



AI-Empowered Methods for Smart Energy Consumption: A Review of Load Forecasting, Anomaly Detection and Demand Response

Xinlin Wang^{1,2,3} · Hao Wang^{4,5} · Binayak Bhandari⁶ · Leming Cheng¹

Received: 5 January 2022 / Revised: 4 July 2023 / Accepted: 5 July 2023 / Published online: 23 September 2023
© The Author(s) 2023

Abstract

This comprehensive review paper aims to provide an in-depth analysis of the most recent developments in the applications of artificial intelligence (AI) techniques, with an emphasis on their critical role in the demand side of power distribution systems. This paper offers a meticulous examination of various AI models and a pragmatic guide to aid in selecting the suitable techniques for three areas: load forecasting, anomaly detection, and demand response in real-world applications. In the realm of load forecasting, the paper presents a thorough guide for choosing the most fitting machine learning and deep learning models, inclusive of reinforcement learning, in conjunction with the application of hybrid models and learning optimization strategies. This selection process is informed by the properties of load data and the specific scenarios that necessitate forecasting. Concerning anomaly detection, this paper provides an overview of the merits and limitations of disparate learning methods, fostering a discussion on the optimization strategies that can be harnessed to navigate the issue of imbalanced data, a prevalent concern in power system anomaly detection. As for demand response, we delve into the utilization of AI techniques, examining both incentive-based and price-based demand response schemes. We take into account various control targets, input sources, and applications that pertain to their use and effectiveness. In conclusion, this review paper is structured to offer useful insights into the selection and design of AI techniques focusing on the demand-side applications of future energy systems. It provides guidance and future directions for the development of sustainable energy systems, aiming to serve as a cornerstone for ongoing research within this swiftly evolving field.

Keywords Power distribution system · Load forecasting · Anomaly detection · Demand response · Artificial intelligence · Machine learning · Deep learning · Reinforcement learning

1 Introduction

The urgent challenge facing our society is the decarbonization of the energy system to mitigate the impact of climate change and achieve a net-zero carbon future. Climate change is affecting over 25,000 species, pushing them towards

extinction, as reported in numerous studies [1–3]. Additionally, traditional energy sources, such as oil and gas, are expected to be depleted by 2050 [2], while the demand for energy continues to increase. In fact, non-OECD economic growth is expected to increase energy demand by over 30% by 2050 [4].

In this context, it is crucial to increase the use of renewable energy sources (RESs) in future energy systems to meet

This paper is an invited paper (Invited Review)

✉ Hao Wang
hao.wang2@monash.edu

✉ Binayak Bhandari
binayak.bhandari@unsw.edu.au

Xinlin Wang
xinlin.wang@csiro.au

Leming Cheng
Lemingc@uci.edu

¹ Donald Bren School of Information and Computer Sciences, University of California, Irvine, CA, USA

² Department of Mechanical and Aerospace Engineering, Seoul National University, Seoul, South Korea

³ Energy, CSIRO, Newcastle, Australia

⁴ Department of Data Science and Artificial Intelligence, Faculty of Information Technology, Monash University, Melbourne, Australia

⁵ Monash Energy Institute, Monash University, Melbourne, Australia

⁶ School of Mechanical and Manufacturing Engineering, University of New South Wales, Sydney, Australia

the rising energy demand and decarbonize the energy sector. Figure 1 illustrates the historical global power consumption data from 1800 to 2019 [5], showing that the proportion of energy provided by RESs is increasing rapidly, although it still remains relatively small [6]. According to the International Energy Agency’s Electricity Market Report 2023, RESs together with nuclear, will on average meet more than 90% of the increase in global demand by 2025 [7].

Although RESs are becoming more prevalent, integrating them directly into current power grids designed for dispatching conventional power generation presents two critical challenges [8–10].

- RESs exhibit intermittency, making it difficult to predict their generation and manage the associated risks of power imbalances and blackouts.
- Current energy systems and markets lack adequate mechanisms to integrate sustainable renewable energy for required emission controls in major decarbonization efforts.

Artificial intelligence (AI) has successfully solved many real-world problems in computer vision and natural language

processing, and is promising in addressing energy challenges. AI can enable a comprehensive framework for effective power system control, management, energy market pricing, and policy recommendations [11–15]. Machine learning (ML) and deep learning (DL) models have been widely employed to optimize energy efficiency, conversion, distribution, and decarbonization in smart grids, showcasing the potential of data-driven approaches in this field [10, 11, 13–15]. These models provide timely feedback, enabling efficient two-way communication between the grid and customers and significantly enhancing the security, reliability, and efficiency of the system [4, 16–18]. With AI, a smart grid can optimize renewable resource utilization, balance electricity production and consumption, improve grid reliability, and ensure security [19]. Smart grid applications have grown rapidly in recent years, with a growing market share [13]. Figure 2 illustrates a typical smart grid structure.

1.1 AI Techniques on Demand Side

The demand side, or consumption side, is one of the crucial parts of future smart energy systems. It’s expected to facilitate low-carbon and net-zero development as energy

Fig. 1 Global power consumption and renewable energy usage

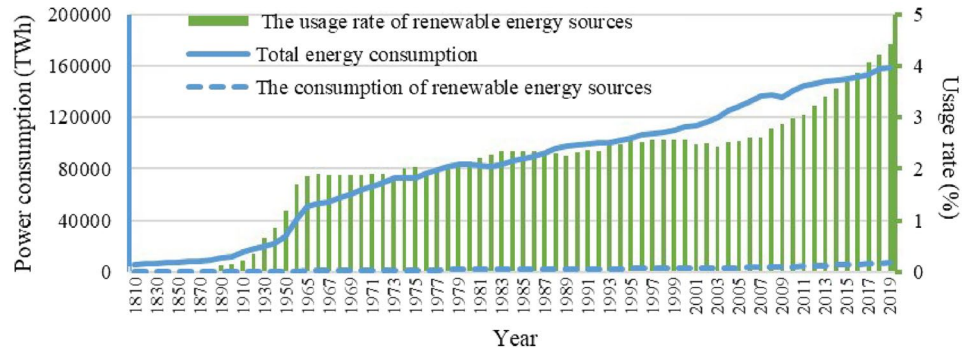
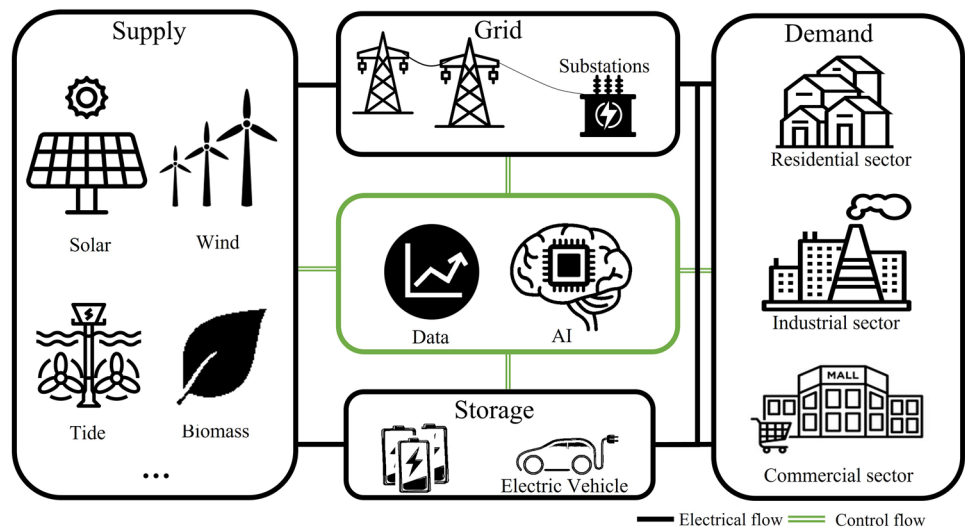


Fig. 2 A typical smart grid structure includes sustainable energy supply, smart energy consumption, advanced grid analytics, stationary and mobile energy storage, and real-time control and management



consumption increases and consumers are empowered by AI techniques [20]. Various AI-based technologies have been applied to enable smarter power consumption. For instance, neural network-based AI methods have been widely employed to predict future power consumption in smart grids, significantly improving power dispatch, load scheduling, and market management [20, 21]. Data-driven classification methods have also been utilized for electrical load anomaly detection, primarily aimed at ensuring the safe operation of power systems [22]. The research scope of anomaly detection ranges from large-scale industrial power system monitoring and smart buildings to generic residential houses. Anomaly detection can also be categorized into system security detection and non-invasive residents' health status monitoring based on the type of detection activities [20, 22, 23]. Previous works on demand response have proposed methods to manage resources efficiently and provide feedback to the energy market [24].

Given the critical role of the consumption side in power systems, numerous reviews have recently been conducted to summarize the applications of AI-based strategies in demand-side management. For instance, Raza et al. presented a review work on the application of load forecasting [25]. The paper discussed the classification of load forecasting, the impact of surrounding environments on predictive results, and the performance of the different artificial neural network (ANN)-based forecasting approaches. It identified key parameters that affect the accuracy of ANN-based forecast models, such as forecast model architecture, input combination, activation functions, and the training algorithm of the network. The review highlighted the potential of AI techniques for effective load forecasting to achieve the concept of smart grid and buildings, providing valuable insights into the importance of accurate load forecasting for efficient energy management and better power system planning. However, the review primarily focused on short-term load forecasting techniques, which may limit its applicability to long-term planning and decision-making. Additionally, it did not provide a comparative analysis of different forecasting techniques or their relative strengths and limitations from the standpoint of real-world applications. Tanveer et al. provided a comprehensive overview of data-driven and large-scale-based approaches for forecasting building energy demand [26]. The review discussed the importance of energy consumption models in energy management and conservation for buildings, categorizing methods for building energy simulations into four level classes. However, the review did not delve into specific examples or approaches of how these AI schemes have been applied in real-world situations. Khan et al. conducted a related work investigation focusing on load forecasting, dynamic pricing, and demand-side management (DSM) [27]. The review discussed the role of load forecasting in the planning and operation of power

systems and how future smart grids will utilize load forecasting and dynamic pricing-based techniques for effective DSM. It also provided a comparative study of forecasting techniques and discussed the challenges of load forecasting and DSM, covering various techniques such as appliance scheduling, dynamic pricing schemes, and optimization of energy consumption to manage energy on the consumer side. Nevertheless, the review did not provide a detailed analysis of the effectiveness or practical implementation issues of these techniques. Additionally, the paper primarily focused on residential energy management systems and may not be directly applicable to other types of smart grid applications.

Himeur et al. presented a survey on AI-enhanced anomaly detection approaches on the consumption side [28]. The review provided a comprehensive taxonomy to classify existing algorithms based on different modules and parameters adopted. It also presented a critical analysis of the state-of-the-art, exploring current difficulties, limitations, and market barriers associated with the development and implementation of anomaly detection systems. The review focused on anomaly detection frameworks for building energy consumption, but it did not provide a comprehensive review of anomaly detection frameworks for other types of energy consumption, such as industrial or transportation. Antonopoulos et al. conducted a review summarizing the applications of ML methods on power consumption [29]. They discussed over 160 related papers, 40 companies, commercial initiatives, and 21 large-scale projects. The review provides insights into the potential benefits of these technologies in improving energy efficiency and sustainability. It discusses various AI techniques that have been used in this field, their advantages and drawbacks, and their real-world applications. The review also highlights the challenges and opportunities for future research in this area. However, it is possible that the work may overlook some real-world challenges and limitations in implementing AI and ML on the demand side. For example, the review may not fully consider the cost-effectiveness of these technologies, the availability of data and infrastructure required for their implementation, and the potential ethical and social implications of their use. Additionally, the review may not fully address the practical challenges of integrating AI with existing energy systems and policies. Wang et al. presented an application-oriented review based on smart meter data [30]. The authors begin by conducting a literature review of smart meter data analytics on the demand side, focusing on the newest developments, particularly over the past five years. They then provide a well-designed clustering of smart meter data analytics applications from the perspective of load analysis, load forecasting, load management, and more. The paper also discusses open research questions for future research directions, including big data issues, new ML technologies, new business models, the transition of energy systems, and data

privacy and security. One major contribution of this paper is its comprehensive overview of current research in smart meter data analytics. The authors provide a detailed taxonomy for different applications of smart meter data analytics and discuss various techniques and methodologies adopted or developed to address each application. Additionally, they identify key research trends such as big data issues and novel ML technologies. However, similar to the previous review works, while the proposed study does discuss some practical challenges facing the implementation of smart meter data analytics in the power industry (such as data privacy and security), it does not provide much information on how these challenges can be addressed in practice.

1.2 The Motivation of Our Study

Despite the growing application of AI techniques in power systems, a significant gap remains between theoretical algorithms and their practical implementation. Prior works, while evaluating the effectiveness of AI in enhancing performance across diverse applications such as load forecasting, often fail to provide a comprehensive guide for the selection, optimization, and construction of various ML models that meet the specific requirements of energy systems. For instance, Fig. 3 illustrates the three crucial components on the demand side of power systems: load forecasting, anomaly detection, and demand response. Each of these components follows the data-driven approach, which begins with data processing to collect input data from various sources such as power distribution systems, previous prediction/detection outputs, markets, and historical datasets. This data is then managed, structured, and in some cases, subjected

to complex processing procedures such as feature extraction and normalization before being used to train the selected AI model for specific scenarios.

From a systemic viewpoint, it is vital to note that load forecasting and anomaly detection are intertwined tasks that output estimated future power usage information and real-time diagnosis results for power systems. These tasks are not isolated, but mutually dependent. The results of load forecasting and anomaly detection provide critical references for demand response, which in turn impacts the energy market, supply and demand sides, and the environment, ultimately influencing future forecasting and detection results. The interconnectedness of these components underscores the importance of considering their interactions when designing data-driven approaches for energy systems. Thus, there is a need for a review that considers not just the effects of each component, but also provides practical insights into their interplay. By synthesizing previous research across domains, our study offers a holistic and strategic perspective on future sustainable development. While emphasizing the importance of each research area, our study bridges these domains and underscores the need to address real-world application challenges, ultimately providing a comprehensive overview of the demand side of power systems. Our contributions focus on providing practical insights for evaluating, selecting, and optimizing various ML and DL models in each component, as well as offering a holistic view for better understanding and meeting the requirements of energy systems. Some practical issues such as energy system sensor/input noises, data labeling errors/cost, the resilience of existing energy infrastructure, data imbalance, data availabilities, operational constraints, etc., are analyzed for better application

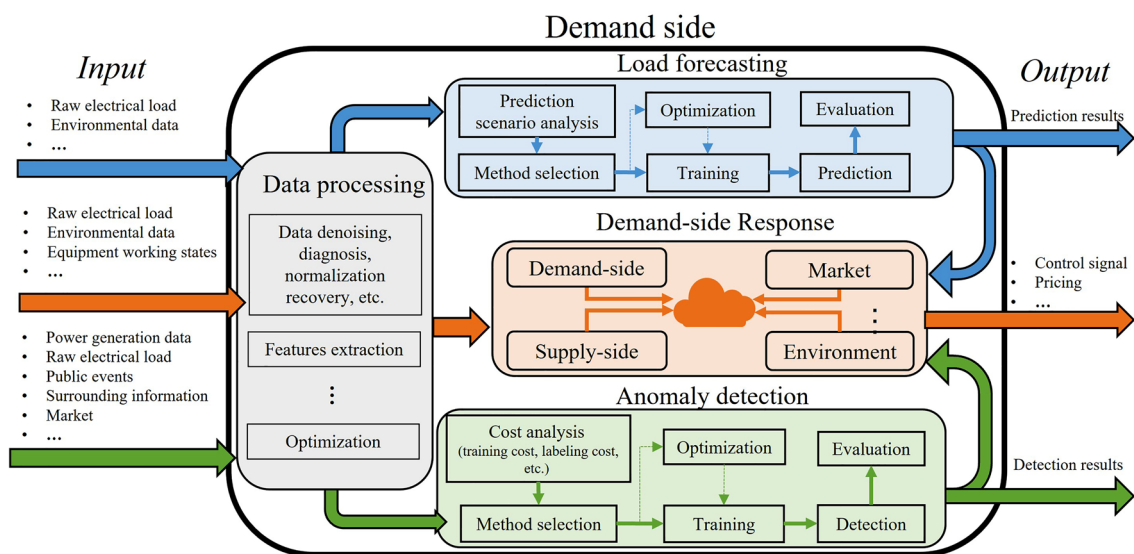


Fig. 3 The diagram of this review, including AI-empowered load forecasting, anomaly detection, and demand response

of different ML/DL models, as these issues can impact the implementation of AI methods in power distribution systems [22, 31–33].

