Review:

Smart grid dispatch powered by deep learning: a survey[*]

Gang HUANG[‡1], Fei WU[2], Chuangxin GUO[3]

1Zhejiang Lab, Hangzhou 311121, China
2College of Computer Science and Technology, Zhejiang University, Hangzhou 310027, China
3College of Electrical Engineering, Zhejiang University, Hangzhou 310027, China
E-mail: huanggang@zju.edu.cn; wufei@zju.edu.cn; guochuangxin@zju.edu.cn

Abstract: Power dispatch is a core problem for smart grid operations. It aims to provide optimal operating points within a transmission network while power demands are changing over space and time. This function needs to be run every few minutes throughout the day; thus, a fast, accurate solution is of vital importance. However, due to the complexity of the problem, reliable and computationally efficient solutions are still under development. This issue will become more urgent and complicated as the integration of intermittent renewable energies increases and the severity of uncertain disasters gets worse. With the recent success of artificial intelligence in various industries, deep learning becomes a promising direction for power engineering as well, and the research community begins to rethink the problem of power dispatch. This paper reviews the recent progress in smart grid dispatch from a deep learning perspective. Through this paper, we hope to advance not only the development of smart grids but also the ecosystem of artificial intelligence.

1 Introduction

Power systems can be traced back to the Pearl Street station and its distribution network built by Thomas Edison (Kundur, 1994). They were first built to support lighting systems only, but soon spread to provide energy for many other applications with the accelerated adoption of electrification technologies (Mai et al., 2018). This revolution is still ongoing. For example, we see more and more electric vehicles running on the road in daily life. In fact, a reliable supply of electricity has already become a cornerstone of our digital economy (Walsh,

2013), and it is of vital importance in many areas, including transportation (Jaller et al., 2020), finance (Oughton et al., 2017), food and agriculture (Zimmerman R et al., 2016), and epidemic prevention and control (Liu ZF, 2020). Its increasing importance indicates increasing complexities, and smart grids are believed to be the next-generation power system (Gharavi and Ghafurian, 2011; Wen GH et al., 2021) for meeting the increasing requirements.

Smart grid will be much more complicated than traditional power systems. It is essentially an electrical grid enhanced by information technology (Hossain et al., 2012), which turns the electrical grid into an intelligent system. The features of smart grids and conventional power grids are summarized in Table 1. The main benefits of smart grids include the following: (1) efficiency—smart grids enable greater utilization of generators and lines; (2) resilience—smart grids ensure a more reliable supply of

Table 1 Smart grid and conventional power grid features (Mohamed and Eltamaly, 2018)

Item	Smart grid	Conventional power grid
1	Cyber-physical system	Electromechanical system
2	Sensors throughout the network	Few sensors
3	Relatively distributed generators	Relatively centralized generators
4	Automatic monitoring	Manual monitoring
5	More reliable and resilient	Suffering more from failures and blackouts
6	Consumers have many choices	Consumers have few choices

electricity, especially against disasters; (3) sustainability—smart grids permit greater penetration of renewable energy sources; (4) market-enabling—smart grids allow more choices and therefore lower energy bills.

Power dispatch is a kernel problem that must be addressed in fulfilling the smart grid promise (Gandhi et al., 2016). With the purpose of deciding a best strategy for power generation, transmission, and even consumption, smart grid dispatch connects different components within the whole system. In particular, demand is always changing over space and time, which requires dispatch methods to be reliable and computationally efficient (Hasan et al., 2020). Furthermore, with the increased integration of intermittent renewable energies (Huang et al., 2015; Wen YF et al., 2016) and the severity of uncertain disasters getting worse (Huang et al., 2017a, 2017b), the generation and transmission parts of grids are now becoming more and more uncertain, which makes the smart grid dispatch problem even more challenging than ever before.

Artificial intelligence (AI) is now a popular technology in various industries (Wu F et al., 2020). It promises to revolutionize almost all kinds of tough problems by switching the original problem to a data-driven one. This is a new paradigm for scientific research (Hey et al., 2009) and has proved its power and efficiency in many real-world applications such as computer vision (Ibrahim et al., 2020), natural language processing (Cambria and White, 2014), and speech recognition (Haridas et al., 2018). Recently, it has also attracted much attention in the power engineering community (Bose, 2017) from both academia and industries.

AI itself has gone through several ups and downs, and interested readers are referred to Buchanan (2005) for more details. The most recent resurgence of AI is powered by deep learning (LeCun et al., 2015; Goodfellow et al., 2016), which is part of a broader family of machine learning methods that are based on artificial neural networks with representation learning. On one hand, deep learning can handle a huge amount of data where other machine learning methods often fail. On the other hand, the power industries are obtaining more and more data due to the wide spread of advanced information and communication technologies (Huang et al., 2019) such as phasor measurement units (PMU), micro phasor measurement units (μPMU), and advanced meter infrastructures. As a result, many scholars have been trying to apply deep learning to problems of smart grid dispatch, with the hope of tackling the issues that remain unsolved by traditional methods. This motivated us to write this paper to summarize what has been done already and where deep learning is headed regarding smart grid dispatch. Note that other machine learning methods, such as random forests (Rahman et al., 2020) and logistic regression (Rudin et al., 2012; Eskandarpour and Khodaei, 2017), will not be the focus of this paper.

2 Smart grid dispatch problems

Mathematically, smart grid dispatch is an optimization problem that can be formulated as follows:

$$\min f(x)$$
$$\text{s.t.} \begin{cases} g(x) = 0, \\ h(x) \leq 0, \end{cases} \tag{1}$$

where x represents decision variables (e.g., outputs of generators), $f(x)$ represents the objective function (e.g., total cost of all the generators), and $g(x)$ and $h(x)$ represent equality constraints (e.g., power balance constraints) and inequality constraints (e.g., output limits of generators), respectively. Under different assumptions or for different requirements, smart grid dispatch problems will have different formulations. Herein, we summarize some popular formulations that are often used in power engineering.

2.1 Economic dispatch

The economic dispatch problem is the process of allocating generation among different generating units to achieve the minimum operating cost (Wood et al., 2013). It is the simplest formulation of smart grid dispatch and is often used for real-time operation. The formulation is as follows:

$$\min c(P_\mathrm{G})$$

$$\text{s.t.} \begin{cases} \sum_G P_\mathrm{G} - \sum_D P_\mathrm{D} = 0, \\ P_\mathrm{G}^\mathrm{min} \le P_\mathrm{G} \le P_\mathrm{G}^\mathrm{max}, \end{cases} \quad (2)$$

where P_G and P_D represent the generator outputs and load demands, respectively. G is the set of power generators and D is the set of load demands. $c(P_\mathrm{G})$ is the total cost function, which can be nonlinear (e.g., quadratic function $c(P_\mathrm{G}) = \sum(\boldsymbol{\alpha} + \boldsymbol{\beta} P_\mathrm{G} + \boldsymbol{\gamma} P_\mathrm{G}^2)$ is often used) or linear (i.e., $c(P_\mathrm{G}) = \sum(\boldsymbol{\alpha} + \boldsymbol{\beta} P_\mathrm{G})$). The first constraint is the power balance constraint, without considering the power flow through transmission lines. The second constraint presents the limits of generator outputs.

The classical economic dispatch problem targets minimization of the operating costs as shown above. With the increasing concern about environmental matters, other costs, such as carbon emission cost, can also be included (Yin et al., 2018). In addition, network loss cost is often considered (Zeng et al., 2014), especially for electricity market applications.

2.2 Optimal power flow

Economic dispatch is quite successful when the flows within transmission lines are at a low level. However, with quickly increasing demand for electricity, transmission lines were soon working close to their capacities, and a fear of line overloading led to the concept of optimal power flow (Carpentier, 1979).

