Net Load Forecasting for Renewable Energy Integrated Power Systems: A Critical Review

Nusraat Nawreen, *Student Member*, *IEEE*, Nima Amjady, *Fellow*, *IEEE*, Rakibuzzaman Shah, *Senior Member*, *IEEE*, and S. M. Muyeen, *Fellow*, *IEEE*

***Abstract*—Net load (NL) refers to the difference between gross demand and variable renewable energy (RE) generation. As the integration of RE continues to grow, the variability and uncertainty of power systems increase, creating significant challenges for power system operators due to the intermittent behavior of RE sources. NL forecasting (NLF) has become crucial to maintaining the reliable and secure operation of power systems. This paper critically reviews the state-of-the-art NLF models in RE-integrated power systems, with a particular focus on emerging techniques such as Transformer models and generative artificial intelligence (Gen-AI) models. This paper systematically explores and analyses various data preprocessing, feature engineering (feature selection and extraction), and hyperparameter tuning methods as well as forecasting engines used in NLF. Furthermore, this paper categorizes the NLF models with corresponding strengths and challenges. In addition, an assessment of the interpretability, scalability, and applicability of NLF models is presented, highlighting their performance under varying levels of RE penetration. This paper also outlines challenges for NLF, e. g., the quality of data and the spatial and temporal scales. Finally, the future research directions are outlined based on the identified gaps in NLF.**

***Index Terms*—Net load forecasting (NLF), data preprocessing, artificial intelligence, feature engineering, feature selection, feature extraction, hyperparameter tuning, forecasting engine, renewable energy.**

I. Introduction

MITIGATING the severe consequences of climate change and limiting the global temperature increase to 1.5 °C by the end of this century has become the focal point. Fossil fuel-based generators are major contributors to greenhouse gas emissions [1]. According to the International Energy Agency (IEA), it is anticipated that the amount of energy generated by renewable energy sources (RESs) will be increased to approximately 17000 TWh by the end of 2030 [2]. While RESs have significantly contributed to the clean energy transition, their increasing penetration into power systems has introduced challenges related to power system security, stability, and quality due to their intermittent nature.

RESs depend heavily on unpredictable weather and cloud conditions, causing rapid fluctuations in their generation that complicate net load forecasting (NLF). Accurate NLF is crucial for operation, demand-side management, and stability of power system [3]. Power system operators often struggle with system operation during midday due to high photovoltaic (PV) generation [4]. Additionally, high levels of renewable energy (RE) penetration can exacerbate these issues, with weather events causing significant demand or generation fluctuations [5]. To ensure reliable and secure operation and prevent power shortages, power system operators need to forecast the power generation and demand. In this regard, various forecasting engines like statistical, machine learning (ML), and hybrid models have been proposed [6]-[8].

Various ML models have been explored to achieve accurate NLF results. Artificial neural networks (ANNs), extreme gradient boosting (XGBoost), *k*-nearest neighbors (KNN), random forest (RF), and recurrent neural networks (RNNs) including long short-term memory (LSTM) and gated recurrent unit (GRU) have demonstrated their effectiveness in providing reliable forecasts and capturing temporal dependencies [6]-[8]. Additionally, Bayesian neural networks (BNNs) add uncertainties to their predictions to make NLF more reliable [9].

To achieve better performance, researchers often use several data preprocessing, feature engineering (feature selection and extraction), and hyperparameter tuning methods as well as forecasting engines. Performance comparison between direct and indirect NLF models in RES-dominated microgrids and power systems has also been reported in [10] and [11]. In [12], both integrated and additive NLF models are examined, with a comparison given across three techniques: configurable forecasting tools, artificial intelligence (AI) approaches, and traditional statistical methods. Although using these techniques for predictive analysis is not new, NLF is far less represented in the literature than load forecasting (LF). Three NLF models under different aggregation levels, i.e., aggregated, partially aggregated, and disaggregated, have been comprehensively analyzed in [13]. Due to the increasing penetration of uncertain RESs, the focus is shifting from point forecasting towards probabilistic forecasting [14], [15].

Nowadays, Transformer models are gaining popularity in NLF due to their ability to capture long-term dependencies

Manuscript received: September 30, 2025; revised: January 13, 2026; accepted: March 16, 2026. Date of CrossCheck: March 16, 2026. Date of online publication: May 15, 2026.

N. Nawreen, N. Amjady (corresponding author), and R. Shah are with the Centre for New Energy Transition Research (CfNETR), Federation University Australia, Mt Helen, Australia (e-mail: nnawreen@students.federation.edu.au; n.amjady@federation.edu.au; m.shah@federation.edu.au).

S. M. Muyeen is with the Department of Electrical Engineering, Qatar University, Doha 2713, Qatar (e-mail: sm.muyeen@qu.edu.qa).

DOI: 10.35833/MPCE.2025.000940

[16]. Extensive research has been conducted in [14], [17], and [18], where various Transformer models have been developed and applied. Some of these models have shown exceptional capability in analyzing multivariate correlations and temporal dependencies for NLF. Nevertheless, research on their performance under various uncertainty factors in the context of NLF continues to be a subject of ongoing investigation.

Although many studies have explored different aspects of NLF, there is a lack of comprehensive review summarizing various NLF models. This paper addresses the research gaps by examining data preprocessing, feature engineering, and hyperparameter tuning methods as well as forecasting engines proposed for NLF. Additionally, this paper suggests future research directions to enhance NLF performance. The main contributions of this paper are summarized as follows.

1) An in-depth bibliometric study of the current literature on NLF is conducted, providing insights into research trends, major contributions, and conceptual developments.

2) A comprehensive analysis of data preprocessing, feature engineering, and hyperparameter tuning methods as well as forecasting engines is addressed, emphasizing the role of each in improving the accuracy of NLF.

3) Various sources of uncertainty are outlined, offering a clear overview of probabilistic forecasting models and generative AI (Gen-AI) models.

4) Challenges related to the accuracy and reliability of NLF in power systems with high penetration of uncertain RE are examined, and studies that focus on the interpretability, scalability, and applicability of NLF models are assessed.

The remainder of this paper is organized as follows. Section II discusses different fundamental aspects of NLF. Section III highlights and discusses data preprocessing methods for NLF. Section IV illustrates feature engineering and hyperparameter tuning methods. In Section V, various forecasting engines for NLF are thoroughly examined. Section VI reviews the interpretability, scalability, and applicability of NLF models. Section VII explicitly discusses challenges arising in NLF and provides necessary future research directions. Finally, the conclusions are drawn in Section VIII. The structure of this paper is illustrated in Supplementary Material A Fig. SA1, which highlights the main sections and their connections and presents a clear overview of paper organization.

II. Fundamental Aspects of NLF

This section provides a comprehensive introduction to the topic, i.e., the key characteristics of NLF and the factors influencing NLF. Furthermore, the distinctions between traditional LF and NLF are evaluated and discussed.

A. Concept and Characterization of Net Load (NL)

The NL in a power system is commonly defined as the difference between the gross demand and the variable RE generation [19]. The NL curve effectively represents fluctuations in the power system and highlights the flexibility required to accommodate the combined effects of changing demand and RE generation [19]. The variability of NL, coupled with the growing demand for electric vehicles (EVs), further intensifies flexibility challenges [20]. Moreover, the significant and abrupt fluctuations in solar PV and wind generation, referred to as ramping events, present critical concerns for power system operators [21].

An NL ramping event (NRE) refers to any variation in the total NL that exceeds 5% of the maximum NL value [22]. Such events pose challenges for power system operators. Notably, downward ramping events are more challenging than upward ones, since the sudden reduction in RE generation requires compensation from conventional generators to maintain system reliability. In contrast, the increase in RE generation can be managed through generator rescheduling or by curtailing the excess supply [22].

In 2013, California Independent System Operator (CAISO) released a load profile known as the "duck chart", illustrating the potential for "overgeneration" as solar PV penetration increases [23]. It emphasizes the risk of solar PV generation surpassing demand, which, if not addressed, could result in overvoltage and congestion. In such scenarios, conventional power plants may struggle to rapidly ramp up or down, hindering the effective integration of surplus solar energy into power system and impairing a stable balance between supply and demand. To prevent this, power system operators must adjust output from conventional generators [23].

1) Differences Between LF and NLF

LF models have been developed to predict future electricity demand of customers for power system planning and operation such as unit commitment, network management, generation dispatch, fuel allocation, and various planning activities. LF enables secure and cost-effective supply by preventing over- or under-utilization of production and transmission capacities. Reference [24] and the references therein critically review different aspects of LF and outline different LF models.

NLF is more crucial than traditional LF due to its strong correlation with intermittent sources in RE-integrated power systems [17]. Consequently, the accurate NLF is important to handle the inherent uncertainties associated with RESs to ensure the smooth and reliable operation of power systems [13]. Figure 1 outlines the significant differences between LF and NLF [5], [17].

2) Different NLF Models

Figure 2 presents a classification for NLF models in terms of forecasting nature, aggregation level, forecasting engine, and forecasting routine.

Several studies have evaluated the performance of direct and indirect NLF models using different frameworks [10], [12], [25]. Supplementary Material A Table SAI gives a thorough comparison between direct and indirect NLF models, outlining their respective strengths, weaknesses, scalability, and suitability for different RE generation. Figure 3 illustrates the concept of direct and indirect NLF models, where FM stands for forecasting method. The key difference between the two NLF models lies in how the relationship between gross demand and RE generation is modeled.

The direct NLF model integrates gross demand and RE generation into a single FM, while the indirect NLF model separately forecasts gross demand and RE generation and then combines the forecasting results.

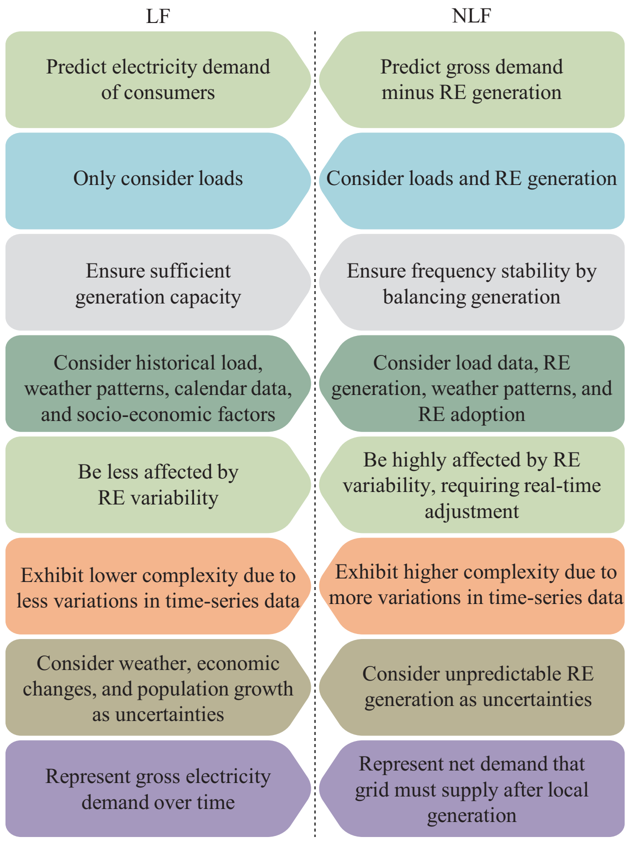

Fig. 1. Significant differences between LF and NLF.

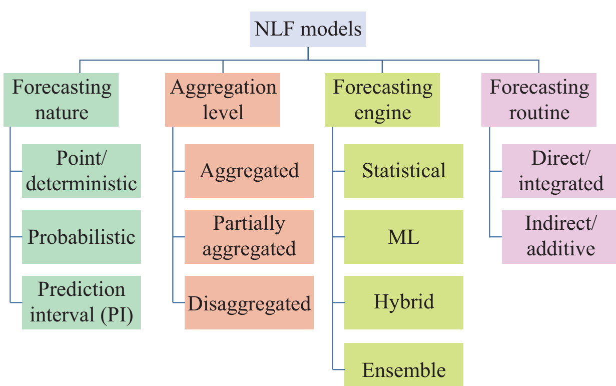

Fig. 2. Classification of NLF models.

B. Different Forecasting Horizons for NLF

Reference [24] discusses the forecasting horizons for LF, and the similar classification could be applied to NLF. NLF is divided into four categories based on the forecasting horizons: very short-term NLF (VSTNLF), short-term NLF (STNLF), medium-term NLF (MTNLF), and long-term NLF (LTNLF).

1) VSTNLF

The volatility of NL makes the NLF crucial for power system operation in shorter periods. VSTNLF predicts the NL from a few minutes to a few hours ahead [26].

It supports utilities in real-time generation dispatch, load frequency control, and demand response, while aiding retail, commercial, and industrial enterprises in business operations [27]. Reference [27] confirms that the ML-based VSTNLF models are effective in microgrid environments, especially focusing on the challenges introduced by high levels of solar PV penetration. The models proposed in [28]-[30] enhance the VSTNLF accuracy for very short time intervals (i.e., 5 min) by addressing uncertainties in RE generation. These models show strong potential due to their ability to handle minute-level fluctuations effectively.

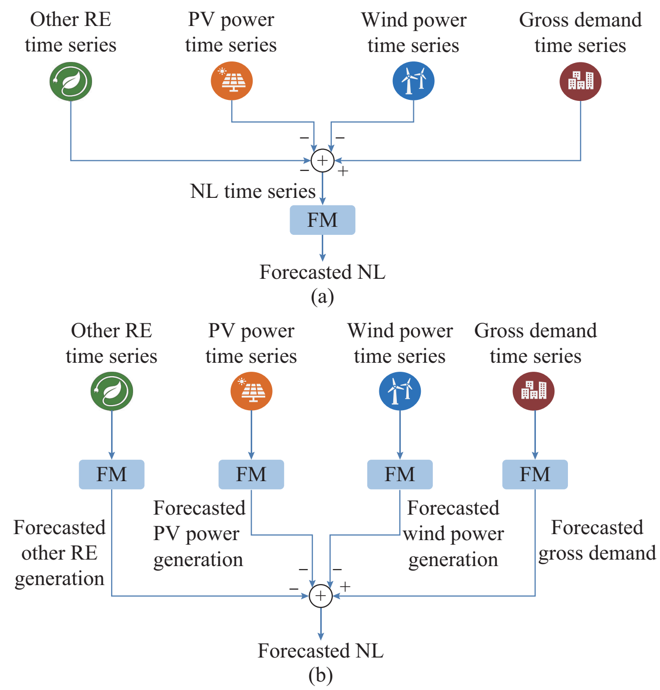

Fig. 3. Concept of direct and indirect NLF models. (a) Direct NLF model. (b) Indirect NLF model.

