

Review of Learning-assisted Power System Optimization

Guangchun Ruan[✉], *Student Member, IEEE*, Haiwang Zhong[✉], *Senior Member, IEEE*, Guanglun Zhang[✉], *Student Member, IEEE*, Yiliu He[✉], *Student Member, IEEE*, Xuan Wang, *Member, IEEE*, and Tianjiao Pu, *Senior Member, IEEE*

Abstract—With dramatic breakthroughs in recent years, machine learning is showing great potential to upgrade the toolbox for power system optimization. Understanding the strength and limitation of machine learning approaches is crucial to decide when and how to deploy them to boost the optimization performance. This paper pays special attention to the coordination between machine learning approaches and optimization models, and carefully evaluates how such data-driven analysis may improve the rule-based optimization. The typical references are selected and categorized into four groups: the boundary parameter improvement, the optimization option selection, the surrogate model, and the hybrid model. This taxonomy provides a novel perspective to elaborate the latest research progress and development. We further compare the design patterns of different categories, and discuss several key challenges and opportunities as well. Deep integration between machine learning approaches and optimization models is expected to become the most promising technical trend.

Index Terms—Artificial intelligence, data-driven, deep learning, machine learning, neural network, smart grid.

I. INTRODUCTION

WITH the advanced computing systems and big data, machine learning has successfully entered a booming period in recent years [1]. Machines are proven to outperform humans in more and more applications, but most experts claim that machine learning is still fast developing and has not yet reached its peak.

Optimization is popular in tremendous applications of power system, and the typical optimization tasks include: optimal dispatch, planning, system identification, dynamic security, and electricity market operations. For these applications, the

major challenge is how to design an efficient and reliable method to meet the increasing requirements for optimization performance. Conventional optimization methods have shown their limitations in complicated and volatile environments, such as future power grids with a high penetration of renewable energy [2]. These conventional methods tend to repeatedly solve similar problems without accumulating any experience. Machine learning, in contrast, is powerful to gain experience from historical data and previous decisions [3]. Empirical studies have shown that the integration of machine learning and power system optimization is able to offer significant benefits [4].

A. Bibliometric Analysis

We conduct a bibliometric analysis on the publications that are indexed in the well-known “Web of Science” database. This is helpful to provide an overview of the research trend. Here, the searching query is formulated as follows:

```
TS=(("power system" OR "smart grid")
AND ("optimization" OR "optimal")
AND ("data-driven" OR "artificial
intelligence" OR "machine learning"
OR "deep learning" OR "reinforcement
learning" OR "neural network" OR
"support vector machine" OR "decision
tree"))).
```

Within the topics of interest, Fig. 1 shows the number of publications as well as the proportion among all publications. To derive this proportion, the number of publications on learning-assisted power system optimization is divided by the number of all power system optimization publications. To count the total number, the query expressions behind the second keyword “AND” are accordingly dropped.

As shown in Fig. 1, the publications of interest only account for a small proportion from 2012 to 2017, but in the next two years, the proportion grows rapidly to 6.93% and 8.61%, respectively. This observation indicates that, in recent years, increasing attention is given to machine learning applications in power system optimization, and the great potential of learning-assisted power system optimization is still under exploration.

Early researches made some preliminary attempts at Hopfield network [5], radial basis network [6], and self-organizing

Manuscript received June 3, 2020; revised September 1, 2020; accepted October 15, 2020. Date of online publication November 20, 2020; date of current version January 8, 2021. This work was supported in part by the National Key Research and Development Program of China (No. 2020YFB0905900) and National Natural Science Foundation of China (No. 51777102, No. U1766212).

G. C. Ruan, H. W. Zhong (corresponding author, e-mail: zhonghw@mail.tsinghua.edu.cn; ORCID: <https://orcid.org/0000-0001-9545-6243>), G. L. Zhang, Y. L. He and X. Wang are with the State Key Lab of Power Systems, Department of Electrical Engineering, Tsinghua University, Beijing 100084, China.

T. J. Pu is with China Electric Power Research Institute, Beijing 100192, China.

DOI: 10.17775/CSEEJPES.2020.03070

network [7], but after 2010, deep learning [8] attracted considerable attentions and gradually became the mainstream method.

This paper pays special attention to the latest developments, and we carefully select a few articles that are the most representative work in the latest three years—24 articles are published in the first half of 2020, and 38 articles are published in either 2018 or 2019. Moreover, these articles are mainly chosen from several prestigious journals in power system domain, including *IEEE Transactions on Power System*, *IEEE Transactions on Smart Grid*, and *Applied Energy*.

B. Comparison with Related Review Articles

This paper shows two essential differences from the existing review articles. First, a different taxonomy is developed in this paper to demonstrate and highlight the methodological features of existing literature. In comparison, existing reviews are more focused on different applications [9]–[14] or different learning methods [15].

Second, this paper limits the scope of references within power system optimization rather than a wide range of machine learning applications. More technical details and discussions are therefore provided in this paper to understand the most typical methodologies.

In summary, this paper intends to provide a novel taxonomy to understand how machine learning approaches may benefit the power system optimization. The selected references will be divided into different categories according to different methodological features. We further discuss the key challenges in practical applications, and recommend some potential solutions as well.

C. Contributions and Paper Structure

The major contributions are summarized as follows.

- 1) We propose a novel and well-designed taxonomy according to different methodological features. This taxonomy helps elaborate the coordination between machine learning approaches and power system optimization models. We also summarize the key technologies for each category from a number of recent publications.

- 2) Major challenges and opportunities in learning-assisted power system optimization are fully discussed, and the latest academic explorations are summarized to provide some possible solutions.

In the rest of this paper, Section II proposes a well-designed taxonomy, and the key technologies for each category are studied in Section III. We further discuss the major challenges and future opportunities in Section IV. At last, Section V concludes this paper.

II. TAXONOMY

This section provides a novel and well-designed taxonomy according to different methodological features. We intend to answer a key question that how the machine learning approaches may change the power system optimization. Here, the potential benefits and risks are both taken into consideration.

Specifically, we focus on the coordination between the machine learning approaches and the optimization models, and make four categories to extract valuable insights from the existing literature. Fig. 2 shows an overview of these four categories: a) **boundary parameter improvement**, b) **optimization option selection**, c) **surrogate model**, and d) **hybrid model**.

Let us revisit the power system optimization models from a full-cycle perspective. Generally, an optimization model contains an objective function and several constraints, whose coefficients (boundary parameters) are important for a high-quality solution. Classical optimization methods are widely used, and some optimization options, e.g., initial value or iterative step size, need to be effectively configured in advance. Additionally, the surrogate machine learning model and well-designed hybrid model are two alternative solvers which can boost the optimization performance by the experience learned from historical data.

Based on these discussions, the basic ideas of the proposed four categories are quite straightforward. The first category, boundary parameter improvement, uses machine learning approaches to improve coefficient estimation. The second category is focused on selecting better optimization options. The reinforcement learning and other special machine learning approaches are working as the surrogate models or the third category. The last category, the hybrid models, includes various hybrid analytical and data-driven frameworks. Note that the last two categories are technically different from the previous two—the coordination pattern changes from a tandem structure to some iterative, coupled or other complex structures. These new structures are more likely to perform better than using optimization or machine learning model alone.

We next analyze an example of demand response from the perspectives of different categories. Reference [16] (Category I) focused on the uncertainty issue of effective participation period, which was one of the critical boundary parameters. Reference [17] (Category II) estimated a better initial ON/OFF status of responsive appliances. While [18] (Category III) applied reinforcement learning for optimal control, [19] designed a hybrid framework to integrate neural network and

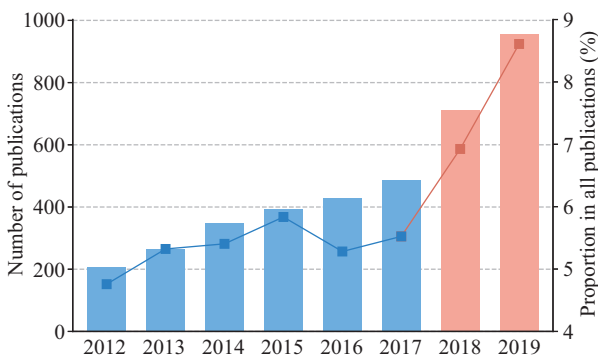


Fig. 1. Research trend of learning-assisted power system optimization in recent years. The number of publications is shown with a bar chart, and the associated proportion of all publications is plotted by a line chart. A significant increase can be observed in 2018 and 2019, indicating a promising research direction. Data Source: Web of Science.

