for structured data," *ACM Trans. Comput. Syst.*, vol. 26, no. 2, pp. 1–26, 2008.

[106] M. Zaharia *et al.*, "Resilient distributed datasets: A fault-tolerant abstraction for in-memory cluster computing," in *Proc. 9th USENIX Symp. Networked Syst. Des. Implement. (NSDI)*, 2012, pp. 15–28.

[107] (2021). *Energy and Utilities, Sustainably Fueling the Future*. [Online]. Available: https://www.ibm.com/U.K.-en/industries/energy

[108] IBM. (2021). *Oncur Puts Data to Work to Prevent Energy Outages*. [Online]. Available: https://mediacenter.ibm.com/id/1_3kwmzfsi

[109] C. Feng, Y. Wang, Q. Chen, Y. Ding, G. Strbac, and C. Kang, "Smart grid encounters edge computing: Opportunities and applications," *Adv. Appl. Energy*, vol. 1, Feb. 2021, Art. no. 100006.

[110] B. McMahan, E. Moore, D. Ramage, S. Hampson, and B. A. Y. Arcas, "Communication-efficient learning of deep networks from decentralized data," in *Proc. 20th Int. Conf. Artif. Intell. Statist. (PMLR)*, 2017, pp. 1273–1282.

[111] C.-C. Liu and T. Dillon, "State-of-the-art of expert system applications to power systems," *Int. J. Electr. Power Energy Syst.*, vol. 14, nos. 2–3, pp. 86–96, Apr. 1992. [Online]. Available: https://www.sciencedirect.com/science/article/pii/0142061592900314

[112] S. Rahman, "Artificial intelligence in electric power systems: A survey of the Japanese industry," *IEEE Trans. Power Syst.*, vol. 8, no. 3, pp. 1211–1218, Aug. 1993.

[113] M. Glavic, R. Fonteneau, and D. Ernst, "Reinforcement learning for electric power system decision and control: Past considerations and perspectives," *IFAC-PapersOnLine*, vol. 50, no. 1, pp. 6918–6927, Jul. 2017. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S2405896317317238

[114] S. Wen, C. Zhang, H. Lan, Y. Xu, Y. Tang, and Y. Huang, "A hybrid ensemble model for interval prediction of solar power output in ship onboard power systems," *IEEE Trans. Sustain. Energy*, vol. 12, no. 1, pp. 14–24, Jan. 2021.

[115] C. Wan, C. Zhao, and Y. Song, "Chance constrained extreme learning machine for nonparametric prediction intervals of wind power generation," *IEEE Trans. Power Syst.*, vol. 35, no. 5, pp. 3869–3884, Sep. 2020.

[116] C. Zhao, C. Wan, and Y. Song, "Operating reserve quantification using prediction intervals of wind power: An integrated probabilistic forecasting and decision methodology," *IEEE Trans. Power Syst.*, vol. 36, no. 4, pp. 3701–3714, Jul. 2021.

[117] Y. Wang, Q. Chen, D. Gan, J. Yang, D. S. Kirschen, and C. Kang, "Deep learning-based socio-demographic information identification from smart meter data," *IEEE Trans. Smart Grid*, vol. 10, no. 3, pp. 2593–2602, May 2019.

[118] J. Wang, H. Zhong, X. Lai, Q. Xia, Y. Wang, and C. Kang, "Exploring key weather factors from analytical modeling toward improved solar power forecasting," *IEEE Trans. Smart Grid*, vol. 10, no. 2, pp. 1417–1427, Mar. 2019.

[119] M. Cui, J. Zhang, Q. Wang, V. Krishnan, and hboxB.-M. Hodge, "A data-driven methodology for probabilistic wind power ramp forecasting," *IEEE Trans. Smart Grid*, vol. 10, no. 2, pp. 1326–1338, Mar. 2019.

[120] L. Cheng, H. Zang, T. Ding, Z. Wei, and G. Sun, "Multi-meteorological-factor-based graph modeling for photovoltaic power forecasting," *IEEE Trans. Sustain. Energy*, vol. 12, no. 3, pp. 1593–1603, Jul. 2021.

[121] H. Long, C. Zhang, R. Geng, Z. Wu, and W. Gu, "A combination interval prediction model based on biased convex cost function and auto-encoder in solar power prediction," *IEEE Trans. Sustain. Energy*, vol. 12, no. 3, pp. 1561–1570, Jul. 2021.

