Machine Learning and Metaheuristic Methods for Renewable Power Forecasting: A Recent Review

Hanin Alkabbani, Ali Ahmadian, Qinqin Zhu and Ali Elkamel*

Department of Chemical Engineering, University of Waterloo, Waterloo, ON, Canada

The global trend toward a green sustainable future encouraged the penetration of renewable energies into the electricity sector to satisfy various demands of the market. Successful and steady integrations of renewables into the microgrids necessitate building reliable, accurate wind and solar power forecasters adopting these renewables' stochastic behaviors. In a few reported literature studies, machine learning- (ML-) based forecasters have been widely utilized for wind power and solar power forecasting with promising and accurate results. The objective of this article is to provide a critical systematic review of existing wind power and solar power ML forecasters, namely artificial neural networks (ANNs), recurrent neural networks (RNNs), support vector machines (SVMs), and extreme learning machines (ELMs). In addition, special attention is paid to metaheuristics accompanied by these ML models. Detailed comparisons of the different ML methodologies and the metaheuristic techniques are performed. The significant drawn-out findings from the reviewed papers are also summarized based on the forecasting targets and horizons in tables. Finally, challenges and future directions for research on the ML solar and wind prediction methods are presented. This review can guide scientists and engineers in analyzing and selecting the appropriate prediction approaches based on the different circumstances and applications.

Keywords: machine learning, artificial neural networks, forecasting, renewable energy, support vector machine, extreme learning machine, metaheuristics

INTRODUCTION

Motivation

In response to the environmental crisis and reducing greenhouse gas emissions, governments and policymakers promoted the penetration of renewable energies into the electricity production sector. According to the report of the international renewable energy agency, the contribution of renewable energy resources to electricity generation is projected to reach 85% by 2050, which is mainly due to the growth of solar-produced and wind-produced power (Global energy transformation: A roadmap to 2050, 2019). Although renewables are highly efficient, pollutant free, and inexpensive to produce and distribute, they lack consistency. Unlike the capability of generating conventional resources (coal and fossil fuels) according to the consumption and at specific and accurate schedules, the production of renewable energies is variable and renewable energies rely on seasonal and weather conditions (e.g., temperature, pressure, wind speed, visibility, etc.; Zerrahn et al., 2018).

These chaotic conditions can change dramatically from time to time, which enforce difficulties in the scheduling and management of optimal electricity generation and impose concerns regarding electricity quality and stability (Zerrahn et al., 2018). In fact, if the integration of renewable energy into the electricity sector is not handled and controlled adequately, it could cause an imbalanced and excess power production, which may increase the expenses of government instead of reducing the expenses (Lara-Fanego et al., 2012). Moreover, this unpredictable stochastic nature of renewables resulted in serious unit commitment issues (Chakraborty et al., 2012). Therefore, an accurate prediction of renewables has become an enduringly worldwide interest in a few literature studies.

Thus far, various research studies have been employed to tackle the problems of unreliability and inaccurateness of renewable power forecasting models. These include persistence models, physical models, statistical models, artificial intelligence (AI) models, and hybrid models consisting of a combination of two or more of these models. Lately, among these forecasting methodologies, AI-based models, particularly machine learning (ML) models, have gained the interest of researchers. Unlike statistical models, ML techniques can capture the non-linearity in power data. They can be applied for several purposes with only minor modifications. Therefore, because of their flexibility and compatibility, ML forecasters could outperform and alternate the conventional forecasters (Chakraborty et al., 2012; Santhosh and Venkaiah, 2019).

Nevertheless, despite the flourishing of studies related to proposed ML-based forecasting models, a review that summarizes forecasting models of these renewables and analytically evaluates their performance from a categorization perspective has not been investigated yet.

Contribution

This paper provides a comprehensive review of the recently published and proposed wind power and solar power forecasting ML-based models. In comparison to the existing studies on the same topic, the contributions of this paper are:

(1) A broad review of ML-based renewable power prediction methodologies and the metaheuristic optimizers of these methodologies is, for the first time, performed from a categorization viewpoint. Categorization, in this paper, is achieved by systematically allocating the ML prediction approaches and optimizers based on their similarities and differences and on the type of forecasted renewable energy. This will provide an analytical review of the current ML renewable power forecasting studies based on the approach and the sorting of renewable energy (wind or solar).

(2) Comparative evaluations of the ML-based renewable prediction methods and their metaheuristic optimizers are carried out. The drawn-out results would help other scholars to decide on the appropriate ML predictors and metaheuristic optimizers for various forecasting situations and purposes.

(3) Highlighting the challenges of ML applications for renewable power forecasting and providing the key directions that would guide other scholars to focus on the potential issues that have not been resolved yet.

In summary, this paper analyzes the renewable power prediction using the ML tools and their optimizers, emphasizes their weaknesses and strengths, and underlines the challenges accompanied by them to direct researchers on the issues that have not been settled yet.

AN OVERVIEW OF RENEWABLE POWER FORECASTING

In this section, recent structures in a few literature studies for renewable power forecasting are reviewed. Various schemes and methodologies plus AI tools are discussed and described.

The term renewable power generally encompasses all types of power gathered and generated from carbon-free renewable resources, such as wind, sunlight, rainfall, and waves. Specifically, wind energy and solar energy are fluctuating resources because their production rates depend on intermittent, unpredictable weather conditions (wind speed and direction and solar irradiation, respectively). Thereby, the renewable power forecasting studies consider either forecasting the generated wind and solar power or the wind speed and solar irradiation that are responsible for this produced power. From that perspective, this paper will review the forecasting methodologies of renewable power, including wind speed and/or output power and solar irradiance and/or output power.

Forecasting Methodologies

Figure 1 illustrates the differences between these horizons and their electricity sector applications (Santhosh and Venkaiah, 2019). With the inclusion of AI approaches, researchers proposed different forecasting structures using various methods and from multiple perspectives. These approaches and related research work for the forecasting of renewables (mainly wind power and solar power) are reviewed in the following sections.

Persistence Methodologies

These methodologies simply assume that the values of power data in a next time step are similar to those values in the current time step. Although these methodologies are not very practical for long-term forecasting, they perform well in very short-term and short-term forecasting (from a few seconds to 6 h-ahead; Nielsen et al., 1998).

Physical Methodologies

In addition to geographical locations and physical characteristics and layouts of wind turbines or solar panels, these methodologies depend on numerical weather predictions (NWPs), such as temperature, pressure, wind speed, wind density, roughness, turbulence intensity, etc. (Lei et al., 2009). Although these methodologies are reliable for medium-term and long-term forecasting, they cannot perform accurately for short-term forecasting (Giebel et al., 2011). In addition, they fail to adopt interferences and are computationally expensive, and require advanced computing machines (Murata et al., 2018). Giebel et al.