Figure 4 shows the hierarchical structure of each component. In this section, we focus on summarizing the previous efforts and discussing the current state-of-the-art in load forecasting, anomaly detection, and demand response using AI techniques. In the load forecasting domain, we summarize previous efforts according to different AI technologies used, discuss promising optimization schemes from the perspective of implementation, and compare the

advantages and limitations of reviewed prediction methods for different applications. Next, we review anomaly detection approaches that identify abnormal load patterns and consumption behaviors, ensuring the security of power grids and reducing unnecessary power usage and CO2 emissions. We provide a holistic summary of promising optimization schemes for addressing data imbalance issues and discuss the associated challenges and trade-offs of these anomaly detection approaches. Finally, we introduce advanced strategies in demand response to comprehensively assess demand-side power usage and facilitate interaction between

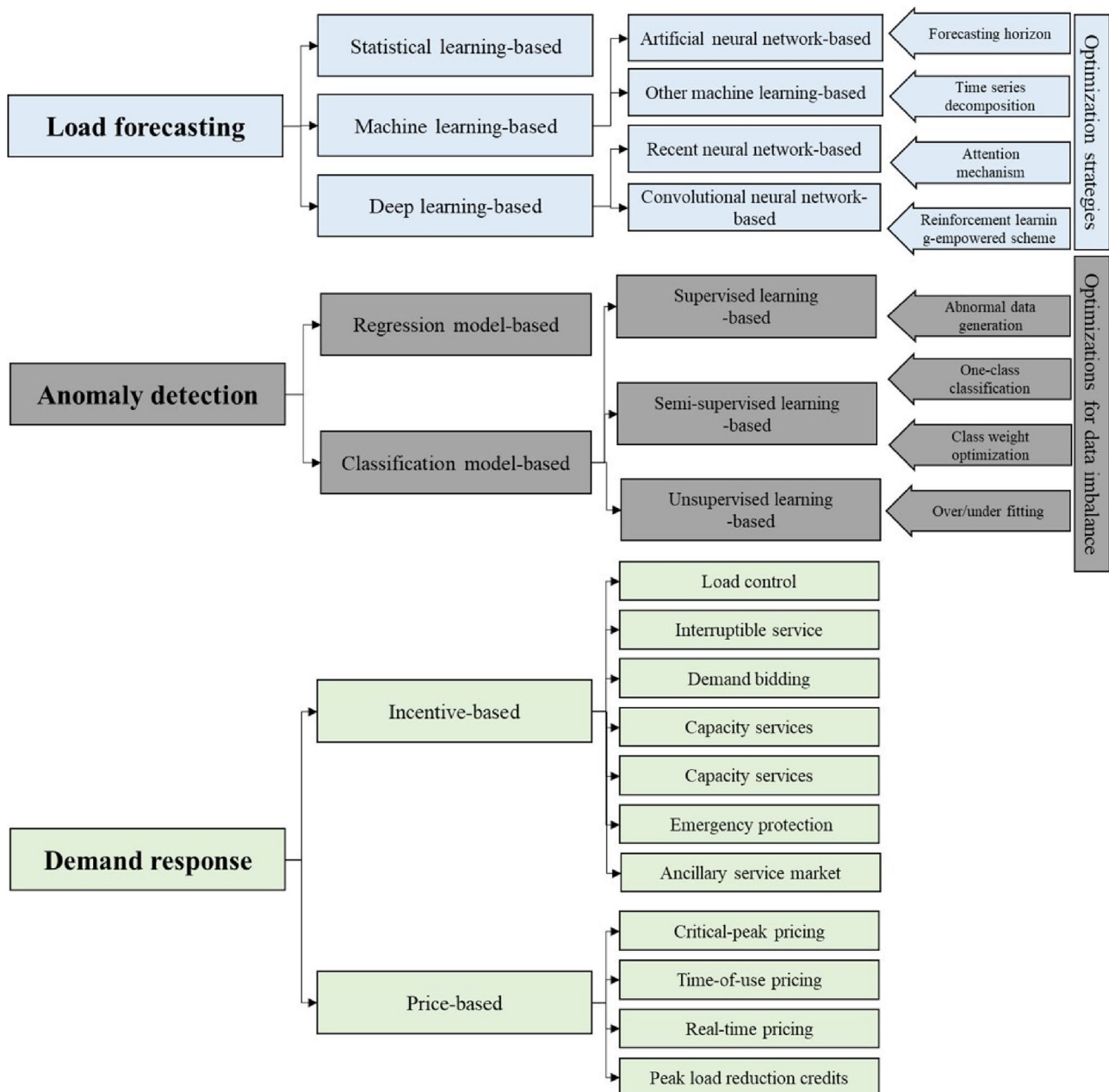


Fig. 4 Taxonomy of load forecasting, anomaly detection, and demand response in the demand side

the system and consumers. Demand response assists in managing consumption to reduce cost, waste, and risks, ensuring the balance between power generation and consumption and the reliability of future power systems. Our comprehensive review aims to provide a roadmap for researchers and practitioners to better understand the capabilities of AI techniques in enhancing power consumption on the demand side. By examining the features and challenges of each field and discussing optimization strategies, this work has the potential to drive innovation and inform the development of practical solutions that can benefit the power industry and society as a whole. The remainder of the paper is organized as follows: Sect. 2 summarizes load forecasting developments, Sect. 3 reviews anomaly detection, Sect. 4 explores demand response, and Sect. 5 concludes our work.

2 Load Forecasting

Load forecasting plays an essential role in energy dispatch and grid operations for energy suppliers and system operators. Effective and accurate load forecasting methods can improve system reliability, load scheduling, energy utilization, and reduce operating costs and risks. Especially given the vulnerability of RESs to environmental factors, the ability to estimate future power consumption is essential to improve the distribution efficiency of RESs [34–38].

Load forecasting methods can be typically categorized based on application scenarios into short-term, medium-term, and long-term forecasting [34–40].

Short-term forecasting typically refers to predictions up to 72 hours ahead, and it is essential for operational planning, such as unit commitment and economic dispatch. Medium-term forecasting, covering a period from one week to one year, is often used for maintenance scheduling and fuel reserve management. Long-term forecasting methods, used for periods exceeding one year, contribute to strategic planning, such as capacity expansion and infrastructure

development [19, 40, 41]. While there is no universally accepted classification based on the predictive horizon, it is important to note that different forecasting scenarios present their unique challenges and advantages, necessitating different modeling strategies [34–36].

Another way to classify load forecasting methods is based on the datadriven techniques employed. In general, these forecasting approaches can be categorized as ML-based and DL-based, each presenting distinct advantages and limitations [42]. In our work, we strive to offer a comprehensive analysis of data-driven model-based load forecasting, illustrating the features, benefits, and limitations of each approach in real-world applications from an AI perspective. To this end, we have classified the previous efforts into three categories: ML-based, DL-based, and statistical learning-based. Given the lack of universally accepted definitions for statistical learning and ML, the classification of forecasting methods can be somewhat subjective. For clarity and a more meaningful analysis, we define the classifications as shown in Table 1. Figure 5 effectively visualizes the different classifications of load forecasting approaches, indicating that ML, DL, and statistical learning-based approaches can be employed across diverse scenarios based on their respective strengths and suitability.

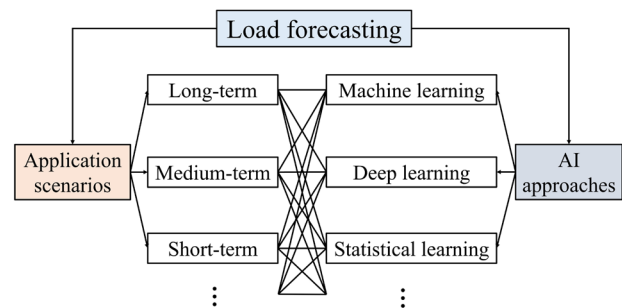


Fig. 5 Different classifications of load forecasting approaches

Table 1 Definition of forecasting approaches in this study

Forecasting method	Definition
Statistical learning-based	This approach operates based on pre-established rules and assumptions, generating a prediction model that characterizes the relationship between variables [43]. It involves fitting probabilities specific to the project to historical data and making load predictions via statistical inference. These methods are often simple, interpretable, and efficient [43]
ML-based	As a subset of AI, this approach enables computers to learn and adapt automatically from experience [44, 45]. It involves training a model on a dataset and using the trained model to make predictions on new data. These methods can model complex relationships and patterns, but they require careful feature selection and parameter tuning [45]
DL-based	This is an advanced form of ML that employs layered algorithms and neural networks with multiple hidden layers for learning and making predictions [44, 45]. DL architectures typically consist of deeply nested networks with more than one hidden layer, enabling them to learn complex representations of the data and make accurate predictions on large, complex datasets. However, they can be computationally expensive and require large amounts of data for training [44, 45]

2.1 Statistical Learning-Based Load Forecasting

Statistical learning-based forecasting using regression models is simple to implement and interpret. Energy consumption is often correlated with exogenous factors and historical consumption, and a regression model can be fitted using the data to predict future loads. Linear regression is one of the most classical and popular methods used in load forecasting [46]. For example, Fan et al. proposed a comparison between linear and nonlinear techniques and found that both have their own best feature set for model development [47]. The proposed work showed that linear models might be better when raw features are used, and nonlinear methods can achieve satisfactory performance when the feature set is well-processed. Ming et al. and Syed et al. explored improved linear models for load forecasting [48, 49]. Compared with typical nonlinear models, the proposed improved linear regression models are much more computationally efficient, and the simplicity of linear methods makes them useful over different time scales [50].

Unlike general regression-based methods that correlate various information with energy consumption data, an autoregression-based method focuses on data points from a time series and correlates the future values of a variable with only its past values. Thus, in autoregression models, the historical load is the only factor that affects future consumption [23]. One of the most used autoregression models is the autoregressive integrated moving average (ARIMA) for non-stationary time series forecasting. Many studies aimed to improve the standard ARIMA. Lee et al. developed an improved ARIMA-based short-term load forecasting model [51], outperforming back-propagation neural networks (BPNN). To overcome the limitations of ARIMA, it is commonly combined with a nonlinear model to yield a hybrid load predictor. Zou et al. developed an ARIMA-based hybrid short-term load predictor [52], in which a BPNN is used to reduce the residual error of ARIMA. The test results demonstrate that the proposed hybrid framework outperforms the original ARIMA. A similar effort was also provided by Wang et al. [23]. To improve the accuracy of real-time forecasting using ARIMA, Wang et al. proposed a method that incorporates an artificial neural network (ANN) model to dynamically learn the forecasting errors of ARIMA [23]. The proposed approach employs an online learning approach, where the final prediction output is a combination of the ARIMA predicted result and an estimated error provided by the ANN model. Comparative results demonstrated that the proposed approach outperforms the single ARIMA model in terms of accuracy.

While linear models can be effective, they are not well-suited for highly nonlinear data and may require hybrid approaches to improve performance. To address this limitation, more effective nonlinear models are needed. Fan et al.

developed a data-mining technology-based one-day-ahead building load predictor and compared it with multiple linear regression (MLR), support vector regression (SVR), multi-layer perceptron (MLP), binary tree, multivariate adaptive regression spline, and k-nearest neighbor (kNN) methods [53]. It was found that although MLR requires the least computational time compared to the other models, it still does not perform very well because smart building related processes are usually nonlinear and complex [53]. Khorsheed et al. developed a nonlinear model for long-term peak load forecasting and demonstrated that this model is more accurate than the linear model [54]. Fan et al. developed two load forecasting models to manage and optimize building cooling and compared them with a few models, such as MLR, autoregression, and BPNN [55, 56]. It was found that MLR accurately predicts the on-site cooling load while requiring minimal training data using simple hardware with low computational complexity. Benefiting from the straightforward structure and fast calculation speed of the statistical models, statistical learning-based methods are widely applied in load forecasting for smart grids. Both linear and nonlinear models play important roles in different forecasting scenarios [57]. Principal component analysis, sensitivity analysis, and stepwise regression were introduced by Yildiz et al. [50] to further enhance the accuracy of regression methods. Regression models are also widely employed to monitor energy consumption, measure and verify energy efficiency [58–60], and identify operation and maintenance problems [60–62].

2.2 Machine Learning-Based Load Forecasting

Recently, neural network-based methods have gained significant attention and are increasingly being applied for load forecasting [63]. Classical ML models are relatively simple, readily interpretable, and computationally efficient compared to DL models.

2.2.1 Artificial Neural Network (ANN)-Based Load Forecasting

Among various ML-based load predictors, artificial neural networks (ANNs) have the advantage of extracting features from data to perform accurate regression, making them widely used in load forecasting [63–69]. Amber et al. proposed a building electric load forecasting method that incorporates various environmental data, and compared it with other methods such as genetic algorithms, support vector machines (SVM), and deep neural networks (DNN) [69]. The results indicated that ANN outperforms all the other methods while maintaining a reasonable level of complexity. Furthermore, the comparison between DNN and ANN revealed that the performance of DNN might be superior when the training dataset is limited. Davut Solyali

proposed an ML-based load forecasting method in their previous work [63], where they compared different models for both long- and short-term predictions. The models included SVM, multiple linear regression (MLR), a neuro-fuzzy inference system, and an ANN. The results indicated that SVM is the best-performing model for long-term forecasting, while ANN showed better performance for short-term forecasting.

Some studies have also focused on enhancing the prediction accuracy of ANN by selecting model inputs and features. For instance, Ding et al. and Patel et al. developed load prediction methods by integrating ANNs and k-nearest neighbors (kNN) [64, 70]. These methods considered the effects of environmental factors and used kNN to cluster input variables. They demonstrated that training ANN on the processed data can lead to better outcomes. Ding et al. explored the effect of input variables on the cooling load prediction accuracy of an office building [64]. The authors used two ML models, ANN and SVM, for prediction. The clustering method was utilized to optimize the selection of input variables for 1-hour-ahead cooling load prediction. The comparison aimed to enhance the prediction performance of ML-based predictors.

2.2.2 Other Machine Learning-Based Load Forecasting

In addition to ANNs, other ML models, such as extreme gradient boost (XGBoost), have been utilized for load forecasting in smart grids. Al-Rakhami et al. developed an XGBoost-based load forecasting model for residential buildings and revealed that XGBoost can effectively address overfitting issues in load forecasting [70]. By comparing the results of different forecasting approaches, the advantage of XGBoost in avoiding overfitting problems is demonstrated. Vantuch et al. studied the computational complexities of prediction and training, showing that random forest regression (RFR) and XGBoost exhibited the lowest complexities, followed by ANN and SVR. The flexible neural tree is the most computationally expensive model, and its prediction could have been better compared to those of RFR and XGBoost [66]. Wang et al. proposed a load forecasting framework that employs an XGBoost-based one-step-ahead forecaster with quantile regression [71]. The XGBoost-based quantile regression model is used to generate a prediction interval for the next step. The proposed strategy dynamically tracks recent prediction results to adjust the parameters of the quantile regression model to optimize the final results. Notably, the proposed approach is compared with various ML/DL models, and the comparison results demonstrate that the proposed forecasting framework is the most accurate, providing reliable and accurate results in real-time. Zhang et al. developed improved grey wolf optimization- and extreme learning

machine-based load predictors hosted in the cloud [72]. The improved grey wolf optimization is used to optimize the parameters of the extreme learning machine. A comparison between the initial extreme learning machine model and the model with the selected optimal parameters validates the effectiveness of the improved grey wolf optimization- and extreme learning machine-based load predictors. Fan et al. developed a data mining-based ensemble load forecasting method to achieve one-day-ahead forecasting [53]. Through introducing the data training and feature selection and elimination progress, eight different ML models are employed to improve the feasibility of the proposed ensemble learning approach. Finally, the study concludes that the proposed comprehensive framework outperforms the initial eight models and achieves better forecasting results.

2.2.3 Recurrent Neural Network (RNN)-Based Load Forecasting

Among various DL methods, recurrent neural networks (RNNs) are commonly used for load forecasting with sequential data. Long short-term memory (LSTM) is a typical RNN-based method widely employed for load forecasting [34, 38, 73–78]. Song et al. and Kumari et al. recently developed two different LSTM-based load predictors [76, 77, 79]. Song et al.'s work focused on verifying the accuracy of the LSTM-based forecasting approach using two real-world datasets and employed model echo state networks for comparison. Kumari et al. tested the robustness of LSTM for time-series forecasting using different datasets. Both LSTM-based approaches demonstrated good performance in load forecasting. Zhang et al. studied a hybrid forecasting approach that integrates Fiber Bragg Grating sensors with LSTM to forecast electrical load [79]. Compared with BPNN, the proposed method reduced the complexity of the network, saved running time, and improved forecasting accuracy. Bouktif et al. developed an advanced LSTM-based load forecasting algorithm [73] using a genetic algorithm (GA) to obtain the optimal lag and layer number, providing guidelines for optimizing LSTM.

Although RNNs are commonly used for time series forecasting, the inherent limitation of the vanishing gradient during training restricts their applications. As a result, many authors have attempted to improve RNNs by integrating hybrid models [34, 74]. For example, Zhang et al. and Cenek et al. developed LSTM/ANN-based predictors [38, 77]. Zhang et al. trained an LSTM to extract features from time-series data, and an ANN was used to analyze the relationship between features and the load [38]. Cenek et al. developed a grid load predictor that considered environmental factors [77]. An LSTM was used to predict future weather outputs for ANN, which forecasted

the final outputs for the load. Both studies found that the hybrid methods outperformed individual ANN and LSTM models in load forecasting.

2.2.4 Convolutional Neural Network (CNN)-Based Load Forecasting

Convolutional neural networks (CNNs) are also widely used in load forecasting [80]. CNNs leverage the unique linear operation called convolution in at least one layer of the network to effectively learn representations and extract features from time series data [81]. The output layer is typically placed after the fully connected layers and resembles a standard ANN layer.

Kuo et al. and Amarasinghe et al. developed CNN-based load predictors [35, 82]. Kuo et al. compared CNN with several benchmark models including SVM, decision tree, MLP, and LSTM, and found that CNN outperformed all the tested models [35]. However, Amarasinghe et al. found that CNN did not provide a clear advantage over other models, and LSTM showed better outcomes in their comparison results [82]. This difference in performance may be due to the complexity and diversity of the building-level power consumption data in the test dataset. As a sequential model, LSTM is better equipped to learn features from such time series data compared to CNN [82]. Dong et al. developed a hybrid load forecasting method that combines CNN with K-means clustering [83]. By applying K-means clustering to the data before training the CNN, they achieved better predictive accuracy compared to other models such as SVR, neural network, linear regression + K-means, SVR + K-means, and neural network + K-means. Another CNN-based hybrid short-term load predictor was presented by Aurangzeb et al. who used a novel pyramidal CNN model to reduce the computational complexity of load forecasting [81]. The proposed method demonstrated improved forecasting accuracy. However, it is worth noting that CNN-based hybrid predictors may have higher training costs compared to other machine learning-based load forecasting approaches. Alhussein developed a CNN-LSTM-based load forecasting method that enhanced forecasting accuracy and outperformed baseline models [84]. While CNN-based hybrid predictors generally improve predictive accuracy, it is important to consider their training costs compared to other machine learning-based load forecasting approaches.

2.3 Optimization Strategies for Improving Learning-based Load Forecasting

Regardless of the different models employed for load forecasting, the forecasting performance is affected by factors

from real-world applications, such as the predictive horizon, feature extraction, and computational resources. In this section, we discuss several recent improvement methods [42].

2.3.1 Forecasting Methods Under Different Horizons

Different models perform differently on short- and long-term load forecasting. It is essential to choose the most suitable model for various forecasting horizons to maximize prediction accuracy and performance [42]. In this work, we point out that:

- For short-term load forecasting, it is essential to use models that can effectively capture the changes and variations in time-series load data. Moreover, the running time of these models is also a crucial consideration, as real-time predictions are often required [42, 67, 84].
- For long-term load forecasting, the models should be robust against the impact of data noise and overfitting or underfitting issues. These models need to have a good generalization capability to be able to forecast load accurately for longer periods [42, 66, 67].