[122] X. Kong, C. Li, F. Zheng, and C. Wang, "Improved deep belief network for short-term load forecasting considering demand-side management," *IEEE Trans. Power Syst.*, vol. 35, no. 2, pp. 1531–1538, Mar. 2020.

[123] Y. Yuan, K. Dehghanpour, F. Bu, and Z. Wang, "A data-driven customer segmentation strategy based on contribution to system peak demand," *IEEE Trans. Power Syst.*, vol. 35, no. 5, pp. 4026–4035, Sep. 2020.

[124] Y. Yuan, K. Dehghanpour, F. Bu, and Z. Wang, "A multi-timescale data-driven approach to enhance distribution system observability," *IEEE Trans. Power Syst.*, vol. 34, no. 4, pp. 3168–3177, Jul. 2019.

[125] H. Zhang, J. Yan, Y. Liu, Y. Gao, S. Han, and L. Li, "Multi-source and temporal attention network for probabilistic wind power prediction," *IEEE Trans. Sustain. Energy*, vol. 12, no. 4, pp. 2205–2218, Oct. 2021.

[126] G. Ruan, D. S. Kirschen, H. Zhong, Q. Xia, and C. Kang, "Estimating demand flexibility using Siamese LSTM neural networks," *IEEE Trans. Power Syst.*, vol. 37, no. 3, pp. 2360–2370, May 2022.

[127] C. Zhao, C. Wan, and Y. Song, "An adaptive bilevel programming model for nonparametric prediction intervals of wind power generation," *IEEE Trans. Power Syst.*, vol. 35, no. 1, pp. 424–439, Jan. 2020.

[128] G. W. Chang and H.-J. Lu, "Integrating gray data preprocessor and deep belief network for day-ahead PV power output forecast," *IEEE Trans. Sustain. Energy*, vol. 11, no. 1, pp. 185–194, Jan. 2020.

[129] F. Bu, K. Dehghanpour, Z. Wang, and Y. Yuan, "A data-driven framework for assessing cold load pick-up demand in service restoration," *IEEE Trans. Power Syst.*, vol. 34, no. 6, pp. 4739–4750, Nov. 2019.

[130] L. Chavarro-Barrera, S. Perez-Londono, and J. Mora-Florez, "An adaptive approach for dynamic load modeling in microgrids," *IEEE Trans. Smart Grid*, vol. 12, no. 4, pp. 2834–2843, Jul. 2021.

[131] J. Simeunovic, B. Schubnel, P.-J. Alet, and R. E. Carrillo, "Spatio-temporal graph neural networks for multi-site PV power forecasting," *IEEE Trans. Sustain. Energy*, vol. 13, no. 2, pp. 1210–1220, Apr. 2022.

[132] B. Goehry, Y. Goude, P. Massart, and J.-M. Poggi, "Aggregation of multi-scale experts for bottom-up load forecasting," *IEEE Trans. Smart Grid*, vol. 11, no. 3, pp. 1895–1904, May 2020.

[133] C. Feng, M. Sun, and J. Zhang, "Reinforced deterministic and probabilistic load forecasting via Q-learning dynamic model selection," *IEEE Trans. Smart Grid*, vol. 11, no. 2, pp. 1377–1386, Jun. 2020.

[134] M. Li, W. Wei, Y. Chen, M.-F. Ge, and J. P. S. Catalao, "Learning the optimal strategy of power system operation with varying renewable generations," *IEEE Trans. Sustain. Energy*, vol. 12, no. 4, pp. 2293–2305, Oct. 2021.

[135] X. Lei, Z. Yang, J. Yu, J. Zhao, Q. Gao, and H. Yu, "Data-driven optimal power flow: A physics-informed machine learning approach," *IEEE Trans. Power Syst.*, vol. 36, no. 1, pp. 346–354, Jan. 2021.

[136] R. Dobbe, O. Sondermeijer, D. Fridovich-Keil, D. Arnold, D. Callaway, and C. Tomlin, "Toward distributed energy services: Decentralizing optimal power flow with machine learning," *IEEE Trans. Smart Grid*, vol. 11, no. 2, pp. 1296–1306, Mar. 2020.

[137] X. Fu, Q. Guo, and H. Sun, "Statistical machine learning model for stochastic optimal planning of distribution networks considering a dynamic correlation and dimension reduction," *IEEE Trans. Smart Grid*, vol. 11, no. 4, pp. 2904–2917, Jul. 2020.

[138] M. Chatzos, T. W. K. Mak, and P. Vanhentenryck, "Spatial network decomposition for fast and scalable AC-OPF learning," *IEEE Trans. Power Syst.*, early access, Nov. 13, 2021, doi: 10.1109/TPWRS.2021.3124726.