FIGURE 1 | Forecasting time horizons (Santhosh and Venkaiah, 2019).
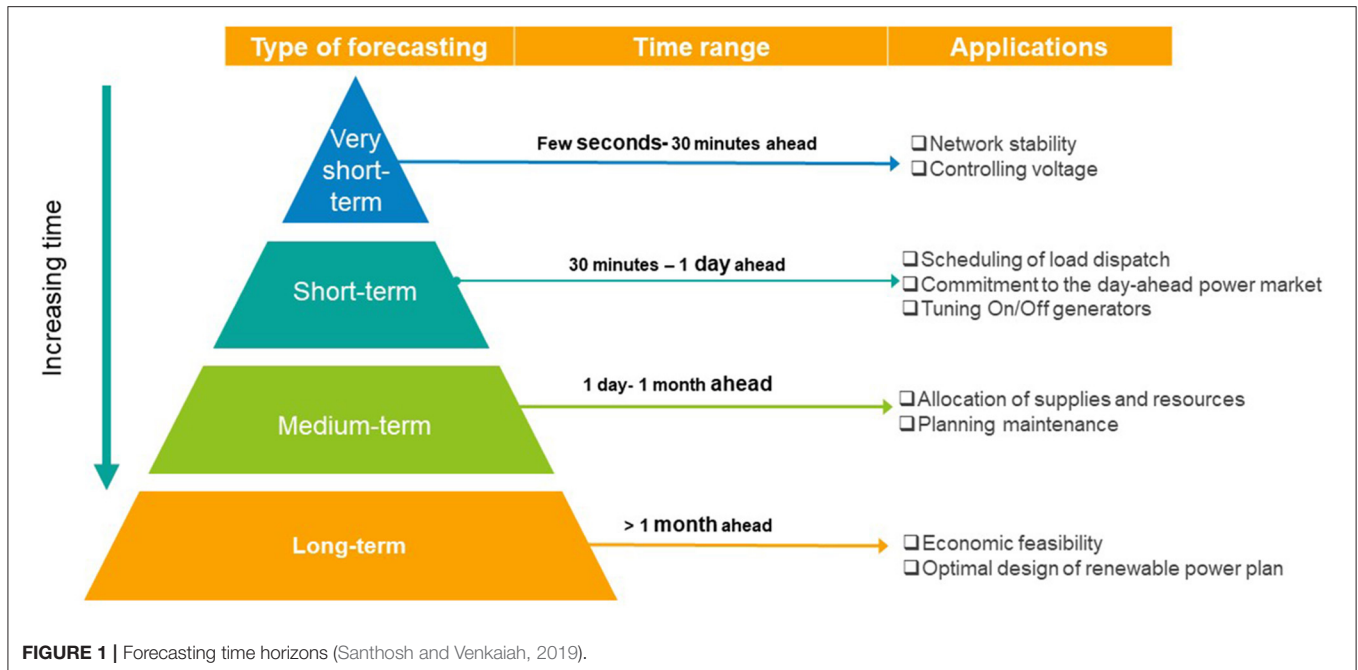
(2018) comprehensively reviewed the published studies tackling short-term forecasting using NWP models. This study was the last to summarize the applications of NWP-based models because the implication of these models was no longer attractive for researchers, and many other recent methodologies started to flourish and outperform physical models (Lei et al., 2009).

Statistical Methodologies

Statistical-based forecasting models are the mathematical models that attempt to map and recognize the relationship between time series historical data and target outputs (Ahmed and Khalid, 2019). They can clearly describe the linear relationship of data with basic simple mathematical equations (Santhosh and Venkaiah, 2019). Furthermore, since they can be formulated easily, they can deliver timely predictions. Thus, in a few literature studies, these forecasters are mainly used for short-term forecasting (Ezzat et al., 2018).

A comprehensive literature review on statistical approaches for time series and renewable energy forecasting was presented by Ghofrani and Alolayan (2018). Autoregressive (AR) and moving average (MA) models are well-known examples of statistical forecasting systems (Jiang et al., 2018). The hybrid integration of these two techniques is known as the autoregressive moving average (ARMA). ARMA is widely used for forecasting and provides models with a high accuracy for different applications. Erdem and Shi (2011) compared four other ARMA-based models for the forecasting of wind speed and direction. Gomes and Castro (2012) presented a comparative study between ARMA and artificial neural networks (ANNs) for the prediction of wind speed and power. They concluded that both approaches provide the similar results; however, the ARMA performance is slightly better. In Fentis et al. (2019), a non-linear autoregressive (NAR) model was suggested for the forecasting of short-term

photovoltaic (PV) power utilizing only historical data of the PV power (without using the NWP data). When comparing the performance of an NAR model with the Auto-regressive with an exogenous input (ARX) model, it was determined that NAR gives better results than ARX. This conclusion contrasts with the result obtained by Bacher et al. (2009) where ARX performed better.

Another robust approach known as auto-regressive integrated moving average (ARIMA) is widely employed for different purposes in a few literature studies to date. For example, Atique et al. (2019) used the ARIMA approach to predict the daily solar energy production. It is noted that the application of ARIMA models requires the utilized data to be stationary; therefore, in their work, the non-static seasonal data are transformed into stationary ones. For longer-term forecasting, Pasari and Shah (2020) used the ARIMA model for 1-year ahead forecasting of wind speed and temperature. According to their conclusion, this generated model is generic and having some minor modifications such as increasing the size of the input data, this model can be applied for 2-year ahead forecasting.

A particular parsimonious type of ARIMA, known as fractional-ARIMA, was studied by Kavasseri and Seetharaman (2009) for wind forecasting. Fractional-ARIMA is computationally simple and can capture time series relations for both long-term and short-term forecasting horizons. In this paper, this model was employed for forecasting an hourly wind speed and up to 2 days ahead. The results were promising and showed that this simple model could improve the forecasting accuracy by 42% compared to persistence models of forecasting.

In general, statistical models are considered attractive to researchers to date because they are inexpensive and straightforward to apply. They presented acceptable accurate results for short-term horizons up to 2 days; however, they fail in forecasting and result in very unstable predictions for

longer-term horizons (Ezzat et al., 2018). In addition, they require preprocessing of time series data (mostly when the data are discontinuous) for reliable performance and for providing accurate prediction models. This preprocessing could cause issues and requires expensive computation machines. Thus, researchers started to use a hybrid combination of these statistical models with AI methods to resolve the preprocessing issues.

Regression Methodologies

This type of model aims to find the best mathematical representation that relates independent variables (generally NWP and some physical properties and operation conditions of the turbines or solar panels) to the dependent variables (wind or solar power) through curve-fitting hyperparameter optimization techniques. Multilinear regression models are the simplest case of regression where the forecast variable is related to the predictors by a simple linear relationship (Ahmed and Khalid, 2019). For example, Abuella and Chowdhury (2015b) utilized multilinear regression to build a solar power probabilistic forecasting model. In addition, a simple linear quantile regression was used by Lauret et al. (2017) to create three different probabilistic models within the day (1–6 h ahead) solar irradiation prediction. For building the three models, the authors utilized historical data of solar irradiance as endogenous inputs and the day-ahead NWP of irradiance as exogenous inputs. The obtained results showed that the presence of NWP as exogenous inputs improved the prediction results. However, similar to the results obtained by Abuella and Chowdhury (2015b), a comparative study of the performance of a model in two different sites showed that the probabilistic models are highly dependent on the regional sky conditions. In a study by Massidda and Marrocu (2017), a multilinear adaptive regression spline method was used with a small size of training samples and a limited number of features to define a model for day-ahead solar power forecasting. This proposed regression model used historical power output data and weather forecasts.

Wang et al. (2016) proposed a novel partial functional linear regression (PFLR) model to forecast the daily output energy of a PV system. PFLR is similar to a multilinear regression but it can also represent a non-linearity structure in solar power data. Unlike statistical models focusing on utilizing historical data and underestimating the importance of the renewables data within the day pattern, PFLR incorporates the intra-day pattern of data and extracts valuable information from them. This work showed that this novel model that involves a few parameter estimates was able to outperform the ANN models and the regular multilinear regression. Another regression technique, known as multitasking Gaussian process regression (MTGP), was used by Cai et al. (2020) as a post-processing step to improve the NWP of wind speed. This additional step tackled the unreliable predictions that yield from the NWP when the behavior of the wind speed data is very complex and intermittent. The MTGP technique in this paper improved the forecasting accuracy of long-term forecasters and shorter-term forecasters. This improvement of NWP resulted in superior prediction results compared to the statistical predictors that are well-known for their accuracy for short-term forecasting. Keshtegar et al. (2018) performed a comparative study to compare four different

heuristic regression techniques, including Kriging, response surface method (RSM), multivariate adaptive regression (MARS), and M5 model tree (M5 Tree) for solar irradiation modeling. Comparative results showed that Kriging executed a better performance in comparison to the other three methods.