As an extension of the classical economic dispatch problem, optimal power flow is an optimization problem that solves the economic dispatch problem and power flow equations simultaneously (Wood et al., 2013). Its original formulation is as follows:

$$\min c(P_\mathrm{G})$$

$$\text{s.t.} \begin{cases} P_\mathrm{G} - P_\mathrm{D} - \boldsymbol{V}(\boldsymbol{G}\boldsymbol{V}\cos\boldsymbol{\Theta} + \boldsymbol{B}\boldsymbol{V}\sin\boldsymbol{\Theta}) = 0, \\ Q_\mathrm{G} - Q_\mathrm{D} - \boldsymbol{V}(\boldsymbol{G}\boldsymbol{V}\sin\boldsymbol{\Theta} - \boldsymbol{B}\boldsymbol{V}\cos\boldsymbol{\Theta}) = 0, \\ -P_\mathrm{L}^\mathrm{max} \le \boldsymbol{X}\boldsymbol{\Theta} \le P_\mathrm{L}^\mathrm{max}, \\ P_\mathrm{G}^\mathrm{min} \le P_\mathrm{G} \le P_\mathrm{G}^\mathrm{max}, \\ Q_\mathrm{G}^\mathrm{min} \le Q_\mathrm{G} \le Q_\mathrm{G}^\mathrm{max}, \end{cases}$$
$$(3)$$

where $P_\mathrm{G}, Q_\mathrm{G}, P_\mathrm{D}, Q_\mathrm{D}, \boldsymbol{V}, \boldsymbol{G}, \boldsymbol{B}$, and $\boldsymbol{\Theta}$ indicate the generator active output, generator reactive output, load active demand, load reactive demand, voltage magnitude, the real part of the node admittance matrix, the imaginary part of the node admittance matrix, and the voltage angle difference, respectively. $\boldsymbol{X}$ is the reactance matrix of transmission lines, and $P_\mathrm{L}^\mathrm{max}$ represents the transmission capacity. The first two constraints address the power flow equations, the third constraint addresses the transmission line constraints, and the last two constraints address the limits of generator outputs.

The above problem is a highly nonlinear and nonconvex optimization problem, which makes it quite difficult to solve for practical usage. As a result, a linear approximation called direct current (DC) optimal power flow is often used in research and industry. Here, decision variables in the DC optimal power flow formulation correspond to active generator outputs only. Generally, it is believed that DC optimal power flow gives a good approximation of the original optimal power flow problem (i.e., alternating current (AC) optimal power flow) and is much faster and easier to apply (Wood et al., 2013). DC optimal power flow has the following formulation:

$$\min c(P_\mathrm{G})$$

$$\text{s.t.} \begin{cases} P_\mathrm{G} - P_\mathrm{D} - \boldsymbol{B}\boldsymbol{\Theta} = 0, \\ -P_\mathrm{L}^\mathrm{max} \le \boldsymbol{X}\boldsymbol{\Theta} \le P_\mathrm{L}^\mathrm{max}, \\ P_\mathrm{G}^\mathrm{min} \le P_\mathrm{G} \le P_\mathrm{G}^\mathrm{max}. \end{cases} \quad (4)$$

The optimal power flow problem provides a building block for many other smart grid dispatch applications (Chatzos et al., 2020), and is almost the most popular formulation for smart grid dispatch.

2.3 Unit commitment

In economic dispatch and optimal power flow, it is assumed that there are N_G generating units working in the power system. Under this assumption for the economic dispatch problem and the optimal

power flow problem, we aim to find an optimal operating point for grids. Nevertheless, the optimal subset of the complete set of generating units remains undecided, which results in the problem of unit commitment (Sheble and Fahd, 1994; Padhy, 2004; Tejada-Arango et al., 2020).

A general formulation of unit commitment is as follows:

$$
\min \sum_t \left(c_1(P_{\mathrm{G},t}) + c_2(u_{\mathrm{G},t}) + c_3(\mathrm{su}_{\mathrm{G},t}) + c_4(\mathrm{sd}_{\mathrm{G},t}) \right)
$$

$$
\mathrm{s.t.} \begin{cases} \sum_G P_{\mathrm{G},t} - \sum_D P_{\mathrm{D},t} = 0, \\ u_{\mathrm{G},t} P_{\mathrm{G}}^{\min} \leq P_{\mathrm{G},t} \leq u_{\mathrm{G},t} P_{\mathrm{G}}^{\max}, \\ u_{\mathrm{G},t} = u_{\mathrm{G},t-1} + \mathrm{su}_{\mathrm{G},t} - \mathrm{sd}_{\mathrm{G},t}, \\ \sum_t \mathrm{su}_{\mathrm{G},t} \geq \mathrm{SU}_{\mathrm{G}}, \\ \sum_t \mathrm{sd}_{\mathrm{G},t} \geq \mathrm{SD}_{\mathrm{G}}, \end{cases}
$$

$$(5)$$

where $c_1(\cdot)$, $c_2(\cdot)$, $c_3(\cdot)$, and $c_4(\cdot)$ indicate the fixed cost, variable cost, startup cost, and shutdown cost of generating units, respectively. $u_{\mathrm{G},t}$, $\mathrm{su}_{\mathrm{G},t}$, and $\mathrm{sd}_{\mathrm{G},t}$ indicate the commitment decision, startup decision, and shutdown decision, respectively. The first constraint addresses the power balance constraints, the second constraint addresses the generator output constraints, the third constraint addresses the unit status, and the last two constraints address the minimum up- and down-time constraints.

Unit commitment is often used for day-ahead generation scheduling. The desired output is the commitment status, which is binary. As a result, the unit commitment problem is a mixed-integer optimization problem, which is more complicated than the economic dispatch problem.

2.4 Security-constrained dispatch

Contingencies are things that might happen to a system for various reasons, for example, natural disasters (Huang et al., 2017a, 2017b) or human attacks (Xiang et al., 2018). The above formulations, however, do not consider contingencies, which can lead to load shedding or even large blackouts when disruptions happen. As a result, it is necessary to take into account contingencies as constraints in smart grid dispatch problems.

Contingency analysis models single-failure events (e.g., one line fails) or multiple equipment failure events (e.g., two transmission lines fail at the same time) (Wood et al., 2013). According to

whether corrective actions are allowed when integrating contingencies in smart grid dispatch formulations, security-constrained dispatch problems can be categorized into preventive security-constrained dispatch and corrective security-constrained dispatch (Xu et al., 2014). Taking the contingencies of transmission line outages in DC optimal power flow problems as an example, the preventive security-constrained optimal power flow problem can be formulated as follows:

$$
\min c(P_{\mathrm{G}})
$$

$$
\mathrm{s.t.} \begin{cases} P_{\mathrm{G}} - P_{\mathrm{D}} - \boldsymbol{B\Theta} = 0, \\ -P_{\mathrm{L},c}^{\max} \leq \boldsymbol{X\Theta} \leq P_{\mathrm{L},c}^{\max}, \\ P_{\mathrm{G}}^{\min} \leq P_{\mathrm{G}} \leq P_{\mathrm{G}}^{\max}, \end{cases}
$$

$$(6)$$

where the subscript c indicates the set of contingencies that will be taken into account.

Similarly, we can present the formulation of corrective security-constrained optimal power flow problems as follows:

$$
\min c(P_{\mathrm{G},c})
$$

$$
\mathrm{s.t.} \begin{cases} P_{\mathrm{G},c} - P_{\mathrm{D}} - \boldsymbol{B\Theta}_c = 0, \\ -P_{\mathrm{L},c}^{\max} \leq \boldsymbol{X\Theta}_c \leq P_{\mathrm{L},c}^{\max}, \\ P_{\mathrm{G}}^{\min} \leq P_{\mathrm{G},c} \leq P_{\mathrm{G}}^{\max}. \end{cases}
$$

$$(7)$$

Note that contingencies will greatly increase the scale of the optimization problem, and therefore the solution of smart grid dispatch problems will be much harder.