Wang et al. in a study of building thermal load forecasting, comprehensively discussed twelve forecasting models, including LSTM and XGBoost [84]. The test results of this work pointed out that the heuristic load forecasting methods are recommended for projects with limited budgets and resources, and LSTM is proven to be robust against input uncertainty and recommended for short-term forecasting. For long-term forecasting, XGBoost is found to be more accurate and trained with the predicted results to enhance the robustness of the model. Yang et al. developed a multi-step-ahead load predictor using an autoencoder neural network with a pre-recurrent feature layer [85]. This method showed satisfactory performance when the forecasting horizon is less than 3 hours. Vantuc et al. compared different ML-based approaches under varying forecasting horizons from one hour to one week [66]. A comparison of ANN, SVR, RFR, XGBoost, and a flexible neural tree showed that ANN provides the highest stability when the forecasting horizon changes, which is also consistent with the findings by Sangrody et al. [67]. Due to the significant impact of the prediction horizon on results, Wang et al. proposed a load forecasting framework that combines an XGBoost-based one-step-ahead forecaster with an LSTM-based one-day-ahead predictor [86]. The framework can dynamically evaluate the step-by-step LSTM-based one-day-ahead load forecasting. If the forecast result is deemed inaccurate, the framework can automatically switch to the XGBoost-based one-step-ahead load forecasting model to improve prediction accuracy.

2.3.2 Decomposition-Based Feature Extraction

Data decomposition is another widely used strategy for enhancing the performance of load forecasting algorithms. A non-stationary series, like electrical load, is characterized by statistical properties that vary over time. Therefore, decomposing a non-stationary time series signal into a set of intrinsic mode functions and a residue can facilitate learning [87].

Empirical mode decomposition (EMD) is an unsupervised data-driven decomposition method commonly used for non-stationary time series data. Qiu et al. and Bedi et al. developed load forecasting methods based on EMD and compared their results to popular DL models such as DNN and LSTM [2, 88]. Both hybrid methods outperformed the single DL models, demonstrating the effectiveness of EMD-based decomposition in improving forecasting performance. Another decomposition method, seasonal-trend decomposition based on Loess (STL), was employed by Fan et al. to enhance load forecasting [89]. After STL decomposition, the subsequences became more regular and easier to learn, leading to improved performance compared to LSTM and LSTM- XGBoost models. Park et al. explored complex multi-user load prediction by developing a characteristic load decomposition-based load predictor [90]. The aggregated load measured at one node was divided using the characteristic load decomposition method to improve forecasting performance.

These studies demonstrate that decomposition-based feature extraction methods can be effective in enhancing the learning performance of load forecasting models. Decomposition provides an alternative way to extract features from time-series data. However, it is important to consider the training cost, efficiency, and flexibility of these methods to determine their practicality for real-world load forecasting applications. Further research is needed to investigate these aspects.

2.3.3 Attention Mechanism for Learning-Based Forecasting Approaches

The attention mechanism, developed from cognitive attention for sequence-to- sequence learning, has proven to be a powerful tool for improving DL-based learning models [91]. By allowing models to focus on the relationship between inputs and outputs during the training process, attention can improve interpretability and learning performance.

In recent studies, attention mechanisms have been applied to load forecasting models with promising results. For example, Li et al. proposed an RNN-based

load predictor with attention and demonstrated its effectiveness through comparative results [92]. Similarly, Jin et al. designed an attention- based encoder-decoder network with Bayesian optimization for load prediction and found that the attention-enhanced approach consistently provided more accurate results [93]. Wang et al. developed a bi-LSTM-based predictor with attention and rolling update for short-term load forecasting [94]. The rolling update is used effectively to update the training dataset in the real-time forecasting process. Attention is employed to assign the influence weights of different input variables. Then a bi-LSTM is used for predicting. Unlike general LSTMs that transmit unidirectionally and only focus on past information, bi-LSTM provides a two-path training method. Both past and future data are taken into account. According to their test results, compared to the traditional bi-LSTM model, the proposed attention and rolling update boosted bi-LSTM model can achieve more accurate forecasting results. Another method was reported by Wu et al. [95], in which a short-term forecasting method is presented using an attention-based approach with CNN, LSTM, and bi-LSTM. The input dataset includes temperature, cooling load, and gas consumption information for the past five days. An attention-based CNN is utilized to extract the features of the input. Then, LSTM and bi-LSTM are combined to forecast the load for the next hour. The outputs indicate that the proposed method is more accurate than single LSTM, bi-LSTM, BPNN, RFR, SVR, and hybrid models: CNN-bi-LSTM, and CNN-LSTM. Sehovac et al. proposed a load forecasting method combining Sequence RNN and attention [96]. Given that the structure is composed of an encoder and decoder, and longer series may increase decoding challenges, attention is introduced to prioritize the input series. Compared with the other attention-empowered studies, the results showed that the proposed method with attention is more accurate. At the same time, this work also notes that the forecasting accuracy decreases as the forecasting horizon increases; moreover, longer input sequences may not always increase the accuracy.

As demonstrated in previous works, the attention mechanism is an effective strategy to improve the performance of predictors. However, it should be noted that incorporating attention mechanisms into models may lead to longer training times, which may limit the performance of predictors, especially for the applications of short-term load forecasting. Nonetheless, the potential benefits of attention mechanisms in terms of interpretability and improved prediction accuracy make it a promising area of research in the field of learning-based forecasting approaches.

2.3.4 Reinforcement Learning (RL)-Empowered Load Forecasting Schemes

Reinforcement learning (RL) is a subset of ML in which an agent learns to make decisions by interacting with the environment and receiving rewards or penalties based on its actions [97]. The agent senses the environment and chooses actions that influence it, seeking to find a policy that maximizes the accumulated rewards [97]. Compared to ML/DL-based load forecasting models that rely solely on training data, RL models can adjust their predictions dynamically based on new inputs, resulting in more accurate and reliable load forecasts, even in the presence of unforeseen events [98–100].

There are two main RL methods: policy-based and value-based RL [98]. In policy-based RL, such as policy gradient, the agent directly learns and updates the policy that maps state to action, repeating the process of “selecting the initial policy, finding the value function, and finding the new policy from the value function” until the optimal policy is found. To improve the forecastability of the entire electrical load, Xie et al. proposed an RL-based data sampling control approach that interacts with smart meter data [101]. The proposed approach can be implemented both offline and online by interacting with real-time data from each household. The test results show that the proposed RL-based algorithm outperforms competing algorithms and delivers superior predictive performance.

In value-based RL, such as Q-learning and Deep Q-Network (DQN), the agent focuses on the action-value function and finds the policy through the value function, repeating the process of “selecting the initial value function, choosing the best action in the state, finding the new value function” until the optimal value function is found. Feng and Zhang proposed a dynamic predictive model selection mechanism based on Q-learning to enhance the accuracy of multi-model load forecasting [98]. The approach dynamically selects the most accurate forecasting output from multiple ML/DL models based on real-time observations, and experimental results using two years of data show that it improves accuracy by approximately 50% compared to using any single ML/DL method alone. Dabbaghjamanesh et al. presented a Q-learning-based load forecasting algorithm for hybrid electric vehicle (EV) charging [102]. The algorithm employs both an ANN and an RNN to forecast the load simultaneously, and a Q-learning agent is utilized to determine which results to use as the final output. The algorithm is tested in three different scenarios, each with varying EV charging characteristics, and the comparison results demonstrate that the Q-learning-based approach can consistently predict EV charging load with greater accuracy and flexibility compared to the ANN and RNN methods. Park et al. proposed a novel approach to improve the accuracy

of short-term load forecasting using a similar day selection model based on DQN [100]. The proposed model dynamically selects the most suitable training data, significantly boosting the performance of the ML model. The proposed method outperforms existing models in terms of accuracy, as demonstrated through extensive experiments on real-world load and meteorological data from Korea. As demonstrated by the aforementioned works, RL technologies have been extensively employed in load forecasting and have shown remarkable forecasting results. However, it should be noted that, at present, most of these advances primarily utilize RL to dynamically select the appropriate model, training data, sampling rate, and other related parameters. The primary prediction models still rely on ML and DL techniques. That is the reason the RL-empowered strategies are listed in Section 2.4. In line with the OpenAI framework, it is expected that more promising RL-based prediction approaches will be developed in the future.

2.4 Discussion of Forecasting Methods

As a review work targeted to discuss each strategy with an eye on real-world applications, it is critical to point out that

- It is intractable to identify the “most accurate” or “fastest” approach [65, 71, 86]. The evaluation of prediction results shows different trade-offs depending on how various metrics prioritize their performance criteria from different perspectives [86].

For example, a variety of evaluation metrics has been proposed to evaluate load forecasting results, such as mean absolute error (MAE), root mean square deviation (RMSE), mean absolute percentage error (MAPE), symmetric mean absolute percent error (SMAPE), R2, and Theil’s inequality coefficient (TIC) [27, 103, 104]. However, it is found that MAPE offers a multi-dimensional perspective of predictive results evaluation. Sometimes, it may misinterpret the results, especially when the length of the data series is not specified [65].

- Various prediction horizons require different considerations to trade-off accuracy, flexibility, and reliability.

As mentioned before, according to the application scenarios, load forecasting methods can be categorized into short-, medium-, and long-term forecasting [34–40]. The short-term is generally defined as up to 72 hours, which can help to fix the time-lag measurements and provide an immediate impact on the operation [37, 38, 40]. However, the implementation of this method is always limited by sensing noise, calculation speed, human behaviors, and over-or-under training problems [37]. The Mid-term refers to the prediction period

as one week, months, or up to one year, while the long-term is the prediction period with a prediction period longer than one year [40]. These works have greater importance for long-term planning, economic growth, policy adjustment, system capacity determination, and maintenance, but suffer from the limitation of prediction accuracy [37, 40, 105]. Although each forecasting approach has its application scenarios and advantages, a long-standing problem with load forecasting is that no classification criteria based on the predictive horizon have yet emerged [34–36]. For example, some previous work further divided the short term into very short term and short term [35]. Where the very short-term is defined as less than one day, short-term is longer than one day and less than one week, med-term is less than one year and long-term is longer than one year. For a better discussion of the applications of each strategy, the predictive horizons in this work are classified into short- and long-term, defined as less than one day and more than one day, respectively [34–36]. In Table 2, this work summarizes the features of each forecasting approach.

Table 2 compares the advantages and limitations of DL-based, ML-based, and statistical learning-based methods and guides the selection of load forecasting models. A statistical learning-based approach is recommended for forecasting scenarios requiring easy implementation, low computational cost, and real-time processing regardless of the prediction horizon. However, according to the features of target data, additional optimization may be needed to provide accurate forecasting outputs [46, 53, 56]. ML-based methods are usually recommended for long-term prediction. Although ML-based methods generally require a longer training time compared to statistical learning-based methods, they still offer a reasonable trade-off between computational complexity and prediction accuracy [54, 56].

DL-based methods are recommended for short-term predictions when accuracy is the priority. The strategy can automatically capture more data features and yield more accurate predictions [81]. Though the interpretability of DL has been questioned, the recent research efforts in explainable AI and attention mechanism shed light on the interpretability of DL, which could improve the interpretability of DL-based load forecasting methods [38]. At the same time, DL methods generally require more training data than ML approaches, which also can be another limitation [69]. Besides, it is also worth noting that, given the complexity of electrical load data over a long period, the features captured by DL-based methods can be less relevant and thus less effective in prediction [84]. To improve the application of DL approaches in power systems, some hybrid models using DL-based methods for extracting data features and ML-based methods for prediction also achieve satisfactory results [38].

However, it is important to emphasize that these observations may not always hold true and may vary depending

Table 2 Comparison of load forecasting methods

Methods	Long-term load forecasting	Short-term load forecasting
DL-based	Advantages	1. High robustness against input uncertainty 2. Flexible and support ensemble methods, for example, popular hybrid models using DL to capture features and ML to improve predictive accuracy
	Limitations	Requires more training data
ML-based	Advantages	Flexible and support ensemble methods, for example, popular hybrid models using DL to capture features and ML to improve predictive accuracy
	Limitations	1. Learning results may not be particularly relevant or helpful; 2. Requires more training data High robustness against overfitting, especially XGBoost and ANN
Statisticsbased	Limitations	Additional feature selection processes are required to ensure high accuracy
	Advantages	Simple, small training dataset, computationally efficient
	Limitations	Generally low predictive accuracy

Table 3 Loss in a power grid

Type of loss	Examples	Features
Technical	Power dissipation, internal electrical resistance	Inevitable
Non-technical	Energy theft, bill fraud, meter failure, illegal power utilization	Evitable

on the data and implementation. For instance, Hosein et al. discovered that although DL-based methods often require longer computational times than other models, they can still be preferable, even with a short epoch [36]. Additionally, RL technology also provides a promising strategy to improve the performance of ML/DL-based forecasting approaches [106]. For load forecasting with large-scale or multi-feature training datasets, RL can be used as a step-by-step feature extraction mechanism to dynamically improve the forecasting accuracy [100].

3 Anomaly Detection

Electricity providers aim to meet the electricity demand while facing both technical and non-technical electricity losses. Technical losses result from the physical limitations of a power system, such as resistance, while non-technical losses arise from electricity theft and leakage, leading to high costs and potential security breaches. It is essential to

avoid non-technical losses. The differences between technical and non-technical losses are presented in Table 3, as reported in [107].

Anomaly detection is a crucial aspect of smart grid, particularly on the demand side. To effectively identify abnormal power usage, various data science techniques are utilized to train an anomaly detection algorithm [108]. The candidate data can be classified as either “general anomalies” or divided into different types of anomalies, depending on the feature extraction method used [109]. Anomaly detection can provide feedback on energy consumption for problem diagnosis, which benefits energy suppliers and ecosystems [23, 110–112]. Electrical load anomaly detection methods can be categorized into two main groups: regression model-based and classification model-based [22]. Figure 6 illustrates the detection mechanism of each strategy.

3.1 Regression-Model-Based Anomaly Detection

Regression model-based anomaly detection is a technique that relies on the principles of forecasting. It utilizes a forecasting model to fit historical data and predict the load for future time slots. The predicted values are then compared to the realized actual data, and large deviations between the two can be identified as anomalies. Zhang et al. used a linear regression model to fit and predict electrical load considering the effects of environmental factors on residential energy consumption [68]. A linear regression model outputs prediction as a baseline and any realized data points that deviate

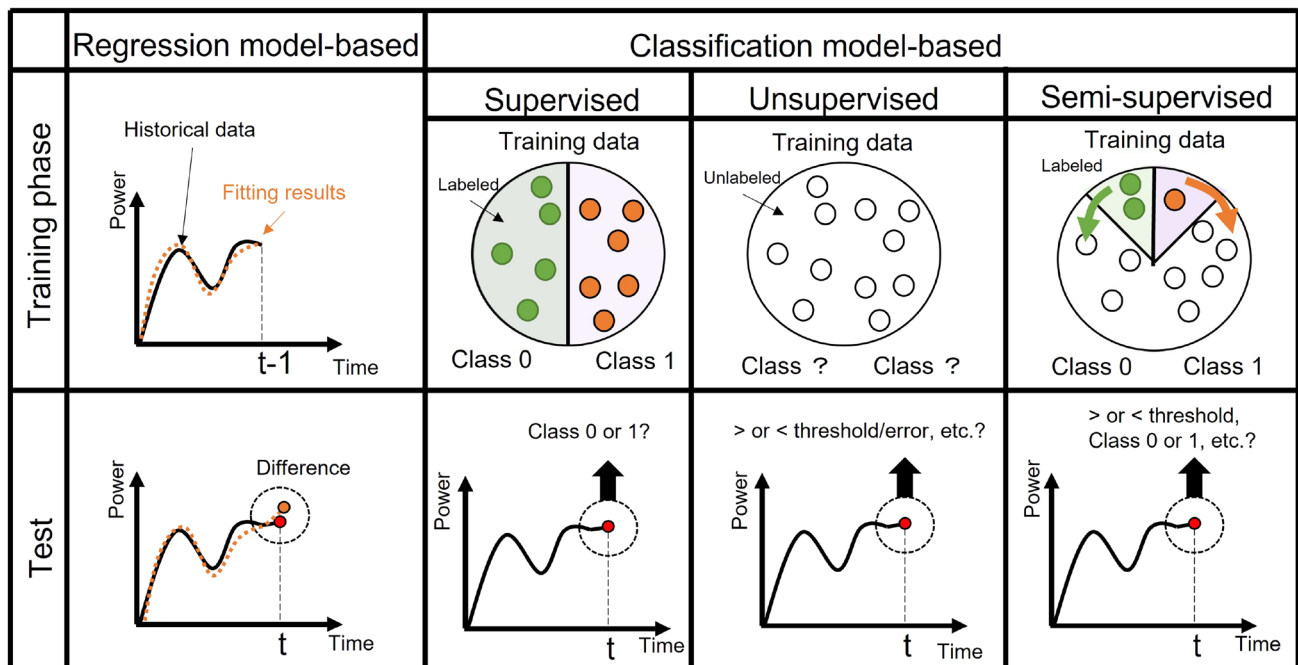


Fig. 6 Detection mechanism of regression-model-based and classification-model-based detection methods

extensively from the baseline are identified to be anomalous. However, from the application perspective, there are two concerns as follows. First, individual differences in temperature sensitivity can restrict the applications of this approach. Second, considering the complexity of the residential load series, linear regression may only adequately capture some essential features, resulting in false detection. Chou et al. and Hollingsworth et al. developed hybrid prediction model-based anomaly detectors capturing the nonlinearity in electrical load series [113, 114]. Both studies employed ARIMA to perform linear regression, then incorporated non-linear ML models, such as ANN and LSTM, to compensate for ARIMA losses. Consistent with previous forecasting models, the hybrid models facilitate better characterization of load time series. Though hybrid methods can improve load forecasting performance, there is no guarantee for perfect anomaly detection. First, a hybrid model may need to forecast a future load with higher accuracy. Second, considering the inaccuracy in predictions, the two-sigma rule may be insufficiently effective in detecting all anomalous data.

To address issues with unsatisfactory detection, Luo et al. developed a dynamic anomaly detector [115]. No fixed or preset threshold is used when exploring differences between predictions and real data. Instead, an active adaptive threshold is employed, and the dynamic detection rule ensured that the design could adapt to time-varying anomalies. However, the dependence on the forecast results remained high. To further improve anomaly detection, Fenza et al. developed a drift-aware method to detect anomalies in smart grids [116]. An LSTM is used to extract historical data features and forecast load. Then the detection algorithm calculated the trend in predictive error and detected consumption anomalies. Some other novel attempts have been recently proposed by Wang et al., Xu et al. and Cui et al. to improve the regression-based detection [20, 23, 117, 118].