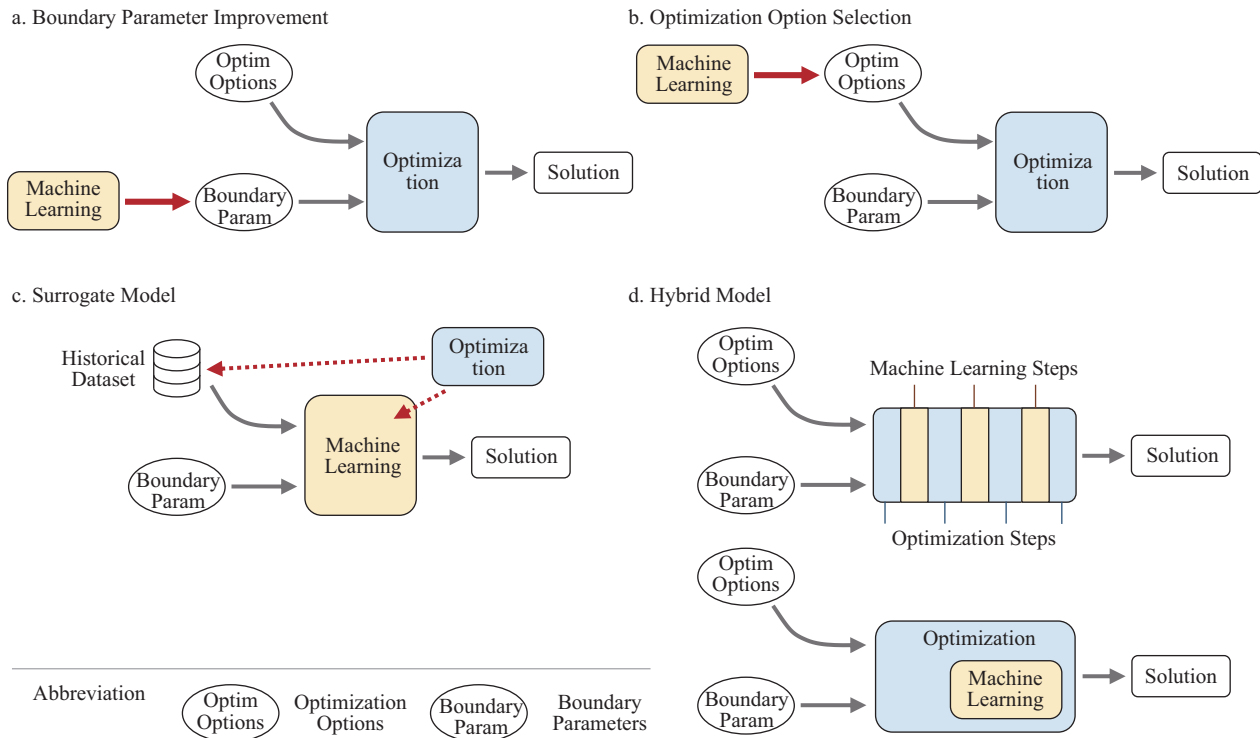


Fig. 2. Overview of the proposed taxonomy. This taxonomy is focused on the methodological features of existing literature, and the main ideas of each category is graphically illustrated. This novel taxonomy contributes to understanding the coordination between machine learning approaches and optimization models.

optimization models together. Both [18] and [19] achieved a significant algorithmic speedup.

Beyond forecasting, machine learning approaches are showing an attractive prospect for potential applications in power system optimization. The taxonomy of this paper clearly demonstrates the current academic progress, and becomes useful to design new-style models.

III. KEY TECHNOLOGIES

According to the proposed taxonomy, this section summarizes the latest research progress as well as key technologies. A list with all reference details can be downloaded from [20]. Fig. 3 gives an overview of the selected references. Here, different colors are assigned according to the application scenarios, and the height of each slice is proportional to the number of selected references. We will next dive into the technical details of each category.

A. Category 1. Boundary Parameter Improvement

This category uses machine learning approaches to improve the estimations of boundary parameters, and thus formulates more accurate feasible regions and better optimal solutions. There are many factors that may deteriorate the accuracy of boundary parameters, including natural variability, human behaviors, and partial observability of power systems.

1) Impact of Natural Variability

With high penetration of renewable energy, power systems are meeting the challenges of increasing uncertainty. Reference [21] applied the generative adversarial networks to

create scenarios of renewable energy, and this method was also applicable to other highly uncertain environments. Conditional generative adversarial networks performed better than classical methods to account for wind power uncertainty [22]. In [23], a deep learning approach was proposed to identify the fault types of transmission lines, and the features could be automatically learned from raw data. Reference [24] considered a wind and storage power plant that took part in a pool market. The authors combined the multivariate clustering technique and a recurrent neural network in order to model the uncertain electricity prices and wind power. Considering the socio-technical complexities in the planning stage, [25] developed a data-driven framework to analyze the distributed energy resources. In [26], the natural variability is captured by joint chance constraints, and the bounds are estimated using support vector machines.

2) Impact of Human Behaviors

Existing literature has used machine learning approaches to eliminate the randomness caused by human behaviors and obtain a better solution. Reference [27] proposed a long short-term memory neural network to adequately consider the uncertain prices in electricity markets. Neural networks played a significant role in balancing between the thermal comfort and energy use in buildings [28], achieving optimal dispatch in ancillary services market [29] and reduced load curtailments [16]. In [30], a data-driven model was used to increase the learning ability for price responsive behaviors. A game theory-based market strategy was designed to increase profit of each market participant with operational security guarantees. Reference [31] designed several local control strategies in

distribution networks by learning historical control schemes.

3) *Impact of Partial Observability*

Partial observability may result in inaccurate models and decisions. Deep learning models were deployed to identify time-varying parameter for composite load model [32], phasor measurement unit data manipulation attacks [33], phase detection in power distribution systems [34], and feature extraction for security rules [35]. To enhance the observability of distribution systems, [36] presented a data-driven method to estimate the daily consumption patterns of the customers without smart meters. In addition, [37] showed that generative adversarial networks could quickly assess the dynamic security with missing data.

B. *Category 2. Optimization Option Selection*

This category intends to boost the optimization performance by selecting better optimization options, e.g., initial values. Practice shows that some options have significant impacts on the convergence feature and speed, but the default settings are sometimes unsatisfactory. Machine learning approaches, in this context, can provide effective guidance from the past experiences.

Many researches use machine learning approaches to estimate a good initial value, which is beneficial for a warm-start algorithm. Reference [38] proposed a “learn to initialize” strategy to improve the Gauss-Newton algorithm. The authors achieved this with a neural network, and designed a special loss function (only penalizing the maximized errors) to improve the overall performance. Reference [39] established a predict-and-reconstruct approach to predicting the generation states (part of the decision variables), and reconstructed the phase angles (other part of the decision variables) using power flow equations. A deep neural network was developed for this task, and the network size was properly configured according to approximation accuracy. In [40], a data-driven approach to reconstructing the solution of a centralized optimal power flow was proposed. The idea was that local controllers could find a near-optimal solution by learning the limited but locally available data.

Some other extensions were discussed in [41]–[43]. The supervised and transfer learning were applied in [41] to estimate the Pareto front that is made up with a series of initial values. This task was indeed more difficult than [38]–[40]. The numerical tests indicated that such estimation might cause large errors under specific conditions, so further validation and fine-tuning were extremely important. Reference [42] proposed a linear power flow model to accelerate and approximate the power flow calculation. This work was further extended in [43] to tackle the challenge of hidden measurement noises. The authors formulated three quadratic programming models with several Jacobian-matrix-guided constraints to achieve this goal.

The potential advantages in discrete optimization are more attractive than expected. An early work [44] introduced a combined approach for unit commitment problems. This approach first used a neural network to determine the discrete variables, and after that, applied the simulated annealing method to optimize the continuous variables. Case studies showed that

the neural network found near-optimal commitment results, and achieved a roughly 50x computing speedup. A recent work [45] made further progress in security-constrained unit commitment. Different machine learning approaches have been adopted to study previously solved instances, and accelerate the computation by predicting redundant constraints, good initial feasible solutions, and active affine subspaces. On average, the authors achieved a 4.3x speedup with optimality guarantees and a 10.2x speedup without optimality guarantees (but with no discernible solution difference). It provided a valuable insight that predicting warm start states or active hyper-planes were significantly harder than estimating redundant constraints. Similar technique was also applicable in demand response. Reference [17] used a neural network to estimate the optimal ON/OFF status of home appliances, which could be regarded as an efficient warm start setting.