3 Deep learning solutions

Many methods have been developed for smart grid dispatch problems. For example, Newton-based methods (Sun et al., 1984) and interior-point methods (Capitanescu and Wehenkel, 2013) have been thoroughly investigated and widely used. However, these methods have to repeatedly solve the challenging dispatch problems, which could hardly meet the fast-growing requirements of reliability and computational efficiency. As a result, in contrast to the above mathematical optimization methods, scholars began to investigate the application of deep learning methods for smart grids. After briefly introducing the deep learning basics in Section 3.1, we review the recent progress. This progress can be divided into

three categories, whose logic, along with the logic of traditional mathematical optimization methods, is summarized in Fig. 1. Herein, type-1, type-2, and type-3 correspond to end-to-end learning, constraint set learning, and iterative process learning, respectively, and they will be reviewed in Sections 3.2–3.4. Different methods provide different pipelines to derive dispatch decisions from a given system state. Our taxonomy of deep learning solutions is shown in Table 2.

3.1 Deep learning basics

We focus on deep learning, especially deep neural networks in this study. Let us use $\vec{x}$ to represent the input data of smart grid dispatch problems (e.g., power system state), and $\vec{y}$ to represent the output data (e.g., grid dispatch decision). The aim of deep learning methods is to derive an optimal model $f(\cdot)$ to capture the relationship between $\vec{x}$ and $\vec{y}$. In other words, the following problem will be minimized to learn the pattern between $\vec{x}$ and $\vec{y}$:

$$\mathcal{L}(f) = E\{\ell(\vec{y}^i, f(\vec{x}^i))\}. \tag{8}$$

Here, $\vec{x}^i$ and $\vec{y}^i$ represent the input and output parts of a given sample i. Function $\ell(\cdot)$ measures the difference between $f(\vec{x})$ and $\vec{y}$, and is also known as the loss function or empirical risk. For classification problems, we often use 0/1-loss, namely, $\ell(\vec{y}^i, f(\vec{x}^i)) = 0$ if $\vec{y}^i = f(\vec{x}^i)$ and $\ell(\vec{y}^i, f(\vec{x}^i)) = 1$ if $\vec{y}^i \neq f(\vec{x}^i)$. For regression problems, square-loss is often used, namely, $\ell(\vec{y}^i, f(\vec{x}^i)) = (\vec{y}^i - f(\vec{x}^i))^2$.

3.2 End-to-end learning

As illustrated in Fig. 1, smart grid dispatch end-to-end learning refers to training a complex machine learning model to directly learn the inputs and outputs of the original optimization problems. The power of the end-to-end learning philosophy has been demonstrated in many tasks, for example, image recognition (He et al., 2016), self-driving cars (Bojarski et al., 2016), and three-dimensional (3D) object detection (Zhou Y and Tuzel, 2018). Thus, the straightforward brute-force technique has been almost the most popular direction in deep learning applications. Typical works on end-to-end learning methods for smart grid dispatch problems are listed in Table 2.

Pan et al. (2019) developed a deep neural network approach called DeepOPF for the DC optimal power flow problem. The DeepOPF approach uses a multi-layer feed-forward neural network to learn the mapping between power demand and generator output. In this paper, generator limit constraints were guaranteed by the Sigmoid activation function, and a penalty function was proposed to deal with transmission line limit constraints. Because the line limits cannot be exactly satisfied, a post-processing procedure was activated when lines got overloaded. Note that power balance equations were there to derive line flow when power demands and power generators were derived.

Based on the above work, Pan et al. (2021a) extended the DeepOPF approach to the security-constrained DC optimal power flow problem. Their case studies showed that the proposed approach can
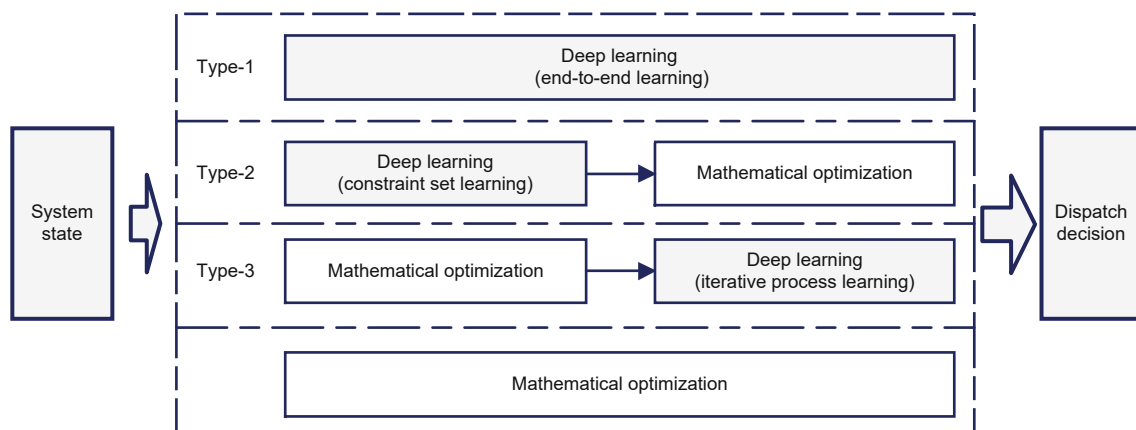

Fig. 1 Logic summary of deep learning solutions

Table 2 Taxonomy and representative publications on deep learning solutions

Category	Subcategory	Publications
End-to-end learning		Yin et al. (2018, 2020), Fioretto et al. (2019), Pan et al. (2019, 2021a, 2021b), Chatzos et al. (2020), Duchesne et al. (2020a), Owerko et al. (2020), Venzke et al. (2020a), Yan and Xu (2020), Zamzam and Baker (2020), Zhao et al. (2020), Zhou YH et al. (2020)
Constraint set learning	Identify constraints	Ardakani and Bouffard (2018), Deka and Misra (2019), Chen YZ and Zhang (2020), Yang et al. (2020b)
	Replace constraints	King et al. (2016), Liu TJ et al. (2020), Venzke et al. (2020b), Yang et al. (2020a), Hu et al. (2021)
Iterative process learning	Initialize solvers	Diehl (2019), Jamei et al. (2019), Biagioni et al. (2020), Chen LJ and Tate (2020)
	Replace solvers	Baker (2020), Robson et al. (2020)

speed up the computation by up to two orders of time, compared to the Gurobi solver (Gurobi Optimization, 2019). As a cost, 18.3% of solutions might be infeasible before the post-processing procedure in the worst case, and 0.2% optimality loss on average would be unavoidable.

Pan et al. (2021b) also investigated the Deep-OPF approach for the AC optimal power flow problem. Because explicit expressions of penalty gradients are unavailable, AC power flow equations make the widely used training methods (e.g., stochastic gradient descent (Le et al., 2011) and Adam (Kingma and Ba, 2015)) fail. Thus, a two-point zero-order optimization method was proposed in Pan et al. (2021b) to compute the estimated gradient during training. The effectiveness of the proposed method was verified through a small-scale test system, but evaluation of DeepOPF for medium- or large-scale AC optimal power flow problems still needs further research.