Overall, although regression methodologies of forecasting are simple and performed promisingly in some applications, they lack generalization and highly depend on the input data. They require too many explanatory variables to increase the accuracy of their predictions (Akhter et al., 2019). Moreover, the linear regression models assume a linear relationship between the independent and dependent variables; this assumption highly limits the application of these models for renewable power forecasting.

ML FORECASTING METHODOLOGIES

Artificial intelligence is a subfield of computer science; in AI, intelligent machines or artifacts are designed and trained to function like humans by following specific commands in computer programming systems. AI-based forecasting models accelerate decision-making, data mining, and clustering problems because they can robustly handle big data fitting and develop good representations. In addition, they can employ complex tasks with moderately short time and without being explicitly programmed. Thereby, AI forecasters have been used for various prediction applications in different areas of engineering, medicine, economy, and agriculture (Mellit and Kalogirou, 2008). Thus, the focus on proposing AI-based forecasting models grew a lot in the past few years and even started to alternate the conventional known predictors (Mellit et al., 2009).
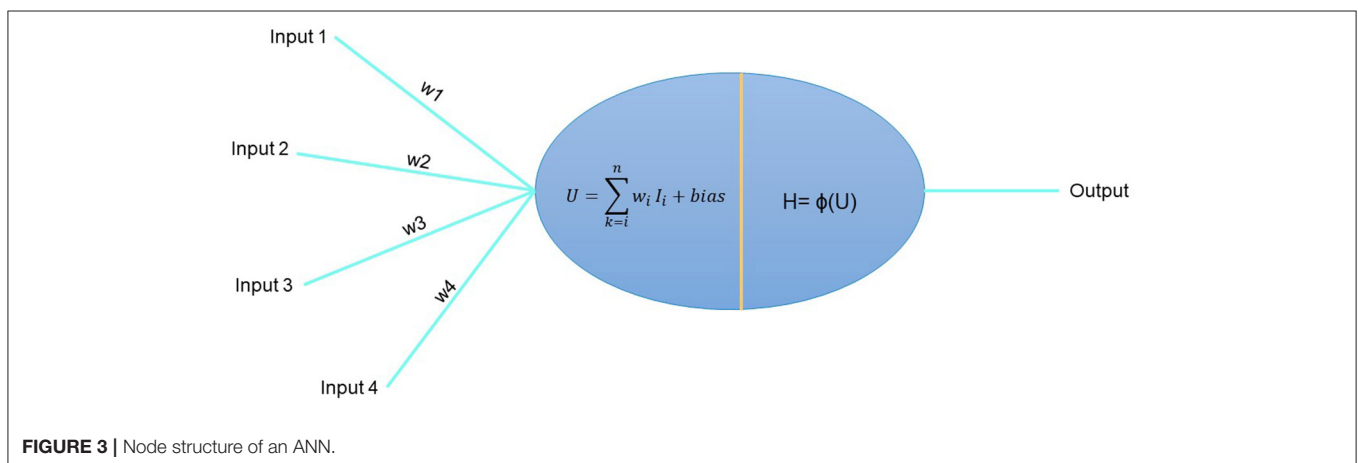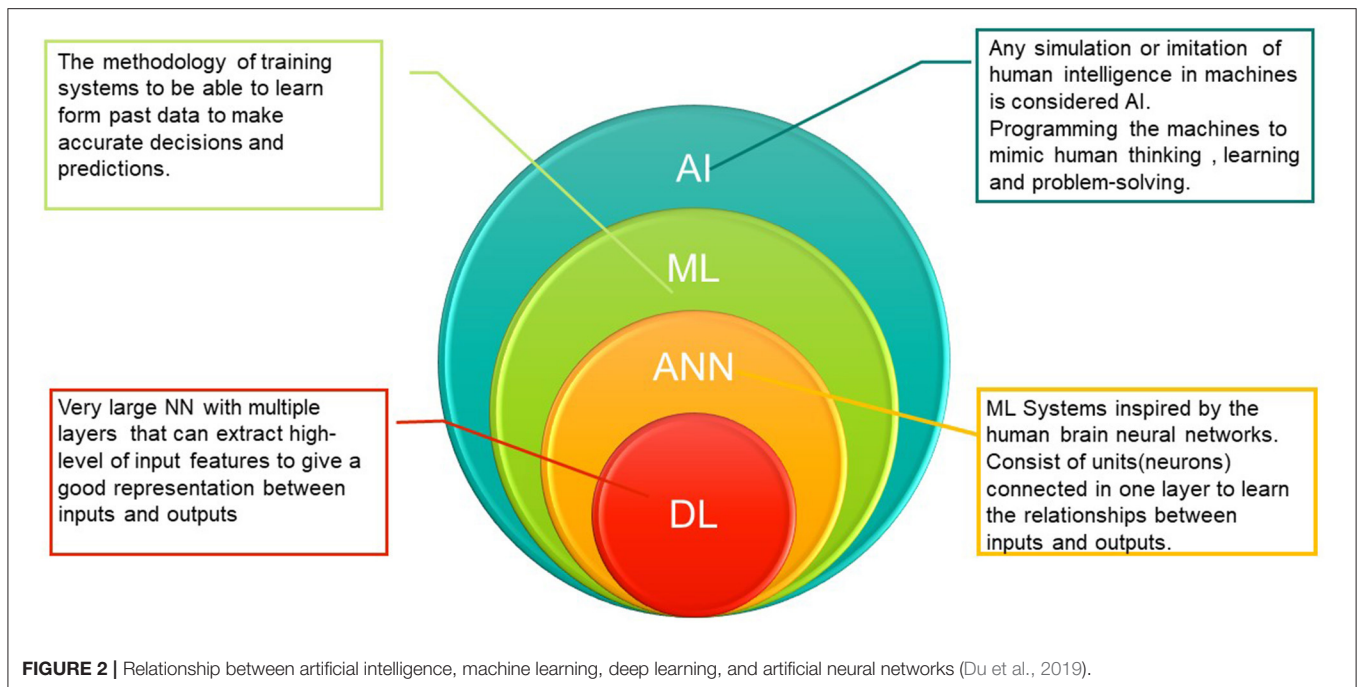
ML, ANN, and deep learning (DL) all are the subsets of AI. **Figure 2** illustrates the differences and relationships between these subsets (Sindhu and Nivedha, 2020). The following sections will review the recent research routes of renewable power forecasting (both wind power and solar power) based on the used ML algorithms for forecasting.

ML is an approach for data analysis, which gives computer systems the power to learn from data through experience. Unlike statistical-based models, ML techniques can generally capture the non-linearity and adapt instability in data, resulting in more reliable predictors (Jiang et al., 2018). Therefore, in the past few decades, ML tools were employed for forecasting various problems, such as renewable energy forecasting.

According to our survey, ANN, recurrent neural network (RNN), support vector machine (SVM), and extreme learning machine (ELM) are the most used ML techniques for renewable energy forecasting. Section ANN-Based Methodologies will review the work of some of the researchers in utilizing ML approaches for wind power and solar forecasting.

ANN-Based Methodologies

All types of ANNs have layers of neurons: the input layer is a layer where the network receives the input features and each neuron in this layer takes an input feature. The output layer is a layer where the final targets are estimated. The hidden layer, a connection between the input layer and the output layer, in which most of the

FIGURE 2 | Relationship between artificial intelligence, machine learning, deep learning, and artificial neural networks (Du et al., 2019).

The methodology of training systems to be able to learn form past data to make accurate decisions and predictions.

Any simulation or imitation of human intelligence in machines is considered AI.
Programming the machines to mimic human thinking , learning and problem-solving.