2) STNLF

STNLF predicts the NL from several hours to several days ahead [26]. Reference [31] incorporates 12-hour, 24-hour, 48-hour, and weekly periods for NLF and compares their performance.

3) MTNLF

MTNLF predicts the NL over a period extending up to several weeks or months [26]. Reference [32] highlights the importance of incorporating solar irradiance data and uses disaggregation techniques to improve the MTNLF accuracy. Reference [33] introduces a decoupling-based monthly NLF framework considering the solar PV generation.

4) LTNLF

LTNLF predicts the NL over a period from a year to several decades [26]. In long-term LF (LTLF), factors such as human habits, environmental conditions, demographics, population distributions, economic indicators, and electricity prices play significant roles [34]. These factors are also important for LTNLF. Reference [35] uses census data and information-theoretic feature selection to develop a 15-year NLF model for predicting the EV and solar PV penetration. A hybrid model using auto-regressive integrated moving average (ARIMA) and least-square support vector machine (LSSVM) is developed for LTNLF in [36].

Supplementary Material A Table SAII compares the VSTNLF, STNLF, MTNLF, and LTNLF in various aspects

[5]-[18], [20], [25]-[94], and provides an overview. Typically, ML models require enough training samples to avoid underfitting or overfitting. For this reason, they are commonly applied to NLF with shorter time horizons, where sufficient historical data are available [26]-[31]. In contrast, for LTNLF, where historical data may be limited, applying ML models can be challenging due to insufficient training samples. Statistical or hybrid models that incorporate statistical models are often preferred in such cases [35], [36]. Based on the available literature, this review mostly focuses on STNLF.

C. Influential Factors

Accurate NLF is inherently linked to the understanding and modeling of various factors that influence the demand and RE generation. These factors can be broadly categorized into meteorological, temporal, socio-economic, and distributed energy resource (DER)-related factors.

Meteorological factors such as solar irradiance, wind speed, temperature, and humidity influence demand and RE generation, thereby impacting the NLF. Multiple studies account for solar fluctuations [31], clear sky solar power [5], and ambient temperature to achieve accurate NLF [37].

Temporal factors such as seasonality, weekly patterns, and intraday variability significantly affect NLF. Day type [37], multi-scale temporal patterns [17], external weather factors [38], weather volatility classification [39], calendar effects [40], and weather-seasonal information [41] have been used to enhance the accuracy of NLF and model adaptability to sudden weather changes.

Socio-economic factors such as economic growth, population growth, urbanization, and electricity prices also significantly affect the long-term trends of demand and DER penetration. The level of DER penetration such as rooftop solar PV installation and battery storage is significantly influenced by electricity prices and government policies.

DER-related factors such as variability from solar PV generation and changes in load profiles from EV charging can cause sudden or disproportionate changes in RE adoption, further introducing non-linear dynamics and uncertainties that require advanced patterns [42]. Furthermore, the increasing penetration of DERs such as rooftop solar PV, EVs, and energy storage systems (ESSs) adds variability and uncertainty to NLF [20].

Researchers in [14] and [43] particularly emphasize how stochastic variations in demand and RE generation affect the accuracy and reliability of NLF. The trade-off among forecasting accuracy, uncertainty calibration, and computational complexity is studied in [43]. Reference [95] highlights the effect of variable RE generation on NREs emphasizing the need for integrated ramp forecasting.

The major influential factors of NLF [7], [96] are outlined in Fig. 4.

In summary, this section explores fundamental concepts of NLF, and distinguishes it from traditional LF. It discusses the significance of accurate NLF in the context of RE integration and highlights various NLF models under the volatility of RE generation. This section also discusses the influential factors affecting the accuracy of NLF and emphasizes the increasing importance of NLF in power systems with high levels of RE penetration.

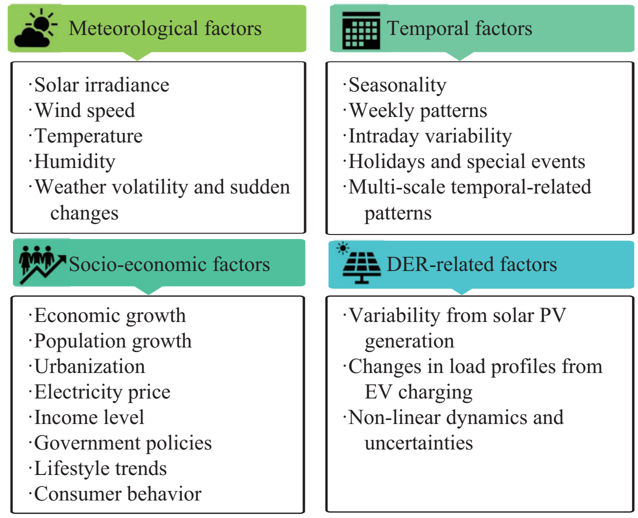

Fig. 4. Major influential factors of NLF.

III. Data Preprocessing Methods

Data preprocessing is a critical step as it prominently influences the performance and reliability of NLF model. It is required to reduce noise, address the missing or inconsistent data, and extract relevant features from input data of NLF. This section briefly explains the data preprocessing methods for NLF.

A. Data Integration

Since NLF involves both demand and RE generation, diverse data, e.g., load data, RE generation data, weather data, electricity market data, and others containing relevant data about NL, must be gathered, cleaned, and integrated to create a unified dataset for accurate forecasting. Supplementary Material A Table SAIII illustrates data sources, data types, and data collection methods for NLF. Different data sources have various time resolutions. It is essential to integrate these sources using a common timestamp through temporal synchronization to ensure consistency [97]. Data from diverse sources can be combined into a consistent repository like a data warehouse. Key considerations in data integration include schema integration, object matching, and redundancy identification, where correlation analysis can be applied to identify redundant attributes [98].

B. Data Consistency

Real-world data often contain missing or noisy values. Erroneous data detection and missing value imputations are often used for data cleaning purposes [98]. Noise and outliers in data significantly affect the accuracy and development of NLF model.

Noise arises from faulty measurements or human errors and it is generally categorized as measurement noise (i. e., feature noise) and label noise (i.e., output noise) [99]. The label noise can be even more harmful to predictive models as it directly affects the target labels used in training, leading to the learning of incorrect relationships between features and outputs [99]. The noisy labels cause the deep learning (DL) models to learn patterns that do not generalize well and cause overfitting, where the model memorizes the noise rather than the true data distribution [100]. Furthermore, the presence of noise distorts the statistical analysis, making it difficult to identify stable and reliable relationships between input features and target labels. As a result, the forecasting accuracy decreases and uncertainty increases, which undermines the decision-making process.

In time-series forecasting, noisy data accumulate over time with compounding errors that make multi-step-ahead predictions less reliable [101]. Thus, the effective data preprocessing and noise filtering techniques are critical for improving model accuracy and ensuring reliable predictions. Reference [98] identifies various sources of label noise, explores their potential impacts, and classifies the data preprocessing methods into three categories: noise-robust, noise-cleansing, and noise-tolerant. Reference [102] reviews various noise identification techniques including ensemble techniques, distance-based algorithms, and single learning-based approaches. It also discusses different noise handling methods (e.g., filtering out noise or addressing noise through data polishing, scrubbing, or relabeling) along with their associated trade-offs. Furthermore, sensor malfunctions, errors in data transmission, operational anomalies (e.g., equipment failure or disturbance in power systems), and external events (e.g., events due to weather behaviors) could cause outliers in the dataset [103]. The manual threshold has been applied in [17] and [18] to detect the outliers in the dataset. Density-based spatial clustering of applications with noise (DBSCAN) based on data density with effective noise isolation is used to identify the outliers in [44]. Automated detection routines in [45] enhance the efficiency by systematically identifying problematic data without manual intervention. Similarly, [46] employs statistical anomaly detection methods that not only identify outliers but also correct them, thereby improving data quality for forecasting. Each method exhibits different trade-offs among computational complexity, adaptability, and accuracy, emphasizing the need to select outlier detection methods based on dataset characteristics and forecasting objectives. Additionally, the auto-regressive model with exogenous input (ARX) is integrated into an online preprocessing-forecasting loop for real-time data integration and processing [79]. This allows each new batch of grid measurements to be promptly cleaned, imputed, and utilized to update the rolling NLF. The online data processing module identifies outliers based on voltage and current thresholds, removes duplicate timestamps, and imputes missing intervals using the ARX. For each missing interval, the available leading and lagging data are employed to implement the ARX.

Supplementary Material A Table SAIV presents various missing data imputation and noise handling methods used in NLF, where grey cells indicate missing data imputation methods and white cells indicate noise handling methods.

C. Data Reduction

Data reduction simplifies a large dataset by retaining essential information while eliminating redundancy, making it easier to process the data. With the advancement of sensing, metering, and communication facilities in power systems [104], high-frequency and high-dimensional data are quite common in NLF. Therefore, data reduction is necessary to improve the computational efficiency and tractability of NLF models. Moreover, it helps focus on the most relevant features, thus enhancing the forecasting accuracy and scalability. Data reduction tasks include instance selection, discretization, feature selection, and feature extraction [105].

Instance selection improves model efficiency by removing irrelevant, noisy, or redundant instances from the training set. It helps reduce the overfitting and speed up the training process [105]. Discretization simplifies continuous numerical data by converting them into categorical intervals to reduce complexity while preserving the data structure [105]. Feature selection and feature extraction are very important for NLF, which will be discussed in Section V.

D. Data Transformation

In NLF, data transformation is essential for handling diverse and inconsistent data from sources like meters and weather sensors. To improve the quality of dataset and keep consistency and compatibility with different algorithms, it is often required to change data structures, formats, or values. Data transformation tasks include smoothing, aggregation, generalization, and normalization [105]. The one-hot encoding (OHE) data transformation is used for NLF [32]. It transforms categorical data into binary vectors, making it suitable for numerical ML models. Other than this, various data transformation techniques have also been implemented, which are discussed below.

1) Data Standardization

The presence of variables at different scales is a critical aspect of data. For instance, temperature may be in the range of [0, 50]°C, while NL values may range from 0 to 10000 MW. Without proper data standardization, features with larger numerical ranges can disproportionately influence the model training and lead to biased and inaccurate forecasting results. Common data standardization methods are min-max scaling and Z-score normalization [106]. Other less common methods include decimal scaling, median normalization, sigmoid normalization, layer normalization, zero-mean normalization, fuzzy systems, and log transformations [43], [47], [48], [90].

Min-max scaling applies a linear transformation to the raw dataset, ensuring that all scaled data fall within a specific range such as [0, 1]. This method is commonly used in datasets intended for distance metric-based learning methods and can help accelerate the training of ANNs by enabling faster weight convergence. The application of min-max scaling is limited due to its sensitivity to outliers and the requirements of prior knowledge of the minimum and maximum values of variables [105]. In Z-score normalization, all the data have a mean of 0 and a standard deviation of 1, which makes it robust to outliers [105]. For each variable, the Z-score normalization can be expressed as [49], [50]:

$$z=\frac{x-\mu}{\sigma} \tag{1}$$

where z is the normalized value; x is the raw data value; and μ and σ are the mean and standard deviation of raw data, respectively.

2) Decomposition of Time Series

Decomposition techniques are essential for decomposing complex time series into simpler components to identify patterns, trends, and anomalies in NLF [48]. In LF, these techniques help reveal the underlying data structure. Empirical mode decomposition (EMD) is a representative technique in this field. Although effective, EMD can suffer from noise sensitivity and mode mixing [41]. This has motivated the development of its advanced variants such as complete ensemble empirical mode decomposition with adaptive noise (CEEMDAN) and improved CEEMDAN (ICEEMDAN) [41]. Furthermore, variational mode decomposition (VMD), reported in [44], offers a complementary alternative by decomposing signals into narrowband modes centered around distinct frequencies. Compared with EMD, VMD demonstrates enhanced robustness to noise and sampling errors. Seasonal-trend decomposition techniques separate cyclical and trend components, which enable time-series models to address inter-feature dependencies and seasonal variations explicitly [18]. Wavelet transformation (WT) is a popular method for NLF by representing signals in the time-frequency domain [51]. It decomposes a time-domain signal into wavelet components by shifting and scaling a mother wavelet [107]. Singular spectrum analysis (SSA) decomposes time series into trend, seasonal, and noise components via eigenvalue decomposition to simplify complex seasonal structures, thereby enabling more focused learning within models like LSTM and multi-layer perceptron (MLP) [48]. Selecting the right data transformation method depends on the dataset, noise levels, and forecasting model, motivating the development of hybrid approaches that combine decomposition techniques for better performance.

This section examines various data preprocessing methods applied to enhance the NLF performance. It highlights the significance of preparing data through methods that ensure consistency, relevancy, and quality. Several missing data imputation and noise handling methods are highlighted with their challenges and limitations to provide a clear direction for accurate and reliable NLF.