As an improvement of DeepOPF, DeepOPF+ was recently proposed (Zhao et al., 2020). It is also based on a multi-layer feed-forward neural network, but tries to deal with the feasibility issue that remains in previous methods. The idea is to calibrate the generation limits and transmission line limits during training dataset preparation. The benefit is that a post-processing procedure is no longer required to guarantee the generation limits and transmission line limits. As a result, further computation acceleration can be guaranteed. It was also reported that a larger calibration magnitude will lead to a larger optimality loss, while a higher feasibility rate can be achieved.

Encouraged by Pan et al. (2019), Zamzam and Baker (2020) investigated how to learn the mapping from system loading to optimal generation setpoints and voltages with deep neural networks for the AC optimal power flow problem. Similar to the limitation in Pan et al. (2019), operational constraints cannot be guaranteed directly by the deep learning method in Zamzam and Baker (2020). To alleviate this issue, a fast iterative procedure based on the power flow problem is required to recover a feasible solution. Note that transmission line constraints were ignored in Zamzam and Baker (2020).

To analyze the robustness of deep neural network methods used in the above works, Venzke et al. (2020a) introduced a framework based on mixed-integer linear programming to obtain worst-case guarantees for DC optimal power flow problems. Three conditions were considered: (1) maximum constraint violations, (2) maximum distances between the predicted and optimal decision variables, and (3) maximum distance between the predicted and optimal solution values. The results on a range of PGLib-OPF networks (Babaeinejadsarookolaee et al., 2021) up to 300 buses showed that the worst-case guarantees can be up to one order of magnitude larger than the empirical lower bounds by conventional methods. Nevertheless, the worst-case guarantees can be reduced to an acceptable level by training on a larger input domain.

Fioretto et al. (2019) explored a deep learning approach called OPF-DNN for highly efficient and accurate solutions of AC optimal power flow problems. The OPF-DNN method consists of a deep neural network architecture and Lagrangian duality. Here, Lagrangian duality was used to deal with the physical and engineering constraints. The

OPF-DNN method was verified on systems up to 300 buses and results showed that it is faster than the DC optimal power flow approximation solved with IPOPT (Wächter and Biegler, 2006) by at least two orders of magnitude.

The OPF-DNN method was revisited in Chatzos et al. (2020). The authors improved the algorithm design and training procedures, and through the case study in a French 3400-bus transmission system, the effectiveness of the proposed framework on large-scale power systems was verified. The biggest finding was that high-fidelity approximations of large-scale optimal power flow problems can be solved within milliseconds.

Because smart grids can be viewed as a graph, Owerko et al. (2020) proposed using graph neural networks in an imitation learning framework to solve AC optimal power flow problems. Graph neural networks (Wu ZH et al., 2021) belong to a class of neural networks that map the inputs to the output with consideration of the underlying graph. Case studies on the IEEE 30-bus system and IEEE 118-bus system were provided. The results showed that graph neural networks can lead to lower root mean squared errors.

Zhou YH et al. (2020) applied deep reinforcement learning to solve the AC optimal power flow problem. A deep neural network agent was trained by the state-of-the-art proximal policy optimization algorithm to derive an optimal policy for the reinforcement learning framework. To further increase the training speed and performance of the deep neural network agent, they used another independent neural network to learn a better initialized policy.

Yan and Xu (2020) applied a Lagrangian-based deep reinforcement learning method to solve real-time optimal power flow problems. The Lagrangian relaxation approach was proposed to deal with the operational constraints, and a policy model based on a deep deterministic policy gradient was used to update parameters of the deep neural network. Simulations were conducted to show that the proposed method can derive decisions continuously in real time for the IEEE 118-bus system.

Duchesne et al. (2020a) dealt with the day-ahead unit commitment and economic dispatch decision ranking problem with deep neural networks. This problem is different from the above problems because it belongs to the field of reliability management for smart grid dispatch. Running a very large number of security-constrained optimal power flow problems for all the day-ahead decisions is impractical, so a deep learning method was used to assess day-ahead decisions and rank them according to next-day operating costs. Multi-task learning was also used to provide more information about different cost values such as preventive redispatch cost, preventive load shedding cost, and preventive wind curtailment cost.

Yin et al. (2018) used a deep learning method to replace the traditional automatic generation control method. The method consists of two deep neural networks connected by a "selector," which was designed to determine the optimal next state in the next states predicted by the first neural network. A "relaxed operator" was used after the second neural network to ensure that the constraints of generating units are met. The feasibility of the proposed method was verified through the IEEE New-England power system, Hainan power grid, and IEEE 118-bus power system.

Yin et al. (2020) extended the work in Yin et al. (2018) to scenarios where the number of generators and the topology of systems are not fixed. An expandable deep learning method was proposed for dynamic expandable future smart grids. The proposed method consists of basic deep neural networks and expanding operation which makes the neural networks expandable. A 118-bus power system and a 13 659-bus system were used to verify the effectiveness of the proposed method.

3.3 Constraint set learning

While end-to-end learning is elegant and powerful, it may be inefficient (Glasmachers, 2017) and the operational constraints in smart grid dispatch problems can hardly be satisfied exactly. An alternative is to use deep learning to predict the constraint set for mathematical optimization models (Type-2 in Fig. 1). The constraint set learning method works as an intermediate step between deep learning and mathematical optimization. As a result, traditional algorithms such as interior-point methods and simplex methods can be significantly accelerated. In fact, some complex smart grid dispatch problems might even be downgraded to have analytic solutions when the filtered constraints are simple.

A hot field is active constraint set learning (Misra et al., 2019), where active constraints mean

the constraints that are satisfied with equality in the corresponding optimization problem. For example, Deka and Misra (2019) took advantage of deep neural network based classifiers to learn optimal active constraints from uncertainty realization. Given an uncertain injection scenario, the proposed classification method can derive the corresponding active constraints, which can then be used to calculate the results of DC optimal power flow problems. It was shown that the deep learning method is superior to the traditional ensemble policy method in both speed and performance.

Chen YZ and Zhang (2020) proposed a deep neural network to learn the mapping between power loads and the optimal system cost, but interpreted the gradient of neural networks as dual variables, which could be used to identify active constraints in the DC optimal power flow problem. The proposed method is error-correcting and proved to be an order of magnitude faster than the state-of-the-art linear programming solvers.

Yang et al. (2020b) proposed a deep learning based pre-screening method to identify active constraints of security-constrained economic dispatch problems. To deal with topology changes that might occur in a system, a fast tuning strategy based on transfer learning was also provided. The authors verified the effectiveness of their method on the IEEE 30-bus system, IEEE 118-bus system, and a practical 661-bus utility system.

Umbrella constraints are different from active constraints. They are the constraints that form the feasible set of solutions, which makes active constraints a subset of umbrella constraints in the same problems (Ardakani and Bouffard, 2018). Regarding the umbrella constraint set, deep learning is not yet applied for smart grids. A related work is Ardakani and Bouffard (2018), in which a single-layer neural network was applied to learn the optimal umbrella constraints for the security-constrained optimal power flow problem. The input is a specific demand scenario and the output is the umbrella set corresponding to that scenario. A customized training algorithm was proposed to replace the traditional backpropagation algorithm to avoid false negative errors. However, the proposed method has weaker learning ability than deep learning methods, and thus 2864 neural network models are required for 2864 constraints.

Both active constraint set learning and umbrella constraint set learning aim to identify effective constraints for mathematical optimization. In addition, deep learning methods can be used to replace some constraints for smart grid dispatch problems. For example, Yang et al. (2020a) and Hu et al. (2021) focused on replacing power flow equations with deep learning methods. Here, Yang et al. (2020a) constructed a deep neural network to learn the probabilistic power flow calculation for smart grid dispatch, and Hu et al. (2021) encoded physical knowledge of the grid to regularize neural networks to solve the overfitting problem.