Three improved regression model-based detection methods were proposed by Wang et al. to overcome the intrinsic limitations of the regression-based approaches [20, 23, 71]. The Bayes information criterion is employed to avoid over-under-fitting during real-time prediction and improve the predictive performance. Then, an independent detection mechanism is developed to analyze the estimated next-step load and the observed real-time load and screen out anomalies. The proposed test results show that the hybrid approach outperformed the introduced alternative ML- and DL-based approaches in terms of both prediction and anomaly detection. Wang et al. also proposed a novel prediction result-based detection strategy, which outperforms traditional ML/DL-based detection methods and saves on training costs [71]. The approach dynamically evaluates real-time power consumption information and identifies if it is consistent with the historical power usage habits, using the estimated

next-step load as a reference. The proposed method offers accurate anomaly detection without relying on labeled data and can be beneficial in scenarios where labeled data is not readily available. Cui et al. also used the outcomes of predictive models as baselines for anomaly detection [117]; however, the predictive results are not simply replayed. Instead, a supervised learning-based classification is employed to perform good detection, thus improving accuracy and enhancing forecasting robustness. Xu et al. developed an RNN-based predictor using quantile regression and Z-scores to detect anomalies [118]. These new approaches employed ML to improve prediction and detection based on regression-based methods. However, the simplicity of regression-based detection has been compromised as the prediction and detection structures become more complex.

3.2 Classification Model-Based Anomaly Detection

In addition to regression model-based anomaly detectors, classification model-based anomaly detectors have also been widely used [23]. Classification-based anomaly detectors can be further divided into supervised learning-based, unsupervised learning-based, and semi-supervised learning-based methods.

3.2.1 Supervised Learning-Based Methods

Various supervised learning-based models are used for anomaly detection, among which SVM has gained popularity due to its good performance in large-scale systems [119]. Since each kernel is non-parametric and operates locally, it is not necessary to have the same functional form for all data, thus reducing computational costs [119, 120]. Nagi et al., Jokar et al. and Depuru et al. have designed different SVM-based anomaly detectors for smart grids [120–122]. Nagi et al. used a feature selection/extraction function to process raw data and trained SVM to determine data correctness, resulting in improved detection accuracy [120]. Jokar et al. developed an SVM-based anomaly detector to identify electricity theft from a grid with an advanced metering infrastructure [122]. They employed K-means to cluster the training data into different groups, and the number of clusters is determined by the Silhouette coefficient. The clustered datasets are used to train an SVM to identify abnormal samples, resulting in higher accuracy for anomaly detection. Depuru et al. developed an SVM-based electricity theft detection algorithm by combining SVM with a rule engine that divides customers into genuine electricity users and thieves [121]. Although the classification results of the original SVM seem to be suboptimal, the hybrid SVM methods effectively detected smart grid anomalies.

Supervised learning-based methods, such as SVM, have shown good performance in large-scale systems due to their ability to handle high-dimensional data and non-linear relationships (Ozay et al. [119]). However, the performance of these models can vary depending on the system size and the data's class distribution. Pinceti et al. [123] compared the effectiveness of KNN, SVM, and replicator neural networks in detecting anomalies in real-world data and found that KNN performed well. However, their test data suffered from class imbalance issues, which may have contributed to the superior performance of the simple clustering method. In contrast, Ozay et al. [119] tested SVM and KNN on various systems and found that KNN outperformed SVM in small sized systems but performed worse in large-sized ones due to its sensitivity to class imbalance. In the context of anomaly detection, hidden Markov models have also been widely used. Makonin et al. [124] proposed an improved non-intrusive load monitoring method using a novel variant of the Viterbi algorithm, the sparse Viterbi algorithm, which can effectively manage large sparse matrices. The proposed method disaggregated a model with many super-states while preserving between-load dependencies in real-time, leading to better load monitoring.

DL-based anomaly detectors are known to achieve accurate detection compared to ML-based methods. Devlin et al. developed a feed-forward neural network-based load monitor that detected anomalies with an average precision of 76.3% using raw meter data for time series decomposition and detection [125]. To discuss the advantages of DL-based methods, Buzau et al. and He et al. developed two novel electrical load anomaly detectors [126, 127]. Buzau et al. used an LSTM and MLP-based classifier that outperformed traditional ML-based methods such as SVM, LR, XGBoost tree, MLP, and CNN. He et al. developed a conditional deep belief network-based electricity theft detector that can detect anomalies in real-time, and its performance is better than ANN- and SVM-based detectors [127]. Furthermore, compared to detection accuracy, another essential aspect of detection method evaluation is robustness [128]. A robust model should consistently provide accurate results under different circumstances [129]. In this regard, Rolnick et al. suggested that DL-based supervised learning methods are more robust to label noise than ML-based supervised learning classification methods [128]. To improve the performance of DL methods in load forecasting, attention mechanisms have been introduced. Javed et al. proposed a new anomaly detection strategy incorporating a combination of an attention mechanism with an LSTM-based CNN to identify erroneous and anomalous readings generated through errors or attacks in Connected-and-Automated Vehicles [130]. According to the test results, the proposed approach significantly improved the detection rate compared to alternative methods based on the Kalman filter and CNN-Kalman filter.

3.2.2 Unsupervised Learning-Based Methods

Supervised learning methods can be effective in anomaly detection, but their reliance on high-quality labeled data can be expensive and limit their practicality [22]. To address this challenge, unsupervised learning methods offer a promising alternative that does not require labeled data and can be more cost-effective [131]. One commonly used approach is the autoencoder, which compresses input data into latent variables through an encoder, then reconstructs the data using a decoder [132, 133]. Anomaly scores are calculated for each observation, and the ones exceeding a threshold are classified as anomalies.

Several studies have demonstrated the effectiveness of unsupervised learning-based anomaly detection. Fan et al. proposed a method that combined spectral density estimation with decision tree classification [131]. After feature extraction, an autoencoder is used to calculate anomaly scores, and observations with scores above a preset threshold are identified as anomalies. Zheng et al. developed an autoencoder-based model that used maximum likelihood estimation to detect anomalies [111]. By calculating the average and correlation variances of the dataset using reconstruction error vectors, anomalies are detected with reference to the distribution of reconstruction errors.

Other studies have used unsupervised learning methods for specific applications. For instance, Zhao et al. used graph signal processing to process power consumption data for low-rate load monitoring purposes [134]. Meanwhile, Husain et al. developed an electricity theft detection method that employed statistical feature extraction, robust principal component analysis, and outlier removal clustering [111].

While unsupervised learning-based anomaly detection methods can achieve accurate results with low labeling costs, they have some limitations. These include low computational efficiency for large datasets, sensitivity to feature extraction, lack of ground truth data for evaluating results, and weak interpretability [31, 131, 135, 136]. Nonetheless, unsupervised learning-based anomaly detection methods offer a cost-effective and efficient alternative to supervised methods, and their effectiveness has been demonstrated in various applications.

3.2.3 Semi-Supervised Learning-Based Methods

Semi-supervised classification is a type of ML that combines both supervised and unsupervised learning methods. It aims to reduce the labeling cost of the supervised classification and improve the interpretability of the unsupervised classification. This approach is particularly useful when external information is scarce [137]. In a training dataset, some data are labeled, and the semi-supervised method can identify unlabeled classes associated with labeled patterns to

determine whether the unlabeled data belong to such clusters [138, 139].

One of the most commonly used machine learning models in semi-supervised learning-based load detectors is SVM. Yan et al. developed a semi-SVM-based anomaly detector to detect faults in air handling units [140]. The classifier iteratively inserts new test samples during semi-supervised learning and compares its classification accuracy to the pre-set confidence level threshold. If the accuracy is higher than the threshold, the training data size is increased. Wang et al. proposed a hybrid semi-supervised learning framework that employs SVM and K-means to achieve detection at low labeling costs [22]. The proposed method first employs K-means for data preprocessing. Then, an SVM-based classifier is proposed to identify the obtained patterns. Finally, the cross-entropy loss function is used to evaluate the classification results of the SVM. The classification result is determined after calculating the loss before and after introducing new samples into the dataset. Similar to Yan et al.'s work [140], if the classification accuracy is higher than a threshold, the test data is adopted into the training pool.

Iwayemi et al. developed a semi-supervised learning-based residential appliance annotator that created two-dimensional feature vectors featuring the dynamic time warping distance and the step changes in the power consumption of an appliance event [135]. The Mahalanobis distance is used to measure these feature vectors and identify the boundaries of appliance groups. After labeling all unlabeled data, semi-supervised learning-based algorithms completed the training. These hybrid models are shown to effectively achieve a good trade-off between model performance and labeling costs.

DL methods can enhance anomaly detection in semi-supervised learning-based approaches. Lu et al. developed a semi-supervised auto-encoder-based load anomaly detector [141]. The semi-supervised auto-encoder generative model consists of an encoder, decoder, discriminator, and classifier. The encoder and decoder are connected in the form of an autoencoder that captures the features of the data. Yang et al. designed a temporal convolutional network-based semi-supervised load monitoring method to evaluate

classification loss and compared the approach to machine learning- and deep learning-based methods [142]. In their work, the proposed temporal convolutional network-based method is found to be practical and applicable in real-time detection tasks when compared with other alternative semi-supervised detection approaches.

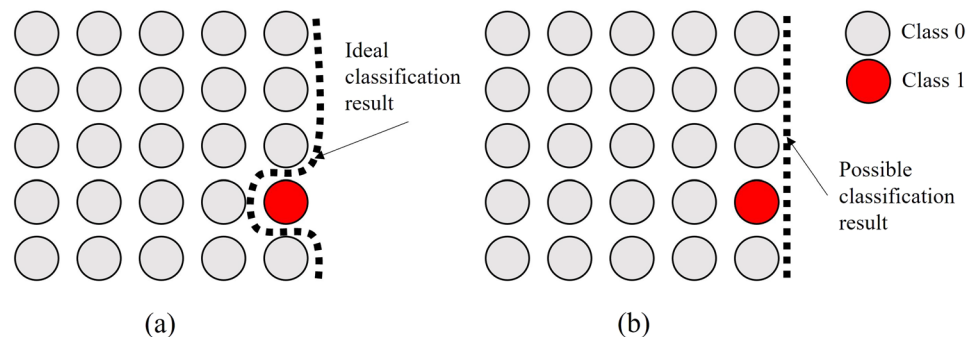
In addition to the advantage of relatively low labeling costs, semi-supervised learning is also robust to data sparsity, thereby reducing the impact of data imbalance on classification [119]. However, from the view of real-world application, semi-supervised learning has some limitations. It relies on accurate prior knowledge about the relationship between labeled and unlabeled data structures [137, 140, 143]. That limits its applications. Furthermore, introducing new unlabeled samples into the training data pool may cause performance degradation [144].

3.2.4 Data Imbalance and Optimization Strategies

The issue of data imbalance is prevalent in real-world power consumption datasets, where abnormal events occur at a relatively low frequency, resulting in a sparse distribution of abnormal data [122, 131]. However, this characteristic of sparse anomalies poses a challenge for neural network-based models, which are a fundamental component of many DL approaches [145]. Neural networks learn the features of each class by optimizing the weights and activations of each node during the learning phase [146]. However, the assumption underlying this learning is that the training data are uniformly distributed among all classes [131, 145, 146]. In other words, if a class is not evenly represented, the neural network-based model may not achieve optimal performance.

Figure 7 illustrates a conceptual view of imbalanced classification, where grey nodes represent training data in Class 0 (the majority class), and red nodes represent Class 1 (the minority class). The ideal classification result is shown in Figure 7a. However, due to the significant difference in the number of data samples in the two classes, the neural network often fails to learn the features of Class 1, resulting in the classification result shown in Figure 7b. To address this

Fig. 7 Impact of data imbalance on classification



issue, several strategies have been proposed to optimize the performance of imbalanced classification models.

a. Abnormal data generation

To address the problem of data imbalance, Jokar et al. developed an SVM- based anomaly detector that can identify electricity theft in a grid [122]. This method involves generating “fake data”—i.e., abnormal power consumption data—from benign data. Specifically, n data samples are randomly selected from the benign dataset to establish a similarly sized but new benign dataset labeled x . Electricity theft is typically associated with reduced paid power consumption, and the electricity theft dataset ‘ y ’ can be represented as ‘ $y = x * a$ ’, where ‘ a ’ is a preset parameter in $[0, 1]$. The method used to learn the optimized dataset is irrelevant, although the training is identical. However, a limitation of this method is that the distributions of all malicious samples are assumed to be consistent with the distributions of the benign samples, which may not be true in all cases.

This method has also been used in Wang et al.’s work [23]. Considering the imbalanced nature of real-world power consumption datasets, their work introduced this method to include some “fake data” in the training data to balance the dataset. Their test results show that the model with optimized training data effectively screens out anomalies. However, it should be noted that this method requires prior knowledge of the types of anomalies that need to be detected and may not be applicable in scenarios where the types of anomalies are unknown, or new types of anomalies are emerging.

b. One-class classification

The one-class classification approach is a special case in supervised learning- based detection and involves a binary rule where data are classified as abnormal if they cannot be classified into a benign dataset [109, 117]. This approach has been shown to save labeling costs and improve the efficiency of supervised- based detectors [147]. Several studies have compared the detection performance of one-class and multi-class classification methods in the presence of training data imbalance. Nguyen et al. and Fu et al. showed that one-class classification methods perform better in such cases [148, 149]. However, Kokar et al. found the opposite result in their study, where an SVM-based one-classification test failed to show promising results compared to multi-class-based classification [122]. Therefore, there is no clear-cut conclusion as to which method is better, but several studies have demonstrated that the one-class classification approach

is promising in reducing labeling costs and addressing data imbalance.

c. Class weights-based optimization

To further address the data imbalance problem, various optimization strategies have been proposed. One such strategy is the two-step optimization approach, which assigns optimal class weights during learning. This approach has been used in many works to improve the performance of classification models against data imbalance [150, 151].

The two-step optimization approach involves two main steps. In the first step, a data survey is conducted to determine which class is the minority class and the proportion of the minority class. In the second step, a proper class weight is set for the minority class, and the classification model can weigh the class more heavily during the training phase. This way, the impact of data imbalance on training can be reduced, and the model can be optimized for better performance [145, 151].

Although the two-step optimization approach can effectively address the data imbalance problem, it is important to note that the optimization process can lead to extra training time and computational costs. Therefore, the trade-off between the improved performance and increased training time and computational costs should be carefully considered when applying this approach.

d. Over-or-under sampling

This approach for addressing imbalanced data problems is based on data sampling techniques [152]. By randomly over-sampling or under-sampling the minority or majority class, the classification method can manage the number of samples in each class to train models in a more balanced way [153]. While this approach has not yet been applied to smart grid anomaly detection, Huan et al. previously proposed an under-sampling-based detection approach for screening abnormal traffic in network management [154]. Their work involved setting the number of clusters in the normal class to the number of abnormal data, and retaining samples closest to the cluster center to achieve under-sampling. The proposed approach uses clustering to avoid oversampling issues and provides a novel approach for effectively selecting samples from the majority class.

3.3 Discussion of Anomaly Detection Methods

We have presented many anomaly detection methods in the previous sections. Each comes with its unique advantages and limitations, making them suitable for different scenarios.

Table 4 Comparison of load anomaly detection methods

Methods type	Advantages	Limitations
Regression-based methods	One framework with load forecasting and anomaly detections	<ol style="list-style-type: none"> 1. Over-reliance on prediction results without independent determination mechanism 2. May not suitable for complex power systems
Classification based methods	High detection accuracy	<ol style="list-style-type: none"> 1. Difficult to define anomalies in power systems 2. High data labeling cost 3. ML methods may sensitive to data labeling noise 4. Customer privacy issue
Supervised learning	<ol style="list-style-type: none"> 1. Relatively high accuracy 2. Medium labeling cost 3. Good interpretability 4. Robust to training data sparsity 	<ol style="list-style-type: none"> 1. Relying on prior knowledge about the relationship between labeled and the unlabeled data structure; 2. Potential performance degradation when new unlabelled data is introduced; 3. Customer privacy issue
Semi-supervised learning	<ol style="list-style-type: none"> 1. No labeling cost 2. Robust to input noise 	<ol style="list-style-type: none"> 1. Low computational efficiency when managing big data; 2. Over-reliance on feature extraction; 3. Weak interpretability, especially in power systems; 4. Lack of ground truth for evaluation results
Unsupervised learning		

To assist in selecting the appropriate anomaly detection method for smart grid implementation, we summarize the advantages and limitations of these methods in Table 4.

It should be noted that various evaluation metrics have been proposed for assessing the performance of anomaly detection methods, similar to load forecasting strategies [155]. However, due to the imbalanced nature of anomalies in real-world datasets, there is ongoing debate regarding the most appropriate evaluation metric. Some research suggests that the G-mean score provides a more balanced interpretation of classification performance, while others claim that the Matthew correlation coefficient (MCC) is more suitable [156]. The choice of metric depends on the specific application and priorities of the stakeholders involved [65].

- Therefore, pointing out the “best” metric for anomaly detection is impractical [155].

In this work, we aim to evaluate the performance of each detection strategy from the perspectives of both ML/DL and real-world implementation.

The regression-based approach is a useful method as it can provide both load forecasting and anomaly detection within a single framework. However, while it can save training data and has a simple detection mechanism, its detection accuracy is limited. This is because the method lacks an independent abnormal data determination design, which causes it to rely too heavily on the prediction outcomes [23].

Moreover, this method may not be suitable for complex power systems with unpredictable loads.

Classification-based detection strategies offer a promising approach for effectively identifying anomalies. However, each approach has its own limitations. The supervised learning-based method, for instance, can provide high detection accuracy but is limited by its data labeling cost [22]. Additionally, defining an anomaly in the context of daily electricity usage variability can be challenging, which makes data labeling quality a potential issue. Moreover, customer privacy concerns must be taken into account as a preliminary data exploration phase is required.

Unsupervised learning-based approaches provide a labeling-free scheme and are robust to input noise [157, 158]. However, their applications are limited due to their high computational cost and weak interpretability [31, 131, 135, 136]. The semi-supervised learning approach strikes a balance between supervised and unsupervised learning approaches. However, it is important to note that the underlying assumption of semi-supervised learning is that all classes are separated clearly, which may make this approach unsuitable for some unknown datasets [137, 140, 143]. Additionally, as with supervised learning-based strategies, customer privacy issues are also involved in this approach.