IV. Feature Engineering and Hyperparameter Tuning Methods

A. Feature Engineering Methods

Feature engineering methods aim to select and extract the high-impact features. Feature selection and extraction are vital to improve the accuracy and efficiency of NLF models [51]. Pearson correlation coefficient is a simple and interpretable feature selection criterion; however, it is limited to linear dependencies and is sensitive to outliers [107]. Mutual information (MI) [107] and maximal information coefficient (MIC) [50] can detect non-linear relationships, but with higher computational burden. Gaussian copula correlation [52] and neural networks such as autoencoders [53] and convolutional neural networks (CNNs) [54] effectively handle non-linear and mixed data but need more resources and show reduced interpretability. Relief and RReliefF show robustness to noise with reasonable computational cost, but their dependency on nearest neighbors reduces their scalability [107].

More advanced techniques like feature construction with multiplex encoders [16], multi-scale Transformer-based attention blocks [17], and feature graph attention modules [18] are used to improve the feature correlation and interpretability in NL data. Furthermore, deep belief networks (DBNs) are applied to extract hidden features from NLF time series [55]. Temporal convolutional network (TCN) combined with extreme feature interaction network and adaptive feature fusion network can capture both local and global features and mitigate feature redundancy [56]. Overall, the feature selection and extraction methods should align with data characteristics, model complexity, and resource availability to optimize NLF. The advantages and disadvantages of the feature selection and extraction methods are presented in Supplementary Material A Table SAV.

B. Hyperparameter Tuning Methods

The performance of ML models is heavily influenced by the selection of their hyperparameters [108]. Hyperparameters, which affect model architecture, regularization level, and learning rates, are settings of an ML model. Moreover, hyperparameters control the system performance when applied to new and unseen data [109]. However, the manual selection of hyperparameters through trial and error is a tedious, biased, and error-prone process. This can also hinder reproducibility and efficiency [108]. To achieve higher accuracy and prevent overfitting, automatic and effective hyperparameter tuning methods are needed.

1) Grid Search (GS)

GS is widely used in ML models for optimizing hyperparameters by exhaustively testing all possible hyperparameter combinations within a defined parameter space [110]. While GS is a straightforward method, it is computationally costly to evaluate every possible hyperparameter combination [14], [20], [38]. It becomes difficult for many researchers to implement this method especially when the dataset is large and there are a large number of hyperparameters.

2) Random Search (RS)

To mitigate the drawbacks of GS, RS is often adopted to reduce the computational burden. Unlike GS, RS samples are set randomly inside the searching space to find the optimal hyperparameter combinations [111], which could reduce the computational burden while saving time and resources. However, when sample sizes are small, RS may fail to find the optimal hyperparameter combinations. Despite this, RS is popular for its effectiveness and simplicity. For instance, [57] proposes a Transformer-based probabilistic residential NLF method. The automated hyperparameter search makes it possible to quickly evaluate a wide range of parameter combinations. It selects hyperparameter values independently without using information from previous trials.

3) Bayesian Optimization (BO)

BO uses the probabilistic models to optimize the hyperparameters, efficiently identifying the optimal settings by balancing exploration and exploitation [110]. Unlike GS and RS, BO usually demands fewer evaluations to find feasible hyperparameter combinations, increasing the computational efficiency [48]. It intelligently searches the parameter space using past evaluation results to predict and select the next set of hyperparameters. It is particularly effective for computationally costly hyperparameter evaluation and can achieve fine-tuned configurations with only a few trials.

Comparisons among these hyperparameter tuning methods are illustrated in Supplementary Material A Table SAVI.

4) Cross Validation (CV) Technique

The CV technique strategically assesses forecasting models by dividing datasets into training and validation sets. This technique helps address the inherent variability of datasets. Reference [58] uses a stacking ensemble prediction model with a k-fold ($k=5$) CV technique. In this process, the dataset is divided into 5 equal-sized subsets or folds. The model is trained on $k-1$ ($k-1=5-1=4$) folds and tested on the remaining one. However, the varying temporal behaviors of NLF datasets may make the k-fold CV technique challenging. An alternative is time-series CV technique, which includes forward chaining and rolling window CV techniques [112]. As NLF model involves sequential data, different time-series CV techniques are widely adopted. The 2-fold rolling window CV technique is popular due to its time-series characteristics [113]. In this technique, a fixed-length training set moves forward in time at each iteration, and the validation set consists of the next immediate data point, also called leave-one-out CV (LOOCV) technique [114].

Overall, the CV technique ensures reliable model evaluation by preventing overfitting and providing a realistic measure of forecasting accuracy [115]. CV techniques test forecasting models on an unseen validation dataset to ensure that they perform well beyond just the training dataset, ultimately avoiding misleading results [112].

In summary, this section reviews various methods for selecting and extracting informative features applied for NLF, highlighting their advantages and disadvantages. It also provides a comparative analysis of hyperparameter tuning methods.

V. Forecasting Engines

Various forecasting engines have been employed for NLF, which can be divided into statistical models, ML models (e.g., AI-based non-deep learning (non-DL) models, AI-based DL models, Gen-AI models), hybrid models, and ensemble models.

Statistical models are often used due to their simplicity and straightforward applications.

ML models offer improved forecasting accuracy by learning patterns from historical data. They are capable of handling non-linear relationships, which is a crucial capability for NLF. Recently, researchers have been focusing on implementing Gen-AI models, hybrid models, and ensemble models. Recently, the emerging technology of Gen-AI models is capable of simulating multiple possible scenarios. It provides advanced capabilities for managing uncertainties.

Hybrid models can combine the strengths of multiple models to improve the prediction performance, while ensemble models aggregate results from various individual models to reduce errors and increase the reliability of the forecasting engine.

In this section, the above-mentioned forecasting engines are reviewed, and their key features are discussed. Additionally, a comparison of different forecasting engines used in NLF is presented in Supplementary Material A Table SAVII, summarizing their performance on real-world datasets with varying levels of RE penetration.

A. Statistical Models

Classical statistical models are widely used in the NLF domain. Reference [91] utilizes the model of seasonal ARIMA with exogenous input (SARIMAX) to investigate NLF under different levels of solar PV and wind power penetration. Their findings indicate a slight reduction in the forecasting accuracy as the proportion of PV power increases, which is attributed to the higher uncertainty driven by local weather conditions and fluctuations in sunlight. Reference [91] also investigates the NLF under different RE portfolios comprising solar PV and wind power, and emphasizes the importance of considering the local energy mix and unique RE characteristics. It concludes that a thorough evaluation of local weather patterns and the balance between solar PV and wind power are critical for ensuring accurate NLF. This effect may vary across different regions. Furthermore, the regression-based models have also been extended to include multiple variables (such as weather conditions and user behaviors) for NLF, known as multiple linear regression (MLR) model. MLR model is a baseline model in [32], which demonstrates that MLR model is useful for understanding the relationship between solar PV generation and NL, but it becomes less effective as solar PV generation increases.

In [59], the kernel density estimation (KDE) has been implemented for probabilistic NLF due to its non-parametric nature, as it estimates the full distribution of future values directly and flexibly. Another probabilistic NLF method is presented in [60] based on the conditional extreme value framework utilizing generalized additive model (GAM) combined with quantile regression (QR) and generalized Pareto distribution (GPD). This method can capture the entire distribution of NL, particularly focusing on extreme tail quantiles to manage risks associated with low-probability events. Moreover, an adaptive QR-based probabilistic NLF method is applied in [61], which uses online gradient descent for the forecasting adaptability. QR adopts inputs from adaptive mean forecasting, which enables the uncertainty quantification around the mean, and is derived from a GAM combined with Kalman filtering. QR plays a crucial role in handling uncertainty by providing a non-parametric approach to model different quantiles of the NL distribution. By using QR, the models in [60] and [61] are able to better handle the uncertainty resulting from sources like weather variability and consumer behaviors.

Another approach using Markov chain Monte Carlo regression (MCMCR) is proposed in [29] for probabilistic NLF, where unobserved variables are sampled via data augmentation. These variables are drawn from their posterior distribution that enables MCMCR. This approach further enhances the forecasting reliability by integrating the C-vine copula-based joint probability distribution for forecasting errors of load, wind generation, and solar PV generation. This enables probabilistic forecasts with upper and lower boundaries to address uncertainties in the input variables. Traditional Monte Carlo sampling faces challenges in independent parameter sampling, which are addressed using sequential time series of random values known as Markov chain. Overall, the MCMCR can generate independent parameter samples, which are crucial for applications like regression. The independent parameter sampling is necessary to explore the full regression plane for uncertainty quantification [29]. In [25], Gaussian processes (GPs) are implemented for the probabilistic forecasting of residential electricity consumption, solar PV generation, and NL for a household.

Forecasting with PIs provides a range of potential outcomes to quantify the forecasting uncertainty. Reference [82] applies the dynamic Gaussian process (DGP) and QR to generate PIs for NLF. It is concluded that the data aggregation enhances the sharpness and reliability of forecasted PIs. The DGP models the uncertainty by using random variables that follow a Gaussian distribution, which is suited for managing uncertainty in probabilistic NLF due to its ability to model complex and non-linear relationships in time-series data. The covariance function in GPs captures prior knowledge about the data variability, while the posterior distribution provides a measure of uncertainty around predictions. By incorporating different sources of uncertainty, GPs can generate probabilistic forecasting with PIs. These PIs widen in response to high uncertainty, such as cloud cover affecting solar PV generation or fluctuating consumer demand.

A summary of statistical models for NLF along with their key capabilities [29], [32], [59], [60], [91] is given in Fig. 5, where SARIMA is short for seasonal ARIMA; and AR is short for auto-regressive.

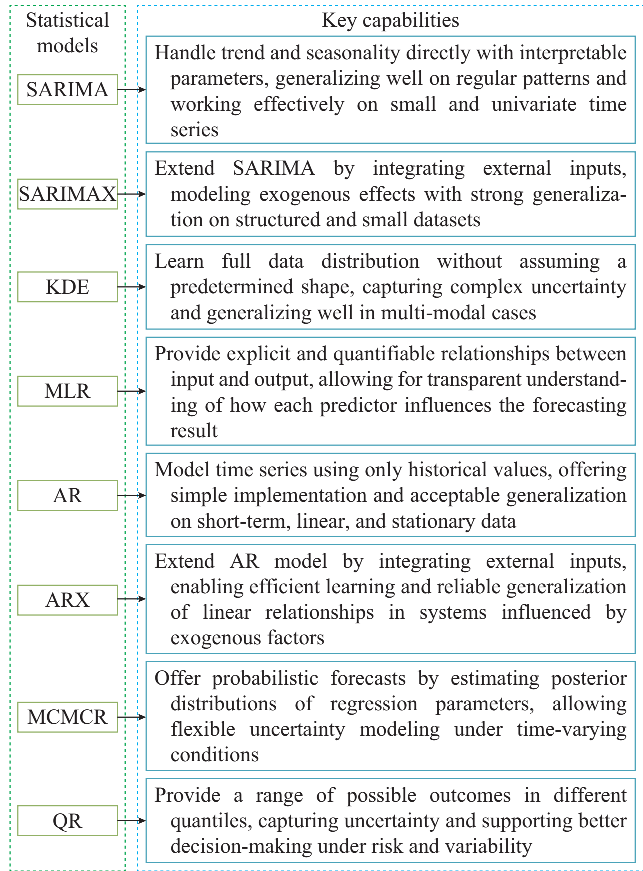

Fig. 5. Summary of statistical models for NLF along with their key capabilities.

B. ML Models

Although statistical models for NLF are easy to use and interpret, their performance is limited under complex and non-linear patterns. Specifically, with the increasing penetration of DERs, the statistical models may fail to effectively handle significant fluctuations and complexities that arise from the non-linear characteristics of variable RE generation. In contrast, ML models can learn from input-output mapping functions of historical data, which could enable better adaptation to the high variability of RE generation.

By learning different patterns and evaluating historical data more deeply, ML models provide higher forecasting flexibility and accuracy. However, these models come with a higher computational burden and may have lower scalability compared with statistical models for NLF. For instance, in [67], an ML model, named Bayesian deep-LSTM (BD-LSTM), achieves a root mean squared error (RMSE) of 17.1698 kW for point forecasting, outperforming QR, which achieves 41.0675 kW. For probabilistic forecasting, BD-LSTM attains pinball and Winkler scores of 4.8852 and 74.7658, respectively, while QR achieves 13.7448 and 189.6120. However, the computational time for BD-LSTM is 2495.13 s, whereas QR takes only 53.60 s.

ML models, e.g., AI-based non-DL models, AI-based DL models, and Gen-AI models, are increasingly used in NLF to overcome the limitations of statistical models.

1) AI-based Non-DL Models

Among AI-based non-DL models, XGBoost demonstrates superior performance for NLF in multiple studies. In [62], it outperforms categorical boosting (CatBoost) using sparrow search algorithm. In [63], it shows higher accuracy than MLP, RF, support vector regression (SVR), linear regression, and KNN for NLF. Its effectiveness in modeling complex data relationships for NLF is further validated in [83]. Reference [84] demonstrates that the light gradient boosting machine (LightGBM), compared with other benchmark models (e.g., RF and LSTM), achieves higher accuracy in NLF, particularly in scenarios with high behind-the-meter (BTM) solar PV generation. SVR and SVR with exogenous input (SVRX), where the exogenous input is solar PV power forecast, are utilized [5], [92], and their effectiveness in NLF is evaluated. All the above-mentioned models are used for point NLF.

For probabilistic NLF, some other AI-based non-DL models have been presented. In [15], a non-parametric Bayesian ML model, named sparse variational Gaussian process regression (SVGPR), is used for NLF. The main drawback of Gaussian process regression (GPR) is its training time, which has been addressed using SVGPR. Another AI-based non-DL model based on conformal MLP is presented in [49] for NLF with PIs. The conformal MLP combines the deterministic MLP with split conformal prediction to transform point forecasts into PIs without relying on strict assumptions about data distributions [49]. A summary of AI-based non-DL models for NLF along with their key capabilities [5], [15], [49], [62], [63], [83], [84] is given in Fig. 6.