[139] F. Mohammadi, M. Sahraei-Ardakani, D. Trakas, and N. Hatziargyriou, "Machine learning assisted stochastic unit commitment during hurricanes with predictable line outages," *IEEE Trans. Power Syst.*, vol. 36, no. 6, pp. 5131–5142, Nov. 2021.

[140] A. Velloso and P. Van Hentenryck, "Combining deep learning and optimization for preventive security-constrained DC optimal power flow," *IEEE Trans. Power Syst.*, vol. 36, no. 4, pp. 3618–3628, Jul. 2021.

[141] X. Pan, T. Zhao, M. Chen, and S. Zhang, "DeepOPF: A deep neural network approach for security-constrained DC optimal power flow," *IEEE Trans. Power Syst.*, vol. 36, no. 3, pp. 1725–1735, May 2021.

[142] R. M. R. Barros, E. G. Da Costa, and J. F. Araujo, "Maximizing the financial return of non-technical loss management in power distribution systems," *IEEE Trans. Power Syst.*, vol. 37, no. 2, pp. 1634–1641, Mar. 2022.

[143] K. Dehghanpour and H. Nehrir, "An agent-based hierarchical bargaining framework for power management of multiple cooperative microgrids," *IEEE Trans. Smart Grid*, vol. 10, no. 99, pp. 514–522, Jan. 2017.

[144] S. Baltaoglu, L. Tong, and Q. Zhao, "Algorithmic bidding for virtual trading in electricity markets," *IEEE Trans. Power Syst.*, vol. 34, no. 1, pp. 535–543, Jan. 2019.

[145] S. Pineda, J. M. Morales, and A. Jimenez-Cordero, "Data-driven screening of network constraints for unit commitment," *IEEE Trans. Power Syst.*, vol. 35, no. 5, pp. 3695–3705, Sep. 2020.

[146] L. Liu and Z. Hu, "Data-driven regulation reserve capacity determination based on Bayes theorem," *IEEE Trans. Power Syst.*, vol. 35, no. 2, pp. 1646–1649, Mar. 2020.

[147] F. Safdarian, A. Kargarian, and F. Hasan, "Multiclass learning-aided temporal decomposition and distributed optimization for power systems," *IEEE Trans. Power Syst.*, vol. 36, no. 6, pp. 4941–4952, Nov. 2021.

[148] Y. Li, N. Yu, and W. Wang, "Machine learning-driven virtual bidding with electricity market efficiency analysis," *IEEE Trans. Power Syst.*, vol. 37, no. 1, pp. 354–364, Jan. 2022.

[149] B. Stappers, N. G. Paterakis, K. Kok, and M. Gibescu, "A class-driven approach based on long short-term memory networks for electricity price scenario generation and reduction," *IEEE Trans. Power Syst.*, vol. 35, no. 4, pp. 3040–3050, Jul. 2020.

[150] S. Chai, Z. Xu, and Y. Jia, "Conditional density forecast of electricity price based on ensemble ELM and logistic EMOS," *IEEE Trans. Smart Grid*, vol. 10, no. 3, pp. 3031–3043, May 2019.

[151] S. Li, T. Ding, C. Mu, C. Huang, and M. Shahidehpour, "A machine learning-based reliability evaluation model for integrated power-gas systems," *IEEE Trans. Power Syst.*, early access, Nov. 13, 2021, doi: 10.1109/TPWRS.2021.3125531.

[152] B. Zhang, M. Wang, and W. Su, "Reliability analysis of power systems integrated with high-penetration of power converters," *IEEE Trans. Power Syst.*, vol. 36, no. 3, pp. 1998–2009, May 2021.

[153] K. Mahmoud and M. Abdel-Nasser, "Fast yet accurate energy-loss-assessment approach for analyzing/sizing PV in distribution systems using machine learning," *IEEE Trans. Sustain. Energy*, vol. 10, no. 3, pp. 1025–1033, Jul. 2019.

[154] Z. S. Hosseini, A. Khodaei, and A. Paaso, "Machine learning-enabled distribution network phase identification," *IEEE Trans. Power Syst.*, vol. 36, no. 2, pp. 842–850, Mar. 2021.

[155] G. Dalal, E. Gilboa, S. Mannor, and L. Wehenkel, "Chance-constrained outage scheduling using a machine learning proxy," *IEEE Trans. Power Syst.*, vol. 34, no. 4, pp. 2528–2540, Jul. 2019.