Machine learning approaches were also useful to configure other optimization options. Reference [46] created an effective algorithm selector by machine learning approaches and found less overloads in power flow management. Similar method was also adopted in a unit commitment problem [47] where a learning model was trained to assign weights according to some heuristic rules. Reference [48] formulated a three-stage framework (mid-term, short-term, and real-time) for outage scheduling, and a nearest neighboring classifier was trained to approximate the intermediate results to speed up the mid-term decision-making.

C. *Category 3. Surrogate Model*

This category seeks to completely replace optimization models by other data-driven models. These surrogate models are extremely powerful when analytical models are unavailable or too computationally expensive. We study the most prevalent method, reinforcement learning, as well as other specific methods in the following paragraphs.

Reinforcement learning is developed for sequential decision making that can be formulated as Markov decision process. It becomes more and more popular as a surrogate method for various complex optimization or control tasks. The basic idea is to construct an agent that takes actions to maximize the cumulative rewards, and the objective is to design an optimal policy for action selection. Many reinforcement learning algorithms have been developed with diverse reward estimations and policy designs. Additionally, deep reinforcement learning further enhanced the intelligent ability of agents to adapt to different applications. Fig. 4 shows the basic structure of (deep) reinforcement learning.

A major advantage of reinforcement learning is the model-free formulation that is independent of prior knowledge. Reference [3] used deep reinforcement learning to study the responsive behaviors of the microgrids in a distribution network. Reference [49] applied fitted Q-iteration to control a cluster of electro-thermal loads with unknown characteristics. A convolutional neural network was constructed to process the high-dimensional inputs and capture the hidden patterns. Reference [50] applied reinforcement learning to control a group of heterogeneous batteries, and their diverse physical characteristics were exploited to improve the overall efficiency.

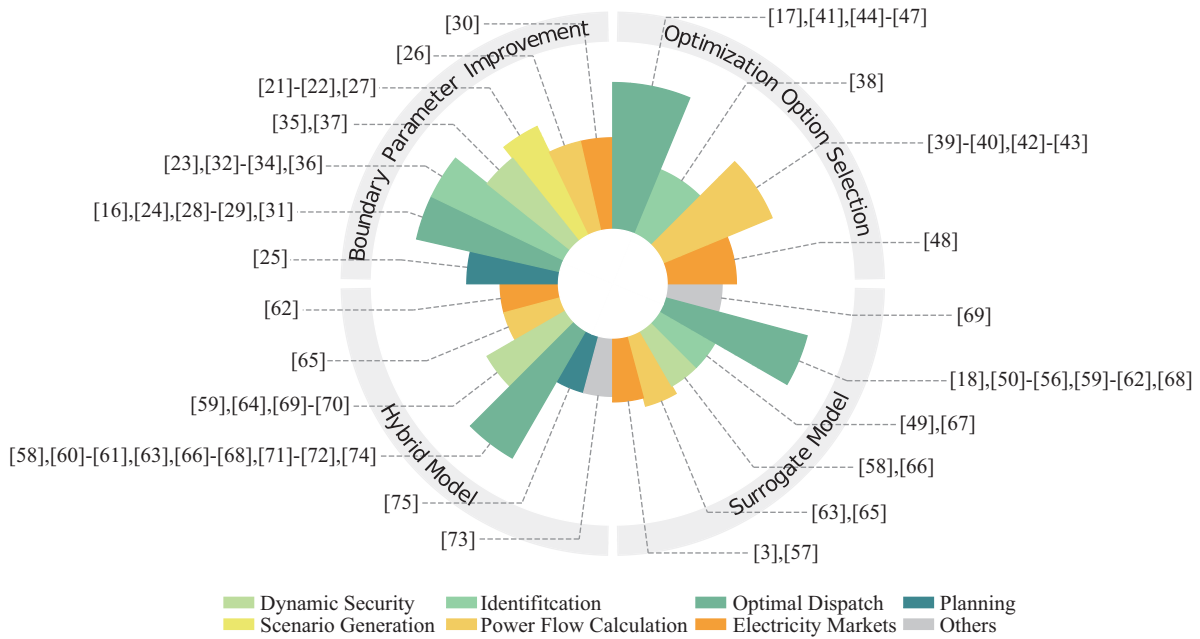


Fig. 3. Overview of all selected references in Section III. These references are classified by different application scenarios and grouped by the proposed taxonomy. Different colors are used for different applications, and the height of each slice is proportional to the number of selected references. This figure is useful for a quick search by applications or the proposed taxonomy.

A coordinated wide-area voltage controller was designed in [51] to optimize the transient process of an electric grid, showing a better dynamic stability than conventional methods. Reference [52] formulated an empirical model to approximate the degradation probabilities of grid components, and applied reinforcement learning to make maintenance decisions. Reference [53] used deep reinforcement learning to navigate the electric vehicles that need recharging, and the total travel time and cost were greatly minimized.

Another advantage of reinforcement learning is its online learning ability to make immediate response to the fluctuations of exogenous factors. For example, [54] solved online optimization to minimize energy cost and users' dissatisfaction with regard to the electricity prices and solar output. Reference [18] used an artificial neural network to predict day-ahead prices, and combined it with a Q-learning model for optimal dispatch of home energy system. Similar methods were also applied to microgrid scheduling, e.g., [55], [56].

Reinforcement learning is also useful in distributed operation or under information asymmetry. Reference [57] designed a multi-agent system to represent independent generation companies. A policy gradient method was applied to make decisions with limited market information. For current and voltage control, [58] integrate the consensus method and deep reinforcement learning to coordinate distributed generators in an island microgrid. The distributed reactive power optimization was solved by the collaborative equilibrium Q-learning to minimize operating cost and carbon emission [59]. Reference [60] proposed a multi-agent architecture to schedule electric vehicle charging based on Q-learning and W-learning.

A major concern of reinforcement learning is that the black box formulation can hardly consider physical characteristics or constraints. Large decision errors may happen in some

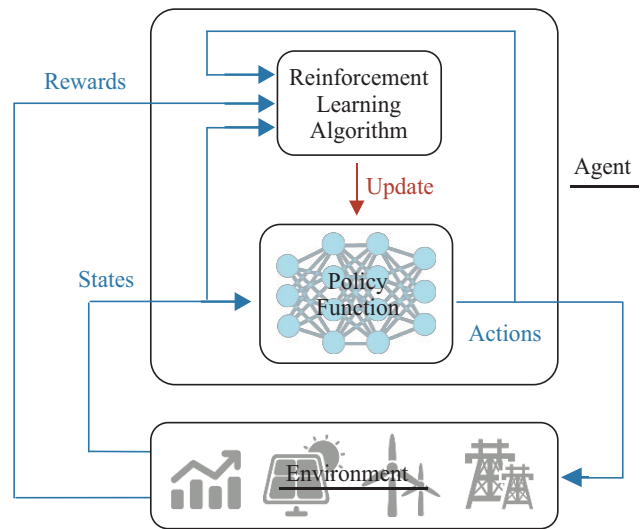


Fig. 4. Basic structure of reinforcement learning. The agent can observe the states of the environment, and take actions in response. After that, the environment gives back a reward according to the actions. The major objective of the agent is maximizing its total reward during a decision period.

unexpected cases because the performance of reinforcement learning heavily depends on the training data.

To solve this problem, safe reinforcement learning was developed to guarantee certain security constraints. The basic idea is to introduce some penalty terms corresponding to the security constraints, and minimize them in priority during the learning process. Reference [61] adopted this idea to consider charging constraints of electric vehicle batteries. Reference [62] optimized voltage and reactive power by a safe off-policy deep reinforcement learning algorithm to avoid voltage violations.