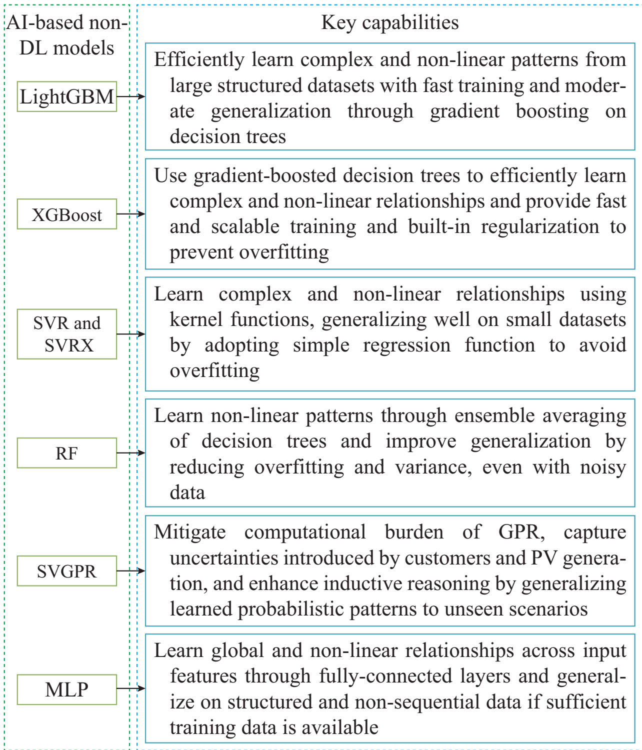

Fig. 6. Summary of AI-based non-DL models for NLF along with their key capabilities.

2) AI-based DL Models

AI-based DL models (e. g., LSTM and its different variants, GRU, CNN, and RNN) are used for point and probabilistic NLF. Among these models, LSTM-based models have been frequently used as they can capture long-term dependencies of time-series data [7], [20], [46], [89]. Bi-directional LSTM (Bi-LSTM) used in [50], [64], and [85] extends LSTM by processing data in both directions. Reference [44] uses a dual-layer LSTM to further enhance the forecasting accuracy by stacking two layers. This modification makes it suitable for capturing complex data. However, it may overfit for simple forecasting tasks or small datasets. Multi-input LSTM in [90] integrates multiple input factors, which improve the forecasting accuracy for data influenced by diverse variables. Additionally, [65] employs osprey optimization algorithm to fine-tune the hyperparameters of LSTM, as its forecasting accuracy heavily depends on the optimal hyperparameter selection. GRU, as an alternative with simpler gating mechanism compared with LSTM, is used for NLF in [8]. Moreover, a sequence-to-sequence RNN with GRU is introduced, which is able to provide explainability for the STNLF, with the capability of handling long-term predictions [116]. This architecture can handle variable-length sequences and effectively use long-term properties from the input data. In [93], an unsupervised NL disaggregation method is used to estimate BTM solar PV generation. Then, deep-auto-regressive (Deep-AR) RNN and time-series dense encoder (TiDE) models are employed for NLF. All the above-mentioned models are used for point NLF.

For probabilistic NLF, BNN-based models are preferred. References [26] and [66] apply the BNN for probabilistic NLF, which uses probabilistic layers to handle uncertainty. Another day-ahead probabilistic NLF is proposed in [67] to capture both epistemic uncertainty and aleatoric uncertainty using BD-LSTM. It integrates Bayesian probability theory with DL models to address aleatoric uncertainty (also known as stochastic uncertainty), which arises from factors such as climate variability, intermittent power generation, and consumer behavior. Additionally, it handles epistemic uncertainty (also known as model uncertainty), which is associated with uncertainty in model parameters and structures. The Bayesian network is a probabilistic model that represents the relationships between variables using conditional probabilities. In this model, nodes denote variables and edges indicate the probabilistic dependencies between variables. The Bayesian network generates outcomes by sampling from probability distributions, enabling it to manage uncertainty. Furthermore, it employs Bayesian inference to update its beliefs as new evidence is integrated [67].

In [117], a visual analytics-based tool is designed for comparing the performance of probabilistic NLF models. This tool helps evaluate DL models by visually analyzing their performance. However, it is sensitive to resolution, seasonality, or input fluctuations. Reference [68] develops a fully parameterized sequence to quantile (FPSe2Q) regression model using deep neural network (DNN) for probabilistic NLF to predict the entire conditional distribution of NL.

Two comparative studies [6] and [38] assess various AI-based non-DL and DL models for STNLF in microgrids with a high level of RE penetration. In [6], the BNN achieves the best performance with a normalized root mean square error (nRMSE) of 4.26%, while the RF in [38] excels with an nRMSE of 4.32%. Both studies highlight the effectiveness of ML models for reliable NLF in RE-integrated microgrids.

A summary of AI-based DL models for NLF along with their key capabilities [6]-[8], [20], [26], [38], [44], [46], [65]-[68], [89], [90], [93], [116] is given in Fig. 7, where QRNN is short for quantile regression neural network. Typically, forecasting steps of 15-min, 30-min, or 1-hour are widely used with a 24-hour forecasting horizon, with a few exceptions where the forecasting horizon is extended to weeks or even months [20].

3) Gen-AI Models

Conventional Transformer models encode the feature variables at a given timestep into a single token, which restricts the model ability to fully exploit temporal information [16]. To overcome this, several novel Gen-AI models are proposed for deterministic NLF, which are typically trained on extensive datasets to extend original content by generating synthetic time series, text, image, or audio by learning underlying patterns [118].

In [16], the iTransformer model effectively handles robust temporal dependencies, while multi-task learning enables the extraction of correlated features in NLF. The effectiveness of iTransformer model is evaluated based on the case study using four datasets from different regions [16], showing a significant reduction in mean absolute error (MAE) values compared with other baseline models.

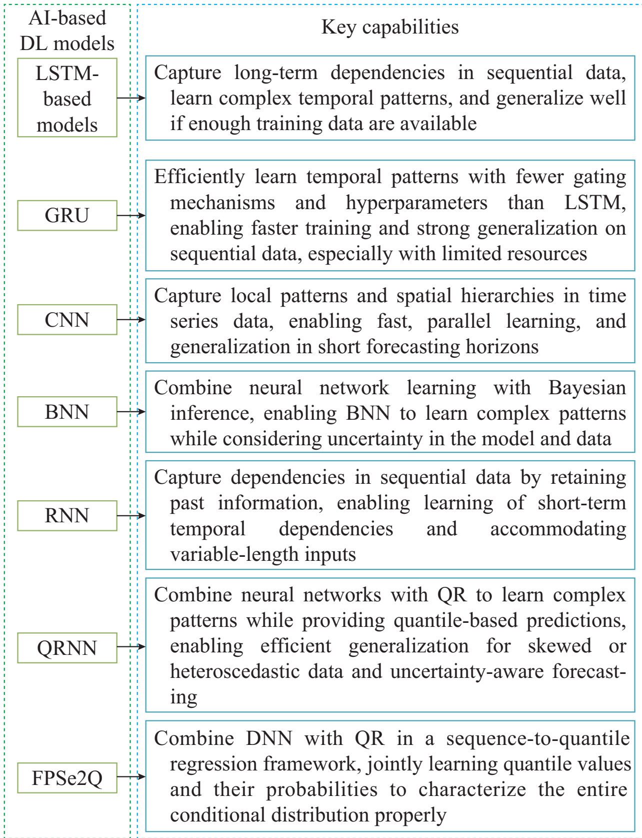

Fig. 7. Summary of AI-based DL models for NLF along with their key capabilities.

In [18], the graph Transformer (GTrans) combined with series decomposition and patch technology (denoted as PSGformer) is proposed for NLF. In that work, a feature graph attention module is embedded in the Transformer encoder to capture inter-feature dependencies at different time segments, while a segment-wise iterative module is designed to extract multiscale trend patterns.

In [47], a novel Transformer model named TransformGraph is proposed for NLF, integrating graph convolutional network (GCN) to encode multivariate sequences and capture correlation information. Considering the uncertainty raised from solar PV and wind power generation, the effectiveness of the TransformGraph is compared with five baseline models on three different datasets. The MAE values of TransformGraph are 0.406 MW, 0.606 MW, and 0.284 MW on the three datasets, with average reductions of 42.53%, 40.54%, 34.21%, 8.67%, and 2.59% compared with ARIMA, RNN, LSTM, Transformer, and Informer, respectively. Moreover, experimental results across different forecasting horizons confirm that TransformGraph exhibits good stability, demonstrating its capability for both STNLF and MTNLF tasks.

In [17], the temporal-channel attention multi-scale Transformer (TCAMS-Trans) model is introduced for direct NLF. It presents a hierarchical encoder-decoder structure with a temporal-channel attention (TCA) block, which extracts multi-scale features and uncovers relationships between NL and its various influential factors. Temporal attention captures temporal patterns, while channel attention identifies dependencies between variables. The TCAMS-Trans model is tested on three different datasets with different prediction horizons (i.e., 24, 48, 96, 168, 336, and 720 hours) and compared with traditional statistical models such as ARIMA. The Friedman test is used to rank the forecasting models in each test scenario, where TCAMS-Trans model consistently outperforms other baseline models, achieving the best performance with the rank of 1, while ARIMA scores an average rank of 9.11.

Temporal fusion Transformer (TFT) is used for disaggregated day-ahead STNLF in [93], which is designed for forecasting multiple time series with varying frequencies and irregular intervals. By combining attention and gating mechanisms, TFT captures long-term dependencies and handles both static and dynamic features in multivariate time series. This enables TFT to provide interpretable features that reveal the impact of each variable on the forecasted values to make it highly suitable for multi-feature time-series forecasting.

Gen-AI models also enhance the accuracy of probabilistic NLF by effectively modeling temporal correlations in the time series. In [69], an attention mechanism-based model, named GTrans network, is proposed, which is built on three architectures (i.e., graph attention network, variational autoencoder, and Transformer) for the day-ahead NLF with the BTM effect. It employs data from adjacent zones and incorporates temporal dependencies in the time series. For a specific zone, the pinball score of GTrans is 48.79, while that of QR is 64.90.

Current research emphasizes the application of Transformers, along with GPR-enabled residual modeling techniques, for enhanced probabilistic NLF, because parametric probabilistic NLF may introduce epistemic uncertainties [14]. The Transformer network is used to model the temporal dependencies in NL data, while GPR is employed to model the residual errors, which makes Transformer + GPR model probabilistic and provides uncertainty quantification. Comparative tests show that this model outperforms several benchmark models by effectively capturing uncertainty and improving the accuracy of NLF.

A Transformer model with QR, named Interformer, is proposed, which contains a local variable selection network to automatically extract important input features to construct an interpretable probabilistic residential NLF [57]. Moreover, Interformer incorporates an interpretable sparse self-attention mechanism to model the long-term dependencies across all input time steps. This self-attention mechanism identifies relevant time steps that enable Interformer to capture persistent temporal patterns in multi-step-ahead predictions. The goal of Interformer is to address the challenges posed by the widespread integration of small-scale solar PV systems into the residential sector.

Diffusion models have recently gained attention for probabilistic forecasting to capture uncertainty and NL dynamics. Denoising diffusion probabilistic model creates new data by simulating a random diffusion process that turns random noise into the desired data distribution. This model has two main steps: adding noise in the forward process and removing noise in the reverse process. It provides better stability and accuracy than traditional models by using a diffusion-denoising process based on Bayesian learning, which improves uncertainty modeling. Reference [119] introduces an enhanced conditional diffusion model (ECDM) that creates multiple scenarios for NLF using adaptive KDE to produce reliable PIs that improve the forecasting accuracy and help power system operators manage uncertainties with high levels of RE penetration. Hence, historical NL data guide the generative process, with weather and calendar information integrated via cross-attention mechanisms. A unique feature engineering method captures periodic characteristics of NL, and an unconditional model ensures diversity in the generated scenarios. NLF error may result from imprecise weather data and epistemic uncertainty. The condition-integrated diffusion model can enhance the accuracy of NLF, but it lacks generative diversity. Therefore, it can be combined with unconditional model that retains the same structure as the conditional model, but without cross-attention module and independent of external factors such as weather data. The accuracy and adaptability of NLF are enhanced by this design.

A couple of Gen-AI models apply 1-hour steps [16], [69] with a 24-hour forecasting horizon, while Transform-Graph [47] and TCAM-Trans [17] apply 1-hour steps with a 7-day forecasting horizon, and PSGformer extends the forecasting horizon to 30 days [18]. Interformer applies a 30-min step with a 14-day forecasting horizon [57]. TFT applies a 15-min step [93].

A summary of Gen-AI models for NLF along with their key capabilities [14], [16]-[18], [47], [57], [93] is given in Fig. 8.

C. Hybrid and Ensemble Models

To increase the accuracy and reliability of NLF, hybrid and ensemble models have gathered significant attention from researchers. These models can incorporate the strengths of multiple models while reducing the limitations of individual ones.

Hybrid models integrate heterogeneous models (e.g., statistical, physical, and mathematical models) in a unified framework to address problems that individual models may not be able to solve alone [120]. For instance, QR-CNN-GRU-based hybrid NLF model is applied in [39], where the limitation of GRU (i.e., weakness in capturing volatility) is addressed using CNN, and QR is used for predicting the conditional quantiles of NL. Similarly, a CNN-GRU-BiLSTM-based hybrid NLF model is proposed in [87] to determine the optimal size of the appropriate ESS for supporting the microgrid.

Reference [73] presents a federated learning (FL) model for NLF in microgrids, using LSTM and a node selection strategy. To efficiently train forecasting models, selecting an optimal subset of participating nodes for FL model is essential. In [73], the node selection problem is formulated as a Markov decision process (MDP) and solved using the asynchronous advantage actor-critic (A3C) deep reinforcement learning algorithm. FL model, as a decentralized ML model, preserves client data privacy and provides better scalability compared with centralized models [73]. These benefits make FL model a promising one for implementing NLF models, ensuring scalability and privacy in decentralized settings.