Very large NN with multiple layers that can extract high-level of input features to give a good representation between inputs and outputs

ML Systems inspired by the human brain neural networks. Consist of units(neurons) connected in one layer to learn the relationships between inputs and outputs.

FIGURE 3 | Node structure of an ANN.

Input 1

Input 2

Input 3

Input 4

$w1$

$w2$

$w3$

$w4$

$$U = \sum_{k=i}^{n} w_i I_i + bias$$

$H = \phi(U)$

Output

required computational operations occur; **Figure 3** represents a structure of the node of hidden layers. As shown in **Figure 4**, the outputs of the nodes of an ANN are determined by passing the input features multiplied by their corresponding weights to an activation function in the nodes of the hidden layer. There are several types of ANNs, **Figure 4** represents the classical structure of an ANN. In this section, the applications of ANNs for wind power and solar power forecasting are reviewed.

ANN for Wind Power Forecasting

A systematic literature review for wind power forecasting by Maldonado-Correa et al. (2019) confirmed that ANNs are considered the most frequently applied intelligence models in the literature studies for wind power forecasting in the past 5 years. These networks provided adequate results because of their ability to capture non-linearity in wind patterns, especially

for short-term and medium-term forecasting (Du et al., 2019; Maldonado-Correa et al., 2019). The simplest type of ANNs is the feed-forward NN (FFNN; Nielson et al., 2020), this network was used to predict a monthly energy production of 2.5 MW of a wind turbine. To train this network and increase forecasting accuracy, Nielson et al. (2020) selected wind speed and density incorporated with the atmospheric stability (represented in turbulence intensity, Richardson number, and wind shear) as input features to this network. This proposed approach reduced mean absolute error (MAE) of the wind power estimation by 59% in comparison to the standard estimation method.

On the other hand, Li and Shi (2010) compared the performance of Feed forward back propagation-ANN (FFBP-ANN) with other two ANNs, namely an adaptive linear element (ADALINE) NN and a radial basis function NN (RBF-NN) for wind speed forecasting. According to the evaluation

FIGURE 4 | Schematic diagram of ANN structure.
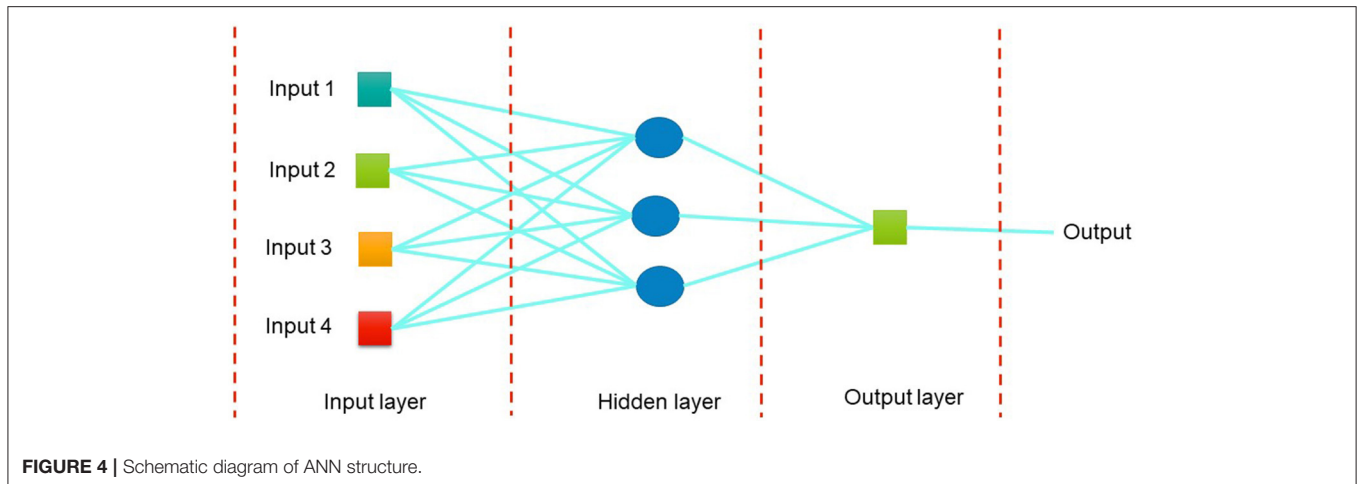
metrics, none of the three networks showed universally superior performance to the other networks. However, RBF-NN resulted in favorably accurate predictions when utilized in the hybrid wind forecasting model (Hong et al., 2019).

Although a logarithmic sigmoid function is the most commonly exploited transfer function, Grassi and Vecchio (2010) showed that building of an ANN with two hidden layers with two different activation functions, a hyperbolic tangent transfer function in the first hidden layer and a sigmoid transfer function in the second hidden layer could improve the wind energy prediction accuracy and exploits the features of data. It is essential to mention that, in this work, the monthly maintenance hours were used with metrological data as inputs to this ANN. Simulation results demonstrated that considering maintenance hours as an input improved the model reliability since they are inconsistent from month to month and directly affect the power production. Another work incorporated a differential polynomial function in ANN to build a wind speed correction model. The findings of this work illustrated that a differential polynomial function could model an existing complex system by solving and forming differential equations. On the other side, wavelet neural networks (WNNs) are also well-known powerful prediction tools when highly accurate predictions and fast convergence are needed (Bashir and El-Hawary, 2000). For instance, the sine activation function was incorporated with a rough concept to build a rough sinusoidal ANN (Jahangir et al., 2020). This work showed that a rough sinusoidal function handled the dramatic changes and the erratic stochastic behaviors in wind speed, especially at the peaks.

Besides integrating the rough concept to the ANN, the fuzzy concept also showed powerful, promising performance in wind prediction approaches. Although training an adaptive neuro-fuzzy inference system (ANFIS) is time consuming and considered complex, it is considered a universal estimator that lowers convergence errors (Marugán et al., 2018). For instance, Liu et al. (2017) employed a hybrid ANFIS approach for 48-h-ahead short-term wind power forecasting. This approach combines the predicted power by three different forecasters and outputs the final forecasted power. By a comprehensive

FIGURE 5 | RNN cell.
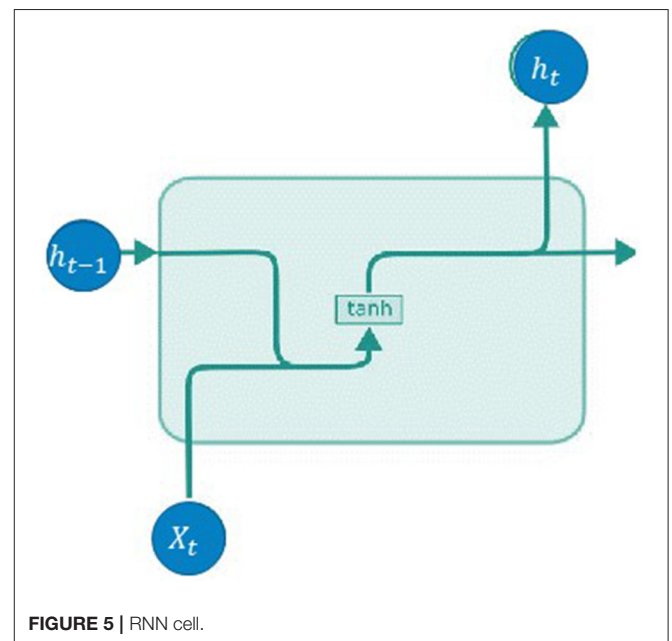
performance comparison between the hybrid proposed model and three individual forecasting models, namely RBF-NN, BPNN, and LSSVM, the authors demonstrated that their hybrid methodology has superior performance with respect to reliance and accuracy. In addition, unlike the three models that their accuracy differs from season to season, the ANFIS model significantly improved the forecasting data throughout the different seasons.