Venzke et al. (2020b) focused on the more challenging AC optimal power flow problems with dynamic security constraints. The authors first used a deep neural network to solve the AC optimal power flow problem including dynamic security constraints. Then, they reformulated the nonlinear activation function with binary constraints and the deep neural network was transformed into mixed-integer constraints. Taking into account the above constraints in a mathematical optimization model, the authors tackled the issue of how to consider dynamic security constraints for optimal power flow problems.

Liu TJ et al. (2020) also investigated how to take advantage of deep learning methods to incorporate dynamic security constraints into optimal power flow problems. First, a Bayesian neural network was trained by the "Bayes by Backprop" method (Blundell et al., 2015) for confidential dynamic security assessment. Then, a security-constrained optimal power flow problem was reformulated by replacing dynamic security constraints with the above neural network model. Finally, the Bayesian optimization method was used to solve the smart grid dispatch model embedded with deep learning constraints.

King et al. (2016) focused on transient stability-constrained optimal power flow problems. The authors proposed to replace the dynamic and transient stability constraints with a critical clearing time constraint. Then, the critical clearing time constraint was approximated by a neural network model. As a result, dynamic simulations during traditional optimization methods were no longer required in the proposed method, and the computational burden of dynamic simulations was therefore removed.

The performances of constraint set learning and end-to-end learning were compared in a simple way

in Guha et al. (2019). The results on the IEEE 30-bus system and IEEE 118-bus system showed that the constraint learning method can be more effective. However, we need to note that a more thorough comparison is required to lead to a general conclusion.

3.4 Iterative process learning

Smart grid dispatch problems are traditionally formulated as mathematical optimization problems and then solved by an iterative process such as Newton-based methods (Sun et al., 1984) or interior-point methods (Capitanescu and Wehenkel, 2013). However, the iterative numerical methods might result in a failure of convergence due to a bad initialization point (Baker, 2019). Moreover, Newton-based methods have to calculate a full Jacobian matrix (e.g., Newton–Raphson method) or an approximate Jacobian matrix (e.g., quasi-Newton method) during iterations, which is time-consuming and might even be numerically unstable (Baker, 2020). To deal with the above issues, deep learning following the formulation of the mathematical optimization problem is also investigated.

Because a good initialization point can avoid poor local optimum and speed up the convergence process, Jamei et al. (2019) proposed a meta-optimizer to initialize interior-point solvers. Inspired by meta learning, the meta-optimizer is based on a neural network model that could be trained to learn the pattern between the dispatch problem formulation and initialization point. The loss function was defined as the number of iterations to reach an optimal solution. Experiments verified that the convergence process could be accelerated.

Based on the results from DC power flow calculation, Chen LJ and Tate (2020) proposed a one-dimensional convolutional neural network method to obtain initial values of voltage magnitude and phase values for Newton–Raphson-based AC power flow problems. Experiments on the IEEE 118-bus system and PEGASE 2869-bus system showed that the proposed method could achieve more than 30% reduction in computation time and 49.54%–66.47% reduction in solution iterations.

Diehl (2019) proposed using a graph neural network method to warm-start the AC optimal power flow solvers. Results on the Illinois 200-bus system showed that the proposed method would achieve 25% reduction in computation time compared to the

IPOPT solver. In fact, the proposed method would even be slightly faster than the IPOPT method warm-started by the DC optimal power flow.

Biagioni et al. (2020) focused on the distributed formulation (i.e., the alternating direction method of multipliers) of DC optimal power flow problems. The authors proposed a recurrent neural network based on the gated recurrent unit to predict the equilibrium values and accelerate the iterations. Case studies on the IEEE 14-bus system, IEEE 118-bus system, and RTE 2848-bus system verified the effectiveness of the proposed method.

Baker (2020) used a deep neural network with feedback to learn the Newton step as follows:

$$x^{k+1} = x^k - \alpha \boldsymbol{J}^{-1}(x^k)\boldsymbol{d}(x^k), \tag{9}$$

where x^k is the solution at iteration k, $\boldsymbol{J}^{-1}(x^k)$ is the Jacobian matrix at iteration k, $\boldsymbol{d}(x^k)$ is a vector of the Karush–Kuhn–Tucker (KKT) conditions at iteration k, and α is an optional step-size parameter. Case studies showed that the deep learning method can avoid singular Jacobian matrices and ill-conditioned matrices in Newton-based methods, and the computational efficiency improvement will be large for large power grids.

Robson et al. (2020) proposed a deep-learning-based method to iteratively solve optimal power flow problems. First, a neural network model was used to generate a reduced optimal power flow model by removing the predicted non-binding constraints. Then, the reduced model was extended by iteratively checking and adding violated constraints during iterations. The parameters of the neural network model were determined by minimizing the meta-loss function which is defined by the computational cost of the warm-stated optimal power flow problems. Because this loss function is not differentiable, particle swarm optimization rather than back-propagation was used to train the model.

4 Challenges

The fast development of smart grid and AI will bring about great opportunities. Much progress has been made in smart grid dispatch via deep learning methods; however, many issues in this field still require further research and some problems even remain untouched. Herein, we discuss several significant challenges based on our observations.

4.1 Data availability, security, and privacy

Deep learning is a data-driven method; so, datasets that could be used to train deep learning models and evaluate their performance are of vital importance (Venzke et al., 2021). In fact, to realize its greater learning ability, deep learning requires more data than traditional machine learning methods. However, while various public datasets are available for fields such as computer vision (Prabhu and Birhane, 2020) and natural language processing (Dror et al., 2017), very few electric power system datasets are publicly available at the present time (Ruan et al., 2021). Although we can easily find the data for many power networks (e.g., in Coffrin et al. (2019)), current research often has to run simulation tools (e.g., MATPOWER (Zimmerman RD et al., 2011)) to generate training samples for deep learning. This raises an obstacle for deep learning experts who are not familiar with how power grids work. The availability of power system datasets, especially real-world power system datasets, is still urgent.

Data security and privacy are other issues that often accompany deep learning methods (Wu C et al., 2019). A promising direction is distributed deep learning, especially federated learning (McMahan and Ramage, 2017; Kairouz et al., 2019; Li et al., 2020; Huang et al., 2021), which could enable collaborative machine learning while the raw data remain distributed. Another direction is to develop small sample learning techniques (Zhang et al., 2018), for example, few-shot learning (Ravi and Larochelle, 2016) and zero-shot learning (Changpinyo et al., 2016). In case we cannot gather enough data to construct a training dataset that we need, data augmentation (Shorten and Khoshgoftaar, 2019) can also help.

4.2 Model robustness and explainability

Deep learning methods have already shown their power in mastering games (Silver et al., 2017), quantifying behaviors (Mathis et al., 2018), and so forth. However, how significantly deep learning will impact real-world power systems still requires investigation. There is no doubt that deep learning methods, although vulnerable (Papernot et al., 2016) sometimes, can be widely used in fields that are not so security-sensitive. However, power systems are critical in-frastructure and are very sensitive, so robust deep learning methods are necessary. This issue is particularly challenging because the vulnerability of deep learning methods for power systems has been shown (Chen YZ et al., 2018; Venzke and Chatzivasileiadis, 2021).

Although traditional numerical methods for smart grid dispatch problems can sometimes fail to converge, deep learning methods will always generate a solution. This is not at all an advantage, however, because deep learning methods are widely viewed as black-box models, which means that we cannot explain the solution they provide. This will hinder the application of deep learning methods, especially in fields like smart grids. In general, there are two directions to alleviate the explainability issue: (1) interpreting the deep learning models with visualization (Chatzimparmpas et al., 2020) and (2) building interpretable deep learning models (Duchesne et al., 2020b).