Some studies have also explored embedding physical char-

acteristics within deep learning methods. A physics-guided neural network approach was proposed in [63] to calculate probabilistic power flow. The training process was improved by the grid characteristics, and the case study showed a great computation speedup. Similar applications were completed by [64], [65] with a fully connected neural network and extreme learning machine, respectively. A graph convolutional network was trained in [66] to capture the topology information, and calculate the optimal load-shedding under contingency. Reference [67] used a convex neural network to approximate the electrothermal characteristics of a building, and then solve a convex optimization to minimize the electricity consumption. In the online control of plug-in electric bus [68], the trip information was represented by the length ratio, and this mapping relationship was learned by a neural network. Reference [69] studied the pump speed control to optimize the waste heat recovery of internal-combustion engines. The authors used the dynamic programming and supervised learning methods, whose inputs were specially designed according to several physical differential equations.

D. Category 4. Hybrid Model

This category combines the machine learning approaches and optimization models together to boost the overall performance. There are two hybrid types shown in Fig. 2(d), and more details are provided in Fig. 5 as well. Here, one type is an iterative structure with machine learning and optimization steps, and the other is a coupled structure that embeds machine learning models into optimization models to replace some inefficient parts. Such combinations are expected to provide a deep integration between two kinds of models, and fully explore their hidden potentials. Some case studies have achieved higher modeling accuracy and optimality property at the same time, e.g., [19], [70].

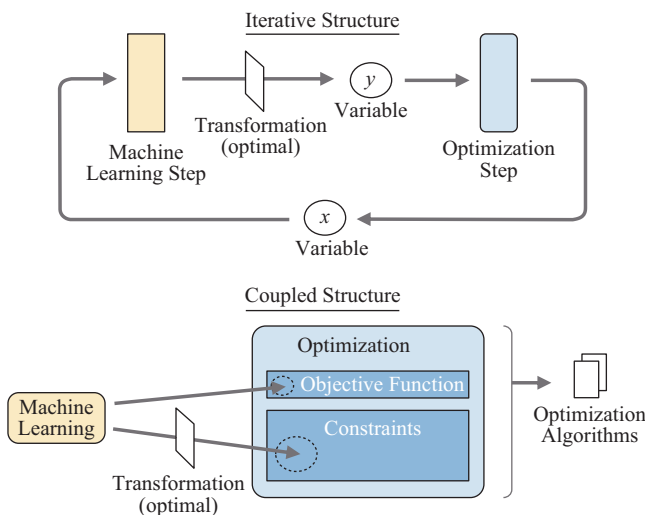


Fig. 5. Structure details of two typical hybrid models. An iterative structure takes learning steps and optimization steps alternately, while a coupled structure replaces certain inaccurate parts in an optimization model with machine learning formulations.

1) Iterative Structure

This structure is very popular in many optimization and optimal control tasks, and the typical flowchart is shown in Fig. 5. Reference [19] proposed an accelerated algorithm for distributed demand response, which applied a neural network to iteratively predict the consumers' price responses. The authors further developed a transformation model to search for better step sizes. The most promising feature of this algorithm was that it could cut off 60-80% iterations but still guarantee optimality property. In addition to step sizes, [71] showed that neural network was also able to improve searching directions of a sequential linear programming model. This paper considered an information asymmetric situation where a retailer company needed to make decisions with limited knowledge of its consumers. The dual neural networks were designed and one network was working to transform and derive searching directions.

Other intelligent optimization algorithms were also implemented in existing researches. Reference [72], [73] formulated a similar optimal dynamic pricing model for retailers but solved the model by the genetic algorithm and mean-variance mapping optimization, respectively. Here, the consumer responsive features were learned and coordinated with the intelligent optimization algorithms later. Reference [74] established a three-stage model predictive control where neural networks were predicting the energy demand and renewable energy supply to construct the optimal decisions. In [75], a hybrid model and data-driven simulation platform was operating in real-time for selecting the power system security features. Within this platform, the analytical models generated the samples, and they were then analyzed to extract several fine-tuned security rules.

2) Coupled Structure

This structure is more complex than the iterative structure mentioned above. As shown in Fig. 5, current researches made effort on embedding machine learning models in the objective function or constraints, and the key technical difficulty is how to design the optimization algorithms and transformation process. Reference [76] extracted a mapping function from the neural network formulation and included it as a constraint in an optimal power flow model. This hybrid model was then solved by a nonlinear programming optimizer. Reference [77], [78] integrated the machine learning models in the constraints of a building energy optimization model. To solve this model, [77] applied particle swarm optimization, while [79] applied a hybrid approach of exhaustive search method and subsequent quadratic programming. Similarly, the particle swarm optimization was also conducted in [80] to solve an optimization whose constraints were modeled by a radial basis function neural network. Reference [81] formulated a preventive control model with a Bayesian neural network that could predict the steady-states. This model was later solved by Bayesian optimization.

We next talk about some typical transformation measures. A piece-wise linear approximation transformed neural networks in [82] to finally derive a mixed integer programming model. Reference [70], [83] chose different machine learning ap-

proaches as an alternative. In [70], the sparse oblique decision tree was applied to learn some accurate, understandable, and linear security rules for economic dispatch. These rules could be embedded in an optimization as several mixed integer linear constraints. In [83], the authors conducted an extreme learning machine to enhance the hydrostatic tidal turbine control. This model was basically a linear model and can be easily transformed to linear constraints. Another special technique was introduced in [84], where the authors designed a sequential approximation method with dynamically trained neural networks. With the consideration of running time, such dynamic training might be more suitable for small networks.

Many researches have formulated objective functions with machine learning elements. Reference [85] trained a neural network to learn the combined heat and power simulation results, and formulated the objective function with this neural network integrated. The dispatch schedule was optimized by genetic algorithm. Reference [86] designed a convolutional-neural-network-based classifier for faulted line localization. The placement problem of phasor measurement units was a hybrid model, and the objective function contained the loss function of convolutional neural network. This problem was actually a special hyperparameter optimization, and was later solved by greedy algorithm.

E. Comparison and Comments

Overall, the selected references have very diverse ideas and features. We will compare all above categories and make further comments on their applications.

1) Difficulty of the Learning Tasks

Considering the same data dimension and precision requirement, the Category 3 contains the most difficult learning tasks. Here, the so-called “difficulty” can be measured by how large a machine learning model is needed to finish the task. The Category 2 usually takes the second place, but the situations for the remaining two categories are uncertain.

Regardless of some exceptions, the surrogate models (Category 3), especially those reinforcement learning models, are usually very complicated to calibrate, so a large amount of data and computing resources are needed. These models might meet great difficulty if the simulation data are not available or effective. Furthermore, the boundary parameter improvement (Category 1) is often easier to achieve, but its potential benefits are strongly related to the specific cases. As for the remaining two categories, some physical knowledge may be relatively helpful to boost the overall performance.

2) Difficulty of the Model Designs

For a similar task, the Category 4 or Category 2 methods are often in need of more efforts or even case-by-case designs. In contrast, the Category 3 is more likely to have some off-the-shelf tools, and little manual processing is needed.

There are two significant features in power system optimization—highly sensitive to decision errors, and dependent on extensive physical knowledge. However, leveraging these physical knowledge to design a dedicated framework or machine learning model is still under exploration. In this aspect, the hybrid model (Category 4) and other variants

in other categories deserve more exploration although more design efforts are needed.

IV. CHALLENGES AND OPPORTUNITIES

Power system optimization always has high requirements for the algorithm reliability because a potential failure could lead to a significant financial loss. However, machine learning is still having troubles in real-world power system applications, and the major concerns should be carefully analyzed and handled [87]. With this purpose, this section points out three major challenges, and shares some possible solutions as well.

A. Data Bottleneck

Collecting clean and reliable data is essential to every machine learning application, and the requirement for modern deep learning models is even higher. Two special features of power system data should be noted here: First, due to privacy concerns or confidentiality requirements [15], very few real-world data sets are public available, simulation data are therefore widely applied as an alternative [88]. Second, real-world data are often imbalanced, and the rare part may be extremely important, e.g., unstable system conditions [11].

Data issues show adverse impacts on all four categories, and thus become the major bottleneck for real-world applications. Apparently, it will bring more risks to those data-intensive models. In addition to expanding data sets with policy support, there are also some emerging technologies that could help tackle these data issues—data augmentation and few-shot learning.

Data augmentation is a strategy that significantly increases data volume by a series of transform operations. Reference [89] collected a list of useful resources, including classical techniques, papers and Github repositories. For example, the mainstream techniques for time-series data augmentation can be classified into simple operations (warping, jittering and perturbing) and advanced operations (embedding space and generative approaches).