ANN for Solar Power Forecasting
Similar to wind forecasting, solar forecasting is widely achieved by the different types of ANN approaches. This section will review some proposed ANN-based methodologies for forecasting solar irradiation and power.

FIGURE 6 | Schematic diagram of RNN structure.
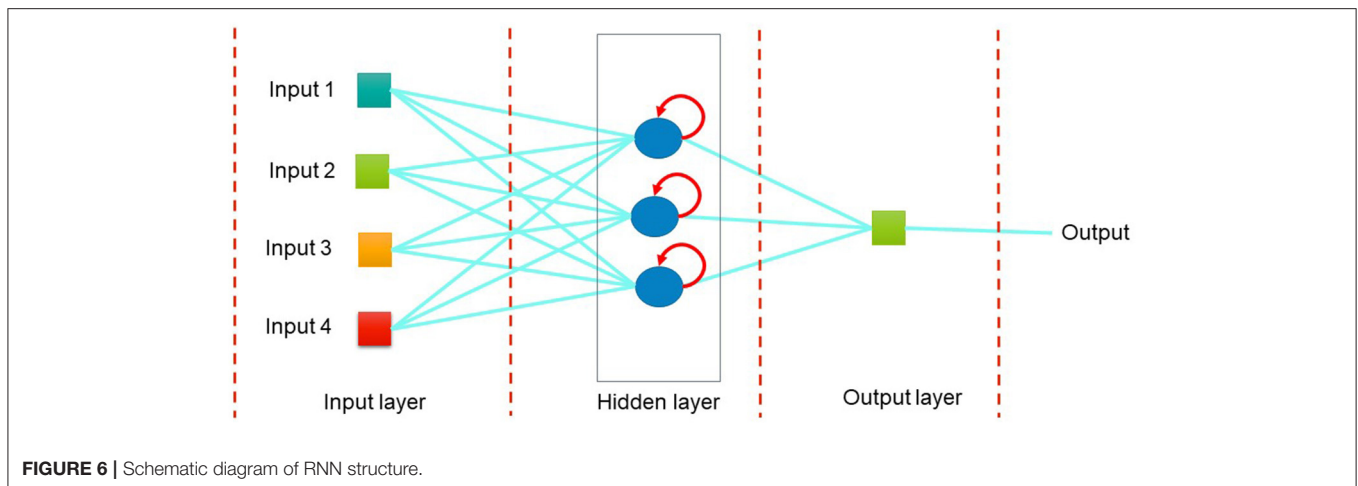
Abuella and Chowdhury (2015a) showed that the 14 input FFNNs could outperform a well-known multilinear regression methodology for hourly solar power forecasting. Despite the importance of the normalization of input data that is always discussed in the literature studies, the analysis of this paper shows that the normalized input data does not significantly improve the accuracy of the forecasted data. Nevertheless, their investigations revealed that data preparation and cleansing significantly affect the results and the ease of training the ANN. Moreover, the findings showed that eliminating the night hours from the input data could slightly improve the performance, and as expected, the predictions for clear sky hours and days were more reliable than cloudy and rainy days. To overcome the issue related to sky conditions for solar forecasting, O'Leary and Kubby (2017) suggest that the input masking technique is used based on the error clustering in the time domain. They categorized time frames into four categories [Night, Sunrise, Day: when solar energy is consistent (on sunny days), and Sunset]. Simulation results showed that input masking could improve the prediction outputs of the ANN by 1.3%. They suggest that the same input masking is performed for different environments and scenarios to confirm the importance of masking.

The correlation factor of the monthly prediction of solar energy improved by 9% in Ozoegwu (2019) when the ANN was hybridized into the NAR method. Besides improving accuracy, this hybridization reduced the size of inputs to the NAR approach, which saves memory. To guarantee and prove the prediction generalization, the method was simulated by using the data from various sites with different climates in Nigeria. In general, this model showed adequate results for longer-term forecasting, which is considered essential for planning and scheduling solar power applications.

With all of the proposed forecasting techniques in the literature studies, it became challenging to choose the most reliable prediction method. To discourse this issue, Yagli et al. (2019) raised an essential question on how to perform a fair comparison that reflects the actual superiority of models concerning the nature of data. This question was addressed by comparing 68 ML and statistical techniques for 1-h ahead global horizontal irradiance (GHI) forecasting, using the data from seven stations in five different climate zones in the USA. This finding of this work contributes to suggest the most appropriate prediction methodology for each specific climate zone.

Ghimire et al. (2019) reinforced a dramatic influence of feature selection of inputs on ML-based methodologies. They used a neighborhood component analysis to select the appropriate inputs from a pool of 85 different inputs in their work. Their selection was based on the regularization and minimization of a specific objective function that gives the most reliable daily solar irradiation forecasting results. The analysis results showed that evaporation rate, maximum air temperatures, albedo, cloud cover, relative humidity at maximum temperature, and specific humidity at 1,000 hPa are the inputs that resulted in more accurate predictions. Afterward, they compared the performance of different ML techniques, including SVM, process Gaussian, and ANN. According to statistical evaluation metrics, a feed-forward backpropagation ANN with Levenberg–Marquardt as a training function shows significantly superior performance in the five different sites in Queensland in Australia.

RNNs-Based Methodologies

Although the FFNN, as discussed earlier, is adequate for presenting the pattern that relates a specific output into a set of inputs, it learns a pattern of the outputs independently without having any context or memory of the previous outputs (Wang et al., 2020). To tackle this issue, RNN was introduced and used for time series forecasting. RNN is a subset of ANN, and it shows a robust performance when the order or a sequence of events or data matters and affects the following predictions (Su et al., 2019). Unlike the ANN, as shown in **Figure 5**, the RNN considers the features from the current time step inputs (xt) and the features from the previous hidden step (h_{t-1}). **Figure 6** shows a simple structure of RNN with respect to node connections where the hidden neurons take two input sets, one from the input layer and the other from the output of the hidden layer of the previous step. Holding and using information from the past time is considered a memory that relates the prior knowledge to the current one.

FIGURE 7 | Structure of GRU.

FIGURE 8 | structure of LSTM unit.
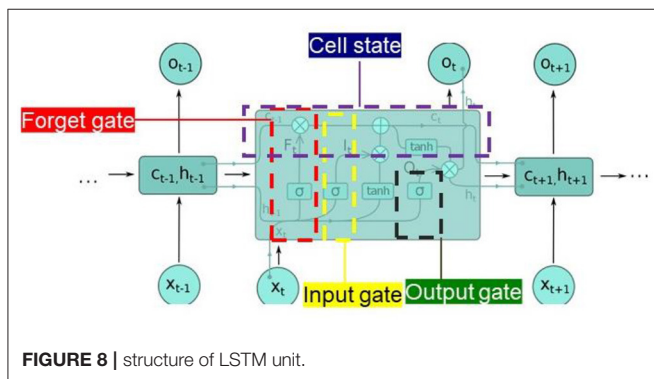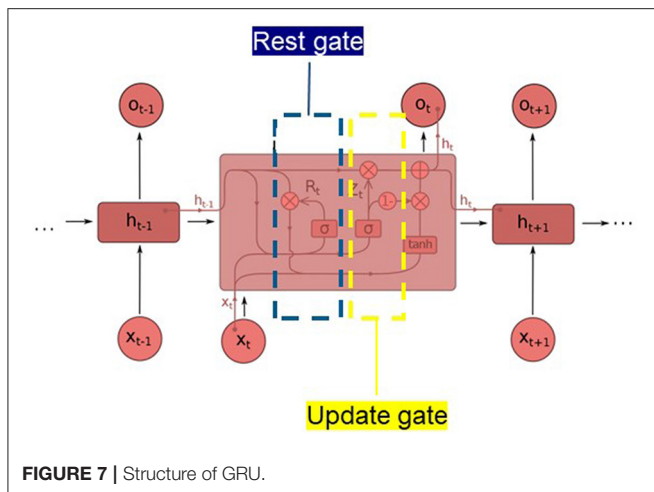
Nevertheless, RNN suffers from short-term memory, i.e., it cannot learn properly to preserve important information for long time sequences (Bianchini et al., 2013). Moreover, during the training process of RNN, the error gradient starts to exponentially fall until it vanishes, which interrupts the training process in the early stages (Kisvari et al., 2021). Two improved types of RNN nodes were proposed to overcome these issues, namely gated recurrent unit (GRU) and long short-term memory unit (LSTM). These two units have inner structures called gates that can control the contribution of information from the previous and current time steps. Using this, they pass significant attributes to long series sequences to predict and ignore unsignificant information (Bianchini et al., 2013). The GRU inputs are similar to the RNN ones; however, the mathematical operation that happens inside the GRU gates is slightly different. As shown in **Figure 7**, the structure of GRU includes two gates, the update and the rest gate. The update gate decides on what previously stored information to remove and what new information to add. While the rest gate decides how much of previous attributes to overlook and forget.