[156] M. Yi and M. Wang, "Bayesian energy disaggregation at substations with uncertainty modeling," *IEEE Trans. Power Syst.*, vol. 37, no. 1, pp. 764–775, Jan. 2022.

[157] J. Yuan and Y. Weng, "Support matrix regression for learning power flow in distribution grid with unobservability," *IEEE Trans. Power Syst.*, vol. 37, no. 2, pp. 1151–1161, Mar. 2022.

[158] Y. Zhang, C. Chen, G. Liu, T. Hong, and F. Qiu, "Approximating trajectory constraints with machine learning—Microgrid islanding with frequency constraints," *IEEE Trans. Power Syst.*, vol. 36, no. 2, pp. 1239–1249, Mar. 2021.

[159] E. Cook, S. Luo, and Y. Weng, "Solar panel identification via deep semi-supervised learning and deep one-class classification," *IEEE Trans. Power Syst.*, early access, Nov. 13, 2021, doi: 10.1109/TPWRS.2021.3125613.

[160] B. Foggo and N. Yu, "Improving supervised phase identification through the theory of information losses," *IEEE Trans. Smart Grid*, vol. 11, no. 3, pp. 2337–2346, May 2020.

[161] W. Gao and D. Gorinevsky, "Probabilistic modeling for optimization of resource mix with variable generation and storage," *IEEE Trans. Power Syst.*, vol. 35, no. 5, pp. 4036–4045, Sep. 2020.

[162] M. Mohammadian, F. Aminifar, N. Amjady, and M. Shahidehpour, "Data-driven classifier for extreme outage prediction based on Bayes decision theory," *IEEE Trans. Power Syst.*, vol. 36, no. 6, pp. 4906–4914, Nov. 2021.

[163] M. Mokhtar *et al.*, "Automating the verification of the low voltage network cables and topologies," *IEEE Trans. Smart Grid*, vol. 11, no. 2, pp. 1657–1666, Mar. 2020.

[164] T. Huang, S. Gao, and L. Xie, "A neural Lyapunov approach to transient stability assessment of power electronics-interfaced networked microgrids," *IEEE Trans. Smart Grid*, vol. 13, no. 1, pp. 106–118, Jan. 2022.

[165] M. Cui, J. Wang, and B. Chen, "Flexible machine learning-based cyberattack detection using spatiotemporal patterns for distribution systems," *IEEE Trans. Smart Grid*, vol. 11, no. 2, pp. 1805–1808, Mar. 2020.

[166] Q. Wang, F. Li, Y. Tang, and Y. Xu, "Integrating model-driven and data-driven methods for power system frequency stability assessment and control," *IEEE Trans. Power Syst.*, vol. 34, no. 6, pp. 4557–4568, Nov. 2019.

[167] L. Zhu and Y. Luo, "Deep feedback learning based predictive control for power system undervoltage load shedding," *IEEE Trans. Power Syst.*, vol. 36, no. 4, pp. 3349–3361, Jul. 2021.

[168] L. Zhu, D. J. Hill, and C. Lu, "Hierarchical deep learning machine for power system online transient stability prediction," *IEEE Trans. Power Syst.*, vol. 35, no. 3, pp. 2399–2411, May 2020.

[169] J. L. Cremer, I. Konstantelos, S. H. Tindemans, and G. Strbac, "Data-driven power system operation: Exploring the balance between cost and risk," *IEEE Trans. Power Syst.*, vol. 34, no. 1, pp. 791–801, Jan. 2019.

[170] S. M. Mazhari, B. Khorramdel, C. Y. Chung, I. Kamwa, and D. Novosel, "A simulation-based classification approach for online prediction of generator dynamic behavior under multiple large disturbances," *IEEE Trans. Power Syst.*, vol. 36, no. 2, pp. 1217–1228, Mar. 2021.

[171] H.-Y. Su and H.-H. Hong, "An intelligent data-driven learning approach to enhance online probabilistic voltage stability margin prediction," *IEEE Trans. Power Syst.*, vol. 36, no. 4, pp. 3790–3793, Jul. 2021.

[172] L. Zhu, D. J. Hill, and C. Lu, "Semi-supervised ensemble learning framework for accelerating power system transient stability knowledge base generation," *IEEE Trans. Power Syst.*, vol. 37, no. 3, pp. 2441–2454, May 2022.

[173] I. Konstantelos, M. Sun, S. H. Tindemans, S. Issad, P. Panciatici, and G. Strbac, "Using vine copulas to generate representative system states for machine learning," *IEEE Trans. Power Syst.*, vol. 34, no. 1, pp. 225–235, Jan. 2019.