Hyperspace Dimensional Computing (HDC) is a novel data-driven classification strategy recently introduced to the energy domain [155]. By mapping the data into a high-dimensional space, the features of each class can be learned

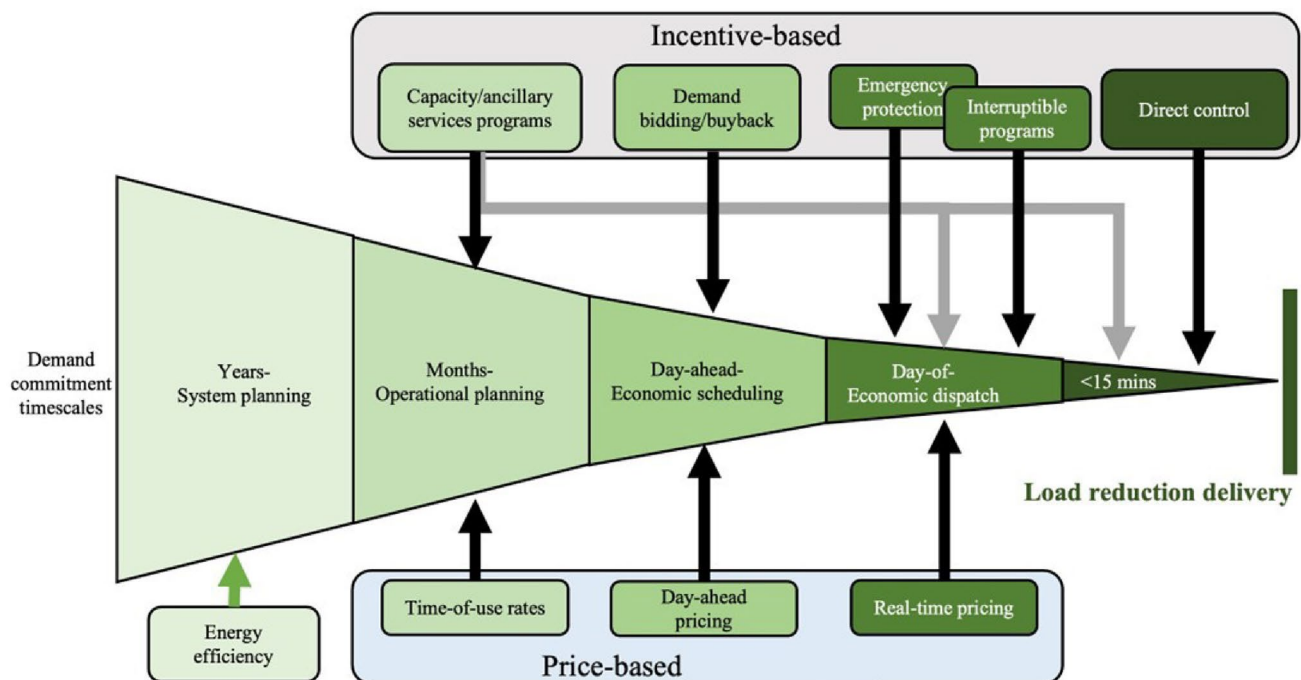


Fig. 8 Classification of demand response strategies

more efficiently. Additionally, HDC is hardware-friendly and can be implemented on a wide range of platforms, including low-power and embedded devices. Wang et al. proposed an HDC-based anomaly detection method, which is the first attempt to introduce this approach in the energy domain [155]. The proposed HDC-based method is compared with various approaches, including ML models such as SVM, KNN, AND, and DL models such as LSTM, DNN, and their class-weight optimized detectors. The test results show that the HDC-based detection method not only provides accurate detection results but also saves model optimization cost and pre-training costs. More importantly, this approach is training time-efficient, saving more time on large-scale dataset-based training. However, it should be noted that the HDC-based approach is based on supervised learning, which may be limited in its application in some cases due to the high data labeling costs.

4 Demand Response

Demand response has gained significant attention and has become an essential component of smart grids [159]. It is an effective way to manage demand by reducing system peak loads, enhancing system reliability, and delaying system upgrades [160, 161]. AI and data science technologies have provided an affordable solution for demand response, which has further propelled the development of power grids [159]. By learning from human power usage behavior, demand response can efficiently control the use of individual appliances, minimize user discomfort, increase the utilization of renewable energy, and reduce energy costs [162].

Demand response can be categorized into two types: incentive-based demand response and price-based demand response, as shown in Figure 8 [8, 160, 161, 163–165]. The incentive-based demand response suggests that users adjust their load profile or adopt some control over their appliances. This approach includes direct load control, interruptible service, demand bidding, capacity services, emergency protection, and ancillary service markets, according to different planning periods [8]. On the other hand, price-based demand response adjusts hourly electricity prices to reflect supply and demand balance or mismatch in real-time, such that users can respond to prices and adjust their load. This approach includes critical-peak pricing, time-of-use pricing, real-time pricing, and peak load reduction credits [8, 166, 167].

Demand response plays an essential role in improving energy efficiency, utilization of renewable energy sources, and reducing emissions from the power consumption side [163, 168].

4.1 AI for Incentive-Based Demand Response

Incentive-based demand response strategies aim to motivate users to adjust their energy consumption behavior to optimize demand. AI-based techniques have been developed to support such strategies. Noor et al. proposed a demand response strategy that aimed to reduce peak load, thus improving grid stability and reducing customer costs [32]. Sharma et al. used a whale optimization algorithm-based demand response approach to adjust the load curve and reshape it through strategic conservation, peak clipping, and load shifting [169]. The method is simple to implement and avoided local minima. Tutkun et al. proposed a load-shifting and valley-filling-based demand response scheme that optimized the daily energy cost for customers in an off-grid microgrid [170]. Werminski et al. developed a Decentralized Active Demand Response (DADR)-based demand response strategy that effectively reduced peak loads without requiring communication with other appliance components [171]. Li et al. proposed a multi-objective demand response optimization scheme that included an upper-layer utility, a middle-layer demand response aggregator, and lower-layer customers [8]. They employed an artificial immune system algorithm to solve the multi-objective problem and showed that all participants could benefit. Finally, Yang and Wang developed a virtual power plant (VPP) using AI techniques to manage residential energy resources and participate in the power system market to provide demand response, feed-in energy, and grid services [172].

Game theory has emerged as a powerful tool for developing demand response methods for intelligent customer agents. Jo et al. proposed an energy trading and operation game for storage charging and discharging to enhance energy efficiency [173]. The simulation results demonstrated a reduction in both the total energy cost and the peak-to-average energy ratio. Zhao et al. investigated the duopoly competition of renewable energy to determine the optimal sizing and operation of energy storage serving consumers in power systems [174]. They developed a three-stage game-theoretic model to model the interactions between suppliers and consumers and studied the equilibrium decisions of storage investment. The use of game theory in demand response optimization is expected to increase as it enables agents to make informed decisions based on strategic reasoning and considerations.

Compared with directly controlling the power systems, ML and DL approaches also play a critical role in understanding consumers' energy consumption patterns for developing more effective and fair demand response programs. Kwac and Rajagopal developed a data-driven method for selecting customers for demand response programs [175]. They propose a scalable method leveraging load data from

individual-level smart meters, and the proposed method is tested using real smart meter data in more than 58k residential households. Wang et al. developed a framework by leveraging ML for consumer activity detection using residential smart meter data, which can inform demand response design to provide customized programs for different consumer activities [176]. In addition to the efficiency of demand response programs, energy equity and fairness have attracted greater attention. Tang et al. identified the drivers of residential energy consumption patterns using both socio-economic and load data [177]. Different ML methods are used to capture the relationship between load patterns and socioeconomic factors, providing insights into how different consumers use energy differently. Wei and Wang adopted the symbolic aggregate approximation method to process load data and used the K-Means method to extract key load patterns. A DNN is designed to analyze the relationship between users' load patterns and their demographic features, which can help operators develop fair demand response programs for different social groups [178]. Wang developed a methodology to study seasonal variations in load patterns and to identify the relationship between seasonal variations in load and the socioeconomic factors of consumers. This can help operators and utilities develop better demand response programs for different seasons [179]. Babar et al. proposed an ML-based demand response strategy to be applied in an IoT-enabled grid, considering the security issue in power systems [180]. The Naive Bayes algorithm is used to classify the current power usage into different safety levels. Then, the effectiveness of IoT in power grids is exhibited. It is pointed out that many factors, such as communication distance requirements, should be comprehensively considered in choosing the communication method for power systems. Zhou et al. presented a predictive results-based demand response approach, in which the building power consumption is forecasted, and four advanced power usage control strategies are compared, including a rule-based controller, predictive controller, interactive feedback controller, and hybrid controller [181]. The proposed forecasting approach is a surrogate model with neural networks, which provides accurate outcomes by effectively discarding the unnecessary connections with small weights, thus overcoming the over-fitting issue and improving the training speed.

RL has been widely used to develop effective strategies for demand response, particularly under uncertain conditions [162]. For example, Qiu et al. proposed an RL-based battery storage scheduling strategy that controls the charge and discharge of batteries to optimize energy use [182]. Similarly, Xiong et al. used a Q-learning algorithm to develop a real-time control strategy for reducing energy loss in electric vehicles (EVs) [183]. By comparing the RL-based approach with traditional rule-based methods, they showed that RL could effectively minimize energy loss.

In addition, Kofinas et al. developed a fuzzy Q-learning algorithm to optimize a standalone microgrid system [184]. The RL approach is modified to ensure supply quality and reliability in the face of uncertainties in renewable energy sources and user demand. Mathew et al. developed a deep RL-based demand response strategy for optimizing home energy use by shifting loads to minimize the aggregate peak load [185]. Their method employed a deep RL algorithm for each agent, which received feedback from the environment after every action.

4.2 AI for Price-Based Demand Response

Price-based demand response methods adjust prices to change users' energy consumption behaviors to improve energy efficiency or reduce peak load. Herter and Song et al. developed demand response schemes to investigate how critical-peak pricing affected power usage [166, 186]. Herter studied the response of different users when critical-peak pricing is applied and found that high-power consumers respond more than low-power customers. Low-income users do not pay more on average when critical-peak pricing is in effect, and there is no significant difference in savings among users with different incomes. Ma et al. used an enhanced Arrow-d'Aspremont-Gerard-Varet (AGV) mechanism to solve the truth-telling problems associated with dynamic pricing in a smart grid [187]. This work proves that the enhanced AGV mechanism could realize the basic conditions of incentive compatibility, individual rationality, and budget balance [188, 189].

Leveraging RL technology, Liu et al. used an RL-based demand response to reduce the energy costs of a passive thermal storage inventory [178]. O'Neill developed an improved Q-learning-based demand response to reduce residents' energy costs [190]. The developed Q-learning algorithm learns the behaviors of users and automatically adapts to changes in behaviors to reschedule energy use. Zhou et al. employed a demand response for peer-to-peer trading of energy storage in a residential community [191]. A Markov decision process is employed to model energy trading, and a fuzzy Q-learning algorithm is used to optimize energy-trading decisions. The developed method facilitates peer-to-peer trading, enabling residents to enjoy cheap renewable energy and reduce energy costs. Lu et al. proposed a deep RL-based demand response method to assist service providers in purchasing energy resources from customers to off-set energy fluctuations and improve grid reliability [192]. DNN is utilized to forecast future power prices and demands to schedule the load. Then, RL is adopted to provide optimal incentive rates for different users, considering the profitability of both service providers and customers. The simulation results suggest that this proposed strategy can yield demand-side participation, benefit both service

providers and customers, and improve system reliability by balancing different resources.

Hamed and Kazemi proposed a multi-objective mixed-integer linear programming-based demand response method to reduce the cost of home power usage [193]. The power consumption activities are divided into three categories: time-shiftable appliances, power-shiftable appliances, and non-shiftable appliances. Their simulation results suggest that by considering different patterns of home power usage, the proposed method can take advantage of lower-cost pricing during mid-peak and off-peak periods, thereby reducing households' peak demand. Zhao et al. [194] designed ToU pricing to incentivize end-users energy storage deployment to help shave the system peak load and reduce the system cost. The proposed ToU pricing can reduce the system cost by over 30% compared to the benchmark with no storage investment. When there is no historical data about consumers, it becomes more challenging to develop an effective demand response algorithm because the consumers' behaviors are uncertain and unknown to the operator. Li et al. developed a joint online learning and pricing algorithm for demand response in [195]. In each time slot, the utility can estimate the cost functions of consumers based on their noisy responses. The proposed work assesses the performance of the algorithm leveraging regret analysis. The results indicate that the proposed method achieves logarithmic regret with respect to the operating horizon.

4.3 Applications of Demand Response in Different Premises

Demand response strategies have been developed for various applications, such as residential home energy usage, building energy consumption, and regional energy management [196–199].

For residential home energy usage, Nadeem et al. proposed a smart home load management and control strategy that integrates the usage of RESs with intelligent optimization algorithms [197]. Their proposed method utilizes a genetic algorithm, binary particle swarm optimization, and wind-driven optimization to achieve a trade-off among power usage reduction, price, and user comfort. The method distinguishes between scheduled and unscheduled power usage activities and adjusts them accordingly. Rocha et al. conducted a similar study on energy demand planning for smart homes, using three AI algorithms [198]. They developed an SVR-based model to forecast day-ahead distributed generations, and used an elitist non-dominated sorting genetic algorithm II to solve a multi-objective optimization problem, balancing user comfort and energy cost.

For building energy consumption, Tronchin et al. proposed a critical analysis of possible paths of renewable energy integration from the perspective of the built environment [196]. The authors suggested that a cross-sectorial data-driven model is necessary in the energy field, and emphasized the importance of investigating the spatial and temporal scalability of modeling techniques utilizing transparent metrics and key performance indicators (KPI). They demonstrated the scalability of inverse modeling techniques for model calibration targeting energy management and the extensibility of techniques for techno-economic optimization. Wang et al. proposed a demand response strategy to improve sustainable development in a living community in Africa, where only solar power is available [24]. The study integrated customer power usage habits and usage time to reduce blackouts even when the total energy consumption from users increases. Finally, Cai et al. presented a realistic case study on a demand response strategy in urban district heating networks [199]. They developed a day-ahead hourly schedule optimization for district heating substations, taking into account circulating pumps, distribution network, building space heating, and domestic hot water demand. Their experimental results indicate that by improving urban district heating operations, an energy cost saving of up to 11% is possible. The authors proposed a sensitivity analysis to quantify the sensitivity of the method to energy cost, comfort cost, and pumping power.

Demand response strategies have been widely applied in different premises, with various approaches and optimization techniques employed to achieve energy efficiency, cost reduction, and user comfort.

Table 5 Comparison of demand response approaches

Methods type	Advantages	Limitations
Incentive-based	<ol style="list-style-type: none"> 1. Results directly execute on power systems; 2. Better flexibility and analytical results are enhanced by AI; 3. Support smart grid integration with renewable energy sources 	Effectiveness is limited by the infrastructure of power systems, etc
Price-based	<ol style="list-style-type: none"> 1. Easy implementation; 2. Low hardware restrictions; 3. Good planning possibilities for customers; 4. Support smart grid integration with renewable energy sources 	<ol style="list-style-type: none"> 1. Customer sensitive; 2. Affected by local development level, environment, etc

4.4 Discussion of Demand Response

The utilization of AI in demand response has facilitated the learning and optimization of intricate power systems that exhibit diverse user and environmental behaviors. This has paved the way for the development of more effective and efficient demand response techniques [200]. These advancements are expected to maintain the balance between supply and demand sides, enhance system reliability, and mitigate supply constraints [32, 33, 201].

Table 5 presents a comparison of the advantages and challenges of two AI-empowered demand response strategies based on control mechanisms. As AI technology advances, both approaches offer better integration of power systems with renewable energy sources [202].

The incentive-based demand response approach provides an efficient way to manage power consumption through an intuitive feedback mechanism [160, 161]. Some AI techniques, such as online learning and RL, further enhance the effectiveness of this method by providing a comprehensive understanding of customers' features and better demand management [162]. However, from the perspective of real-world applications, this approach is still limited by the infrastructure of power systems, among other factors [159, 203–205].

In contrast, the price-based demand response approach proposes flexible solutions that are hardware-friendly, thereby increasing the utilization of renewable energy and addressing climate change [200, 202, 206]. The introduction of AI methods improves the performance of price-based demand response strategies by providing a more comprehensive and strategic approach to assist customers in scheduling their power usage [178]. However, the implementation of this approach needs to consider various factors that may affect its application, such as the environment, local development, and customer sensitivity to prices [206].

In conclusion, the choice of demand response strategy should consider trade-offs such as hardware limitations, software algorithm complexity, and local government policies and regulations. Incentive-based strategies are recommended for advanced power systems without hardware limitations, while price-based strategies provide a promising solution in real-world implementations but require a considerable amount of data sources to ensure its holistic performance.

5 Conclusion

Traditional energy sources are becoming exhausted, and the world is facing numerous unprecedented challenges, such as climate change. There is a growing need to bridge the gap between theoretical algorithms and their practical

implementation in power systems. This paper presents a comprehensive overview of the demand side of power systems, emphasizes the importance of considering real-world application challenges, and focuses on three crucial components: load forecasting, anomaly detection, and demand response. Our main contributions are as follows:

- We provide practical insights for evaluating, selecting, and optimizing various machine learning and deep learning models in each component, as well as offering a holistic view for better understanding and meeting the requirements of energy systems. At the same time, we analyze practical issues such as energy system sensor/input noises, data labeling errors/costs, the resilience of existing energy infrastructure, data imbalance, data availability, and operational constraints for better applications of different machine learning and deep learning models in power systems.
- In the load forecasting domain, we summarize previous efforts according to different data-driven technologies used, discuss promising optimization schemes considering implementation, and compare the advantages and limitations of reviewed prediction methods for different applications.
- We review anomaly detection approaches, provide a holistic summary of promising optimization schemes for addressing data imbalance issues, and discuss the associated challenges and trade-offs of these anomaly detection approaches.
- We introduce advanced strategies in demand response to comprehensively assess demand-side power usage and facilitate interaction between the system and consumers, ensuring the balance between power generation and consumption and the reliability of future power systems.