On the other hand, as illustrated in **Figure 8**, the LSTM has four different gates (forget gate, input gate, cell state, and output gate). The forget gate is similar to the update gate in the GRU. The input gate takes the same inputs as the forget gate and

processes them into sigmoid and tanh functions. The sigmoid function decides what information should be updated, and the tanh pounds the information between -1 and 1 to regulate the flow of information. The outputs of the sigmoid and tanh functions are then multiplied by each other to generate the output of an input gate. Afterward, the input gate outputs and the forget gate outputs are added to give a new cell state. Finally, the result passes to the output gate, which calculates the following hidden state.

The following section will review some published literature papers proposing wind power and solar power forecasting models utilizing RNN, including the regular one units and LSTM.

RNN for Wind Power Forecasting

Since a recursive structure of RNN can handle the complex non-linearity in time series wind data, RNN has been employed in various references to manage the forecasting of wind power and wind speed. This section will review the different classes of RNN proposed for wind forecasting.

Syu et al. (2020) performed an ultra-short-term (15-min ahead) wind speed forecasting utilizing the GRU network. To determine an optimal input size required for training the GRU models, various input sizes were used. The results showed a considerable drop in the MAE when the input is the previous 30 time steps. Nevertheless, the values of MAE and root mean square error (RMSE) start to fluctuate after the 30-input length. Thus, the 30 previous time steps were considered adequate for forecasting in this paper. Afterward, to validate the accuracy of a model, its performance was compared to the simple RNN and LSTM. Although LSTM was always known for its robust execution for time series forecasting, it did not perform better than the GRU approach. In fact, the GRU requires less parameter tuning and can be trained in a considerably shorter time. In addition, as expected, a simple RNN with the fastest training time performed poorly, especially at peaks where wind speed made a severe change.

Consequently, it is more reasonable to consider using GRU when both the performance and training time are essential for forecasting wind speed. In fact, also for wind power forecasting, the same results confirming that the implementation of GRU is similar to the LSTM with faster convergence and less tuning were obtained by Kisvari et al. (2021). To speed up the convergence, i.e., the training time of the LSTM, Yu et al. (2019) proposed an enhancement technique known as LSTM-enhancement forget gate (LSTM-EFG). In this approach, four modifications on the classical LSTM are performed: (1) two peepholes are added, (2) the tanh function is changed into soft sign, (3) the input gate is completely removed, and (4) the data update value is determined by subtracting the output of the forget gate from one matrix. These modifications directly affect the forget gate that, in its role, accelerates the convergence. It is also important to mention that in order to maximize the execution of an LSTM-EFG approach, a clustering technique combined with a temporal feature extraction methodology was incorporated into the system. The conclusions of the work verified the surpassing performance of the LSTM-EFG compared to the classical one and other benchmarking models.

Niu et al. (2020) claimed that the single wind power and speed predictions in some cases fail to be sufficient for electricity grid managing and scheduling. From this perspective, they suggested a multiple-input multiple-output (MIMO) model that forecasts wind power at different time horizons by a one-step simulation. In this model, an attention mechanism GRU coupled with a sequence-to-sequence technique is employed to select features. Unlike the classical feature selectors, which are applied once to discover the dependency of a target on the inputs, the attention mechanism estimates all inputs relevant to the target wind power outputs and creates weights representing these dependencies. Besides that, for each time step, hidden activations of the GRU blocks can extract both the spatial and temporal features, which contributes to improving the accuracy. Conclusions drawn from simulations confirmed that these two proposed strategies enhance the stability and accuracy of forecasting wind power simultaneously at different time horizons. In addition, the attention mechanism GRU lessened the error accumulation problem that was always coupled to the recursive models of forecasting. In general, this proposed model resulted in the competitive performance of the LSTM with faster convergence.

RNN for Solar Power Forecasting

Since the recursive construction of the RNN validated its ability to learn the patterns of time sequence data with seasonal and unstable trends, utilizing RNN for solar power/irradiance forecasting also recently attracted the interest of researchers (Yona et al., 2013). For instance, a comparative study by Aslam et al. (2019) was carried out to compare different methodologies for long-term solar radiation forecasting (1-year interval). The simple RNN network and the RNN with GRU and LSTM units proved their effectiveness in learning temporal dynamic behavior between the inputs and outputs for this case study. The comparison results showed that these methods could accurately generate highly accurate outcomes with low mean squared error (MSE) compared to the traditional forecasting techniques, i.e., random forest regression (RFR) and the conventional shallow FFNN. Another recurrent network known as Elman-RNN was trained by the cooperative neuro-evolution algorithm (Rana et al., 2016) to forecast the half-hourly PV power output. In this paper, the suggested approach considered both univariate and multivariate models. The evaluation results, as expected, highlighted the improvement of the accuracy when training a multivariate model and verified the effectiveness of the proposed model by comparing it to three different persistence forecasting methodologies. Internal memory in the Elman network that can deal with the variability of the PV data is considered as a direct result of this promising performance.

Hosseini et al. (2020) chose to utilize the recurrent networks, namely GRU and LSTM, to compare the univariate and multivariate approaches for direct normal irradiance hourly forecasting. They detected that computational-wise, GRU exhibited a better performance than the LSTM because LSTM is computationally time consuming with no significant superiority, especially for the multivariate approaches. In addition, to confirm the importance of incorporating wind speed and direction and the cloud coverage data to the input layer of networks, they trained the networks with and without these inputs and compared the accuracy of a model. The comparison results reinforced the significance and effectiveness of incorporating these inputs for irradiance forecasting where the accuracy increased by 23.32 and 8.91% for the simulations with the wind coverage and cloud coverage data, respectively.

Commonly, the metrological stations categorically report the daily sky condition without considering the variations from an area to area throughout the day. These data, when used for forecasting solar power, negatively affect the accuracy of the forecasters. Aiming to address this issue and increasing the reliance on solar power forecasters, Hossain and Mahmood (2020) proposed an LSTM-RNN-based approach as a forecasting step. In their model, after performing a statistical correlation analysis to choose the most suitable predictors for the LSTM, a k-mean analysis approach was used to tackle the sky type issue. In this approach, solar irradiance was dynamically clustered for each hour of the day according to the type of sky. These clusters create an hourly numerical approximation of the solar irradiances. Unlike classical sky type information represented for the entire day, the clustering technique makes an hourly synthetic weather forecast. These synthetic data are coupled with weather variables such as humidity, temperature, wind speed, and historical PV data fed as an input to the deep-LSTM. When constructing a comparison simulation between the LSTM with the proposed approach and the other two LSTM networks with the hourly and daily categorical sky type data, the findings verified the effectiveness of the proposed approach to increase the precision of forecasting. Finally, to verify the promising performance of LSTM, it was compared to a simple RNN, a generalized regression NN, and an ELM, all of which had the same synthetic input data. The LSTM followed by RNN outperformed the other two methodologies; this also supports the usefulness of utilizing recursive structures for forecasting.