Few-shot learning intends to feed machine learning models with limited amount of data. The basic idea is to use prior knowledge to avoid the unreliable performance of empirical risk minimizer. In this aspect, the authors of [90] reviewed some model-based methods (constrain the model complexity) and algorithm-based methods (constrain the search strategy for optimal parameters).

B. Robustness and Prediction Errors

Power system optimization imposes high requirements on accuracy as well as robustness. In this context, it is crucial to take care of the robust vulnerability—In extreme cases, a small input change may lead to a significant drop in accuracy.

There are two perspectives to analyze and understand the robustness issue:

- What is the output change of machine learning models with a fluctuation in model inputs?
- What is the change of optimal solution with a fluctuation in machine learning model outputs?

For Category 3, these two perspectives merge into one, while for other categories, the second perspective is important

but seldom discussed. We will next introduce some latest research explorations in the two perspectives.

From the first perspective, adversarial examples are fairly helpful to examine the robustness of machine learning approaches. Reference [91] studied the worst-case adversarial perturbations and found the robustness might be badly harmed. A recent work [92] argued that, in fact, these adversarial examples were features rather than bugs. The authors further demonstrated the robust and non-robust features, and the non-robust ones were the main cause for robust vulnerability.

The second perspective is highly related to the specific characteristics of optimization model. We can divide the model parameters and optimization options into two parts: error-tolerant parameters and error-sensitive parameters. Fig. 6 gives an illustrative example to show the difference. Our main focus is on the overall optimization performance when machine learning models make positive and negative prediction errors. It is shown that an error-tolerant parameter can robustly guarantee a shorter running time for optimization. For Category 1, 2 and 4, choosing error-tolerant parameters to coordinate between machine learning approaches and optimization models could improve the robustness of the whole system.

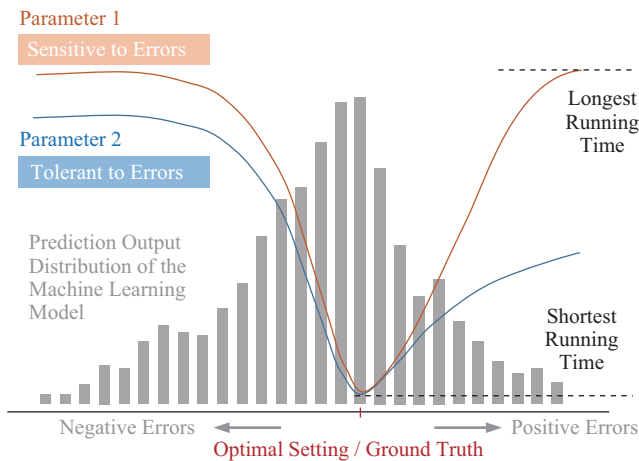


Fig. 6. Comparison between error-tolerant parameters and error-sensitive parameters. In this example, error-tolerant parameters can robustly ensure a shorter running time than error-sensitive parameters, even though their best performances are nearly the same.

C. Interpretability

Interpretability describes how much people can understand the decisions made by machine learning approaches. Many machine learning models, e.g., neural networks, are widely regarded as “black box” models [15]. Ensemble learning makes the interpretability even worse by combining several black box models together. As a consequence, low interpretability severely hinders the wide applications of machine learning approaches in the power industry.

There are two perspectives of the interpretability:

- How are the optimal weights found by machine learning models?
- How can machine learning model outputs improve the optimization performance?

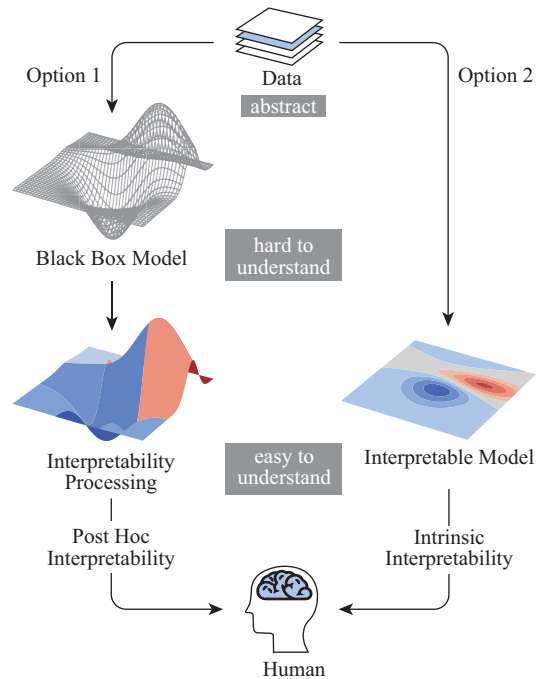


Fig. 7. Two options to achieve machine learning interpretability. These two options, applying an interpretable model or making further processing on a black box model, are illustrated to translate the original data to some easy-to-understand explanations.

Similar as the previous subsection, for Category 3, these two perspectives merge into one, while for Category 1, the second perspective is very intuitive—closer to the ground truth is better. In the remaining two categories, both two perspectives are important and should be carefully discussed.

The first perspective is a conventional topic that have been discussed for years in machine learning community. Reference [93] introduced most of the important progresses in this domain which are also shown in Fig. 7. Basically, there are two options to achieve interpretability: applying an interpretable model or making further processing on a black box model. Model-agnostic interpretation methods, with rapid advances in recent years, are able to extract many human-friendly features and visualization results.

The second perspective is dedicated for learning-assisted optimization, and can be probably demonstrated with the help of optimization theory. There often exist a set of optimal configurations to boost the optimization performance, but they may be computing-intensive or even intractable. Machine learning approaches, however, can approximate these configurations efficiently. A typical example is the selection of near-optimal step sizes in [19].

V. CONCLUSION

This paper conducts a comprehensive review of learning-assisted power system optimization. A novel and well-designed taxonomy is proposed in this paper to categorize the existing articles by their methodological features. The latest research progress and key technologies are thoroughly summarized, along with further comments on the key challenges in real-world applications.

We strongly realize that the deep integration of machine learning and power system optimization is a promising future trend. This review is expected to offer some useful information and new insights in this domain.