5 Conclusions

Artificial intelligence has attracted extensive attention recently, and it is beginning to revolutionize power grids, a typical representation of critical infrastructure systems. As a core problem for power system operation, smart grid dispatch has been reviewed in this paper from the perspective of deep learning, which is, so far, the most advanced AI technique. We have provided our taxonomy of all the deep learning solutions which can be divided into three categories (end-to-end learning, constraint set learning, and iterative process learning). Representative works of each type have been carefully reviewed, and the main challenges regarding deep learning methods for smart grid dispatch have also been discussed. It can be concluded that much more effort is required to promote deep learning methods for smart grid dispatch, and we expect that this paper will provide a helpful review for researchers who are entering this interdisciplinary area of smart grids and AI.

Contributors

Gang HUANG drafted the paper. Gang HUANG, Fei WU, and Chuangxin GUO revised and finalized the paper.

Compliance with ethics guidelines

Gang HUANG, Fei WU, and Chuangxin GUO declare that they have no conflict of interest.

References

Ardakani AJ, Bouffard F, 2018. Prediction of umbrella constraints. Power Systems Computation Conf, p.1-7. https://doi.org/10.23919/PSCC.2018.8450586

Babaeinejadsarookolaee S, Birchfield A, Christie RD, et al., 2021. The power grid library for benchmarking AC optimal power flow algorithms. https://arxiv.org/abs/1908.02788

Baker K, 2019. Learning warm-start points for AC optimal power flow. IEEE 29th Int Workshop on Machine Learning for Signal Processing, p.1-6. https://doi.org/10.1109/MLSP.2019.8918690

Baker K, 2020. A learning-boosted quasi-Newton method for AC optimal power flow. Workshop on Machine Learning for Engineering Modeling, Simulation and Design, p.1-7.

Biagioni D, Graf P, Zhang XY, et al., 2020. Learning-accelerated ADMM for distributed DC optimal power flow. *IEEE Contr Syst Lett*, 6:1-6. https://doi.org/10.1109/LCSYS.2020.3044839

Blundell C, Cornebise J, Kavukcuoglu K, et al., 2015. Weight uncertainty in neural networks. Proc 32nd Int Conf on Machine Learning, p.1613-1622.

Bojarski M, Del Testa D, Dworakowski D, et al., 2016. End to end learning for self-driving cars. https://arxiv.org/abs/1604.07316v1

Bose BK, 2017. Artificial intelligence techniques in smart grid and renewable energy systems—some example applications. *Proc IEEE*, 105(11):2262-2273. https://doi.org/10.1109/JPROC.2017.2756596

Buchanan BG, 2005. A (very) brief history of artificial intelligence. *AI Mag*, 26(4):53-60.

Cambria E, White B, 2014. Jumping NLP curves: a review of natural language processing research. *IEEE Comput Intell Mag*, 9(2):48-57. https://doi.org/10.1109/MCI.2014.2307227

Capitanescu F, Wehenkel L, 2013. Experiments with the interior-point method for solving large scale optimal power flow problems. *Electr Power Syst Res*, 95:276-283. https://doi.org/10.1016/j.epsr.2012.10.001

Carpentier J, 1979. Optimal power flows. *Int J Electr Power Energy Syst*, 1(1):3-15. https://doi.org/10.1016/0142-0615(79)90026-7

Changpinyo S, Chao WL, Gong BQ, et al., 2016. Synthesized classifiers for zero-shot learning. Proc IEEE Conf on Computer Vision and Pattern Recognition, p.5327-5336. https://doi.org/10.1109/CVPR.2016.575

Chatzimparmpas A, Martins RM, Jusufi I, et al., 2020. A survey of surveys on the use of visualization for interpreting machine learning models. *Inform Visual*, 19(3):207-233. https://doi.org/10.1177/1473871620904671

Chatzos M, Fioretto F, Mak TWK, et al., 2020. High-fidelity machine learning approximations of large-scale optimal power flow. https://arxiv.org/abs/2006.16356

Chen LJ, Tate JE, 2020. Hot-starting the AC power flow with convolutional neural networks. https://arxiv.org/abs/2004.09342

Chen YZ, Zhang BS, 2020. Learning to solve network flow problems via neural decoding. https://arxiv.org/abs/2002.04091

Chen YZ, Tan YS, Deka D, 2018. Is machine learning in power systems vulnerable? IEEE Int Conf on Communications, Control, and Computing Technologies for Smart Grids, p.1-6. https://doi.org/10.1109/SmartGridComm.2018.8587547

Coffrin C, Gordon D, Scott P, 2019. NESTA, the NICTA energy system test case archive. https://arxiv.org/abs/1411.0359

Deka D, Misra S, 2019. Learning for DC-OPF: classifying active sets using neural nets. IEEE Milan PowerTech, p.1-6. https://doi.org/10.1109/PTC.2019.8810819

Diehl F, 2019. Warm-starting AC optimal power flow with graph neural networks. Proc 33rd Conf on Neural Information Processing Systems, p.1-6.

Dror R, Baumer G, Bogomolov M, et al., 2017. Replicability analysis for natural language processing: testing significance with multiple datasets. *Trans Assoc Comput Linguist*, 5:471-486. https://doi.org/10.1162/tacl_a_00074

Duchesne L, Karangelos E, Sutera A, et al., 2020a. Machine learning for ranking day-ahead decisions in the context of short-term operation planning. *Electr Power Syst Res*, 189:106548. https://doi.org/10.1016/j.epsr.2020.106548

Duchesne L, Karangelos E, Wehenkel L, 2020b. Recent developments in machine learning for energy systems reliability management. *Proc IEEE*, 108(9):1656-1676. https://doi.org/10.1109/JPROC.2020.2988715

Eskandarpour R, Khodaei A, 2017. Machine learning based power grid outage prediction in response to extreme events. *IEEE Trans Power Syst*, 32(4):3315-3316. https://doi.org/10.1109/TPWRS.2016.2631895

Fioretto F, Mak TWK, van Hentenryck P, 2019. Predicting AC optimal power flows: combining deep learning and Lagrangian dual methods. https://arxiv.org/abs/1909.10461

Gandhi O, Rodríguez-Gallegos CD, Srinivasan D, 2016. Review of optimization of power dispatch in renewable energy system. IEEE Innovative Smart Grid Technologies-Asia, p.250-257. https://doi.org/10.1109/ISGT-Asia.2016.7796394

Gharavi H, Ghafurian R, 2011. Smart grid: the electric energy system of the future. *Proc IEEE*, 99(6):917-921. https://doi.org/10.1109/JPROC.2011.2124210

Glasmachers T, 2017. Limits of end-to-end learning. *Proc Mach Learn Res*, 77:17-32.

Goodfellow I, Bengio Y, Courville A, et al., 2016. Deep Learning. MIT Press, Cambridge, USA.

Guha N, Wang ZC, Wytock M, et al., 2019. Machine learning for AC optimal power flow. Climate Change Workshop at Int Conf on Machine Learning, p.1-4.

Gurobi Optimization, 2019. Gurobi optimizer reference manual. Available from https://www.gurobi.com/wp-content/plugins/hd_documentations/documentation/9.0/refman.pdf [Accessed on Dec. 24, 2020].

Haridas AV, Marimuthu R, Sivakumar VG, 2018. A critical review and analysis on techniques of speech recognition: the road ahead. *Int J Knowl-Based Intell Eng Syst*, 22(1):39-57. https://doi.org/10.3233/KES-180374

Hasan F, Kargarian A, Mohammadi A, 2020. A survey on applications of machine learning for optimal power flow. IEEE Texas Power and Energy Conf, p.1-6. https://doi.org/10.1109/TPEC48276.2020.9042547

He KM, Zhang XY, Ren SQ, et al., 2016. Deep residual learning for image recognition. Proc IEEE Conf on Computer Vision and Pattern Recognition, p.770-778. https://doi.org/10.1109/CVPR.2016.90

Hey T, Tansley S, Tolle K, 2009. The Fourth Paradigm: Data-Intensive Scientific Discovery. Microsoft Research, Redmond, WA, USA.