[174] X. Yin and X. Zhao, "Sensorless maximum power extraction control of a hydrostatic tidal turbine based on adaptive extreme learning machine," *IEEE Trans. Sustain. Energy*, vol. 11, no. 1, pp. 426–435, Jan. 2020.

[175] G. Ravikumar and M. Govindarasu, "Anomaly detection and mitigation for wide-area damping control using machine learning," *IEEE Trans. Smart Grid*, early access, May 18, 2020, doi: 10.1109/TSG.2020.2995313.

[176] M. Jalali, V. Kekatos, N. Gatsis, and D. Deka, "Designing reactive power control rules for smart inverters using support vector machines," *IEEE Trans. Smart Grid*, vol. 11, no. 2, pp. 1759–1770, Mar. 2020.

[177] L. Zhu and D. J. Hill, "Networked time series shapelet learning for power system transient stability assessment," *IEEE Trans. Power Syst.*, vol. 37, no. 1, pp. 416–428, Jan. 2022.

[178] Y. Zhao, J. Chen, and H. V. Poor, "A Learning-to-infer method for real-time power grid multi-line outage identification," *IEEE Trans. Smart Grid*, vol. 11, no. 1, pp. 555–564, Jan. 2020.

[179] S. K. Azman, Y. J. Isbeih, M. S. E. Moursi, and K. Elbassioni, "A unified online deep learning prediction model for small signal and transient stability," *IEEE Trans. Power Syst.*, vol. 35, no. 6, pp. 4585–4598, Nov. 2020.

[180] A. S. Zamzam, X. Fu, and N. D. Sidiropoulos, "Data-driven learning-based optimization for distribution system state estimation," *IEEE Trans. Power Syst.*, vol. 34, no. 6, pp. 4796–4805, Nov. 2019.

[181] S. Karagiannopoulos, P. Aristidou, and G. Hug, "Data-driven local control design for active distribution grids using off-line optimal power flow and machine learning techniques," *IEEE Trans. Smart Grid*, vol. 10, no. 6, pp. 6461–6471, Nov. 2019.

[182] P. Wang and M. Govindarasu, "Multi-agent based attack-resilient system integrity protection for smart grid," *IEEE Trans. Smart Grid*, vol. 11, no. 4, pp. 3447–3456, Jul. 2020.

[183] W. Li, D. Deka, M. Chertkov, and M. Wang, "Real-time faulted line localization and PMU placement in power systems through convolutional neural networks," *IEEE Trans. Power Syst.*, vol. 34, no. 6, pp. 4640–4651, Nov. 2019.

[184] M. Sun, I. Konstantelos, and G. Strbac, "A deep learning-based feature extraction framework for system security assessment," *IEEE Trans. Smart Grid*, vol. 10, no. 5, pp. 5007–5020, Sep. 2019.

[185] Y. Su *et al.*, "An adaptive PV frequency control strategy based on real-time inertia estimation," *IEEE Trans. Smart Grid*, vol. 12, no. 3, pp. 2355–2364, May 2021.

[186] M. Alzayed, H. Chaoui, and Y. Farajpour, "Maximum power tracking for a wind energy conversion system using cascade-forward neural networks," *IEEE Trans. Sustain. Energy*, vol. 12, no. 4, pp. 2367–2377, Oct. 2021.

[187] D. Li and S. Dick, "Residential household non-intrusive load monitoring via graph-based multi-label semi-supervised learning," *IEEE Trans. Smart Grid*, vol. 10, no. 4, pp. 4615–4627, Jul. 2019.

[188] C. Feng, A. Mehmani, and J. Zhang, "Deep learning-based real-time building occupancy detection using AMI data," *IEEE Trans. Smart Grid*, vol. 11, no. 5, pp. 4490–4501, Sep. 2020.

[189] A. Jaech, B. Zhang, M. Ostendorf, and D. S. Kirschen, "Real-time prediction of the duration of distribution system outages," *IEEE Trans. Power Syst.*, vol. 34, no. 1, pp. 773–781, Jan. 2019.

[190] S. Wang, L. Li, and P. Dehghanian, "Distributed intelligence for online situational awareness in power grids," *IEEE Trans. Power Syst.*, early access, Nov. 18, 2021, doi: 10.1109/TPWRS.2021.3128951.

[191] O. Boyaci, M. R. Narimani, K. R. Davis, M. Ismail, T. J. Overbye, and E. Serpedin, "Joint detection and localization of stealth false data injection attacks in smart grids using graph neural networks," *IEEE Trans. Smart Grid*, vol. 13, no. 1, pp. 807–819, Jan. 2022.