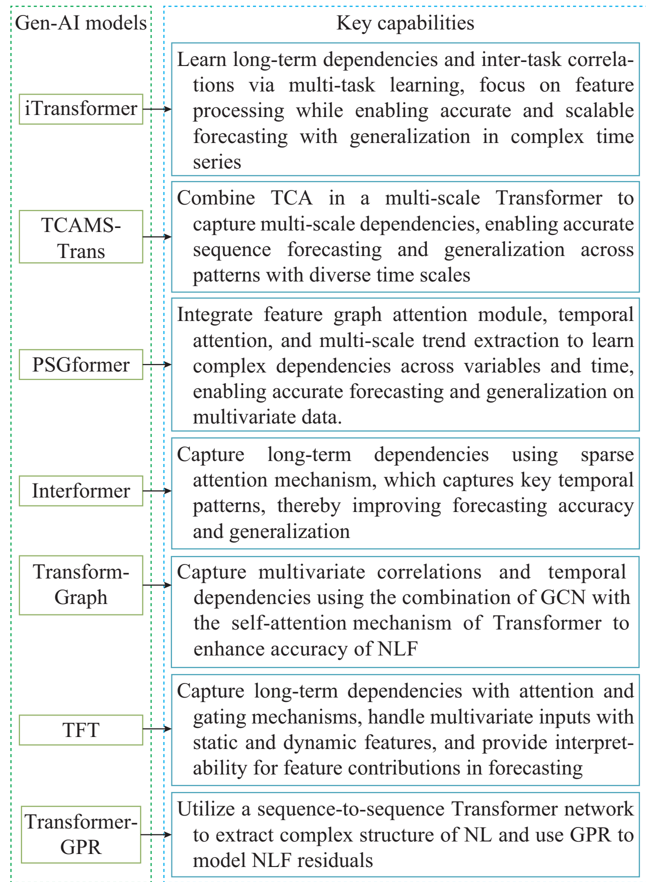

Fig. 8. Summary of Gen-AI models for NLF along with their key capabilities.

Reference [88] uses the feature extraction capability of CNN and sequential processing of Bi-LSTM, which is enhanced with a residual attention (RA) mechanism to improve the overall accuracy of NLF. Similarly, the advantages of CNN and LSTM are combined in [74] with QR to determine the probabilistic NLF. Reference [75] reconstructs the phase space of the NL time series, and the reconstructed data are used for DNN for NLF. This hybrid model performs well with high accuracy under high levels of solar PV penetration and different weather conditions. Moreover, [86] introduces a hybrid model combining CNN and encoder-decoder architectures for STNLF and PV forecasting. It uses dilated causal convolutions, residual blocks, and probabilistic forecasting with Monte Carlo dropout (MCD) and KDE. This model outperforms other traditional and DL models in point and probabilistic forecasting and uncertainty handling in smart grids.

Reference [11] employs an adaptive neuro fuzzy inference system (ANFIS) to investigate the performance of different

NLF models using two real-world datasets under varying levels of solar PV and wind power penetration, while accounting for seasonal fluctuations. This study highlights that specific forecasting strategies may not consistently outperform others across different power systems or time periods.

Furthermore, ensemble models aggregate multiple forecasting engines to improve the overall performance and stability while reducing the variance and bias of individual models. Depending on the combination, it can be classified as bagging, boosting, and stacking ensemble models [120]. For instance, [58] proposes a cluster-based stacking ensemble model including extreme learning machine (ELM), LSTM, XGBoost, support vector machine (SVM), LSSVM, and RF as base learners, and RIME optimization algorithm with local enhancement (LRIME) is applied to optimize the parameters of base learners. In addition, a bootstrap method is employed in [58] to generate PIs for various forecasting scenarios, quantifying uncertainty in residential NL. Reference [76] evaluates the performance of tree ensemble models using least-squares boosting (LSBoost) for day-ahead NLF. This study focuses on the impact of different retraining schedules. Furthermore, an ensemble model using decision tree (DT) regression, SVR, XGBoost, and lasso regression (LR) is applied in [46] for NLF. This model enhances accuracy by combining model strengths, offering more reliable predictions, particularly in situations where NL patterns fluctuate significantly over time.

A summary of hybrid and ensemble models for NLF along with their key capabilities [11], [32], [39], [46], [53], [54], [58], [72], [73], [75]-[78], [80], [81], [86], [88], [121] is given in Fig. 9, where WPE is short for weighted permutation entropy; PSR is short for phase space reconstruction; BPNN is short for back propagation neural network; WNN is short for wavelet neural network; ENN is short for Elman neural network; LMBP is short for Levenberg-Marquardt back propagation; PSLP is short for personalized standard load profile; ASBO is short for autoencoder-based SVM optimized with Bayesian optimization (BO); and EnB is short for ensemble bagged.

To conclude, this section reviews various forecasting engines and evaluates their performance in both real-world and simulated test cases, with a focus on their forecasting capabilities and uncertainty management, especially in scenarios with varying levels of RE penetration. ML models are recognized for their ability to adapt to the complex fluctuations of RE generation, while hybrid and ensemble models combine multiple methods to improve the accuracy and reliability of NLF. This section emphasizes the role of Gen-AI models, particularly state-of-the-art Transformer models, in providing deterministic and probabilistic forecasts while managing the inherent uncertainties in systems with high levels of RE penetration.

VI. Assessment of Interpretability, Scalability, and Applicability of NLF Models

This section assesses the interpretability, scalability, and applicability of NLF models. These characteristics are important for understanding how the NLF models function in real-world applications, especially with high levels of RE penetration.

A. Interpretability

DL models have achieved state-of-the-art results in many forecasting tasks, but they face several challenges, e.g., limited interpretability, poor uncertainty handling, high data and computational demands, optimization complexities, and vulnerability to adversarial attacks [67]. Recent works have focused on addressing these limitations. For instance, [67] specifically targets to resolve the issues with uncertainty and interpretability in NLF using Bayesian DL model. Additionally, [94] incorporates an explainable AI model, namely Shapley additive explanations with LSTM model, to ensure interpretability and transparency of NLF model. In recent years, attention-based mechanisms have been studied to open the black box. For example, TFT is known for its ability to show long-term dependencies and provide interpretability through self-attention values [93]. However, the relationship between attention weights and model outputs is not always clear in this model [116]. Also, the time-series forecasting often involves various input features, and a single temporal attention mechanism can identify correlations over time but may not offer insights into the most important features. To address such issues, Interformer in [57] focuses more on feature-level interpretability by using a local variable selection network. Furthermore, a model-agnostic framework is presented in [116] using sequence-to-sequence RNN, which is able to visualize the informativeness of multiple features over the time dimensions on a two-dimensional color map. This framework provides explainability for a forecasting model, without the need for additional data mining or data labelling.

B. Scalability

The accuracy of NLF varies by region due to varying levels of RE penetration, and the scalability of NLF models depends on their ability to manage larger datasets, higher computational demands, and local environmental variations. Reference [26] discusses the limitations of BNN model in terms of scalability in different climates, network levels, and customer types.

However, these issues are emphasized and overcome in several studies for NLF. Reference [94] proposes a lightweight and adaptive NLF model using GRU and LSTM. This model demonstrates high computational efficiency and is well-suited for on-device deployment in real-time energy monitoring and management applications. Reference [43] explores the scalability of an MLP-based approach in low-voltage distribution networks with high levels of PV penetration to highlight its computational efficiency and suitability for large-scale applications. This approach balances accuracy, uncertainty, calibration, and complexity, providing a more efficient alternative to complex models. References [49] and [68] also apply this scalable approach for predicting NL. Reference [50] proposes a feature reconstruction-based NLF model using Bi-LSTM to handle large datasets, reducing computational complexity and enhancing training efficiency. This model aims to improve the accuracy of NLF in regions with large-scale distributed PV systems.

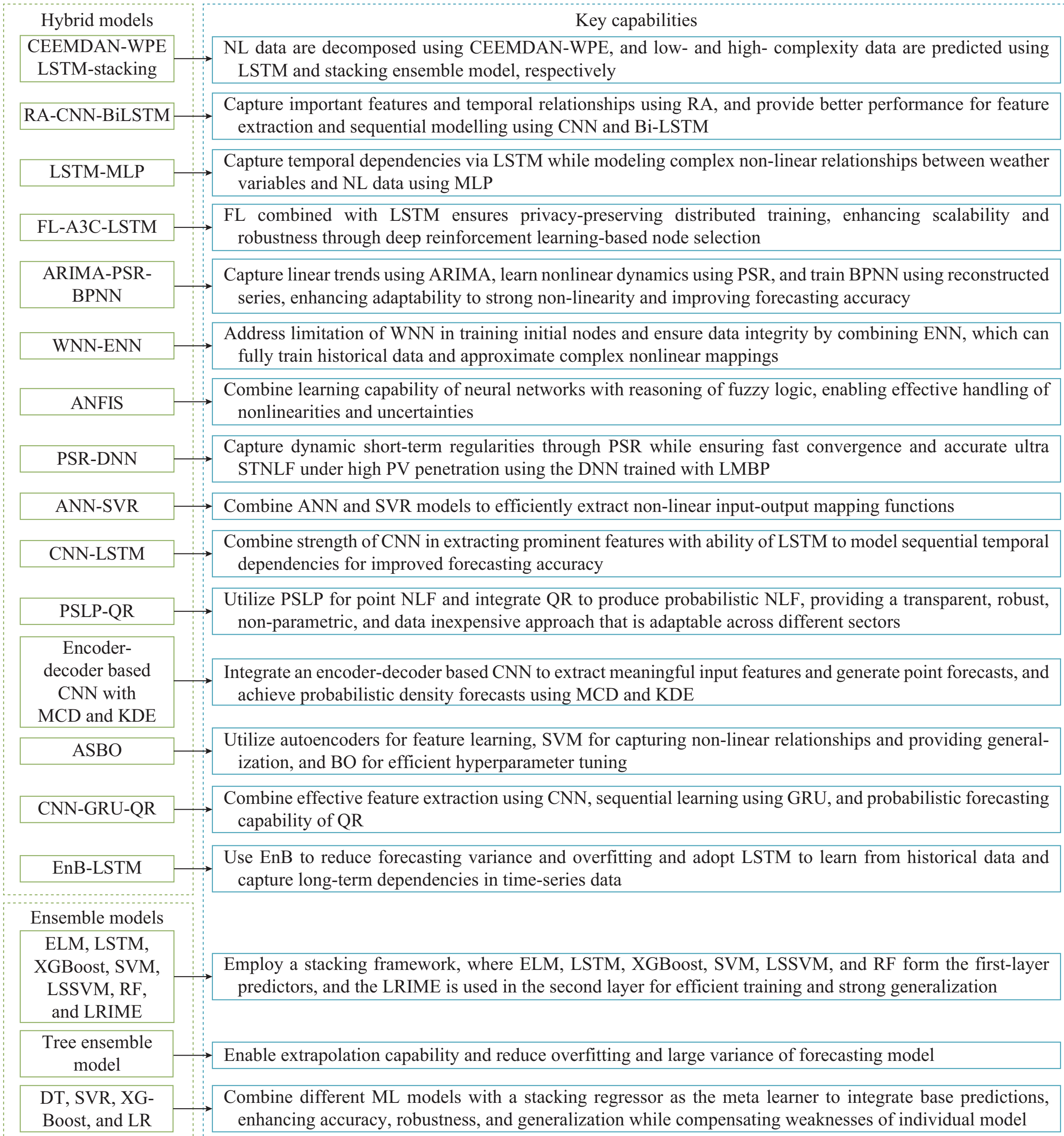

Fig. 9. Summary of hybrid and ensemble models for NLF along with their key capabilities.

C. Applicability

Several studies demonstrate the applicability of their NLF models with varying levels of RE penetration. For instance, [5] presents an NLF model based on statistical models such as AR, SVR, ARX, and SVRX for microgrids with high levels of solar PV penetration. It suggests that the NLF model can also be applied to areas with high levels of wind power penetration, though its performance may be limited due to a lack of high-resolution wind forecasts at the local level. Reference [90] proposes an estimation framework using multi-input LSTM with DWT and a fuzzy system that improves the accuracy of NLF, regardless of the non-linearity and volatility of the energy management problem. Furthermore, [122] compares the LSTM with two incremental learning models, i.e., fully online sequential-extreme learning machine (FOS-ELM) and online sequential-extreme learning machine (OS-ELM), and concludes that these incremental learning models are more adaptable to changes in solar PV penetration. In particular, the OS-ELM, which is properly trained with synthetic training samples, yields the most accurate NLF results. References [11] and [91] incorporate ANFIS and SARIMAX models, respectively, and report that the accuracy of these models reduces with the increasing levels of RE penetration. However, BD-LSTM is tested in [67] with increasing levels of solar PV penetration, and the results show that the accuracy of BD-LSTM improves as PV visibility increases. Furthermore, [60] uses GAM combined with QR and GPD, and tests this combined approach on 14 different regions in Great Britain to evaluate its applicability. However, the obtained results are specific to those regions and need calibration and validation in the case of transferability to other systems.

Overall, this section emphasizes the ongoing efforts to enhance the interpretability, scalability, and adaptability of NLF models, highlighting their importance in addressing the challenges posed by high levels of RE penetration and the need to manage large datasets and diverse environmental conditions.

VII. Challenges and Future Research Directions of NLF

This section intends to investigate the challenges and future research directions.

Variability and intermittency of RE, particularly with high levels of solar PV penetration, create significant challenges for NLF. BTM energy resources add uncertainty due to the lack of generation visibility, and stochastic factors like human activities and climate variability increase complexity. These factors could significantly impact the accuracy of NLF [67]. Figure 10 highlights the main challenges associated with NLF.