Hossain E, Han Z, Poor HV, 2012. Smart Grid Communications and Networking. Cambridge University Press, Cambridge, UK.

Hu XY, Hu HJ, Verma S, et al., 2021. Physics-guided deep neural networks for power flow analysis. IEEE Trans Power Syst, 36(3):2082-2092. https://doi.org/10.1109/TPWRS.2020.3029557

Huang G, Wen YF, Bao YK, et al., 2015. Comprehensive decoupled risk-limiting dispatch. IEEE Power & Energy Society General Meeting, p.1-5. https://doi.org/10.1109/PESGM.2015.7286064

Huang G, Wang JH, Chen C, et al., 2017a. Integration of preventive and emergency responses for power grid resilience enhancement. IEEE Trans Power Syst, 32(6):4451-4463. https://doi.org/10.1109/Tpwrs.2017.2685640

Huang G, Wang JH, Chen C, et al., 2017b. System resilience enhancement: smart grid and beyond. Front Eng Manag, 4(3):271-282. https://doi.org/10.15302/J-FEM-2017030

Huang G, Wang JH, Chen C, et al., 2019. Cyber-constrained optimal power flow model for smart grid resilience enhancement. IEEE Trans Smart Grid, 10(5):5547-5555. https://doi.org/10.1109/TSG.2018.2885025

Huang G, Wu C, Hu YF, et al., 2021. Serverless distributed learning for smart grid analytics. Chinese Phys B, 30: 088802. https://doi.org/10.1088/1674-1056/abe232

Ibrahim MR, Haworth J, Cheng T, 2020. Understanding cities with machine eyes: a review of deep computer vision in urban analytics. Cities, 96:102481. https://doi.org/10.1016/j.cities.2019.102481

Jaller M, Otero-Palencia C, Pahwa A, 2020. Automation, electrification, and shared mobility in urban freight: opportunities and challenges. Transport Res Procedia, 46:13-20. https://doi.org/10.1016/j.trpro.2020.03.158

Jamei M, Mones L, Robson A, et al., 2019. Meta-optimization of optimal power flow. Climate Change Workshop at Int Conf on Machine Learning, p.1-3.

Kairouz P, McMahan HB, Avent B, et al., 2019. Advances and open problems in federated learning. https://arxiv.org/abs/1912.04977v3

King RTFA, Tu XP, Dessaint LA, et al., 2016. Multi-contingency transient stability-constrained optimal power flow using multilayer feedforward neural networks. IEEE Canadian Conf on Electrical and Computer Engineering, p.1-6. https://doi.org/10.1109/CCECE.2016.7726774

Kingma DP, Ba J, 2015. Adam: a method for stochastic optimization. Int Conf for Learning Representations, p.1-15.

Kundur P, 1994. Power System Stability and Control. McGraw-Hill, New York, USA.

Le QV, Ngiam J, Coates A, et al., 2011. On optimization methods for deep learning. Proc 28th Int Conf on Machine Learning, p.265-272.

LeCun Y, Bengio Y, Hinton G, 2015. Deep learning. Nature, 521(7553):436-444. https://doi.org/10.1038/nature14539

Li T, Sahu AK, Talwalkar A, et al., 2020. Federated learning: challenges, methods, and future directions. IEEE Signal Process Mag, 37(3):50-60. https://doi.org/10.1109/MSP.2020.2975749

Liu TJ, Liu YB, Liu JY, et al., 2020. A Bayesian learning based scheme for online dynamic security assessment and preventive control. IEEE Trans Power Syst, 35(5):4088-4099. https://doi.org/10.1109/TPWRS.2020.2983477

Liu ZF, 2020. Research on emergency response processing model of thermal power enterprise based on epidemic situation. Int Conf on Wireless Communications and Smart Grid, p.171-174. https://doi.org/10.1109/ICWCSG50807.2020.00045

Mai TT, Jadun P, Logan JS, et al., 2018. Electrification Futures Study: Scenarios of Electric Technology Adoption and Power Consumption for the United States. NREL/TP-6A20-71500, National Renewable Energy Laboratory, United States.

Mathis A, Mamidanna P, Cury KM, et al., 2018. DeepLabCut: markerless pose estimation of user-defined body parts with deep learning. Nat Neurosci, 21(9):1281-1289. https://doi.org/10.1038/s41593-018-0209-y

McMahan HB, Ramage D, 2017. Federated learning: collaborative machine learning without centralized training data. Google Research Blog. Available from https://ai.googleblog.com/2017/04/federated-learning-collaborative.html [Accessed on Dec. 24, 2020].

Misra S, Roald L, Ng Y, 2019. Learning for constrained optimization: identifying optimal active constraint sets. https://arxiv.org/abs/1802.09639

Mohamed MA, Eltamaly AM, 2018. Modeling and Simulation of Smart Grid Integrated with Hybrid Renewable Energy Systems. Springer, Switzerland.

Oughton EJ, Skelton A, Horne RB, et al., 2017. Quantifying the daily economic impact of extreme space weather due to failure in electricity transmission infrastructure. Space Wea, 15(1):65-83. https://doi.org/10.1002/2016SW001491

Owerko D, Gama F, Ribeiro A, 2020. Optimal power flow using graph neural networks. IEEE Int Conf on Acoustics, Speech and Signal Processing, p.5930-5934. https://doi.org/10.1109/ICASSP40776.2020.9053140

Padhy NP, 2004. Unit commitment—a bibliographical survey. IEEE Trans Power Syst, 19(2):1196-1205. https://doi.org/10.1109/TPWRS.2003.821611

Pan X, Zhao TY, Chen MH, 2019. DeepOPF: deep neural network for DC optimal power flow. IEEE Int Conf on Communications, Control, and Computing Technologies for Smart Grids, p.1-6. https://doi.org/10.1109/SmartGridComm.2019.8909795

Pan X, Zhao TY, Chen MH, et al., 2021a. DeepOPF: a deep neural network approach for security-constrained DC optimal power flow. IEEE Trans Power Syst,

36(3):1725-1735.
https://doi.org/10.1109/TPWRS.2020.3026379

Pan X, Chen MH, Zhao TY, et al., 2021b. DeepOPF: a feasibility-optimized deep neural network approach for AC optimal power flow problems.
https://arxiv.org/abs/2007.01002

Papernot N, McDaniel P, Jha S, et al., 2016. The limitations of deep learning in adversarial settings. IEEE European Symp on Security and Privacy, p.372-387.
https://doi.org/10.1109/EuroSP.2016.36

Prabhu VU, Birhane A, 2020. Large image datasets: a pyrrhic win for computer vision?
https://arxiv.org/abs/2006.16923

Rahman J, Feng C, Zhang J, 2020. Machine learning-aided security constrained optimal power flow. IEEE Power & Energy Society General Meeting, p.1-5.
https://doi.org/10.1109/PESGM41954.2020.9281941

Ravi S, Larochelle H, 2016. Optimization as a model for few-shot learning. Int Conf for Learning Representations, p.1-11.