[192] Z. Yang and J. R. Marti, "Real-time resilience optimization combining an AI agent with online hard optimization," *IEEE Trans. Power Syst.*, vol. 37, no. 1, pp. 508–517, Jan. 2022.

[193] L. Ma, L. Wang, and Z. Liu, "Topology identification of distribution networks using a split-EM based data-driven approach," *IEEE Trans. Power Syst.*, vol 37, no. 3, pp. 2019–2031, May 2022.

[194] Y. Wang, H. Silva-Saravia, and H. Pulgar-Painemal, "Actuator placement for enhanced grid dynamic performance: A machine learning approach," *IEEE Trans. Power Syst.*, vol. 34, no. 4, pp. 3119–3128, Jul. 2019.

[195] A. Ahmed, K. S. Sajan, A. Srivastava, and Y. Wu, "Anomaly detection, localization and classification using drifting synchrophasor data streams," *IEEE Trans. Smart Grid*, vol. 12, no. 4, pp. 3570–3580, Jul. 2021.

[196] Q. Huang, R. Huang, W. Hao, J. Tan, R. Fan, and Z. Huang, "Adaptive power system emergency control using deep reinforcement learning," *IEEE Trans. Smart Grid*, vol. 11, no. 2, pp. 1171–1182, Mar. 2020.

[197] Y. Zhang, X. Wang, J. Wang, and Y. Zhang, "Deep reinforcement learning based volt-VAR optimization in smart distribution systems," *IEEE Trans. Smart Grid*, vol. 12, no. 1, pp. 361–371, Jan. 2021.

[198] E. Mocanu *et al.*, "On-line building energy optimization using deep reinforcement learning," *IEEE Trans. Smart Grid*, vol. 10, no. 4, pp. 3698–3708, Jul. 2019.

[199] C. Yang, A. A. Thatte, and L. Xie, "Multitime-scale data-driven spatio-temporal forecast of photovoltaic generation," *IEEE Trans. Sustain. Energy*, vol. 6, no. 1, pp. 104–112, Jan. 2015.

[200] M. S. Modarresi *et al.*, "Scenario-based economic dispatch with tunable risk levels in high-renewable power systems," *IEEE Trans. Power Syst.*, vol. 34, no. 6, pp. 5103–5114, Nov. 2019.

[201] T. Huang, N. M. Freris, P. R. Kumar, and L. Xie, "A synchrophasor data-driven method for forced oscillation localization under resonance conditions," *IEEE Trans. Power Syst.*, vol. 35, no. 5, pp. 3927–3939, Sep. 2020.

[202] S. A. N. Sarmadi, V. Venkatasubramanian, and A. Salazar, "Analysis of November 29, 2005 western American oscillation event," *IEEE Trans. Power Syst.*, vol. 31, no. 6, pp. 5210–5211, Nov. 2016.

[203] E. J. Candès, X. Li, Y. Ma, and J. Wright, "Robust principal component analysis?" *J. ACM*, vol. 58, no. 1, pp. 1–37, 2009.

[204] D. Wu, D. Kalathil, M. Begovic, and L. Xie, "Deep reinforcement learning-based robust protection in DER-rich distribution grids," 2020, *arXiv:2003.02422*.

[205] D. Wu, D. Kalathil, M. Begovic, and L. Xie, "PyProD: A machine learning-friendly platform for protection analytics in distribution systems," 2021, *arXiv:2109.05802*.

[206] R. Killick and I. A. Eckley, "Changepoint: An R package for changepoint analysis," *J. Stat. Softw.*, vol. 58, no. 3, pp. 1–19, 2014. [Online]. Available: http://www.jstatsoft.org/v58/i03/

[207] J. Redmon. (2013–2016). *DarkNet: Open Source Neural Networks in C*. [Online]. Available: http://pjreddie.com/darknet/

[208] T.-Y. Lin *et al.*, "Microsoft COCO: Common objects in context," in *Proc. Eur. Conf. Comput. Vis.*, 2015, pp 740–755.

[209] T. Hong and S. Fan, "Probabilistic electric load forecasting: A tutorial review," *Int. J. Forecasting*, vol. 32, no. 3, pp. 914–938, 2016.

[210] Y. Sun, W. Hao, Y. Chen, and B. Liu, "Data-driven occupant-behavior analytics for residential buildings," *Energy*, vol. 206, Sep. 2020, Art. no. 118100.