REFERENCES

- [1] D. X. Zhang, X. Q. Han, and C. Y. Deng, "Review on the research and practice of deep learning and reinforcement learning in smart grids," *CSEE Journal of Power and Energy Systems*, vol. 4, no. 3, pp. 362–370, Sep. 2018.
- [2] Y. Wang, N. Zhang, Q. X. Chen, D. S. Kirschen, P. Li, and Q. Xia, "Data-driven probabilistic net load forecasting with high penetration of behind-the-meter PV," *IEEE Transactions on Power Systems*, vol. 33, no. 3, pp. 3255–3264, May 2018.
- [3] Q. Z. Zhang, K. Dehghanpour, Z. Y. Wang, and Q. H. Huang, "A learning-based power management method for networked microgrids under incomplete information," *IEEE Transactions on Smart Grid*, vol. 11, no. 2, pp. 1193–1204, Mar. 2020.
- [4] J. X. Wang, H. W. Zhong, X. W. Lai, Q. Xia, Y. Wang, and C. Q. Kang, "Exploring key weather factors from analytical modeling toward improved solar power forecasting," *IEEE Transactions on Smart Grid*, vol. 10, no. 2, pp. 1417–1427, Mar. 2019.
- [5] T. Yalcinoz and M. J. Short, "Neural networks approach for solving economic dispatch problem with transmission capacity constraints," *IEEE Transactions on Power Systems*, vol. 13, no. 2, pp. 307–313, May 1998.
- [6] B. C. Paucar, J. L. R. Ortiz, K. S. L. Collazos, L. C. Leite, and J. O. P. Pinto, "Power operation optimization of photovoltaic stand alone system with variable loads using fuzzy voltage estimator and neural network controller," in *Proceedings of 2007 IEEE Lausanne Power Tech*, 2007, pp. 544–548.
- [7] M. Wu and P. Rastgoufard, "Optimum decision by artificial neural networks for reactive power control equipment to enhance power system stability and security performance," in *Proceedings of IEEE Power Engineering Society General Meeting*, 2004, pp. 2120–2125.
- [8] E. D. Varga, S. F. Beretka, C. Noce, and G. Sapienza, "Robust real-time load profile encoding and classification framework for efficient power systems operation," *IEEE Transactions on Power Systems*, vol. 30, no. 4, pp. 1897–1904, Jul. 2015.
- [9] T. Yang, L. Y. Zhao, and C. S. Wang, "Review on application of artificial intelligence in power system and integrated energy system," *Automation of Electric Power Systems*, vol. 43, no. 1, pp. 2–14, May 2019.
- [10] Z. D. Zhang, D. X. Zhang, and R. C. Qiu, "Deep reinforcement learning for power system applications: an overview," *CSEE Journal of Power and Energy Systems*, vol. 6, no. 1, pp. 213–225, Mar. 2020.
- [11] Q. Y. Sun and L. X. Yang, "From independence to interconnection—a review of AI technology applied in energy systems," *CSEE Journal of Power and Energy Systems*, vol. 5, no. 1, pp. 21–34, Mar. 2019.
- [12] M. S. Ibrahim, W. Dong, and Q. Yang, "Machine learning driven smart electric power systems: current trends and new perspectives," *Applied Energy*, vol. 272, pp. 115237, Aug. 2020.
- [13] Y. H. Zhang, R. Qiu, F. Yang, S. W. Xu, X. Shi, and X. He, "Overview of application of deep learning with image data and spatio-temporal data of power grid," *Power System Technology*, vol. 43, no. 6, pp. 1865–1873, Jun. 2019.
- [14] A. P. Marugán, F. P. G. Márquez, J. M. P. Perez, and D. Ruiz-Hernández, "A survey of artificial neural network in wind energy systems," *Applied Energy*, vol. 228, pp. 1822–1836, Oct. 2018.
- [15] L. F. Cheng, T. Yu, X. S. Zhang, and L. F. Yin, "Machine learning for energy and electric power systems: state of the art and prospects," *Automation of Electric Power Systems*, vol. 43, no. 1, pp. 15–31, May 2019.
- [16] M. Kamruzzaman, M. Benidris, and S. Commuri, "An artificial neural network based approach to electric demand response implementation," in *Proceedings of 2018 North American Power Symposium*, 2018, pp. 1–5.
- [17] M. S. Ahmed, A. Mohamed, H. Shareef, R. Z. Homod, and J. Abd Ali, "Artificial neural network based controller for home energy management considering demand response events," in *Proceedings of 2016 International Conference on Advances in Electrical, Electronic and Systems Engineering*, 2016, pp. 506–509.
- [18] R. Z. Lu, S. H. Hong, and M. M. Yu, "Demand response for home energy management using reinforcement learning and artificial neural network," *IEEE Transactions on Smart Grid*, vol. 10, no. 6, pp. 6629–6639, Nov. 2019.
- [19] G. C. Ruan, H. W. Zhong, J. X. Wang, Q. Xia, and C. Q. Kang, "Neural-network-based Lagrange multiplier selection for distributed demand response in smart grid," *Applied Energy*, vol. 264, pp. 114636, Apr. 2020.
- [20] G. Ruan, H. Zhong, G. Zhang, Y. He, X. Wang, and T. Pu. Supplementary spreadsheet. [Online]. 2020. Available: <https://github.com/GuangchunRuan/MLOPT>.
- [21] Y. Z. Chen, Y. S. Wang, D. Kirschen, and B. S. Zhang, "Model-free renewable scenario generation using generative adversarial networks," *IEEE Transactions on Power Systems*, vol. 33, no. 3, pp. 3265–3275, May 2018.
- [22] X. C. Dong, Y. Y. Sun, and T. J. Pu, "Day-ahead scenario generation of renewable energy based on conditional GAN," *Proceedings of the CSEE*, vol. 40, no. 17, pp. 5527–5535, Sep. 2020.
- [23] S. W. Xu, C. M. Qiu, D. X. Zhang, X. He, L. Chu, and H. S. Yang, "A deep learning approach for fault type identification of transmission line," *Proceedings of the CSEE*, vol. 39, no. 1, pp. 65–74, Jan. 2019.
- [24] J. L. Crespo-Vazquez, C. Carrillo, E. Diaz-Dorado, J. A. Martinez-Lorenzo, and M. Noor-E-Alam, "A machine learning based stochastic optimization framework for a wind and storage power plant participating in energy pool market," *Applied Energy*, vol. 232, pp. 341–357, Dec. 2018.
- [25] R. K. Jain, J. J. Qin, and R. Rajagopal, "Data-driven planning of distributed energy resources amidst socio-technical complexities," *Nature Energy*, vol. 2, no. 8, pp. 17112, Jul. 2017.
- [26] K. Baker and A. Bernstein, "Joint chance constraints in AC optimal power flow: improving bounds through learning," *IEEE Transactions on Smart Grid*, vol. 10, no. 6, pp. 6376–6385, Nov. 2019.
- [27] 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 Transactions on Power Systems*, vol. 35, no. 4, pp. 3040–3050, Jul. 2020.
- [28] T. Chaudhuri, Y. C. Soh, H. Li, and L. H. Xie, "A feedforward neural network based indoor-climate control framework for thermal comfort and energy saving in buildings," *Applied Energy*, vol. 248, pp. 44–53, Aug. 2019.
- [29] B. Canizes, J. Soares, P. Faria, and Z. Vale, "Mixed integer non-linear programming and artificial neural network based approach to ancillary services dispatch in competitive electricity markets," *Applied Energy*, vol. 108, pp. 261–270, Aug. 2013.
- [30] T. G. Lu, Z. Y. Wang, J. H. Wang, Q. Ai, and C. Wang, "A data-driven Stackelberg market strategy for demand response-enabled distribution systems," *IEEE Transactions on Smart Grid*, vol. 10, no. 3, pp. 2345–2357, May 2019.
- [31] 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 Transactions on Smart Grid*, vol. 10, no. 6, pp. 6461–6471, Nov. 2019.
- [32] M. J. Cui, M. Khodayar, C. Chen, X. N. Wang, Y. Zhang, and M. E. Khodayar, "Deep learning-based time-varying parameter identification for system-wide load modeling," *IEEE Transactions on Smart Grid*, vol. 10, no. 6, pp. 6102–6114, Nov. 2019.
- [33] J. Y. Wang, D. Y. Shi, Y. H. Li, J. F. Chen, H. F. Ding, and X. Z. Duan, "Distributed framework for detecting PMU data manipulation attacks with deep autoencoders," *IEEE Transactions on Smart Grid*, vol. 10, no. 4, pp. 4401–4410, Jul. 2019.
- [34] B. Foggo and N. P. Yu, "Improving supervised phase identification through the theory of information losses," *IEEE Transactions on Smart Grid*, vol. 11, no. 3, pp. 2337–2346, May 2020.
- [35] M. Y. Sun, I. Konstantelos, and G. Strbac, "A deep learning-based feature extraction framework for system security assessment," *IEEE Transactions on Smart Grid*, vol. 10, no. 5, pp. 5007–5020, Sep. 2019.
- [36] Y. X. Yuan, K. Dehghanpour, F. K. Bu, and Z. Y. Wang, "A multi-timescale data-driven approach to enhance distribution system observability," *IEEE Transactions on Power Systems*, vol. 34, no. 4, pp. 3168–3177, Jul. 2019.
- [37] C. Ren and Y. Xu, "A fully data-driven method based on generative adversarial networks for power system dynamic security assessment with missing data," *IEEE Transactions on Power Systems*, vol. 34, no. 6, pp. 5044–5052, Nov. 2019.
- [38] A. S. Zamzam, X. Fu, and N. D. Sidiropoulos, "Data-driven learning-based optimization for distribution system state estimation," *IEEE Transactions on Power Systems*, vol. 34, no. 6, pp. 4796–4805, Nov. 2019.
- [39] X. Pan, T. Y. Zhao, M. H. Chen, and S. Y. Zhang, "DeepOPF: a deep neural network approach for security-constrained DC optimal power flow," arXiv preprint arXiv:1910.14448, 2019.