SVM-Based Methodologies

Support vector machine is a powerful supervised ML technique based on a kernel-learning method that resolves the local minima issue that appears when training ANN (He and Xu, 2019). Through a kernel function in SVM, the input data sets are mapped into linear features with a higher-dimensional space. This data mapping gives the SVM the ability to capture the non-linearity in data and accurately predict erratic estimates such as wind power and solar power (He and Xu, 2019). In general, SVM is highly efficient in high dimensional spaces, comparatively memory effective, and resolves the local optimization problems in training ANN. However, in addition to its poor performance when the training data sets are relatively large, constrained optimization of SVM is computationally expensive. To overcome these drawbacks, a least-square-SVM (LSSVM) was recently introduced as a type of SVM with a loss function incorporating the SSE and transforming the inequality constraints to equality ones. This particular loss function of the LSSVM speeds up the training process and reduces the computational complexity of SVM (Huang et al., 1999). Consideration of the appropriate kernel function has a significant impact on the performance of both SVM and LSSVM. Linear kernel function, polynomial

kernel function, radial basis kernel function, and wavelet kernel function are the most commonly employed functions in assembling the SVM.

SVM for Wind Power Forecasting

As explained earlier, the choice of the proper kernel function and tuning its parameters is a significant key when employing SVM models for forecasting. Thus, He and Xu (2019) suggested a new kernel function that can be incorporated into the SVM by holding the advantages of SVM and at the same time improving its accuracy in forecasting. This hybrid kernel function is a combination of a wavelet kernel function and a polynomial kernel function. The authors claim that this combined kernel function will preserve the good local interpolation ability in the wavelet function and, at the time, improve its extrapolation by combining it with a polynomial function. The claim was verified by training the SVM for ultra-short-term wind speed forecasting using this integrated kernel function. The cross-validation technique was used for evaluating the performance, and the results showed that the hybrid function reduced the mean error by 3.94%. In this article, the dimensionality of the inputs to SVM was reduced by using the PCA approach, and the historical wind data were clustered according to their trend. This preprocessing step also contributed to improving the reliability of the proposed method. From a similar perspective, in a study conducted by Wang and Chen (2020), the density-based spatial clustering of applications with noise (DBSCAN) clustering technique was employed after the reduction of dimensionality before using the SVM for wind speed forecasting. This study also highlighted and reinforced the importance of clustering the data before implementing the SVM for forecasting, where the clustering decreased the MAE by 54%.

Because SVM models provide the generalization, utilizing SVM for wind speed and power forecasting attracted the interest of the scholars. Nevertheless, optimizing the performance and tuning the SVM parameters remains challenging, and no specific optimization algorithm has been highlighted to have superiority over the other. Accordingly, SVM models are accompanied mainly by various optimization techniques creating hybrid forecasting models. More SVM forecasting approaches hybridized with optimization procedures will be reviewed later in section Metaheuristic Optimization for Tuning ML Model Parameters.

SVM for Solar Power Forecasting

In general, SVM models can positively tolerate the noise and the volatility in data, and they can, in most cases, outperform the other ML techniques (Tabari et al., 2012). This superiority was also proven (Quej et al., 2017), where SVM was compared to ANFIS and ANN for estimating global solar irradiance in humid areas. The solar irradiance in damp locations is very chaotic and affected by the cloud coverage and the rainfall, and, in fact, it is not tackled enough in the literature. Therefore, the research conducted by Quej et al. (2017) incorporated the rainfall as an input to the three different ML techniques and tested their performance. As mentioned earlier, the results confirmed the superiority of SVM and illustrated the importance of considering the rain precipitation when forecasting the irradiance in such

humid areas. The performance of ANFIS and ANN was almost similar, and no considerable supremacy was investigated.

To reduce the uncertainty in PV power generation forecasting and maintain the appropriate unit commitment in power plants, Ahmad et al. (2020) suggested four different SVM forecasting models. Based on the seasons, four SVM models were trained to predict power generation and PV module parameters independently. Weather and PV power historical data were used as inputs to the SVM models. RBF kernel and the polynomial kernel were tested to determine a suitable kernel function for each model. According to accuracy, reported simulations in this paper and comparisons showed that the RBF kernel performs better for PV module parameters forecasting than the polynomial kernel. In contrast, the polynomial kernel resulted in lower MSE and MAE for PV power production forecasting. In fact, the work in this paper can provide beneficial guidance for future work related to the management and scheduling of PV power plant.

As explained earlier, the time horizon of forecasting can also affect the accuracy of ML models. For example, Hamamy and Omar (2019) applied LSSVM with RBF kernel to forecast the solar irradiance at different time horizons. Among the different inputs, they utilized sunshine durations and other weather data as inputs to build the models. The results showed that LSSVM performs better for short-term forecasting, and the accuracy of models decreases for longer-term forecasting. In fact, these results go along with the conclusions obtained by Liu et al. (2017), where the LSSVM did not result in an adequate model for 48-h ahead of forecasting. Malvoni and Hatziargyriou (2019) addressed the weakness of LSSVM in longer-term forecasting by hybridizing it with the three dimensional (3D) wavelet transform for a 24-h-ahead PV power forecast. Their proposed approach handled and reduced the high dimensionality of the inputs to the LSSVM and considered both the spatial and temporal features, which improved the long-term forecasting results.

ELM-Based Methodologies

Extreme learning machines are special types of single-layer FFNN that do not require the backpropagation algorithm for updating training and weights. Instead, the ELM uses the Moore–Penrose generalized inverse for estimating the target outputs (Akhter et al., 2019). Unlike the FFNN, this unique ELM structure reduces computational complexity and cuts the need for manually optimizing and tuning multiple parameters (Li et al., 2019). Nevertheless, since the loss function of ELM is based on second-order statistics, it fails to perform with the non-linear or non-Gaussian data. Most of the wind power-related and solar power-related forecasting models are built based on the chaotic and non-linear data. Therefore, an individual ELM approach for both wind power and solar power forecasting is limited in the literature studies. Generally, when the ELM models are used, an optimization algorithm or another forecasting technique is combined with ELM to improve the reliance of prediction models and increase their accuracy.

ELM for Wind Power Forecasting

To improve the ability of ELM in capturing the non-linear pattern in data and increase the accuracy of a forecasting

model, Li et al. (2019) proposed a wind power forecasting model based on ELM with a modified loss function. They incorporated kernel mean p-power error loss instead of the classical MAE loss function in ELM. When authors conducted comparative experiments, they concluded that, from a performance perspective, this adjustment in loss function improved the accuracy and provided reliable results compared to the classical ELM. Nevertheless, it resulted in losing an extreme computational speed of the ELM, which is considered a primary advantage when using the ELM. Therefore, as explained earlier, generally in the literature studies, to preserve the benefit of rapid learning in ELM and at the same time to generate reliable models, hybridizing the ELM with optimization algorithm is necessary and will be discussed later in section Metaheuristic Optimization for Tuning ML Model Parameters.

ELM for Solar Power Forecasting

Hossain et al. (2017) conducted a comparative study for the hourly and daily PV power forecasting of three different grids using various ML techniques. Solar radiation, wind speed ambience, module temperature, and PV power output data were used to train the models. RBF kernel SVM, sigmoid ANN trained with the Levenberg–Marquardt algorithm, and the ELM all were trained and evaluated. Reported experimental simulations illustrated that ELM could perform better for longer-term forecasting and has the highest learning speed than the other two ML techniques. Nevertheless, the authors highlighted that this ELM model could not adopt exogenous input data and suggested addressing this issue in future work.