[211] G. Gross and F. D. Galiana, "Short-term load forecasting," *Proc. IEEE*, vol. 75, no. 12, pp. 1558–1573, Dec. 1987.

[212] D. C. Park, M. A. El-Sharkawi, R. J. Marks, L. E. Atlas, and M. J. Damborg, "Electric load forecasting using an artificial neural network," *IEEE Trans. Power Syst.*, vol. 6, no. 2, pp. 442–449, May 1991.

[213] L. Breiman, J. H. Friedman, R. A. Olshen, and C. J. Stone, *Classification Regression Trees*. Evanston, IL, USA: Routledge, 2017.

[214] A. Rosenberg and J. Hirschberg, "V-measure: A conditional entropy-based external cluster evaluation measure," in *Proc. Joint Conf. Empirical Methods Natural Lang. Process. Comput. Natural Lang. Learn. (EMNLP-CoNLL)*, 2007, pp. 410–420.

[215] P. J. Rousseeuw, "Silhouettes: A graphical aid to the interpretation and validation of cluster analysis," *J. Comput. Appl. Math.*, vol. 20, no. 1, pp. 53–65, 1987.

[216] A. Tayyebi and N. Day. (2021). *Segmentation of LiDAR Point Clouds: LiDAR Application in Vegetation Management*. [Online]. Available: https://www.esri.com/content/dam/esrisites/en-us/events/conferences/2022/user-conference/user-sessions/12796-ai-training-and-inferencing-processes-for-lidar-classification.pdf

[217] A. Tayyebi, K. Jury, and S. Turner. (2021). *Using Data Science to Minimize Power Outage Due to Vegetation*. [Online]. Available: https://www.esri.com/content/dam/esrisites/en-us/events/conferences/2022/user-conference/user-sessions/12955-using-data-science-to-minimize-power-outages-due-to-vegetation.pdf

[218] J. Zhao, S. Maslennikov, E. Litvinov, and X. Geng, "An enhanced transmission operating guide creation framework using machine learning techniques," in *Proc. IEEE Power Energy Soc. Gen. Meeting (PESGM)*, Aug. 2018, pp. 1–5.

[219] J. Gao, P. Chongfuangprinya, Y. Ye, and B. Yang, "A three-layer hybrid model for wind power prediction," in *Proc. IEEE Power Energy Soc. Gen. Meeting (PESGM)*, Aug. 2020, pp. 1–5.

[220] S. Maslennikov *et al.*, "A test cases library for methods locating the sources of sustained oscillations," in *Proc. IEEE Power Energy Soc. Gen. Meeting (PESGM)*, Jul. 2016, pp. 1–5.

[221] X. Zheng *et al.*, "A multi-scale time-series dataset with benchmark for machine learning in decarbonized energy grids," 2021, *arXiv:2110.06324*.

[222] IEEE PES, NASPI. (2021). *IEEE-NASPI Oscillation Source Location Contest*. [Online]. Available: http://web.eecs.utk.edu/~kaisun/Oscillation/2021Contest/

[223] A. Marot *et al.*, "Learning to run a power network challenge for training topology controllers," *Electr. Power Syst. Res.*, vol. 189, Dec. 2020, Art. no. 106635.

[224] A. Marot *et al.*, "Learning to run a power network challenge: A retrospective analysis," 2021, *arXiv:2103.03104*.

[225] C. Ebert, G. Gallardo, J. Hernantes, and N. Serrano, "DevOps," *IEEE Softw.*, vol. 33, no. 3, pp. 94–100, 2016.

[226] L. Zhu, L. Bass, and G. Champlin-Scharff, "DevOps and its practices," *IEEE Softw.*, vol. 33, no. 3, pp. 32–34, May/Jun. 2016.

[227] D. Sculley *et al.*, "Hidden technical debt in machine learning systems," in *Proc. Adv. Neural Inf. Process. Syst.*, vol. 28, C. Cortes, N. Lawrence, D. Lee, M. Sugiyama, and R. Garnett, Eds. New York, NY, USA: Curran Associates, 2015. [Online]. Available: https://proceedings.neurips.cc/paper/2015/file/86df7dcfd896fcaf2674f757a2463eba-Paper.pdf

[228] M. Treveil *et al.*, *Introducing MLOps*. Sebastopol, CA, USA: O'Reilly Media, 2020.

[229] G. E. Karniadakis *et al.*, "Physics-informed machine learning," *Nat. Rev. Phys.*, vol. 3, no. 6, pp. 422–440, 2021.