Robson A, Jamei M, Ududec C, et al., 2020. Learning an optimally reduced formulation of OPF through meta-optimization. https://arxiv.org/abs/1911.06784

Ruan GC, Zhong HW, Zhang GL, et al., 2021. Review of learning-assisted power system optimization. *CSEE J Power Energy Syst*, 7(2):221-231.
https://doi.org/10.17775/CSEEJPES.2020.03070

Rudin C, Waltz D, Anderson RN, et al., 2012. Machine learning for the New York City power grid. *IEEE Trans Patt Anal Mach Intell*, 34(2):328-345.
https://doi.org/10.1109/TPAMI.2011.108

Sheble GB, Fahd GN, 1994. Unit commitment literature synopsis. *IEEE Trans Power Syst*, 9(1):128-135.
https://doi.org/10.1109/59.317549

Shorten C, Khoshgoftaar TM, 2019. A survey on image data augmentation for deep learning. *J Big Data*, 6(1):60.
https://doi.org/10.1186/s40537-019-0197-0

Silver D, Schrittwieser J, Simonyan K, et al., 2017. Mastering the game of go without human knowledge. *Nature*, 550(7676):354-359.
https://doi.org/10.1038/nature24270

Sun DI, Ashley B, Brewer B, et al., 1984. Optimal power flow by Newton approach. *IEEE Power Eng Rev*, PER-4(10):39. https://doi.org/10.1109/MPER.1984.5526285

Tejada-Arango DA, Lumbreras S, Sánchez-Martín P, et al., 2020. Which unit-commitment formulation is best? A comparison framework. *IEEE Trans Power Syst*, 35(4):2926-2936.
https://doi.org/10.1109/TPWRS.2019.2962024

Venzke A, Chatzivasileiadis S, 2021. Verification of neural network behaviour: formal guarantees for power system applications. *IEEE Trans Smart Grid*, 12(1):383-397.
https://doi.org/10.1109/TSG.2020.3009401

Venzke A, Qu GN, Low S, et al., 2020a. Learning optimal power flow: worst-case guarantees for neural networks. IEEE Int Conf on Communications, Control, and Computing Technologies for Smart Grids, p.1-7.
https://doi.org/10.1109/SmartGridComm47815.2020.9302963

Venzke A, Viola DT, Mermet-Guyennet J, et al., 2020b. Neural networks for encoding dynamic security-constrained optimal power flow to mixed-integer linear programs.
https://arxiv.org/abs/2003.07939

Venzke A, Molzahn DK, Chatzivasileiadis S, 2021. Efficient creation of datasets for data-driven power system applications. *Electr Power Syst Res*, 190:106614.
https://doi.org/10.1016/j.epsr.2020.106614

Wächter A, Biegler LT, 2006. On the implementation of an interior-point filter line-search algorithm for large-scale nonlinear programming. *Math Program*, 106(1):25-57.
https://doi.org/10.1007/s10107-004-0559-y

Walsh B, 2013. The surprisingly large energy footprint of the digital economy. Time Magazine. Available from https://science.time.com/2013/08/14/power-drain-the-digital-cloud-is-using-more-energy-than-you-think/ [Accessed on Dec. 24, 2020].

Wen GH, Yu XH, Liu ZW, 2021. Recent progress on the study of distributed economic dispatch in smart grid: an overview. *Front Inform Technol Electron Eng*, 22(1):25-39. https://doi.org/10.1631/FITEE.2000205

Wen YF, Li WY, Huang G, et al., 2016. Frequency dynamics constrained unit commitment with battery energy storage. *IEEE Trans Power Syst*, 31(6):5115-5125.
https://doi.org/10.1109/TPWRS.2016.2521882

Wood AJ, Wollenberg BF, Sheblé GB, 2013. Power Generation, Operation, and Control (3rd Ed.). John Wiley & Sons, New York, USA.

Wu C, Xiao J, Huang G, et al., 2019. Galaxy learning—a position paper. https://arxiv.org/abs/1905.00753

Wu F, Lu CW, Zhu MJ, et al., 2020. Towards a new generation of artificial intelligence in China. *Nat Mach Intell*, 2(6):312-316.
https://doi.org/10.1038/s42256-020-0183-4

Wu ZH, Pan SR, Chen FW, et al., 2021. A comprehensive survey on graph neural networks. *IEEE Trans Neur Netw Learn Syst*, 32(1):4-24.
https://doi.org/10.1109/TNNLS.2020.2978386

Xiang YM, Wang LF, Liu N, 2018. A robustness-oriented power grid operation strategy considering attacks. *IEEE Trans Smart Grid*, 9(5):4248-4261.
https://doi.org/10.1109/TSG.2017.2653219

Xu Y, Dong ZY, Zhang R, et al., 2014. Solving preventive-corrective SCOPF by a hybrid computational strategy. *IEEE Trans Power Syst*, 29(3):1345-1355.
https://doi.org/10.1109/TPWRS.2013.2293150

Yan ZM, Xu Y, 2020. Real-time optimal power flow: a Lagrangian based deep reinforcement learning approach. *IEEE Trans Power Syst*, 35(4):3270-3273.
https://doi.org/10.1109/TPWRS.2020.2987292

Yang Y, Yang ZF, Yu J, et al., 2020a. Fast calculation of probabilistic power flow: a model-based deep learning approach. *IEEE Trans Smart Grid*, 11(3):2235-2244.
https://doi.org/10.1109/TSG.2019.2950115

Yang Y, Yang ZF, Yu J, et al., 2020b. Fast economic dispatch in smart grids using deep learning: an active constraint screening approach. *IEEE Int Things J*, 7(11):11030-11040. https://doi.org/10.1109/JIOT.2020.2993567

Yin LF, Yu T, Zhang XS, et al., 2018. Relaxed deep learning for real-time economic generation dispatch and control with unified time scale. *Energy*, 149:11-23.
https://doi.org/10.1016/j.energy.2018.01.165

Yin LF, Gao Q, Zhao LL, et al., 2020. Expandable deep learning for real-time economic generation dispatch and control of three-state energies based future smart grids. *Energy*, 191:116561.
https://doi.org/10.1016/j.energy.2019.116561

Zamzam AS, Baker K, 2020. Learning optimal solutions for extremely fast AC optimal power flow. IEEE Int Conf on Communications, Control, and Computing Technologies for Smart Grids, p.1-6. https://doi.org/10.1109/SmartGridComm47815.2020.9303008

Zeng B, Ge SY, Kong XY, et al., 2014. Study for economic dispatch considering network loss in power pool market. Int Conf on Power System Technology, p.1754-1759. https://doi.org/10.1109/POWERCON.2014.6993569

Zhang DX, Han XQ, Deng CY, 2018. Review on the research and practice of deep learning and reinforcement learning in smart grids. *CSEE J Power Energy Syst*, 4(3):362-370. https://doi.org/10.17775/cseejpes.2018.00520

Zhao TY, Pan X, Chen MH, et al., 2020. DeepOPF+: a deep neural network approach for DC optimal power flow for ensuring feasibility. IEEE Int Conf on Communications, Control, and Computing Technologies for Smart Grids, p.1-6.

https://doi.org/10.1109/SmartGridComm47815.2020.9303017

Zhou Y, Tuzel O, 2018. VoxelNet: end-to-end learning for point cloud based 3D object detection. Proc IEEE Conf on Computer Vision and Pattern Recognition, p.4490-4499.

Zhou YH, Zhang B, Xu CL, et al., 2020. Deriving fast AC OPF solutions via proximal policy optimization for secure and economic grid operation. https://arxiv.org/abs/2003.12584v1

Zimmerman R, Zhu QY, Dimitri C, 2016. Promoting resilience for food, energy, and water interdependencies. *J Environ Stud Sci*, 6(1):50-61. https://doi.org/10.1007/s13412-016-0362-0

Zimmerman RD, Murillo-Sánchez CE, Thomas RJ, 2011. MATPOWER: steady-state operations, planning, and analysis tools for power systems research and education. *IEEE Trans Power Syst*, 26(1):12-19. https://doi.org/10.1109/TPWRS.2010.2051168