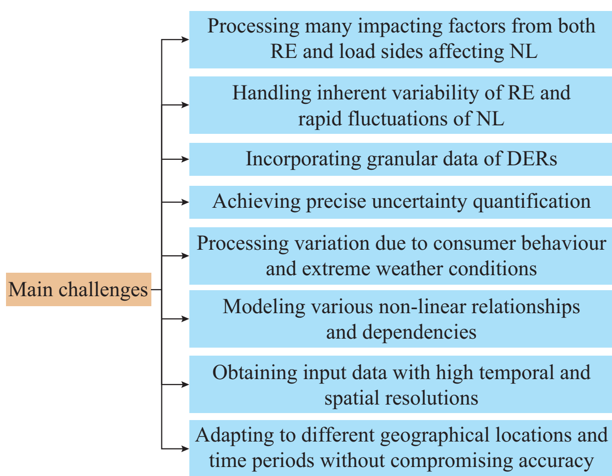

Fig. 10. Main challenges associated with NLF.

Future research should focus on improving the TCA mechanism to help NLF models handle uncertainty by focusing on the most relevant time steps and features [17]. Precise uncertainty quantification is another critical issue in NLF. Exploring adaptive conformal prediction could allow the NLF models to adjust prediction based on data, thereby improving the accuracy of NLF [49]. The transferability of NLF models across different geographical locations and seasons remains a significant challenge. To ensure the generalization of NLF models under varying levels of RE penetration, future research should focus on their performance in diverse geographical areas and seasonal variations.

Additionally, the performance of NLF models should be evaluated for sufficiently long periods to capture evolving NL distributions. Moreover, a deeper exploration of the correlation between RE penetration and NL is necessary to improve the robustness of NLF models.

The influence of various factors from both the RE and load sides on NL needs more attention. However, the application of instance selection to acquire the most informative instances is rarely explored in NLF. Therefore, developing context-aware instance selection algorithms would help address this issue.

Incorporating granular data of DERs (e.g., PVs, EVs, and batteries) is essential for an effective NLF. Hence, the scenario-based forecasting and enhanced data quality management, including anomaly detection, missing data imputation, and feature engineering, should be prioritized. Furthermore, studies on NLF generally perform data preprocessing and feature engineering on historical datasets [16]-[18], [55], [56], with limited attention to real-time data integration [79]. Future work could explore adapting these methods for real-time data integration and data consistency management. Obtaining input data with high temporal and spatial resolutions is necessary for accurate NLF, but not easily available. In such cases, the transfer learning model can be utilized to improve cold start predictions in power systems with limited or no historical data. It is a promising avenue for future research [16]. Additionally, investigating Gen-AI models for synthetic NL data generation while maintaining temporal patterns and trends can be beneficial.

Extreme weather conditions and other influential factors of dynamic market and grid stability require robust NLF models that can adapt to different grid configurations and external conditions. Therefore, improving the reliability and adaptability of NLF models under these changing conditions continues to be a critical priority [71]. Researchers can utilize the theoretical concepts of adaptive game-theoretic models for energy systems to develop advancements in this area [123]. Additionally, the integration of RE further complicates bidding strategies in electricity market [124]. Adaptive NLF model can offer deeper insights into dynamic bidding strategies and improve the stability of NLF in renewable-rich power systems.

Although papers like [32], [92], and [102] point that the varying levels of RE penetration affect the accuracy of NLF and model performance, to strengthen such findings, conducting a sensitivity analysis could offer deeper insights into the model robustness under various conditions such as extreme weather events or high levels of DER penetration. Rapidly changing demand-side management strategies present challenges to NLF. To improve scheduling and operational planning with high levels of RE penetration, NLF models need to capture both temporal and spatial dependencies [70]. Future research should also integrate demand response programs and energy prices to improve the accuracy of NLF and demand-side flexibility [86]. Additionally, incorporating the data of various customer-side demand-impacting factors (e. g., energy efficiency, battery storage, EV charging, and electric heating and cooling systems of buildings) into the NLF process can further improve its accuracy. Therefore, the integration of these factors with demand-side management and the operational planning of ESSs remains an important area for research to enhance the accuracy of NLF [125].

Moreover, incorporating LF into predictive maintenance could improve resource scheduling, minimize outages during peak demand periods, and reduce reliability issues [126]. By accounting for NLF, it is anticipated that predictive maintenance will be even more effective, as NLF considers not only the impact of the gross demand of smart grid but also the effect of RE generation. These issues present potential directions for future research on NLF and its applications.

Limited interpretability of DL models for NLF hampers their effectiveness. Future research should focus on improving the interpretability and transparency of DL and Gen-AI models for NLF, especially by clarifying the roles of individual attention heads to build trust with power system operators [69].

Lastly, privacy-preserving NLF is essential to protect sensitive data such as consumer behavior and energy usage patterns, while ensuring regulatory compliance. Without privacy protection, data misuse and unauthorized access could undermine trust in NLF systems. Although FL-based NLF has been introduced in [73], further work is needed to fully utilize its potential. Future research should prioritize collaborative forecasting by addressing data-sharing and cybersecurity issues, ensuring secure and privacy-preserving data utilization. Figure 11 illustrates the research gaps and outlines potential future research directions of NLF.

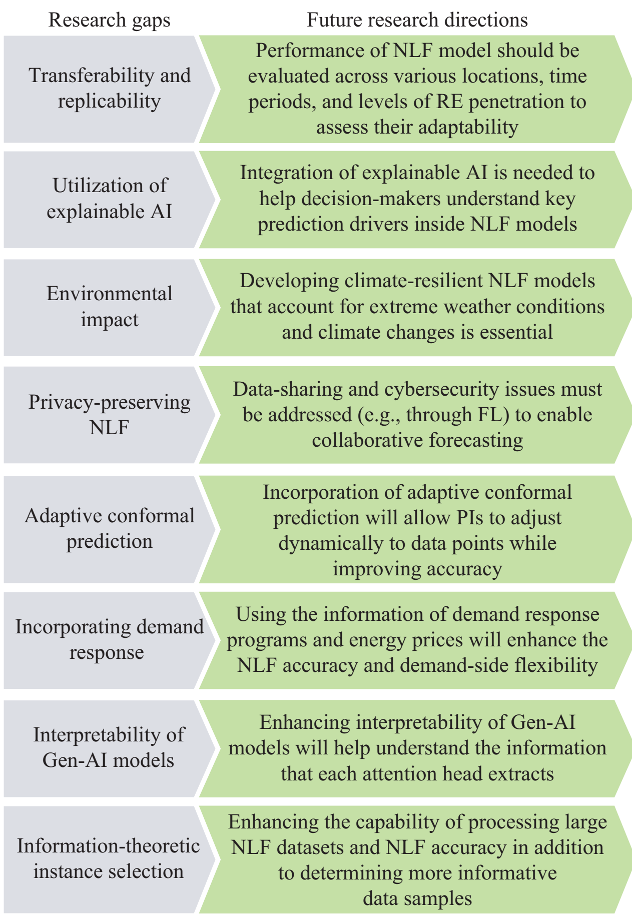

Fig. 11. Research gaps and potential future research directions of NLF.

VIII. Conclusion

This paper critically investigates the current state of NLF for RE-integrated power systems. This literature review has enabled a detailed investigation of the diverse techniques applied in NLF. The critical data preprocessing methods required to shape the input dataset, the feature engineering methods to select and extract the high-impact input variables, and the necessary hyperparameter tuning methods to enhance the performance and accuracy of NLF are critically evaluated.

Furthermore, an extensive review of forecasting engines for NLF is conducted, evaluating their capability and applicability under diverse forecasting scenarios. Despite significant progress in NLF research, the review identifies several key challenges in data quality and model generalization, real-time prediction capabilities, uncertainty in consumer behavior and RE generation, and complexities in non-linear modeling. Also, current NLF models predominantly focus on feature selection and extraction to enhance model accuracy by identifying relevant input variables. However, instance selection and generation remain underexplored in this domain.

In conclusion, the findings of this review emphasize the importance of ongoing research in developing robust, reliable, accurate, and real-time NLF to meet the evolving demands of RE integration and highlight the need for continued innovation in NLF models.

References

[1] National Grid. (2024, Aug.). What is the clean energy transition? [Online]. Available: https://www.nationalgrid.com/stories/energy-explained/what-clean-energy-transition

[2] International Energy Agency (IEA). (2025, Apr.). Renewables—Energy System. [Online]. Available: https://www.iea.org/energy-system/renewables

[3] S. Zachary, S. H. Tindemans, M. P. Evans *et al.*, "Scheduling of energy storage," *Philosophical Transactions of the Royal Society A*, vol. 379, p. 20190435, Jul. 2021.

[4] Y. Zhang, N. Rahbari-Asr, J. Duan *et al.*, "Day-ahead smart grid cooperative distributed energy scheduling with renewable and storage integration," *IEEE Transactions on Sustainable Energy*, vol. 7, no. 4, pp. 1739-1748, Jun. 2016.

[5] A. Kaur, L. Nonnenmacher, and C. F. M. Coimbra, "Net load forecasting for high renewable energy penetration grids," *Energy*, vol. 114, pp. 1073-1084, Nov. 2016.

[6] G. Tziolis, J. Lopez-Lorente, M. I. Baka *et al.*, "Comparative analysis of machine learning models for short-term net load forecasting in renewable integrated microgrids," in *Proceedings of 2nd International Conference on Energy Transition in the Mediterranean Area*, Thessaloniki, Greece, Oct. 2022, pp. 1-5.

[7] A. Stratman, T. Hong, M. Yi *et al.*, "Net load forecasting with disaggregated behind-the-meter PV generation," *IEEE Transactions on Industry Applications*, vol. 59, no. 5, pp. 5341-5351, May 2023.

[8] W. Li, H. Zhong, K. Wand *et al.*, "Short-term net load forecasting model based on GRU neural network optimized by improved arithmetic optimization algorithm," in *Proceedings of IEEE 7th International Conference on Information Systems and Computer Aided Education*, Dalian, China, Sept. 2024, pp. 52-56.

[9] G. Tziolis, A. Livera, J. Montes-Romero *et al.*, "Direct short-term net load forecasting based on machine learning principles for solar-integrated microgrids," *IEEE Access*, vol. 11, pp. 102038-102049, Sept. 2023.

[10] G. Tziolis, A. Livera, A. Michail *et al.*, "Direct against indirect short-term net load forecasting using machine learning principles for renewable microgrids," in *Proceedings of IEEE International Smart Cities Conference*, Bucharest, Romania, Sept. 2023, pp. 1-5.

[11] G. Aburiyana, H. Aly, and T. Little, "Analyzing the performance of direct and indirect net load forecasting strategies under varying penetration levels of PV and wind power," in *Proceedings of ASU International Conference in Emerging Technologies for Sustainability and Intelligent Systems*, Manama, Bahrain, Jan. 2024, pp. 1502-1506.

[12] S. Watts and I. Macgill, "Comparing short-term net load forecasting methods for solar homes," in *Proceedings of IEEE PES GTD International Conference and Exposition*, Istanbul, Turkey, May 2023, pp. 404-408.

[13] M. Beichter, K. Phipps, M. M. Frysztacki *et al.*, "Net load forecasting using different aggregation levels," *Energy Informatics*, vol. 5, no. 1,

pp. 1-22, Sept. 2022.

[14] J. Hu, W. Hu, D. Cao *et al.*, "Probabilistic net load forecasting based on transformer network and Gaussian process-enabled residual modeling learning method," *Renewable Energy*, vol. 225, p. 120253, May 2024.

[15] W. Feng, B. Deng, T. Chen *et al*., "Probabilistic net load forecasting based on sparse variational Gaussian process regression," *Frontiers in Energy Research*, vol. 12, p. 1429241, Sept. 2024.

[16] J. Pei, N. Liu, J. Shi *et al.*, "Tackling the duck curve in renewable power system: a multi-task learning model with iTransformer for net-load forecasting," *Energy Conversion and Management*, vol. 326, p. 119442, Feb. 2025.

[17] Q. Zhang, S. Zhou, B. Xu *et al.*, "TCAMS-Trans: efficient temporal-channel attention multi-scale transformer for net load forecasting," *Computers and Electrical Engineering*, vol. 118, p. 109415, Sept. 2024.

[18] Q. Zhang, S. Zhou, B. Xu *et al.*, "PSGformer: a novel multivariate net load forecasting model for the smart grid," *Journal of Computational Science*, vol. 78, p. 102288, Jun. 2024.

[19] S. Hu, J. Yang, Y. Guo *et al.*, "Study on the evolutionary process and balancing mechanism of net load in renewable energy power systems," *Energies*, vol. 17, no. 18, p. 4654, Sept. 2024.

[20] E. T. Rizi, M. Rastegar, and A. Forootani, "Power system flexibility analysis using net-load forecasting based on deep learning considering distributed energy sources and electric vehicles," *Computers and Electrical Engineering*, vol. 117, p. 109305, Jul. 2024.

[21] J. Zhang, X. Zhu, Y. Xie *et al.*, "Detection and prediction of wind and solar photovoltaic power ramp events based on data-driven methods: a critical review," *Energies*, vol. 18, no. 13, p. 3290, Jun. 2025.

[22] M. Cui, J. Zhang, C. Feng *et al.*, "Characterizing and analyzing ramping events in wind power, solar power, load, and netload," *Renewable Energy*, vol. 111, pp. 227-244, Oct. 2017.

[23] National Renewable Energy Laboratory. (2015, Nov.). Overgeneration from solar energy in California: a field guide to the duck chart. [Online]. Available: http://www.nrel.gov/docs/fy16osti/65453.pdf

[24] M. Hasan, Z. Mifta, S. J. Papiya *et al.*, "A state-of-the-art comparative review of load forecasting methods: characteristics, perspectives, and applications," *Energy Conversion and Management: X*, vol. 26, p. 100922, Apr. 2025.

[25] D. W. van der Meer, M. Shepero, A. Svensson *et al.*, "Probabilistic forecasting of electricity consumption, photovoltaic power generation and net demand of an individual building using Gaussian processes," *Applied Energy*, vol. 213, pp. 195-207, Mar. 2018.