By integrating previous research in these domains, we offer a more comprehensive and strategic view of future sustainable development, elucidating the significance of each research area. We highlight the interconnected nature of these components through a feedback loop, emphasizing the importance of considering their interactions when designing data-driven approaches for energy systems. Our comprehensive review aims to serve as a roadmap for researchers and practitioners to better understand the capabilities of AI techniques in enhancing power consumption on the demand side. By examining the features and challenges of each field and discussing optimization strategies, this work could potentially drive innovation and inform the development of more sustainable and efficient power systems, ultimately benefiting the environment and society as a whole.

Funding Open Access funding enabled and organized by CAUL and its Member Institutions. Australian Research Council, DE230100046, Hao Wang.

Declarations

Conflict of interest On behalf of all authors, the corresponding authors state that there is no conflict of interest.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

References

- Omer, A. M. (2008). Energy, environment and sustainable development. *Renewable and Sustainable Energy Reviews*, 12(9), 2265–2300. <https://doi.org/10.1016/j.rser.2007.05.001>
- Ritchie, H. (2019). How long before we run out of fossil fuels? <https://ourworldindata.org/how-long-before-we-run-out-of-fossil-fuels>
- Climate change is accelerating the sixth extinction. <https://www.iberdrola.com/sustainability/climate-change-endangered-species>. Accessed July 2022
- Risteska Stojkoska, B. L., & Trivodaliev, K. V. (2017). A review of internet of things for smart home: Challenges and solutions. *Journal of Cleaner Production*, 140, 1454–1464. <https://doi.org/10.1016/j.jclepro.2016.10.006>
- Ritchie, H., Roser, M. (2020). Energy. Our World in Data. <https://ourworldindata.org/energy>. Accessed Mar 2020
- Panwar, N. L., Kaushik, S. C., & Kothari, S. (2011). Role of renewable energy sources in environmental protection: A review. *Renewable and Sustainable Energy Reviews*, 15(3), 1513–1524. <https://doi.org/10.1016/j.rser.2010.11.037>
- Agency, T.I.E. The international energy agency's electricity market report 2023. <https://iea.blob.core.windows.net/assets/255e9cba-da84-4681-8c1f-458ca1a3d9ca/ElectricityMarketReport2023.pdf>. Accessed Apr 2023
- Li, D., Chiu, W.-Y., Sun, H., & Poor, H. V. (2018). Multiobjective optimization for demand side management program in smart grid. *IEEE Transactions on Industrial Informatics*, 14(4), 1482–1490. <https://doi.org/10.1109/TII.2017.2776104>
- Thoubboron, K. (2021). Advantages and disadvantages of renewable energy. <https://news.energysage.com/advantages-and-disadvantages-of-renewable-energy>. Accessed Mar 2020
- Zhou, J., He, L., Li, C., Cao, Y., Liu, X., Geng, Y. (2013). What's the difference between traditional power grid and smart grid? from dispatching perspective. In: 2013 IEEE PES Asia-Pacific Power and Energy Engineering Conference (APPEEC), pp. 1–6. 10.1109/APPEEC.2013.6837107
- Wang, X., Ha, B., Lee, G.-Y., Kim, H., Yu, J., Rhee, H., Njau, K., Jande, Y., & Ahn, S.-H. (2020). Low-cost far-field wireless electrical load monitoring system applied in an off-grid rural area of tanzania. *Sustainable Cities and Society*, 59, 102209. <https://doi.org/10.1016/j.scs.2020.102209>
- Wang, X.-L., Ha, B., Manongi, F. A., Jung, W.-K., Jande, Y. A. C., & Ahn, S.-H. (2020). Arduino-based low-cost electrical load tracking system with a long-range mesh network. *Advances in Manufacturing*, 9, 47–63. <https://doi.org/10.1007/s40436-020-00310-5>
- Leitao, P., Karnouskos, S., Ribeiro, L., Lee, J., Strasser, T., & Colombo, A. W. (2016). Smart agents in industrial cyber-physical systems. *Proceedings of the IEEE*, 104(5), 1086–1101. <https://doi.org/10.1109/JPROC.2016.2521931>
- Ali, S. S., & Choi, B. J. (2020). State-of-the-art artificial intelligence techniques for distributed smart grids: A review. *Electronics*. <https://doi.org/10.3390/electronics9061030>
- Saecker, M., Markl, V. (2013). Big data analytics on modern hardware architectures: A technology survey, 138, 125–149. https://doi.org/10.1007/978-3-642-36318-4_6
- Markovic, D. S., Branovic, I., & Popovic, R. (2015). Smart grid and nanotechnologies: a solution for clean and sustainable energy. *Energy and Emission Control Technologies*, 2015, 1–13. <https://doi.org/10.2147/EECT.S48124>
- Yu, W., Wen, G., Yu, X., Wu, Z., & Lu, J. (2014). Bridging the gap between complex networks and smart grids. *Journal of Control and Decision*, 1(1), 102–114. <https://doi.org/10.1080/23307706.2014.885293>
- Tuballa, M. L., & Abundo, M. L. (2016). A review of the development of smart grid technologies. *Renewable and Sustainable Energy Reviews*, 59, 710–725. <https://doi.org/10.1016/j.rser.2016.01.011>
- Bose, B. K. (2017). Artificial intelligence techniques in smart grid and renewable energy systems—some example applications. *Proceedings of the IEEE*, 105(11), 2262–2273. <https://doi.org/10.1109/JPROC.2017.2756596>
- Wang, X., Rhee, H., & Ahn, S.-H. (2020). Off-grid power plant load management system applied in a rural area of africa. *Applied Sciences*, 10, 4171. <https://doi.org/10.3390/app10124171>
- Digest of UK Energy Statistics (DUKES) 2021. Department for Business, Energy & Industrial Strategy, London, UK. <https://www.gov.uk/government/statistics/digest-of-uk-energy-statistics-dukes-2021> (2021). Accessed Mar 2020
- Wang, X., Yang, I., & Ahn, S.-H. (2019). Sample efficient home power anomaly detection in real time using semi-supervised learning. *IEEE Access*, 7, 139712–139725. <https://doi.org/10.1109/ACCESS.2019.2943667>
- Wang, X., & Ahn, S.-H. (2020). Real-time prediction and anomaly detection of electrical load in a residential community. *Applied Energy*, 259, 114145. <https://doi.org/10.1016/j.apenergy.2019.114145>
- Wang, X., Wang, H., & Ahn, S.-H. (2021). Demand-side management for off-grid solar-powered microgrids: A case study of rural electrification in tanzania. *Energy*, 224, 120229. <https://doi.org/10.1016/j.energy.2021.120229>
- Raza, M. Q., & Khosravi, A. (2015). A review on artificial intelligence based load demand forecasting techniques for smart grid and buildings. *Renewable and Sustainable Energy Reviews*, 50, 1352–1372.
- Ahmad, T., Chen, H., Guo, Y., & Wang, J. (2018). A comprehensive overview on the data driven and large scale based approaches for forecasting of building energy demand: A review. *Energy and Buildings*, 165, 301–320.
- Khan, A. R., Mahmood, A., Safdar, A., Khan, Z. A., & Khan, N. A. (2016). Load forecasting, dynamic pricing and dsm in smart grid: A review. *Renewable and Sustainable Energy Reviews*, 54, 1311–1322. <https://doi.org/10.1016/j.rser.2015.10.117>

28. Himeur, Y., Ghanem, K., Alsalemi, A., Bensaali, F., & Amira, A. (2021). Artificial intelligence based anomaly detection of energy consumption in buildings: A review, current trends and new perspectives. *Applied Energy*, 287, 116601. <https://doi.org/10.1016/j.apenergy.2021.116601>
29. Antonopoulos, I., Robu, V., Couraud, B., Kirli, D., Norbu, S., Kiprakis, A., Flynn, D., Elizondo-Gonzalez, S., & Wattam, S. (2020). Artificial intelligence and machine learning approaches to energy demand-side response: A systematic review. *Renewable and Sustainable Energy Reviews*, 130, 109899. <https://doi.org/10.1016/j.rser.2020.109899>
30. Wang, Y., Chen, Q., Hong, T., & Kang, C. (2018). Review of smart meter data analytics: Applications, methodologies, and challenges. *IEEE Transactions on Smart Grid*, 10(3), 3125–3148.
31. Li, D., & Dick, S. (2019). Residential household non-intrusive load monitoring via graph-based multi-label semi-supervised learning. *IEEE Transactions on Smart Grid*, 10(4), 4615–4627. <https://doi.org/10.1109/TSG.2018.2865702>
32. Noor, S., Yang, W., Guo, M., van Dam, K. H., & Wang, X. (2018). Energy demand side management within micro-grid networks enhanced by blockchain. *Applied Energy*, 228, 1385–1398. <https://doi.org/10.1016/j.apenergy.2018.07.012>
33. Logenthiran, T., Srinivasan, D., & Shun, T. Z. (2012). Demand side management in smart grid using heuristic optimization. *IEEE Transactions on Smart Grid*, 3(3), 1244–1252. <https://doi.org/10.1109/TSG.2012.2195686>
34. Agrawal, R.K., Muchahary, F., Tripathi, M.M. (2018). Long term load forecasting with hourly predictions based on long-short-term-memory networks. In: 2018 IEEE Texas Power and Energy Conference (TPEC), pp. 1–6. <https://doi.org/10.1109/TPEC.2018.8312088>
35. Kuo, P.-H., & Huang, C.-J. (2018). A high precision artificial neural networks model for short-term energy load forecasting. *Energies*, 11(1), 213. <https://doi.org/10.3390/en11010213>
36. Hosein, S., Hosein, P. (2017). Load forecasting using deep neural networks. In: 2017 IEEE Power Energy Society Innovative Smart Grid Technologies Conference (ISGT), pp. 1–5. <https://doi.org/10.1109/ISGT.2017.8085971>
37. Vrablcova, P., Bou Ezzeddine, A., Rozinajova, V., Sarik, S., & Sangaiah, A. K. (2018). Smart grid load forecasting using online support vector regression. *Computers & Electrical Engineering*, 65, 102–117. <https://doi.org/10.1016/j.compeleceng.2017.07.006>
38. Zhang, C., Li, J., Zhao, Y., Li, T., Chen, Q., & Zhang, X. (2020). A hybrid deep learning-based method for short-term building energy load prediction combined with an interpretation process. *Energy and Buildings*, 225, 110301. <https://doi.org/10.1016/j.enbuild.2020.110301>
39. Bedi, J., & Toshniwal, D. (2018). Empirical mode decomposition based deep learning for electricity demand forecasting. *IEEE Access*, 6, 49144–49156. <https://doi.org/10.1109/ACCESS.2018.2867681>
40. Ahmad, T., Zhang, H., & Yan, B. (2020). A review on renewable energy and electricity requirement forecasting models for smart grid and buildings. *Sustainable Cities and Society*, 55, 102052. <https://doi.org/10.1016/j.scs.2020.102052>
41. Liu, K., Shang, Y., Ouyang, Q., & Widanage, W. D. (2021). A data-driven approach with uncertainty quantification for predicting future capacities and remaining useful life of lithium-ion battery. *IEEE Transactions on Industrial Electronics*, 68(4), 3170–3180. <https://doi.org/10.1109/TIE.2020.2973876>
42. Almalaq, A., Edwards, G. (2017). A review of deep learning methods applied on load forecasting. In: 2017 16th IEEE International Conference on Machine Learning and Applications (ICMLA), pp. 511–516. <https://doi.org/10.1109/ICMLA.2017.0-110>
43. Ij, H. (2018). Statistics versus machine learning. *Nature Methods*, 15(4), 233.
44. Ongsulee, P. (2017). Artificial intelligence, machine learning and deep learning. In: 2017 15th International Conference on ICT and Knowledge Engineering (ICT&KE), pp. 1–6. IEEE
45. Chahal, A., & Gulia, P. (2019). Machine learning and deep learning. *International Journal of Innovative Technology and Exploring Engineering*, 8(12), 4910–4914.
46. Espinoza, M., Joye, C., Belmans, R., & De Moor, B. (2005). Short-term load forecasting, profile identification, and customer segmentation: A methodology based on periodic time series. *IEEE Transactions on Power Systems*, 20(3), 1622–1630. <https://doi.org/10.1109/TPWRS.2005.852123>
47. Fan, C., Xiao, F., & Zhao, Y. (2017). A short-term building cooling load prediction method using deep learning algorithms. *Applied Energy*, 195, 222–233. <https://doi.org/10.1016/j.apenergy.2017.03.064>
48. Syed, D., Refaat, S.S., Abu-Rub, H. (2020). Performance evaluation of distributed machine learning for load forecasting in smart grids. In: 2020 Cybernetics Informatics (K I), pp. 1–6. <https://doi.org/10.1109/KI48306.2020.9039797>
49. Ming, Y., Cao, J. (2018). Electrical load prediction in energy internet via linear correlation coefficient approach. In: 2018 IEEE International Conference on Energy Internet (ICEI), pp. 157–162. <https://doi.org/10.1109/ICEI.2018.00036>
50. Yildiz, B., Bilbao, J. I., & Sproul, A. B. (2017). A review and analysis of regression and machine learning models on commercial building electricity load forecasting. *Renewable and Sustainable Energy Reviews*, 73, 1104–1122. <https://doi.org/10.1016/j.rser.2017.02.023>
51. Lee, C.-M., & Ko, C.-N. (2011). Short-term load forecasting using lifting scheme and arima models. *Expert Systems with Applications*, 38(5), 5902–5911. <https://doi.org/10.1016/j.eswa.2010.11.033>
52. Zou, Z., Wu, X., Zhao, Z., Wang, Q., bie, Y., Zhou, M. (2018). Prediction of short term electric load based on bp neural networks amp; arima combination. In: 2018 IEEE 4th Information Technology and Mechatronics Engineering Conference (ITOEC), pp. 1671–1674. <https://doi.org/10.1109/ITOEC.2018.8740553>
53. Fan, C., Xiao, F., & Wang, S. (2014). Development of prediction models for nextday building energy consumption and peak power demand using data mining techniques. *Applied Energy*, 127, 1–10. <https://doi.org/10.1016/j.apenergy.2014.04.016>
54. Khorshed, E. (2018). Long-term energy peak load forecasting models: A hybrid statistical approach. In: 2018 Advances in Science and Engineering Technology International Conferences (ASET), pp. 1–6. <https://doi.org/10.1109/ICASET.2018.8376792>
55. Fan, C., & Ding, Y. (2019). Cooling load prediction and optimal operation of hvac systems using a multiple nonlinear regression model. *Energy and Buildings*, 197, 7–17. <https://doi.org/10.1016/j.enbuild.2019.05.043>
56. Fan, C., Ding, Y., & Liao, Y. (2019). Analysis of hourly cooling load prediction accuracy with data-mining approaches on different training time scales. *Sustainable Cities and Society*, 51, 101717. <https://doi.org/10.1016/j.scs.2019.101717>
57. Cheng, M.-Y., & Cao, M.-T. (2014). Accurately predicting building energy performance using evolutionary multivariate adaptive regression splines. *Applied Soft Computing*, 22, 178–188. <https://doi.org/10.1016/j.asoc.2014.05.015>
58. Server, F., Kissock, J. K., Brown, D., & Mulqueen, S. (2011). Hestimating industrial building energy savings using inverse simulation. *Mechanical and Aerospace Engineering Faculty Publications*, 156, 1–8.
59. Kissock, J. K., Reddy, T. A., & Claridge, D. E. (1998). Ambient-temperature regression analysis for estimating retrofit savings