[230] A. Nagabandi *et al.*, "Learning to adapt in dynamic, real-world environments through meta-reinforcement learning," 2018, *arXiv:1803.11347*.

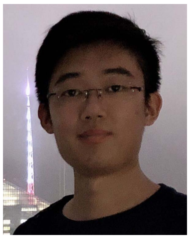
[231] K. Arndt, M. Hazara, A. Ghadirzadeh, and V. Kyrki, "Meta reinforcement learning for sim-to-real domain adaptation," in *Proc. IEEE Int. Conf. Robot. Autom. (ICRA)*, May 2020, pp. 2725–2731.

[232] S. L. Brunton, J. L. Proctor, and J. N. Kutz, "Discovering governing equations from data by sparse identification of nonlinear dynamical systems," *Proc. Nat. Acad. Sci. USA*, vol. 113, no. 15, pp. 3932–3937, 2015.

ABOUT THE AUTHORS

Le Xie (Fellow, IEEE) received the B.E. degree in electrical engineering from Tsinghua University, Beijing, China, in 2004, the M.S. degree in engineering sciences from Harvard University, Cambridge, MA, USA, in 2005, and the Ph.D. degree from Carnegie Mellon University, Pittsburgh, PA, USA, in 2009.

He is currently a Professor with the Department of Electrical and Computer Engineering, Texas A&M University, College Station, TX, USA. His research interests include modeling and control of large-scale complex systems, smart grids application with renewable energy resources, and electricity markets.

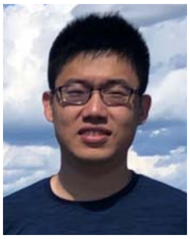


Xiangtian Zheng (Graduate Student Member, IEEE) received the B.E. degree from Tsinghua University, Beijing, China, in 2017. He is currently working toward the Ph.D. degree in electrical engineering at Texas A&M University, College Station, TX, USA.

His industry experience includes an internship with PJM, Valley Forge, PA, USA, in 2019 and an internship with the Mitsubishi Electric Research Laboratories, Cambridge, MA, USA, in 2021. His research interest includes domain knowledge-informed machine learning for power system security.

Yannan Sun (Senior Member, IEEE) received the B.S. degree in mathematics from the University of Science and Technology of China, Hefei, China, in 2004, and the M.S. degree in statistics and the Ph.D. degree in mathematics from Washington State University, Pullman, WA, USA, in 2007 and 2010, respectively.

She was a Scientist/Senior Scientist with the Electricity Infrastructure Group, Pacific Northwest National Laboratory, Richland, WA, USA, from 2010 to 2017. She is currently a Data Scientist with Oncor Electric Delivery, Dallas, TX, USA. Her expertise lies in data analytics and machine learning using power system data, which she has employed to develop many data-driven algorithms for load forecasting, anomaly detection, connectivity correction, and equipment preventive maintenance.



Tong Huang (Member, IEEE) received the B.E. degree in electrical engineering from North China Electric Power University, Beijing, China, in 2013, and the M.S. and Ph.D. degrees in electrical engineering from Texas A&M University, College Station, TX, USA, in 2017 and 2021, respectively.

He was a Visiting Student with the Laboratory for Information and Decision Systems, Massachusetts Institute of Technology (MIT), Cambridge, MA, USA, in 2018. His industry experience includes an internship with ISO New England, Holyoke, MA, USA, in 2018 and an internship with the Mitsubishi Electric Research Laboratories, Cambridge, in 2019. He was a Postdoctoral Researcher with Texas A&M University in 2021. He is currently a Postdoctoral Researcher with MIT.

Dr. Huang received the Best Paper Award at the 2020 IEEE PES General Meeting, the Best Paper Award at the 54th Hawaii International Conference on System Sciences, the Thomas W. Powell'62 and Powell Industries Inc. Fellowship, and the Texas A&M Graduate Teaching Fellowship.

Tony Bruton (Member, IEEE) received the B.S. degree in electrical engineering from Texas Tech University, Lubbock, TX, USA, in 2000.

He began working for Oncor, Dallas, TX , USA, as a Substation Design Engineer. For several years, he managed Oncor's High Voltage Grid in East Texas and then managed the group that designed and built high-voltage transmission lines. He also managed the routing and acquisition of right-of-way for new transmission lines in West Texas. His current role as the Director of T&D Services involves ensuring an accurate system that monitors and controls Oncor's Transmission and Distribution Systems, Energy Management System (EMS), and Advanced Distribution Management System (ADMS), as well as data analytics and operator training.