[26] G. Tziolis, C. Spanias, M. Theodoride *et al.*, "Short-term electric net load forecasting for solar-integrated distribution systems based on Bayesian neural networks and statistical post-processing," *Energy*, vol. 271, p. 127018, May 2023.

[27] A. Rafati, M. Joorabian, E. Mashhour *et al.*, "Machine learning-based very short-term load forecasting in microgrid environment: evaluating the impact of high penetration of PV systems," *Electrical Engineering*, vol. 104, no. 4, pp. 2667-2677, Feb. 2022.

[28] S. Sreekumar, K. C. Sharma, and R. Bhakar, "Gumbel copula based aggregated net load forecasting for modern power systems," *IET Generation, Transmission & Distribution*, vol. 12, no. 19, pp. 4348-4358, Oct. 2018.

[29] S. Sreekumar, N. U. Khan, A. S. Rana *et al.*, "Aggregated net-load forecasting using Markov-chain Monte-Carlo regression and C-vine copula," *Applied Energy*, vol. 328, p. 120171, Dec. 2022.

[30] S. Sreekumar, K. C. Sharma, and R. Bhakar, "Grey system theory based net load forecasting for high renewable penetrated power systems," *Technology and Economics of Smart Grids and Sustainable Energy*, vol. 5, no. 21, pp. 1-14, Oct. 2020.

[31] M. Alipour, J. Aghaei, M. Norouzi *et al.*, "A novel electrical net-load forecasting model based on deep neural networks and wavelet transform integration," *Energy*, vol. 205, p. 118106, Aug. 2020.

[32] J. Blain and D. L. Donaldson, "The effect of increasing solar photovoltaic generation on medium-term net-load forecasting" in *Proceedings of Renewable Power Generation and Future Power Systems Conference 2023*, Glasgow, UK, Nov. 2023, pp. 94-100.

[33] X. Chen, F. Xu, G. He *et al.*, "Decoupling based monthly net electricity consumption prediction model considering high penetration of distributed solar PV systems," *Sustainable Energy, Grids, and Networks*, vol. 32, p. 100870, Dec. 2022.

[34] G. Nalcaci, A. Özmen, and G. W. Weber, "Long-term load forecasting: models based on MARS, ANN and LR methods," *Central European Journal of Operations Research*, vol. 27, no. 4, pp. 1033-1049, Dec. 2019.

[35] F. Heymann, J. Silva, V. Miranda *et al.*, "Distribution network planning considering technology diffusion dynamics and spatial net-load behavior," *International Journal of Electrical Power & Energy Systems*, vol. 106, pp. 254-265, Mar. 2019.

[36] F. Kaytez, "A hybrid approach based on autoregressive integrated moving average and least-square support vector machine for long-term forecasting of net electricity consumption," *Energy*, vol. 197, p. 117200, Apr. 2020.

[37] Y. Wang, N. Zhang, Q. Chen *et al.*, "Data-driven probabilistic net load forecasting with high penetration of behind-the-meter PV," *IEEE Transactions on Power Systems*, vol. 33, no. 3, pp. 3255-3264, May 2018.

[38] G. Tziolis, J. Lopez-Lorente, M. I. Baka *et al.*, "Direct short-term net load forecasting in renewable integrated microgrids using machine learning: a comparative assessment," *Sustainable Energy, Grids and Networks*, vol. 37, p. 101256, Mar. 2023.

[39] R. Liu, J. Shi, G. Sun *et al.*, "A short-term net load hybrid forecasting method based on VW-KA and QR-CNN-GRU," *Electrical Power Systems Research*, vol. 232, p. 110384, Jul. 2024.

[40] J. Son, J. Cha, H. Kim *et al.*, "Day-ahead short-term load forecasting for holidays based on modification of similar days' load profiles," *IEEE Access*, vol. 10, pp. 17864-17880, Feb. 2022.

[41] O. Rubasinghe, T. Zhang, X. Zhang *et al.*, "Highly accurate peak and valley prediction short-term net load forecasting approach based on decomposition for power systems with high PV penetration," *Applied Energy*, vol. 333, p. 120641, Mar. 2023.

[42] L. Cheng, R. Sun, K. Wang *et al.*, "Advancing sustainable electricity markets: evolutionary game theory as a framework for complex systems optimization and adaptive policy design," *Complex & Intelligent Systems*, vol. 11, no. 7, p. 320, Jun. 2025.

[43] A. Faustine, N. J. Nunes, and L. Pereira, "Efficiency through simplicity: MLP-based approach for net-load forecasting with uncertainty estimates in low-voltage distribution networks," *IEEE Transactions on Power Systems*, vol. 40, no. 1, pp. 46-56, May 2024.

[44] Y. Peng, M. Fan, N. Li *et al.*, "Net load forecasting method for distribution substation based on an improved MHGSA mechanism," in *Proceedings of 2024 China International Conference on Electricity Distribution*, Hangzhou, China, Sept. 2024, pp. 1354-1358.

[45] N. Giamarelos, E. N. Zois, M. Papadimitrakis *et al.*, "Short-term electric load forecasting with sparse coding methods," *IEEE Access*, vol. 9, pp. 102847-102861, Jul. 2021.

[46] I. H. Laouali, N. Italiano, Â. Casaleiro *et al.*, "Short term net load forecasting using computational intelligence techniques," *IET Conference Proceedings*, vol. 2023, no. 18, pp. 139-145, Nov. 2023.

[47] Q. Zhang, J. Chen, G. Xiao *et al.*, "TransformGraph: a novel short-term electricity net load forecasting model," *Energy Reports*, vol. 9, pp. 2705-2717, Dec. 2023.

[48] A. Stratigakos, A. Bachoumis, V. Vita *et al*., "Short-term net load forecasting with singular spectrum analysis and LSTM neural networks," *Energies*, vol. 14, no. 14, p. 4107, Jul. 2021.

[49] A. Faustine and L. Pereira, "Conformal multilayer perceptron-based probabilistic net-load forecasting for low-voltage distribution systems with photovoltaic generation," in *Proceedings of IEEE International Conference on Communications, Control, and Computing Technologies for Smart Grid*, Oslo, Norway, Sept. 2024, pp. 59-64.

[50] X. Zheng, M. Yang, Y. Yu *et al.*, "Short-term net load forecasting for regions with distributed photovoltaic systems based on feature reconstruction," *Applied Sciences*, vol. 13, no. 16, p. 9064, Aug. 2023.

[51] O. Abedinia and N. Amjady, "Net demand prediction for power systems by a new neural network-based forecasting engine," *Complexity*, vol. 21, no. S2, pp. 296-308, Jul. 2016.

[52] K. Zhu, G. Gao, Y. Li *et al.*, "Short-term net-load forecasting based on LSTM enhanced by adaptive adjustment and multi period prediction mechanism," in *Proceedings of IEEE 7th International Conference on Information Systems and Computer Aided Education*, Dalian, China, Sept. 2024, pp. 7-10.

[53] A. H. Baharvand, S. H. B. Fard, M. Doostizadeh *et al.*, "An autoencoder-based support vector machine optimized with Bayesian optimization for net load forecasting with uncertainty estimation," in *Proceedings of 28th International Electrical Power Distribution Conference*, Zanjan, Iran, Apr. 2024, pp. 1-6.

[54] Z. Yang, X. Li, X. Kong *et al*., "A method of short-term load prediction of renewable energy power system based on CNNLSTM," in *Proceedings of 25th International Conference on Electrical Machines and Systems*, Chiang Mai, Thailand, Nov. 2022, pp. 1-5.

[55] B. Zhou, Y. Meng, W. Huang *et al.*, "Multi-energy net load forecasting for integrated local energy systems with heterogeneous prosum-

ers," *International Journal of Electrical Power & Energy Systems*, vol. 126, no. PA, p. 106542, Mar. 2021.

[56] S. Tian and J. Zhang, "Microgrid net load forecasting method based on extreme value feature interaction," in *Proceedings of 2024 4th International Conference on Smart Grid and Energy Internet*, Shenyang, China, Dec. 2024, pp. 733-737.

[57] C. Xu and G. Chen, "Interpretable transformer-based model for probabilistic short-term forecasting of residential net load," *International Journal of Electrical Power & Energy Systems*, vol. 155, p. 109515, Jan. 2024.

[58] Y. He, H. Zhang, Y. Dong *et al.*, "Residential net load interval prediction based on stacking ensemble learning," *Energy*, vol. 296, p. 131134, Jun. 2024.

[59] M. Reis, A. Garcia, and R. J. Bessa, "A scalable load forecasting system for low voltage grids," in *Proceedings of 2023 IEEE PES Grid Edge Technologies Conference & Exposition*, San Diego, USA, Jun. 2017, pp. 1-6.

[60] J. Browell and M. Fasiolo, "Probabilistic forecasting of regional netload with conditional extremes and gridded NWP," *IEEE Transactions on Smart Grid*, vol. 12, no. 6, pp. 5011-5019, Aug. 2021.

[61] J. de Vilmarest, J. Browell, M. Fasiolo *et al.*, "Adaptive probabilistic forecasting of electricity (net-)load," *IEEE Transactions on Power Systems*, vol. 39, no. 2, pp. 4154-4163, Aug. 2024.

[62] X. Li, Z. Wang, C. Yang *et al.*, "An advanced framework for net electricity consumption prediction: incorporating novel machine learning models and optimization algorithms," *Energy*, vol. 296, p. 131259, Jun. 2024.

[63] P. Herodotou, G. Tziolis, G. Makrides *et al.*, "Net load prediction using machine learning methods: a case study of residential households," in *Proceedings of 3rd International Conference on Energy Transition in the Mediterranean Area*, Limassol, Cyprus, Oct. 2024, pp. 1-6.

[64] P. Shi, L. Geng, M. Zhang, *et al.*, "A short-term regional net load prediction method based on parallel fragment attention-bi-LSTM," in *Proceedings of IEEE 6th International Conference on Automation, Electronics and Electrical Engineering*, Shenyang, China, Dec. 2023, pp. 624-629.

[65] A. Kakati, S. Intodia, C. Singh *et al.*, "Net load forecasting using LSTM_OOA model," in *Proceedings of 1st International Conference on Trends in Engineering Systems and Technologies*, Kochi, India, Jul. 2024, pp. 1-6.

[66] G. Tziolis, A. Koumis, S. Theocharides *et al.*, "Advanced short-term net load forecasting for renewable-based microgrids," in *Proceedings of IEEE International Smart Cities Conference*, Pafos, Cyprus, Sept. 2022, pp. 1-6.

[67] M. Sun, T. Zhang, Y. Wang *et al.*, "Using Bayesian deep learning to capture uncertainty for residential net load forecasting," *IEEE Transactions on Power Systems*, vol. 35, no. 1, pp. 188-201, Jan. 2020.

[68] A. Faustine and L. Pereira, "FPSeq2Q: fully parameterized sequence to quantile regression for net-load forecasting with uncertainty estimates," *IEEE Transactions on Smart Grid*, vol. 13, no. 3, pp. 2440-2451, May 2022.

[69] J. Liang, W. Tang, S. Zhai *et al.*, "Attention mechanism based probabilistic day-ahead net-load forecasting with behind-the-meter solar," in *Proceedings of IEEE 7th Conference on Energy Internet and Energy System Integration*, Hangzhou, China, May 2023, pp. 3465-3470.

[70] V. Gioia, M. Fasiolo, J. Browell *et al.,* "Additive covariance matrix models: modeling regional electricity net-demand in Great Britain," *Journal of the American Statistical Association*, vol. 120, no. 549, pp. 107-119, Nov. 2025.

[71] Y. Yuan, X. Yuan, H. Wang *et al.*, "Net load forecasting method in distribution grid planning based on LSTM network," *Science and Technology for Energy Transition*, vol. 79, p. 57, Aug. 2024.

[72] H. Liang, G. Li, L. Xu *et al.*, "Short-term load forecasting for a power supplying district based on CEEMDAN-WPE-LSTM-stacking methods," in *Proceedings of 7th International Conference on Energy, Electrical and Power Engineering*, Yangzhou, China, Jul. 2024, pp. 436-441.

[73] Z. Qin, S. Shao, S. Guo *et al.*, "A federated learning approach for net load forecasting in microgrids," in *Proceedings of 24th Asia-Pacific Network Operations and Management Symposium*, Sejong, Korea, Sept. 2023, pp. 298-301.

[74] M. R. Zarei-Jeliani, M. Fotuhi-Firuzabad, and N. Pourghaderi, "Artificial intelligence-based prediction of flexibility requirements in power systems," in *Proceedings of 32nd International Conference on Electrical Engineering*, Tehran, Iran, Sept. 2024, pp. 1-5.

[75] F. Mei, Q. Wu, T. Shi *et al.*, "An ultrashort-term net load forecasting model based on phase space reconstruction and deep neural network," *Applied Science*, vol. 9, no. 7, p. 1487, Apr. 2019.

[76] A. Micallef, M. Apap, and J. Licari, "Comparison of retraining schedules in ensemble models for enhanced day-ahead electricity net-demand predictions," in *Proceedings of IEEE 22nd Mediterranean Electrotechnical Conference*, Porto, Portugal, Jun. 2024, pp. 1054-1059.

[77] J. Zhang, X. Zhu, W. He *et al.*, "The bus-bar net load forecasting based on ARIMA and PSR-BPNN," in *Proceedings of 13th IEEE International Conference on Electronic Measurement & Instruments*, Yangzhou, China, Oct. 2017, pp. 247-252.

[78] M. Cao, L. Shi, Z. Wang *et al.*, "Node net load forecasting of distributed photovoltaic considering different user characteristics," *in Proceedings of IEEE Innovative Smart Grid Technologies-Asia*, Chengdu, China, May 2019, pp. 1766-1771.