- in commercial buildings. *Journal of Solar Energy Engineering*, 120(3), 168–176. <https://doi.org/10.1115/1.2888066>
60. Walter, T., & Sohn, M. (2016). A regression-based approach to estimating retrofit savings using the building performance database. *Applied Energy*, 179, 996–1005. <https://doi.org/10.1016/j.apenergy.2016.07.087>
 61. Katipamula, S., Reddy, T. A., & Claridge, D. E. (1998). Multivariate regression modeling. *Journal of Solar Energy Engineering*, 120(3), 177–184. <https://doi.org/10.1115/1.2888067>
 62. Reddy, T. A., Katipamula, S., Kissock, J. K., & Claridge, D. E. (1995). The functional basis of steady-state thermal energy use in air-side HVAC equipment. *Journal of Solar Energy Engineering*, 117(1), 31–39. <https://doi.org/10.1115/1.2847720>
 63. Solyali, D. (2020). A comparative analysis of machine learning approaches for short-/long-term electricity load forecasting in cyprus. *Sustainability*. <https://doi.org/10.3390/su12093612>
 64. Ding, Y., Zhang, Q., Yuan, T., & Yang, F. (2017). Effect of input variables on cooling load prediction accuracy of an office building. *Applied Thermal Engineering*, 128, 225–234. <https://doi.org/10.1016/j.applthermaleng.2017.09.007>
 65. Jurado, S., Nebot, A., Mugica, F., & Avellana, N. (2015). Hybrid methodologies for electricity load forecasting: Entropy-based feature selection with machine learning and soft computing techniques. *Energy*. <https://doi.org/10.1016/j.energy.2015.04.039>
 66. Vantuch, T., Vidal, A.G., Ramallo-Gonzalez, A.P., Skarmeta, A.F., Misa'k, S. (2018). Machine learning based electric load forecasting for short and long-term period. In: 2018 IEEE 4th World Forum on Internet of Things (WF-IoT), pp. 511–516. <https://doi.org/10.1109/WF-IoT.2018.8355123>
 67. Sangrody, H., Zhou, N., Tutun, S., Khorramdel, B., Motalleb, M., Sarailoo, M. (2018). Long term forecasting using machine learning methods. In: 2018 IEEE Power and Energy Conference at Illinois (PECI), pp. 1–5. <https://doi.org/10.1109/PECI.2018.8334980>
 68. Zhang, Y., Chen, W., Black, J. (2011). Anomaly detection in premise energy consumption data. In: 2011 IEEE Power and Energy Society General Meeting, pp. 1–8. <https://doi.org/10.1109/PES.2011.6039858>
 69. Amber, K. P., Ahmad, R., Aslam, M. W., Kousar, A., Usman, M., & Khan, M. S. (2018). Intelligent techniques for forecasting electricity consumption of buildings. *Energy*, 157, 886–893. <https://doi.org/10.1016/j.energy.2018.05.155>
 70. Patel, M., Dabhi, D., Patel, R., & Patel, J. (2019). Long term electrical load forecasting considering temperature effect using multi-layer perceptron neural network and k-nearest neighbor algorithms. *International Journal of Research in Electronics and Computer Engineering*. <https://doi.org/10.13140/RG.2.2.29592.65288>
 71. Wang, X., Yao, Z., & Papaefthymiou, M. (2023). A real-time electrical load forecasting and unsupervised anomaly detection framework. *Applied Energy*, 330, 120279.
 72. Zhang, S., An, D., He, Z. (2019). Research on load prediction based on improve gwo and elm in cloud computing. In: 2019 IEEE 5th International Conference on Computer and Communications (ICCC), pp. 102–105. <https://doi.org/10.1109/ICCC47050.2019.9064097>
 73. Bouktif, S., Fiaz, A., Ouni, A., & Serhani, M. A. (2018). Optimal deep learning lstm model for electric load forecasting using feature selection and genetic algorithm: Comparison with machine learning approaches. *Energies*. <https://doi.org/10.3390/en11071636>
 74. Zheng, J., Xu, C., Zhang, Z., Li, X. (2017). Electric load forecasting in smart grids using long-short-term-memory based recurrent neural network. In: 2017 51st Annual Conference on Information Sciences and Systems (CISS), pp. 1–6. <https://doi.org/10.1109/CISS.2017.7926112>
 75. Kumari, A., Vekaria, D., Gupta, R., Tanwar, S. (2020). Redills: Deep learning-based secure data analytic framework for smart grid systems. In: 2020 IEEE International Conference on Communications Workshops (ICC Workshops), pp. 1–6. <https://doi.org/10.1109/ICCWorkshops49005.2020.9145448>
 76. Song, B., Yu, Y., Zhou, Y., Wang, Z., & Du, S. (2018). Host load prediction with long short-term memory in cloud computing. *The Journal of Supercomputing*, 74, 6554–6568. <https://doi.org/10.1007/s11227-017-2044-4>
 77. Cenek, M., Haro, R., Sayers, B., & Peng, J. (2018). Climate change and power security: Power load prediction for rural electrical microgrids using long short term memory and artificial neural networks. *Applied Sciences*. <https://doi.org/10.3390/app8050749>
 78. Kwon, B. S., Park, R. J., & Song, K. B. (2020). Short-term load forecasting based on deep neural networks using lstm layer. *Journal of Electrical Engineering & Technology*, 15, 1501–1509. <https://doi.org/10.1007/s42835-020-00424-7>
 79. Zhang, Y., & Song, X. (2019). Load prediction of space deployable structure based on fbg and lstm. *IEEE Access*, 7, 13715–13722. <https://doi.org/10.1109/ACCESS.2019.2893364>
 80. LeCun, Y., & Bengio, Y. (1998). *Convolutional Networks for Images, Speech, and Time Series* (pp. 225–258). Cambridge, MA, USA: MIT Press.
 81. Aurangzeb, K., Alhussein, M., Javaid, K., & Haider, S. I. (2021). A pyramid-cnn based deep learning model for power load forecasting of similar-profile energy customers based on clustering. *IEEE Access*, 9, 14992–15003. <https://doi.org/10.1109/ACCESS.2021.3053069>
 82. Amarasinghe, K., Marino, D.L., Manic, M. (2017). Deep neural networks for energy load forecasting. In: 2017 IEEE 26th International Symposium on Industrial Electronics (ISIE), pp. 1483–1488. <https://doi.org/10.1109/ISIE.2017.8001465>
 83. Dong, X., Qian, L., Huang, L. (2017). Short-term load forecasting in smart grid: A combined cnn and k-means clustering approach. In: 2017 IEEE International Conference on Big Data and Smart Computing (BigComp), pp. 119–125. <https://doi.org/10.1109/BIGCOMP.2017.7881726>
 84. Wang, Z., Hong, T., & Piette, M. A. (2020). Building thermal load prediction through shallow machine learning and deep learning. *Applied Energy*, 263, 114683. <https://doi.org/10.1016/j.apenergy.2020.114683>
 85. Yang, Q., Zhou, Y., Yu, Y., Yuan, J., Xing, X., & Du, S. (2015). Multi-step-ahead host load prediction using autoencoder and echo state networks in cloud computing. *The Journal of Supercomputing*, 71(8), 3037–3053. <https://doi.org/10.1007/s11227-015-1426-8>
 86. Wang, X., Papaefthymiou, M. (2022). A dual-mode real-time electrical load forecasting framework. In: 2022 IEEE Power & Energy Society Innovative Smart Grid Technologies Conference (ISGT), pp. 1–5. IEEE
 87. Nason, G.P. (2006). Stationary and non-stationary time series. In: *Statistics in volcanology*, (vol 60, pp. 137–139)
 88. Qiu, X., Ren, Y., Suganthan, P. N., & Amaratunga, G. A. J. (2017). Empirical mode decomposition based ensemble deep learning for load demand time series forecasting. *Applied Soft Computing*, 54, 246–255. <https://doi.org/10.1016/j.asoc.2017.01.015>
 89. Fan, M., Hu, Y., Zhang, X., Yin, H., Yang, Q., Fan, L. (2019). Short-term load forecasting for distribution network using decomposition with ensemble prediction. In: 2019 Chinese Automation Congress (CAC), pp. 152–157. <https://doi.org/10.1109/CAC48633.2019.8997169>
 90. Park, K., Yoon, S., & Hwang, E. (2019). Hybrid load forecasting for mixed-use complex based on the characteristic load

- decomposition by pilot signals. *IEEE Access*, 7, 12297–12306. <https://doi.org/10.1109/ACCESS.2019.2892475>
91. Bahdanau, D., Cho, K., Bengio, Y. (2014). Neural machine translation by jointly learning to align and translate. arXiv preprint [arXiv:1409.0473](https://arxiv.org/abs/1409.0473). Accessed Mar 2020
 92. Li, A., Xiao, F., Zhang, C., & Fan, C. (2021). Attention-based interpretable neural network for building cooling load prediction. *Applied Energy*, 299, 117238. <https://doi.org/10.1016/j.apenergy.2021.117238>
 93. Jin, X.-B., Zheng, W.-Z., Kong, J.-L., Wang, X.-Y., Bai, Y.-T., Su, T.-L., & Lin, S. (2021). Deep-learning forecasting method for electric power load via attention-based encoder-decoder with bayesian optimization. *Energies*, 14(6), 1596.
 94. Wang, S., Wang, X., Wang, S., & Wang, D. (2019). Bi-directional long short-term memory method based on attention mechanism and rolling update for short-term load forecasting. *International Journal of Electrical Power & Energy Systems*, 109, 470–479.
 95. Wu, K., Wu, J., Feng, L., Yang, B., Liang, R., Yang, S., & Zhao, R. (2021). An attention-based cnn-lstm-bilstm model for short-term electric load forecasting in integrated energy system. *International Transactions on Electrical Energy Systems*, 31(1), 12637.
 96. Sehovac, L., & Grolinger, K. (2020). Deep learning for load forecasting: Sequence to sequence recurrent neural networks with attention. *IEEE Access*, 8, 36411–36426.
 97. Sutton, R. S., & Barto, A. G. (2018). *Reinforcement Learning: An Introduction*. Cambridge: MIT Press.
 98. Feng, C., Zhang, J. (2019). Reinforcement learning based dynamic model selection for short-term load forecasting. In: 2019 IEEE Power & Energy Society Innovative Smart Grid Technologies Conference (ISGT), pp. 1–5. IEEE
 99. Dayan, P., & Niv, Y. (2008). Reinforcement learning: The good, the bad and the ugly. *Current opinion in neurobiology*, 18(2), 185–196.
 100. Park, R.-J., Song, K.-B., & Kwon, B.-S. (2020). Short-term load forecasting algorithm using a similar day selection method based on reinforcement learning. *Energies*, 13(10), 2640.
 101. Xie, G., Chen, X., & Weng, Y. (2021). Enhance load forecastability: Optimize data sampling policy by reinforcing user behaviors. *European Journal of Operational Research*, 295(3), 924–934.
 102. Dabbaghjamanesh, M., Moeini, A., & Kavousi-Fard, A. (2021). Reinforcement learning-based load forecasting of electric vehicle charging station using q-learning technique. *IEEE Transactions on Industrial Informatics*, 17(6), 4229–4237. <https://doi.org/10.1109/TII.2020.2990397>
 103. Callaghan, P. M., & Kunz, D. L. (2021). Evaluation of unmanned aircraft flying/handling qualities using a stitched learjet model. *Journal of Guidance, Control, and Dynamics*, 44(4), 842–853.
 104. Massaoudi, M., Refaat, S. S., Chihi, I., Trabelsi, M., Oueslati, F. S., & Abu-Rub, H. (2021). A novel stacked generalization ensemble-based hybrid lgbm-xgb-mlp model for short-term load forecasting. *Energy*, 214, 118874.
 105. Koprinska, I., Rana, M., & Agelidis, V. G. (2015). Correlation and instance based feature selection for electricity load forecasting. *Knowledge-Based Systems*, 82, 29–40.
 106. Berenji, H. R., & Vengerov, D. (2003). A convergent actor-critic-based frl algorithm with application to power management of wireless transmitters. *IEEE Transactions on Fuzzy Systems*, 11(4), 478–485.
 107. Feng, L., Xu, S., Zhang, L., Wu, J., Zhang, J., Chu, C., Wang, Z., & Shi, H. (2020). Anomaly detection for electricity consumption in cloud computing: framework, methods, applications, and challenges. *EURASIP Journal on Wireless Communications and Networking*. <https://doi.org/10.1186/s13638-020-01807-0>
 108. Rashid, H., Singh, P., Stankovic, V., & Stankovic, L. (2019). Can non-intrusive load monitoring be used for identifying an appliance’s anomalous behaviour? *Applied Energy*, 238, 796–805. <https://doi.org/10.1016/j.apenergy.2019.01.061>
 109. De Santis, E., Rizzi, A., Sadeghian, A., Frattale Mascioli, F.M. (2015). A learning intelligent system for fault detection in smart grid by a one-class classification approach. In: 2015 International Joint Conference on Neural Networks (IJCNN), pp. 1–8. <https://doi.org/10.1109/IJCNN.2015.7280756>
 110. Zeifman, M., & Roth, K. (2011). Nonintrusive appliance load monitoring: Review and outlook. *IEEE Transactions on Consumer Electronics*, 57(1), 76–84. <https://doi.org/10.1109/TCE.2011.5735484>
 111. Hussain, S., Mustafa, M. W., Jumani, T. A., Baloch, S. K., & Saeed, M. S. (2020). A novel unsupervised feature-based approach for electricity theft detection using robustpca and outlier removal clustering algorithm. *International Transactions on Electrical Energy Systems*. <https://doi.org/10.1002/2050-7038.12572>
 112. Spagnuolo, A., Petraglia, A., Vetromile, C., Formosi, R., & Lubritto, C. (2015). Monitoring and optimization of energy consumption of base transceiver stations. *Energy*. <https://doi.org/10.1016/j.energy.2014.12.040>
 113. Chou, J.-S., & Telaga, A. S. (2014). Real-time detection of anomalous power consumption. *Renewable and Sustainable Energy Reviews*, 33, 400–411. <https://doi.org/10.1016/j.rser.2014.01.088>
 114. Hollingsworth, K., Rouse, K., Cho, J., Harris, A., Sartipi, M., Sozer, S., Enevoldson, B. (2018). Energy anomaly detection with forecasting and deep learning. In: 2018 IEEE International Conference on Big Data (Big Data), pp. 4921–4925. <https://doi.org/10.1109/BigData.2018.8621948>
 115. Luo, J., Hong, T., & Yue, M. (2018). Real-time anomaly detection for very short-term load forecasting. *Journal of Modern Power Systems and Clean Energy*, 6(2), 235–243. <https://doi.org/10.1007/s40565-017-0351-7>
 116. Fenza, G., Gallo, M., & Loia, V. (2019). Drift-aware methodology for anomaly detection in smart grid. *IEEE Access*, 7, 9645–9657. <https://doi.org/10.1109/ACCESS.2019.2891315>
 117. Cui, M., Wang, J., & Yue, M. (2019). Machine learning-based anomaly detection for load forecasting under cyberattacks. *IEEE Transactions on Smart Grid*, 10(5), 5724–5734. <https://doi.org/10.1109/TSG.2018.2890809>
 118. Xu, C., & Chen, H. (2020). A hybrid data mining approach for anomaly detection and evaluation in residential buildings energy data. *Energy and Buildings*, 215, 109864. <https://doi.org/10.1016/j.enbuild.2020.109864>
 119. Ozay, M., Esnaola, I., Yarman Vural, F. T., Kulkarni, S. R., & Poor, H. V. (2016). Machine learning methods for attack detection in the smart grid. *IEEE Transactions on Neural Networks and Learning Systems*, 27(8), 1773–1786. <https://doi.org/10.1109/TNNLS.2015.2404803>
 120. Nagi, J., Yap, K. S., Tiong, S. K., Ahmed, S. K., & Mohamad, M. (2010). Nontechnical loss detection for metered customers in power utility using support vector machines. *IEEE Transactions on Power Delivery*, 25(2), 1162–1171. <https://doi.org/10.1109/TPWRD.2009.2030890>
 121. Depuru, S. S. R., Wang, L., Devabhaktuni, V., & Green, R. C. (2013). High performance computing for detection of electricity theft. *International Journal of Electrical Power & Energy Systems*, 47, 21–30. <https://doi.org/10.1016/j.ijepes.2012.10.031>
 122. Jokar, P., Arianpoo, N., & Leung, V. C. M. (2016). Electricity theft detection in ami using customers’ consumption patterns. *IEEE Transactions on Smart Grid*, 7(1), 216–226. <https://doi.org/10.1109/TSG.2015.2425222>
 123. Pinceti, A., Sankar, L., Kosut, O. (2018). Load redistribution attack detection using machine learning: A data-driven approach. In: 2018 IEEE Power Energy Society General Meeting

- (PESGM), pp. 1–5. <https://doi.org/10.1109/PESGM.2018.8586644>
124. Makonin, S., Popowich, F., Bajic, I. V., Gill, B., & Bartram, L. (2016). Exploiting hmm sparsity to perform online real-time nonintrusive load monitoring. *IEEE Transactions on Smart Grid*, 7(6), 2575–2585. <https://doi.org/10.1109/TSG.2015.2494592>
 125. Devlin, M., Hayes, B.P.: Non-intrusive load monitoring using electricity smart meter data: A deep learning approach. In: 2019 IEEE Power Energy Society General Meeting (PESGM), pp. 1–5 (2019). <https://doi.org/10.1109/PESGM40551.2019.8973732>
 126. Buzau, M.-M., Tejedor-Aguilera, J., Cruz-Romero, P., & Gomez-Exposito, A. (2020). Hybrid deep neural networks for detection of non-technical losses in electricity smart meters. *IEEE Transactions on Power Systems*, 35(2), 1254–1263. <https://doi.org/10.1109/TPWRS.2019.2943115>
 127. He, Y., Mendis, G. J., & Wei, J. (2017). Real-time detection of false data injection attacks in smart grid: A deep learning-based intelligent mechanism. *IEEE Transactions on Smart Grid*, 8(5), 2505–2516. <https://doi.org/10.1109/TSG.2017.2703842>
 128. Rolnick, D., Veit, A., Belongie, S., Shavit, N. (2017). Deep learning is robust to massive label noise. arXiv preprint [arXiv:1705.10694](https://arxiv.org/abs/1705.10694). Accessed Mar 2020
 129. Brown, O., Curtis, A., Goodwin, J. (2021). Principles for evaluation of ai/ml model performance and robustness. arXiv preprint [arXiv:2107.02868](https://arxiv.org/abs/2107.02868). Accessed Mar 2020
 130. Javed, A. R., Usman, M., Rehman, S. U., Khan, M. U., & Haghghi, M. S. (2021). Anomaly detection in automated vehicles using multistage attention-based convolutional neural network. *IEEE Transactions on Intelligent Transportation Systems*, 22(7), 4291–4300. <https://doi.org/10.1109/TITS.2020.3025875>
 131. Fan, C., Xiao, F., Zhao, Y., & Wang, J. (2018). Analytical investigation of autoencoder-based methods for unsupervised anomaly detection in building energy data. *Applied Energy*, 211, 1123–1135. <https://doi.org/10.1016/j.apenergy.2017.12.005>
 132. Pereira, J., Silveira, M. (2018). Unsupervised anomaly detection in energy time series data using variational recurrent autoencoders with attention. In: 2018 17th IEEE International Conference on Machine Learning and Applications (ICMLA), pp. 1275–1282. <https://doi.org/10.1109/ICMLA.2018.00207>
 133. Zheng, R., Gu, J., Jin, Z., Peng, H., & Zhu, Y. (2020). Load forecasting under data corruption based on anomaly detection and combined robust regression. *International Transactions on Electrical Energy Systems*, 30(7), 12103. <https://doi.org/10.1002/2050-7038.12103>
 134. Zhao, B., Stankovic, L., & Stankovic, V. (2016). On a training-less solution for non-intrusive appliance load monitoring using graph signal processing. *IEEE Access*, 4, 1784–1799. <https://doi.org/10.1109/ACCESS.2016.2557460>
 135. Iwayemi, A., & Zhou, C. (2017). Saraa: Semi-supervised learning for automated residential appliance annotation. *IEEE Transactions on Smart Grid*, 8(2), 779–786. <https://doi.org/10.1109/TSG.2015.2498642>
 136. Proctor, T., Shaw, R. (2017). You work for them Our Machines now have knowledge we'll never understand. <https://www.wired.com/story/our-machines-now-have-knowledge-well-never-understand/>
 137. Barsim, K.S., Yang, B. (2015). Toward a semi-supervised non-intrusive load monitoring system for event-based energy disaggregation. In: 2015 IEEE Global Conference on Signal and Information Processing (GlobalSIP), pp. 58–62. <https://doi.org/10.1109/GlobalSIP.2015.7418156>
 138. Bair, E. (2013). Semi-supervised clustering methods. *WIREs Computational Statistics*, 5(5), 349–361. <https://doi.org/10.1002/wics.1270>
 139. Elhamifar, E., Sastry, S. (2015). Energy disaggregation via learning ‘power-lets’ and sparse coding. In: Proceedings of the Twenty-Ninth AAAI Conference on Artificial Intelligence. AAAI’15, pp. 629–635
 140. Yan, K., Zhong, C., & Huang, J. (2018). Semi-supervised learning for early detection and diagnosis of various air handling unit faults. *Energy and Buildings*. <https://doi.org/10.1016/j.enbuild.2018.10.016>
 141. Lu, X., Zhou, Y., Wang, Z., Yi, Y., Feng, L., & Wang, F. (2019). Knowledge embedded semi-supervised deep learning for detecting non-technical losses in the smart grid. *Energies*, 12, 3452. <https://doi.org/10.3390/en12183452>
 142. Yang, Y., Zhong, J., Li, W., Gulliver, T. A., & Li, S. (2020). Semisupervised multilabel deep learning based nonintrusive load monitoring in smart grids. *IEEE Transactions on Industrial Informatics*, 16(11), 6892–6902. <https://doi.org/10.1109/TII.2019.2955470>
 143. Lu, T. (2009). *Fundamental Limitations of Semi-supervised Learning*. Waterloo: University of Waterloo.
 144. van Engelen, J. E., & Holger, H. H. (2020). A survey on semi-supervised learning. *Machine Learning*, 109, 373–440.
 145. Shrivastava, I. (2020). Handling class imbalance by introducing sample weighting in the loss function. <https://medium.com/gungum-tech/handling-class-imbalance-by-introducing-sample-weighting-in-the-loss-function-3bd>. Accessed Mar 2020
 146. Hanlon, J. (2017). Why is so much memory needed for deep neural networks? <https://www.graphcore.ai/posts/why-is-so-much-memory-needed-for-deep-neural-networks>. Accessed Mar 2020
 147. Miao, X., Liu, Y., Zhao, H., & Li, C. (2019). Distributed online one-class support vector machine for anomaly detection over networks. *IEEE Transactions on Cybernetics*, 49(4), 1475–1488. <https://doi.org/10.1109/TCYB.2018.2804940>
 148. Fu, S., Liu, J., Pannu, H. (2012) A hybrid anomaly detection framework in cloud computing using one-class and two-class support vector machines, 7713, 726–738. <https://doi.org/10.1007/978-3-642-35527-160>
 149. Nguyen, X.N., Nguyen, D.T., Vu, L.H. (2016). Pocad: A novel pay load-based one-class classifier for anomaly detection. In: 2016 3rd National Foundation for Science and Technology Development Conference on Information and Computer Science (NICS), pp. 74–79. <https://doi.org/10.1109/NICS.2016.7725671>
 150. Karunaratne, G., Le Gallo, M., Cherubini, G., Benini, L., Rahimi, A., & Sebastian, A. (2020). In-memory hyperdimensional computing. *Nature Electronics*, 3(6), 327–337. <https://doi.org/10.1038/s41928-020-0410-3>
 151. Sun, Y., Wong, A. K. C., & Kamel, M. S. (2009). Classification of imbalanced data: A review. *International Journal of Pattern Recognition and Artificial Intelligence*, 23(04), 687–719. <https://doi.org/10.1142/S0218001409007326>
 152. Branco, P., Torgo, L., Ribeiro, R. (2015). A survey of predictive modelling under imbalanced distributions. <https://arxiv.org/abs/1505.01658>
 153. Cui, Y., Jia, M., Lin, T.-Y., Song, Y., Belongie, S. (2019). Class-balanced loss based on effective number of samples. In: IEEE/CVF Conference on Computer Vision and Pattern Recognition
 154. Huan, W., Lin, H., Li, H., Zhou, Y., Wang, Y. (2020). Anomaly detection method based on clustering undersampling and ensemble learning. In: 2020 IEEE 5th Information Technology and Mechatronics Engineering Conference (ITOEC), pp. 980–984 <https://doi.org/10.1109/ITOEC49072.2020.9141897>
 155. Wang, X., Flores, R., Brouwer, J., & Papaefthymiou, M. (2022). Real-time detection of electrical load anomalies through hyperdimensional computing. *Energy*, 261, 125042.