- [40] 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 Transactions on Smart Grid*, vol. 11, no. 2, pp. 1296–1306, Mar. 2020.
- [41] A. T. D. Perera, P. U. Wickramasinghe, V. M. Nik, and J. L. Scartezzini, "Machine learning methods to assist energy system optimization," *Applied Energy*, vol. 243, pp. 191–205, Jun. 2019.
- [42] Y. X. Liu, N. Zhang, Y. Wang, J. W. Yang, and C. Q. Kang, "Data-driven power flow linearization: a regression approach," *IEEE Transactions on Smart Grid*, vol. 10, no. 3, pp. 2569–2580, May 2019.
- [43] Y. X. Liu, Y. Wang, N. Zhang, D. Lu, and C. Q. Kang, "A data-driven approach to linearize power flow equations considering measurement noise," *IEEE Transactions on Smart Grid*, vol. 11, no. 3, pp. 2576–2587, May 2020.
- [44] R. Nayak and J. D. Sharma, "A hybrid neural network and simulated annealing approach to the unit commitment problem," *Computers & Electrical Engineering*, vol. 26, no. 6, pp. 461–477, Aug. 2000.
- [45] A. S. Xavier, F. Qiu, and S. Ahmed, "Learning to solve large-scale security-constrained unit commitment problems," arXiv preprint arXiv:1902.01697, 2019.
- [46] J. E. King, S. C. E. Jupe, and P. C. Taylor, "Network state-based algorithm selection for power flow management using machine learning," *IEEE Transactions on Power Systems*, vol. 30, no. 5, pp. 2657–2664, Sep. 2015.
- [47] K. P. V. Doan and K. P. Wong, "Shapes: a novel approach for learning search heuristics in under-constrained optimization problems," *IEEE Transactions on Knowledge and Data Engineering*, vol. 9, no. 5, pp. 731–746, Sep./Oct. 1997.
- [48] G. Dalal, E. Gilboa, S. Mannor, and L. Wehenkel, "Chance-constrained outage scheduling using a machine learning proxy," *IEEE Transactions on Power Systems*, vol. 34, no. 4, pp. 2528–2540, Jul. 2019.
- [49] B. J. Claessens, P. Vrancx, and F. Ruelens, "Convolutional neural networks for automatic state-time feature extraction in reinforcement learning applied to residential load control," *IEEE Transactions on Smart Grid*, vol. 9, no. 4, pp. 3259–3269, Jul. 2018.
- [50] X. Qiu, T. A. Nguyen, and M. L. Crow, "Heterogeneous energy storage optimization for microgrids," *IEEE Transactions on Smart Grid*, vol. 7, no. 3, pp. 1453–1461, May 2016.
- [51] R. Yousefian and S. Kamalasadani, "Design and real-time implementation of optimal power system wide-area system-centric controller based on temporal difference learning," *IEEE Transactions on Industry Applications*, vol. 52, no. 1, pp. 395–406, Jan./Feb. 2016.
- [52] R. Rocchetta, L. Bellani, M. Compare, E. Zio, and E. Patelli, "A reinforcement learning framework for optimal operation and maintenance of power grids," *Applied Energy*, vol. 241, pp. 291–301, May 2019.
- [53] T. Qian, C. C. Shao, X. L. Wang, and M. Shahidehpour, "Deep reinforcement learning for EV charging navigation by coordinating smart grid and intelligent transportation system," *IEEE Transactions on Smart Grid*, vol. 11, no. 2, pp. 1714–1723, Mar. 2020.
- [54] E. Mocanu, D. C. Mocanu, P. H. Nguyen, A. Liotta, M. E. Webber, M. Gibescu, and J. G. Slootweg, "On-line building energy optimization using deep reinforcement learning," *IEEE Transactions on Smart Grid*, vol. 10, no. 4, pp. 3698–3708, Jul. 2019.
- [55] V. H. Bui, A. Hussain, and H. M. Kim, "Double deep Q-learning-based distributed operation of battery energy storage system considering uncertainties," *IEEE Transactions on Smart Grid*, vol. 11, no. 1, pp. 457–469, Jan. 2020.
- [56] S. Y. Zhou, Z. J. Hu, W. Gu, M. Jiang, and X. P. Zhang, "Artificial intelligence based smart energy community management: a reinforcement learning approach," *CSEE Journal of Power and Energy Systems*, vol. 5, no. 1, pp. 1–10, Mar. 2019.
- [57] Y. J. Ye, D. W. Qiu, J. Li, and G. Strbac, "Multi-period and multi-spatial equilibrium analysis in imperfect electricity markets: a novel multi-agent deep reinforcement learning approach," *IEEE Access*, vol. 7, pp. 130515–130529, Sep. 2019.
- [58] Z. F. Liu, Y. Luo, R. Q. Zhuo, and X. L. Jin, "Distributed reinforcement learning to coordinate current sharing and voltage restoration for islanded DC microgrid," *Journal of Modern Power Systems and Clean Energy*, vol. 6, no. 2, pp. 364–374, Sep. 2018.
- [59] M. Tan, C. J. Han, X. S. Zhang, L. X. Guo, and T. Yu, "Hierarchically correlated equilibrium Q-learning for multi-area decentralized collaborative reactive power optimization," *CSEE Journal of Power and Energy Systems*, vol. 2, no. 3, pp. 65–72, Sep. 2016.
- [60] F. L. Da Silva, C. E. H. Nishida, D. M. Roijers, and A. H. R. Costa, "Coordination of electric vehicle charging through multiagent reinforcement learning," *IEEE Transactions on Smart Grid*, vol. 11, no. 3, pp. 2347–2356, May 2020.
- [61] H. P. Li, Z. Q. Wan, and H. B. He, "Constrained EV charging scheduling based on safe deep reinforcement learning," *IEEE Transactions on Smart Grid*, vol. 11, no. 3, pp. 2427–2439, May 2020.
- [62] W. Wang, N. P. Yu, Y. Q. Gao, and J. Shi, "Safe off-policy deep reinforcement learning algorithm for volt-VAR control in power distribution systems," *IEEE Transactions on Smart Grid*, vol. 11, no. 4, pp. 3008–3018, Jul. 2020.
- [63] Y. Yang, Z. F. Yang, J. Yu, B. S. Zhang, Y. Q. Zhang, and H. X. Yu, "Fast calculation of probabilistic power flow: a model-based deep learning approach," *IEEE Transactions on Smart Grid*, vol. 11, no. 3, pp. 2235–2244, May 2020.
- [64] M. X. Xiang, J. Yu, Z. F. Yang, Y. Yang, H. X. Yu, and H. He, "Probabilistic power flow with topology changes based on deep neural network," *International Journal of Electrical Power & Energy Systems*, vol. 117, pp. 105650, May 2020.
- [65] X. Y. Lei, Z. F. Yang, J. Yu, J. B. Zhao, Q. Gao, and H. X. Yu, "Data-driven optimal power flow: a physics-informed machine learning approach," *IEEE Transactions on Power Systems*, 2020, to be published.
- [66] C. Kim, K. Kim, P. Balaprakash, and M. Anitescu, "Graph convolutional neural networks for optimal load shedding under line contingency," in *2019 IEEE Power & Energy Society General Meeting*, 2019, pp. 1–5.
- [67] Y. Z. Chen, Y. Y. Shi, and B. S. Zhang, "Optimal control via neural networks: a convex approach," in *International Conference on Learning Representations*, 2019. arXiv:1805.11835.
- [68] H. Tian, Z. W. Lu, X. Wang, X. L. Zhang, Y. Huang, and G. Y. Tian, "A length ratio based neural network energy management strategy for online control of plug-in hybrid electric city bus," *Applied Energy*, vol. 177, pp. 71–80, Sep. 2016.
- [69] B. Xu, D. Rathod, A. Yebi, and Z. Filipi, "Real-time realization of dynamic programming using machine learning methods for IC engine waste heat recovery system power optimization," *Applied Energy*, vol. 262, pp. 114514, Mar. 2020.
- [70] Q. C. Hou, N. Zhang, D. S. Kirschen, E. S. Du, Y. H. Cheng, and C. Q. Kang, "Sparse oblique decision tree for power system security rules extraction and embedding," arXiv preprint arXiv:2004.09579, 2020.
- [71] G. C. Ruan, H. W. Zhong, Q. Xia, Q. F. Huang, and C. Zhou, "Embed neural network in optimization model: an application of demand response aggregation under information asymmetry," in *Proceedings of 2019 IEEE Power & Energy Society General Meeting*, 2019.
- [72] F. L. Meng and X. J. Zeng, "A profit maximization approach to demand response management with customers behavior learning in smart grid," *IEEE Transactions on Smart Grid*, vol. 7, no. 3, pp. 1516–1529, May 2016.
- [73] T. Holschneider and I. Erlich, "Optimization of electricity pricing considering neural network based model of consumers' demand response," in *Proceedings of 2013 IEEE Computational Intelligence Applications in Smart Grid*, 2013, pp. 154–160.
- [74] J. Reynolds, M. W. Ahmad, Y. Rezgui, and J. L. Hippolyte, "Operational supply and demand optimisation of a multi-vector district energy system using artificial neural networks and a genetic algorithm," *Applied Energy*, vol. 235, pp. 699–713, Feb. 2019.
- [75] T. N. Huang, Q. L. Guo, H. B. Sun, N. Y. Zhao, B. Wang, and W. X. Guo, "Hybrid model and data driven concepts for power system security feature selection and knowledge discovery: key technologies and engineering application," *Automation of Electric Power Systems*, vol. 43, no. 1, pp. 95–101, 208, May 2019.
- [76] V. J. Gutierrez-Martinez, C. A. Cañizares, C. R. Fuente-Esquivel, A. Pizano-Martinez, and X. P. Gu, "Neural-network security-boundary constrained optimal power flow," *IEEE Transactions on Power Systems*, vol. 26, no. 1, pp. 63–72, Feb. 2011.
- [77] D. Zhang, S. H. Li, M. Sun, and Z. O'Neill, "An optimal and learning-based demand response and home energy management system," *IEEE Transactions on Smart Grid*, vol. 7, no. 4, pp. 1790–1801, Jul. 2016.
- [78] S. Y. Yang, M. P. Wan, W. Y. Chen, B. F. Ng, and S. Dubey, "Model predictive control with adaptive machine-learning-based model for building energy efficiency and comfort optimization," *Applied Energy*, vol. 271, pp. 115147, Aug. 2020.
- [79] B. Zhou, K. W. Chan, T. Yu, and C. Y. Chung, "Equilibrium-inspired multiple group search optimization with synergistic learning for multiobjective electric power dispatch," *IEEE Transactions on Power Systems*, vol. 28, no. 4, pp. 3534–3545, Nov. 2013.
- [80] E. M. Voumvoulakis and N. D. Hatzigryiouri, "A particle swarm optimization method for power system dynamic security control," *IEEE Transactions on Power Systems*, vol. 25, no. 2, pp. 1032–1041, May 2010.
- [81] T. J. Liu, Y. B. Liu, J. Y. Liu, L. F. Wang, L. X. Xu, G. Qiu, and H. J. Gao, "A Bayesian learning based scheme for online dynamic