[79] M. Jalali and O. Alizadeh-Mousavi, "Application of real-time distribution grid monitoring for grid forecasting and control considering incomplete information of resources behind-the-meter," *IEEE Open Access Journal of Power & Energy*, vol. 9, pp. 308-318, Aug. 2022.

[80] Y. Chu, H. T. C. Pedro, A. Kaur *et al.*, "Net load forecasts for solar-integrated operational grid feeders," *Solar Energy*, vol. 158, pp. 236-246, Dec. 2017.

[81] M. Farhoumandi, A. Bahrami, M. Shahidehpour *et al.*, "Short-term forecasting of non-conforming net load using a fusion model with machine learning and deep learning methods," *in Proceeding of 9th International Conference on Control, Decision and Information Technologies*, Rome, Italy, Jul. 2023, pp. 2414-2419.

[82] D. W. van der Meer, J. Munkhammar, and J. Widén, "Probabilistic forecasting of solar power, electricity consumption and net load: investigating the effect of seasons, aggregation and penetration on prediction intervals," *Solar Energy*, vol. 171, pp. 397-413, Sept. 2018.

[83] S. Kerkau, S. Sepasi, H. O. R. Howlader *et al.*, "Day-ahead net load forecasting for renewable integrated buildings using XGBoost," *Energies*, vol. 18, no. 6, p. 1518, Mar. 2025.

[84] S. Lei, X. Liang, X. Wang *et al.*, "A short-term net load forecasting method based on two-stage feature selection and LightGBM with hyperparameter auto-tuning," in *Proceedings of IEEE/IAS 59th Industrial and Commercial Power Systems Technical Conference*, Las Vegas, USA, May 2023, pp. 1-6.

[85] Y. Sai, M. Fei, T. Yu *et al.*, "Bus net load forecast considering photovoltaic output correlation," in *Proceedings of IEEE International Conference on Energy Internet*, Shenyang, China, Mar. 2024, pp. 182-187.

[86] M. Jurado, M. Samper, and R. Rosés, "An improved encoder-decoder-based CNN model for probabilistic short-term load and PV forecasting," *Electric Power Systems Research*, vol. 217, p. 109153, Apr. 2023.

[87] G. Memarzadeh, H. Noori, R. Memarzadeh *et al.*, "Optimal energy storage system management in an uncertain smart grid," *Journal of Energy Storage*, vol. 117, p. 116125, May 2025.

[88] S. Shen, D. He, and W. Yu, "Net load forecasting based on residual attention CNN-BiLSTM model," in *Proceedings of 2nd International Conference on Smart Grids and Energy Systems*, Guangzhou, China, Aug. 2023, pp. 218-225.

[89] J. Silva-Rodriguez, E. Raffoul, and X. Li, "LSTM-based net load forecasting for wind and solar power-equipped microgrids," in *Proceedings of 56th North American Power Symposium*, El Paso, USA, Oct. 2024, pp. 1-6.

[90] M. J. Mokarram, R. Rashiditabar, M. Gitizadeh *et al.*, "Net-load forecasting of renewable energy systems using multi-input LSTM fuzzy and discrete wavelet transform," *Energy*, vol. 275, p. 127425, Jul. 2023.

[91] X. Zhang, F. Du, R. Hu *et al.*, "Net load forecasting and energy storage demand analysis for renewable energy integration," *IET Conference Proceedings*, vol. 2023, no. 15, pp. 1202-1206, Oct. 2023.

[92] A. Falces, C. Capellan-Villacian, M. Mendoza-Villena *et al.*, "Short-term net load forecast in distribution networks with PV penetration behind the meter," *Energy Reports*, vol. 9, no. 3, pp. 115-122, May 2023.

[93] C. Thepprom, N. Nupairoj, and P. Vateekul, "A deep learning framework for net load forecasting with unsupervised behind-the-meter disaggregated data," *IEEE Access*, vol. 12, pp. 94958-94971, Jul. 2024.

[94] A. N. Ndife, D. Steen, and L. A. Tuan, "Scalable and lightweight machine learning based load forecast: netload versus disaggregated forecast," in *Proceedings of IEEE PES Innovative Smart Grid Technologies Europe*, Dubrovnik, Croatia, Oct. 2024, pp. 1-5.

[95] B. Huang, V. Krishnan, and B. M. Hodge, "Analyzing the impacts of variable renewable resources on California net-load ramp events," in *Proceedings of IEEE PES General Meeting*, Portland, USA, Aug. 2018, pp. 1-5.

[96] Z. Qu, X. Ge, J. Lu *et al.*, "Unsupervised disaggregation of aggregated net load considering behind-the-meter PV based on virtual PV sample construction," *Applied Energy*, vol. 381, p. 125007, Mar. 2025.
[97] X. Wang, M. Zhang, and J. Su, "Robust temporal alignment for multivariate time series forecasting," *Expert Systems with Applications,* vol. 289, p. 128299, Sept. 2025.
[98] S. A. Alasadi and W. S. Bhaya, "Review of data preprocessing techniques in data mining," *Journal of Engineering and Applied Sciences*, vol. 12, no. 16, pp. 4102-4107, Sept. 2017.
[99] B. Frenay and M. Verleysen, "Classification in the presence of label noise: a survey," *IEEE Transactions on Neural Networks and Learning Systems*, vol. 25, no. 5, pp. 845-869, May 2014.
[100] D. Karimi, H. Dou, S. K. Warfield *et al.*, "Deep learning with noisy labels: Exploring techniques and remedies in medical image analysis," *Medical Image Analysis*, vol. 65, p. 101759, Oct. 2020
[101] R. R. Chandra, S. Goyal, and R. Gupta, "Evaluation of deep learning models for multi-step ahead time series prediction," *IEEE Access*, vol. 9, pp. 83105-83123, May 2021.
[102] S. Gupta and A. Gupta, "Dealing with noise problem in machine learning data-sets: a systematic review," *Procedia Computer Science*, vol. 161, pp. 466-474, Jan. 2019.
[103] A. S. Alsalehy and M. Bailey, "Improving time series data quality: identifying outliers and handling missing values in a multilocation gas and weather dataset," *Smart Cities*, vol. 8, no. 82, pp. 1-39, May 2025.
[104] H. Akhavan-Hejazi and H. Mohsenian-Rad, "Power systems big data analytics: an assessment of paradigm shift barriers and prospects," *Energy Reports*, vol. 4, pp. 91-100, Nov. 2018.
[105] S. García, J. Luengo, and F. Herrera, *Data Preprocessing in Data Mining.* Cham: Springer International Publishing, 2015.
[106] R. Kern, T. Al-Ubaidi, V. Sabol *et al.*, "Astro- and geoinformatics-visually guided classification of time series data," in *Knowledge Discovery in Big Data from Astronomy and Earth Observation*. Amsterdam: Elsevier, 2020.
[107] N. Amjady, "Short-term electricity price forecasting," in *Electric Power Systems: Advanced Forecasting Techniques and Optimal Generation Scheduling*. Boca Raton: CRC Press, 2017.
[108] B. Bischl, M. Binder, M. Lang *et al.*, "Hyperparameter optimization: foundations, algorithms, best practices, and open challenges," *Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery*, vol. 13, no. 2, p. 1484, Jan. 2023.
[109] L. Li, K. Jamieson, G. DeSalvo *et al.*, "Hyperband: a novel bandit-based approach to hyperparameter optimization," *Journal of Machine Learning Research*, vol. 18, no. 185, pp. 1-52, Apr. 2018.
[110] Z. S. Kadhim, H. S. Abdullah, and K. I. Ghathwan, "Artificial neural network hyperparameters optimization: a survey," *International Journal of Online and Biomedical Engineering*, vol. 18, no. 15, pp. 59-87, Dec. 2022.
[111] A. R. M. Rom, N. Jamil, and S. Ibrahim, "Multi objective hyperparameter tuning via random search on deep learning models," *Telkomnika (Telecommunication Computing, Electronics and Control)*, vol. 22, no. 4, pp. 956-968, Aug. 2024.
[112] M. Bhagat and B. Bakariya, "A comprehensive review of cross-validation techniques in machine learning," *International Journal of Science and Technology*, vol. 16, no. 1, pp. 1-4, Jan. 2025.
[113] F. V. Scheidt, X. Dong, A. Bartos *et al.*, "Probabilistic forecasting of household loads: effects of distributed energy technologies on forecast quality," in *Proceedings of ACM International Conference on Future Energy Systems*, Banff, Canada, Jun. 2021, pp. 231-238.
[114] P. M. Fonte, C. Monteiro, and F. M. Barbosa, "Net load forecasting in presence of renewable power curtailment," in *Proceedings of 13th International Conference on the European Energy Market*, Porto, Portugal, Jul. 2016, pp. 1-5.
[115] E. Bokhari and L. Hubert, "The lack of cross-validation can lead to inflated results and spurious conclusions: a re-analysis of the Macarthur violence risk assessment study," *Journal of Classification*, vol. 35, no. 1, pp. 147-171, Mar. 2018.
[116] G. Gürses-Tran, T. A. Körner, and A. Monti, "Introducing explainability in sequence-to-sequence learning for short-term load forecasting," *Electric Power Systems Research*, vol. 212, p. 108366, Nov. 2022.
[117] K. Bhattacharjee, S. Kundu, I. Chakraborty *et al.*, "Who should I trust? A visual analytics approach for comparing net load forecasting models," in *Proceedings of IEEE PES Grid Edge Technologies Conference & Exposition*, San Diego, USA, Jan. 2025, pp. 1-5.
[118] P. Eigenschink, T. Reutterer, S. Vamosi *et al.*, "Deep generative models for synthetic data: a survey," *IEEE Access*, vol. 11, pp. 47304-47320, May 2023.
[119] Y. Huang, J. Pei, L. Chen *et al*. (2025, Jun.). Probabilistic net load forecasting for high-penetration RES grids utilizing enhanced conditional diffusion model. [Online]. Available: https://arxiv.org/abs/2503.17770
[120] V. H. Rondón-Cordero, L. Montuori, M. Alcázar-Ortega *et al.*, "Advancements in hybrid and ensemble ML models for energy consumption forecasting: results and challenges of their applications," *Renewable and Sustainable Energy Reviews*, vol. 224, p. 116095, Dec. 2025.
[121] J. S. Telle, A. Upadhaya, P. Schönfeldt *et al.*, "Probabilistic net load forecasting framework for application in distributed integrated renewable energy systems," *Energy Reports*, vol. 11, pp. 2535-2553, Jun. 2024.
[122] C. Chupong, N. Junhuathon, S. Dangeam *et al.*, "Comparison of deep learning and incremental learning model for net load forecasting," *ECTI Transactions on Electrical Engineering, Electronics & Communications*, vol. 22, no. 1, pp. 1-12, Feb. 2024.
[123] L. Cheng, M. Li, C. Tan *et al.*, "Computational game-theoretic models for adaptive urban energy systems: a comprehensive review of algorithms, strategies, and engineering applications," *Archives of Computational Methods in Engineering*, vol. 2025, pp. 1-78, Sept. 2025.
[124] L. Cheng, P. Peng, W. Lu *et al.*, "The evolutionary game equilibrium theory on power market bidding involving renewable energy companies," *International Journal of Electrical Power & Energy Systems*, vol. 167, p. 110588, Jun. 2025.
[125] Energy Systems Integration Group. (2025, Aug.). Long-term load and DER forecasting. [Online]. Available: https://www.esig.energy/wp-content/uploads/2025/08/ESIG-Long-Term-Load-DER-Forecasting-report-2025.pdf
[126] N. K. Gupta, V. Upadhyaya, and V. Gali, "An overview of predictive maintenance and load forecasting," in *Intelligent Techniques for Predictive Data Analytics*. Piscataway: Wiley-IEEE Press, 2024.

Nusraat Nawreen received the B.Sc. degree in electrical and electronic engineering from Ahsanullah University of Science and Technology, Dhaka, Bangladesh, in 2021. She is currently pursuing the Ph.D. degree with the Center for New Energy Transition Research (CfNETR), Federation University Australia, Mt Helen, Australia. Her research interests include net demand forecasting, data mining, artificial intelligence, renewable energy system, and energy management and planning.

Nima Amjady is a Professor of Renewable Energy and the Director of the Centre for New Energy Transition Research (CfNETR) at Federation University Australia, Mt Helen, Australia. Before joining Federation University Australia, he was a Professor at Semnan University, Semnan, Iran, where he was selected as the youngest Professor of Iran, in 2007. He is Senior Editor of IEEE Transactions on Sustainable Energy. He is also the Editorial Board Member for IEEE Wiley Book Series. He received the Best paper award of the IEEE Transactions on Power Systems in 2018. His research interests include artificial intelligence and its application in renewable energy, as well as forecasting, operation, and planning of renewable energy-integrated power system.

Rakibuzzaman Shah received the M. Eng degree from the Asian Institute of Technology, Bangkok, Thailand and the Ph.D. degree from the University of Queensland, Brisbane, Australia. He is an Associate Professor in smart power system engineering at Federation University Australia, Mt Helen, Australia. His research interests include future power grid, asynchronous grid connection through voltage source converter based high-voltage direct current (VSC-HVDC), power system stability and dynamics, application of data mining in power system, application of control theory in power system, and distribution system.

S. M. Muyeen received the B.Sc. degree from Rajshahi University of Engineering and Technology (RUET), Rajshahi, Bangladesh, in 2000, and M. Eng. and Ph.D. degrees from Kitami Institute of Technology, Kitami, Japan, in 2005 and 2008, respectively, all in electrical and electronic engineering. He is a Full Professor in the Electrical Engineering Department of Qatar University, Doha, Qatar. He served as Editor/Associate Editor for many prestigious Journals including Journal of Modern Power Systems and Clean Energy, IEEE Transactions on Energy Conversion. He is a Fellow of IEEE, Chartered Professional Engineers, Australia, and a Fellow of Engineers Australia. His research interests include power system stability and control, electrical machine, flexible AC transmission system (FACTS), energy storage system, renewable energy, and high-voltage direct current (HVDC) system.