156. Chicco, D., & Jurman, G. (2020). The advantages of the Matthews correlation coefficient (MCC) over F1 score and accuracy in binary classification evaluation. *BMC Genomics*, 21(1), 1–13.
157. Karadayı, Y., Aydin, M. N., & Oğrenci, A. S. (2020). A hybrid deep learning framework for unsupervised anomaly detection in multivariate spatio-temporal data. *Applied Sciences*, 10(15), 5191.
158. Zhang, C., Song, D., Chen, Y., Feng, X., Lumezanu, C., Cheng, W., Ni, J., Zong, B., Chen, H., Chawla, N.V. (2019). A deep neural network for unsupervised anomaly detection and diagnosis in multivariate time series data. In: Proceedings of the AAAI Conference on Artificial Intelligence, vol. 33, pp. 1409–1416
159. Palensky, P., & Dietrich, D. (2011). Demand side management: Demand response, intelligent energy systems, and smart loads. *IEEE Transactions on Industrial Informatics*, 7(3), 381–388. <https://doi.org/10.1109/TII.2011.2158841>
160. Short term operating reserve: General description of the service. National Grid, London, UK (2017)
161. Lambert, Q. (2012). *Business Models for an Aggregator: Is an Aggregator Economically Sustainable on Gotland*. Stockholm, Sweden: Royal Institute of Technology (KTH). <https://www.diva-portal.org/smash/get/diva2:537356/FULLTEXT01.pdf>
162. Vazquez-Canteli, J. R., & Nagy, Z. (2019). Reinforcement learning for demand response: A review of algorithms and modeling techniques. *Applied Energy*, 235, 1072–1089. <https://doi.org/10.1016/j.apenergy.2018.11.002>
163. Kennel, F., Gorges, D., & Liu, S. (2013). Energy management for smart grids with electric vehicles based on hierarchical mpc. *IEEE Transactions on Industrial Informatics*, 9(3), 1528–1537. <https://doi.org/10.1109/TII.2012.2228876>
164. Siano, P. (2014). Demand response and smart grids—a survey. *Renewable and Sustainable Energy Reviews*, 30, 461–478. <https://doi.org/10.1016/j.rser.2013.10.022>
165. Lee, K., et al. (2010). Us department of energy office of electricity delivery and energy reliability. NSTB ICCP Security Assessment
166. Herter, K. (2007). Residential implementation of critical-peak pricing of electricity. *Energy Policy*, 35, 2121–2130. <https://doi.org/10.1016/j.enpol.2006.06.019>
167. Centolella, P. (2010). The integration of price responsive demand into regional transmission organization (rto) wholesale power markets and system operations. *Energy*, 35(4), 1568–1574.
168. Wu, Y., Wu, Y., Guerrero, J. M., & Vasquez, J. C. (2022). Decentralized trans-active energy community in edge grid with positive buildings and interactive electric vehicles. *International Journal of Electrical Power & Energy Systems*, 135, 107510. <https://doi.org/10.1016/j.ijepes.2021.107510>
169. Sharma, A. K., & Saxena, A. (2019). A demand side management control strategy using whale optimization algorithm. *SN Applied Sciences*, 1(8), 870. <https://doi.org/10.1007/s42452-019-0899-0>
170. Tutkun, N., Ungören, F., Alpagut, B. (2017). Improved load shifting and valley filling strategies in demand side management in a nano scale off-grid wind-pv system in remote areas. In: 2017 IEEE 14th International Conference on Networking, Sensing and Control (ICNSC), pp. 13–18. <https://doi.org/10.1109/ICNSC.2017.8000060>
171. Werminski, S., Jarnut, M., Benysek, G., & Bojarski, J. (2017). Demand side management using dadr automation in the peak load reduction. *Renewable and Sustainable Energy Reviews*, 67, 998–1007. <https://doi.org/10.1016/j.rser.2016.09.049>
172. Yang, Q., Wang, H., Wang, T., Zhang, S., Wu, X., & Wang, H. (2021). Blockchain-based decentralized energy management platform for residential distributed energy resources in a virtual power plant. *Applied Energy*, 294, 117026. <https://doi.org/10.1016/j.apenergy.2021.117026>
173. Jo, J., & Park, J. (2020). Demand-side management with shared energy storage system in smart grid. *IEEE Transactions on Smart Grid*, 11(5), 4466–4476. <https://doi.org/10.1109/TSG.2020.2980318>
174. Zhao, D., Wang, H., Huang, J., & Lin, X. (2020). Storage or no storage: Duopoly competition between renewable energy suppliers in a local energy market. *IEEE Journal on Selected Areas in Communications*, 38(1), 31–47. <https://doi.org/10.1109/JSAC.2019.2951970>
175. Wang, H., Henri, G., Tan, C.-W., Rajagopal, R. (2020). Activity detection and modeling using smart meter data: Concept and case studies. In: 2020 IEEE Power Energy Society General Meeting (PESGM), pp. 1–5. <https://doi.org/10.1109/PESGM41954.2020.9281746>
176. Kwac, J., Kim, J. I., & Rajagopal, R. (2019). Efficient customer selection process for various objectives. *IEEE Transactions on Smart Grid*, 10(2), 1501–1508. <https://doi.org/10.1109/TSG.2017.2768520>
177. Tang, W., Wang, H., Lee, X.-L., & Yang, H.-T. (2021). Machine learning approach to uncovering residential energy consumption patterns based on socio-economic and smart meter data. *Energy*. <https://doi.org/10.1016/j.energy.2021.122500>
178. Wei, Z., Wang, H. (2021). Characterizing residential load patterns by household demographic and socioeconomic factors. In: Proceedings of the Twelfth ACM International Conference on Future Energy Systems. e-Energy 21, pp. 244–248. Association for Computing Machinery, New York, NY, USA. <https://doi.org/10.1145/3447555.3464867>
179. Wang, Z., Wang, H. (2021). Identifying the relationship between seasonal variation in residential load and socioeconomic characteristics. In: Proceedings of the 8th ACM International Conference on Systems for Energy-Efficient Buildings, Cities, and Transportation. BuildSys 21, pp. 160–163. Association for Computing Machinery, New York, NY, USA. <https://doi.org/10.1145/3486611.3486645>
180. Babar, M., Tariq, M. U., & Jan, M. A. (2020). Secure and resilient demand side management engine using machine learning for iot-enabled smart grid. *Sustainable Cities and Society*, 62, 102370.
181. Zhou, Y., & Zheng, S. (2020). Machine-learning based hybrid demand-side controller for high-rise office buildings with high energy flexibilities. *Applied Energy*, 262, 114416.
182. Qiu, X., Nguyen, T. A., & Crow, M. L. (2016). Heterogeneous energy storage optimization for microgrids. *IEEE Transactions on Smart Grid*, 7(3), 1453–1461. <https://doi.org/10.1109/TSG.2015.2461134>
183. Xiong, R., Cao, J., & Yu, Q. (2018). Reinforcement learning-based real-time power management for hybrid energy storage system in the plug-in hybrid electric vehicle. *Applied Energy*, 211, 538–548. <https://doi.org/10.1016/j.apenergy.2017.11.072>
184. Kofinas, P., Dounis, A. I., & Vouros, G. A. (2018). Fuzzy q-learning for multi-agent decentralized energy management in microgrids. *Applied Energy*, 219, 53–67. <https://doi.org/10.1016/j.apenergy.2018.03.017>
185. Mathew, A., Roy, A., & Mathew, J. (2020). Intelligent residential energy management system using deep reinforcement learning. *IEEE Systems Journal*, 14(4), 5362–5372. <https://doi.org/10.1109/JSYST.2020.2996547>
186. Song, L., Xiao, Y., & van der Schaar, M. (2014). Demand side management in smart grids using a repeated game framework.

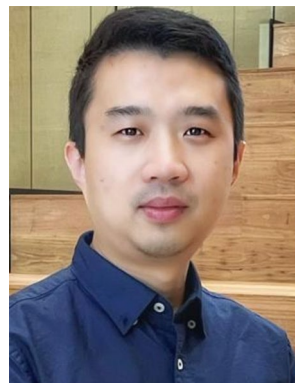
- IEEE Journal on Selected Areas in Communications*, 32(7), 1412–1424. <https://doi.org/10.1109/JSAC.2014.2332119>
187. Ma, J., Deng, J., Song, L., & Han, Z. (2014). Incentive mechanism for demand side management in smart grid using auction. *IEEE Transactions on Smart Grid*, 5(3), 1379–1388. <https://doi.org/10.1109/TSG.2014.2302915>
 188. Rasheed, M., Javaid, N., Ahmad, A., Awais, M., Khan, Z., Qasim, U., & Alrajeh, N. (2016). Priority and delay constrained demand side management in real-time price environment with renewable energy source. *International Journal of Energy Research*. <https://doi.org/10.1002/er.3588>
 189. Campillo, J., Dahlquist, E., Wallin, F., & Vassileva, I. (2016). Is real-time electricity pricing suitable for residential users without demand-side management? *Energy*, 109, 310–325. <https://doi.org/10.1016/j.energy.2016.04.105>
 190. O'Neill, D., Levorato, M., Goldsmith, A., Mitra, U. (2010). Residential demand response using reinforcement learning. In: 2010 First IEEE International Conference on Smart Grid Communications, pp. 409–414. <https://doi.org/10.1109/SMARTGRID.2010.5622078>
 191. Zhou, S., Hu, Z., Gu, W., Jiang, M., & Zhang, X.-P. (2019). Artificial intelligence based smart energy community management: A reinforcement learning approach. *CSEE Journal of Power and Energy Systems*, 5(1), 1–10. <https://doi.org/10.17775/CSEEJES.2018.00840>
 192. Lu, R., & Hong, S. H. (2019). Incentive-based demand response for smart grid with reinforcement learning and deep neural network. *Applied Energy*, 236, 937–949.
 193. Shakouri, H., & Kazemi, A. (2017). Multi-objective cost-load optimization for demand side management of a residential area in smart grids. *Sustainable cities and society*, 32, 171–180.
 194. Zhao, D., Wang, H., Huang, J., & Lin, X. (2021). Time-of-use pricing for energy storage investment. *IEEE Transactions on Smart Grid*. <https://doi.org/10.1109/TSG.2021.3136650>
 195. Li, P., Wang, H., & Zhang, B. (2019). A distributed online pricing strategy for demand response programs. *IEEE Transactions on Smart Grid*, 10(1), 350–360. <https://doi.org/10.1109/TSG.2017.2739021>
 196. Tronchin, L., Manfren, M., & Nastasi, B. (2018). Energy efficiency, demand side management and energy storage technologies – a critical analysis of possible paths of integration in the built environment. *Renewable and Sustainable Energy Reviews*, 95, 341–353. <https://doi.org/10.1016/j.rser.2018.06.060>
 197. Javaid, N., Hafeez, G., Iqbal, S., Alrajeh, N., Alabed, M. S., & Guizani, M. (2018). Energy efficient integration of renewable energy sources in the smart grid for demand side management. *IEEE Access*, 6, 77077–77096.
 198. Rocha, H. R. O., Honorato, I. H., Fiorotti, R., Celeste, W. C., Silvestre, L. J., & Silva, J. A. L. (2021). An artificial intelligence based scheduling algorithm for demand-side energy management in smart homes. *Applied Energy*, 282, 116145. <https://doi.org/10.1016/j.apenergy.2020.116145>
 199. Cai, H., Ziras, C., You, S., Li, R., Honore, K., & Bindner, H. W. (2018). Demand side management in urban district heating networks. *Applied Energy*, 230, 506–518. <https://doi.org/10.1016/j.apenergy.2018.08.105>
 200. Groppi, D., Pfeifer, A., Garcia, D. A., Krajacic, G., & Duic, N. (2021). A review on energy storage and demand side management solutions in smart energy islands. *Renewable and Sustainable Energy Reviews*, 135, 110183. <https://doi.org/10.1016/j.rser.2020.110183>
 201. World Population Prospects – Population Division (2015)
 202. Steen, D., Le, T., Bertling, L. (2012). Price-based demand-side management for reducing peak demand in electrical distribution systems—with examples from gothenburg. In: NORDAC 2012
 203. Ng, K.-H., & Sheble, G. B. (1998). Direct load control—a profit-based load management using linear programming. *IEEE Transactions on Power Systems*, 13(2), 688–694. <https://doi.org/10.1109/59.667401>
 204. Giordano, V., Meletiou, A., Covrig, C.F., Mengolini, A., Ardelean, M., Fulli, G. (2013). Smart Grid projects in Europe: Lessons learned and current developments 2012 update. Publications Office of the European Union. <https://ses.jrc.ec.europa.eu/publications/reports/smart-grid-projects-europe-lessons-learned-and-current-developments-2012-update>. Accessed Mar 2020
 205. Shahidepour, M., Yamin, H., & Li, Z. (2002). *Market Operations in Electric Power Systems: Forecasting, Scheduling, and Risk Management*. Wiley.
 206. Strbac, G. (2008). Demand side management: Benefits and challenges. *Energy Policy*, 36(12), 4419–4426. <https://doi.org/10.1016/j.enpol.2008.09.030>

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Xinlin Wang earned his Ph.D. in 2019 from Seoul National University in South Korea. Following that, he served as a Postdoctoral Fellow at the University of California, Irvine, in the United States. He is currently a Senior Postdoctoral Scholar at CSIRO, Australia's national research agency, located in Newcastle, Australia. His research focuses on machine learning applications in smart buildings and smart grids, as well as power systems optimization, demand-side management, and demand response.

He received the 2020 Outstanding Ph.D. Dissertation Award from Seoul National University.



Hao Wang earned his Ph.D. from The Chinese University of Hong Kong in 2016. He was a Postdoctoral Research Fellow at Stanford University and a Washington Research Foundation Innovation Fellow at the University of Washington. He is currently a Senior Lecturer in the Department of Data Science and AI, Faculty of Information Technology, Monash University. He is a recipient of the Australian Research Council (ARC) Discovery Early Career Researcher Award (DECRA). His research

interests are in optimization, machine learning, and data analytics for power and energy systems.



Binayak Bhandari is currently a Research Associate at University of New South Wales in Sydney, Australia. Prior to relocating in Australia, he earned his Ph.D. in 2014 from Seoul National University in South Korea and served as Assistant Professor and department chair at Woosong University in Daejeon, South Korea for seven years. His contributions span a wide array of fields, including renewable energy, advanced machining, artificial intelligence, composite materials, structural health monitoring,

robotics, railroad engineering, mechanical systems, and appropriate technology. He is a member of the Korean Society for Precision Engineering, the Korean Society of Railway, and Engineers Australia.



Leming Cheng is currently pursuing his Ph.D. at the University of California, Irvine, in the United States. His research interests include embedded systems, FPGA accelerators, autonomous driving, and data-driven approaches-based decision-making and dynamic control.