- security assessment and preventive control,” *IEEE Transactions on Power Systems*, vol. 35, no. 5, pp. 4088–4099, Sep. 2020.
- [82] Y. J. Kim, “A supervised-learning-based strategy for optimal demand response of an HVAC system in a multi-zone office building,” *IEEE Transactions on Smart Grid*, vol. 11, no. 5, pp. 4212–4226, Sep. 2020.
- [83] X. X. Yin and X. W. Zhao, “Sensorless maximum power extraction control of a hydrostatic tidal turbine based on adaptive extreme learning machine,” *IEEE Transactions on Sustainable Energy*, vol. 11, no. 1, pp. 426–435, Jan. 2020.
- [84] Y. L. Hsu, S. G. Wang, and C. C. Yu, “A sequential approximation method using neural networks for engineering design optimization problems,” *Engineering Optimization*, vol. 35, no. 5, pp. 489–511, Oct. 2003.
- [85] M. J. Kim, T. S. Kim, R. J. Flores, and J. Brouwer, “Neural-network-based optimization for economic dispatch of combined heat and power systems,” *Applied Energy*, vol. 265, pp. 114785, May 2020.
- [86] W. T. Li, D. Deka, M. Chertkov, and M. Wang, “Real-time faulted line localization and PMU placement in power systems through convolutional neural networks,” *IEEE Transactions on Power Systems*, vol. 34, no. 6, pp. 4640–4651, Nov. 2019.
- [87] C. Rudin, D. Waltz, R. N. Anderson, A. Boulanger, A. Sallab-Aouissi, M. Chow, H. Dutta, P. N. Gross, B. Huang, S. Jerome, D. F. Isaac, A. Kressner, R. J. Passonneau, A. Radeva, and L. Wu, “Machine learning for the new york city power grid,” *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 34, no. 2, pp. 328–345, Feb. 2012.
- [88] M. J. Li, H. Z. Tao, H. Q. Xu, J. B. Liu, Q. Zhang, and W. Zhang, “The technical framework and application prospect of artificial intelligence application in the field of power grid dispatching and control,” *Power System Technology*, vol. 44, no. 2, pp. 393–400, Feb. 2020.
- [89] A. Mikolajczyk. List of useful data augmentation resources. Available: <https://github.com/AgaMiko/data-augmentation-review> [Online]. 2020.
- [90] Y. Q. Wang, Q. M. Yao, J. T. Kwok, and L. M. Ni, “Generalizing from a few examples: a survey on few-shot learning,” *ACM Computing Surveys*, vol. 53, no. 3, pp. 63, Jun. 2020.
- [91] L. Rice, E. Wong, and J. Z. Kolter, “Overfitting in adversarially robust deep learning,” arXiv preprint arXiv:2002.11569, 2020.
- [92] A. Ilyas, S. Santurkar, D. Tsipras, L. Engstrom, B. Tran, and A. Madry, “Adversarial examples are not bugs, they are features,” in *Advances in Neural Information Processing Systems*, 2019, pp. 125–136.
- [93] C. Molnar, *Interpretable Machine Learning: A Guide for Making Black Box Models Explainable*, Lulu Press, 2019.



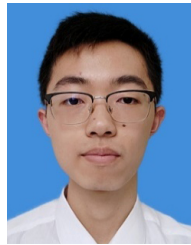
Guangchun Ruan received the B.S. degree in Electrical Engineering from Tsinghua University, Beijing, China, in 2016, where he is currently pursuing the Ph.D. degree. He is a visiting student with University of Washington, Seattle, WA, USA, in 2019, and with Texas A&M University, College Station, TX, USA, in 2020. His research interests include electricity markets, demand response, and data science.



Haiwang Zhong received the B.S. and Ph.D. degrees in Electrical Engineering from Tsinghua University, where he is currently an Associate Professor with the Department of Electrical Engineering. He was a recipient of the ProSPER.Net Young Scientist Award in 2014, the Outstanding Ph.D. Graduate and Outstanding Post-Doctor of Tsinghua University. He serves as the Editor for the *International Transactions on Electrical Energy Systems*, an Associate Editor for the *Journal of Modern Power System and Clean Energy*, and the *CSEE Journal of Power and Energy Systems*. He currently serves as the Vice-Chair of the Demand Response Working Group under the IEEE PES Power System Operations, Planning and Economics Committee, the Secretary of the IEEE PES SBLC Architecture Subcommittee. His research interests include smart grids, power system operations, electricity markets, and demand response.



Guanglun Zhang received the B.S. degree in Electrical Engineering from Tsinghua University, Beijing, China in 2018, where he is currently pursuing the Ph.D. degree. His research interests include electricity markets, energy and electricity economics, and power system under extreme weather events.



Yiliu He is currently working toward the B.S. degree in Electrical Engineering in Tsinghua University, Beijing, China. His research interests include power system operations and demand-side management.



Xuan Wang received the Ph.D. degree in Electrical Engineering from North China Electric Power University, Beijing, China, in 2020. She is currently the Post-Doctor in the Department of Electrical Engineering from Tsinghua University. Her research interests include power system operation, electricity market, and risk assessment of power grid disaster.



Tianjiao Pu received the B.S. degree from the Department of Electric Power Systems and Automation of Tianjin University in 1997 and has worked ever since at the China Electric Power Research Institute. He is a Professorate Senior Engineer, the director of the Artificial Intelligence Application Research Department of China Electric Power Research Institute (CEPRI), a core member of the “Science and Technology Tackling Team of Energy-saving Economic Dispatching” in the State Grid Corporation of China, a Fellow of IET, a senior member of IEEE, a senior member of the Chinese Society for Electrical Engineering, a CIGRE member, Secretary General of the AI Committee of CSEE, and a Deputy Editor-in-Chief of IET Smart Grid. He has been engaged in the research and management of power dispatching automation, smart grid simulation, active distribution network, artificial intelligence and other fields.