METAHEURISTIC OPTIMIZED ML FORECASTING METHODOLOGIES

Generally, metaheuristic algorithms are implemented as a search guide to find the near-optimal approximate solutions that can improve the performance of specific systems with moderate computational costs (Cohoon et al., 2003). Based on the search strategy, metaheuristic algorithms are mainly classified into two main algorithm classes: (1) trajectory-based algorithms and (2) population-based algorithms. Commonly, population-based approaches are favorable for global optimization since they can adopt linear and non-linear, fixed and transitioned, and continuous and discrete objective functions. Thereby, our survey will focus on the population nature-based metaheuristic (namely, evolutionary and swarm) algorithms integrated into ML systems for renewable power forecasting.

According to what has been discussed and reviewed in the previous sections, it is clear that although single ML forecasters can be trained to forecast renewable power, in some cases, ML models are inadequate to fulfill the accuracy required for electricity sector applications. For example, these models can easily fall in the issues of optimal local values and fail to generate generalized forecasting models. In addition, determining the optimal structure of networks and tuning their parameters can be time consuming and requires an enormous number of trial-and-error experiments (Sindhu and Nivedha, 2020). Thereby, to build

computationally inexpensive effective ML networks and reliable prediction results, scholars supplemented various ML approaches and metaheuristic optimization techniques together.

The metaheuristics are used with ML networks for two different purposes: (1) tuning and estimating the model parameters during a training process and (2) tuning the hyperparameters related to the structure of a network (Yang and Shami, 2020). The difference between these two purposes and applications for optimizing them through a metaheuristic optimization will be considered in sections Metaheuristic Optimization for Tuning ML Model Parameters and Metaheuristic Optimization of the Network Parameters of the ML Systems.

METAHEURISTIC OPTIMIZATION FOR TUNING ML MODEL PARAMETERS

The optimization of the performance of ML forecasting methodologies is achieved through optimizing the model parameters. Weights, biases, and/or penalties of kennel functions all are examples of ML model parameters. The parameters of a model are related to the network's training approach and how the attributes of these networks change to increase the accuracy of fitting the targets (minimizing the cost function that evaluates the error between the fitted and actual targets). Since most of the ML optimization problems are non-convex, the choice of an unsuitable optimization approach for training the ML forecasting system could result in estimating the optimum local minimum parameters instead of the global (Yang and Shami, 2020). For example, the Gradient descent algorithm is the most frequently used algorithm for optimizing the parameters of ML models (Sun et al., 2019a). Nevertheless, if the objective function is non-convex, the gradient descent will fail to reach the globally optimum values in some cases (Sun et al., 2019a). Therefore, several metaheuristics were tested and incorporated for optimizing the parameters of ML systems.

The following section will review some applications of optimizing the parameters of ML models applied for renewable power forecasting using two algorithms, the first being evolutionary optimization algorithms and the second being swarm-based optimization algorithms.

Evolutionary Optimization for Tuning ML Model Parameters

Evolutionary optimization techniques utilize a population of solutions from the solution space to determine the approximate optimal solution (Alba, 2005). These techniques imitate the biological evolution in their working mechanism. Reproduction, mutation, recombination, and selection are the steps followed by evolutionary optimization approaches to determine the solutions. Their performance is considered suitable for various problems because no particular assumptions are made regarding fitness functions (Cohoon et al., 2003). Therefore, integrating evolutionary optimization into ML

approaches has been a research hotspot and attracted the attention of researchers.

Evolutionary Optimization Algorithms and ANN-Based Forecasting Methodologies

To forecast the monthly wind power generation of a wind plant in Iran, Jafarian-Namin et al. (2019) performed a comparative study to choose the most suitable adequate modeling methodology. In this study, weather conditions with sunshine hours and precipitations were considered as inputs to regular standalone ANN, hybrid ANN with genetic algorithm (GA) in one model, and particle swarm optimization (PSO) in another model to forecast the monthly wind power. Using the MATLAB software, the models were tested, and according to statistics metric evaluations, it was demonstrated that the hybridized ANN performs better than the individual ANN. The reported RMSE of the GA-ANN and PSO-ANN models was 0.4213 and 0.4250, respectively, whereas the RMSE for the regular ANN was 0.4385. Nevertheless, when applying the ARIMA model for prediction, it surpassed the ANN and the hybridized versions of it. This result supports the drawn-out conclusion from the previous sections that ML approaches perform poorly for longer-term forecastings, such as the case of this study. Pedro and Coimbra (2012) tested the GA-ANN approach for 1 h-ahead and 2 h-ahead average output power forecasting of a solar plant without incorporating any exogenous inputs. The GA-ANN system surpassed the other forecasting approaches, such as ARIMA, k-nearest neighbors (KNN), and the ANN. Nevertheless, similar to the drawn conclusion by Jafarian-Namin et al. (2019), a reduction in forecasting accuracy was reported when testing the model for longer forecasting horizons.

On the other hand, the study conducted by Flores et al. (2019) tested various forecasting approaches [including MLP-ANN, Nearest neighbor (NN), Fuzzy forecasting, Evolving Directed Acyclic Graph (EVOdag), and ARIMA] for 1-day-ahead wind speed forecasting. To improve the performance of the AI-based forecasters, the authors accompanied each technique with an evolutionary optimization approach. For example, MLP-ANN was accompanied by a compact GA (C-GA) for estimating the optimal size of inputs, the number of hidden neurons, and the appropriate training algorithm. The DE algorithm was used for estimating the optimum time lag, embedding dimension, and neighborhood radius size for the NN approach. Forecasting results for 20 different stations were compared and evaluated; the NN-DE system shows a superior prediction performance for most of the stations considered in the study.

Evolutionary Optimization Algorithms and SVM-Based Forecasting Methodologies

Tuning of the kernel parameters of the SVMs is one of the previously mentioned drawbacks of this robust ML approach. Various studies in the literature were conducted for tackling this issue, specifically through accompanying the SVM with metaheuristic approaches. For instance, Salcedo-Sanz et al. (2011) incorporated the Evolutionary Programming (EP) algorithm and a PSO approach into SVM wind speed forecasting systems in Spain. This incorporation aims to explore the kernel function hyperparameter that minimizes the prediction errors and increases the precision of forecasting. When comparing the results obtained from these incorporations, no specific superiority of one of the metaheuristics over the other was reported. However, both of them clearly increased the precision of prediction when compared to MLP prediction systems. These results not only highlight the importance of tuning the hyperparameters of SVMs but also support the previously discussed robustness of SVM methodologies. From a similar perspective of the superiority of SVM, Tian et al. (2020) conducted a study to recommend the most suitable optimization algorithm to estimate the optimal structure and parameters of LSSVM. GA, PSO, and brainstorm optimization algorithm (BSOA) all are individually supplemented by LSSVM, and their performances were evaluated and compared. The findings illustrated that the estimated MAE of the LSSVM-BSOA is significantly lower than the one for GA-LSSVM and PSO-LSSVM. When comparing the GA and PSO, MAE of GA is almost half the MAE for the PSO, and this could indicate that the evolutionary metaheuristics are superior to the swam ones when complemented by the LSSVM.

Similarly, in Wang et al. (2015), when the evolutionary cuckoo optimization algorithm (COA) was used for tuning the penalties factor and gamma of the kernel function RBF in SVM, it resulted in better predictions of wind speed when compared to the PSO-SVM approaches. This superiority can be observed mainly for bouncing samples. Furthermore, the authors concluded that this COA-SVM can actually perform promisingly for multistep-ahead forecasting and can be beneficial for wind station applications.

Evolutionary Optimization Algorithms and ELM-Based Forecasting Methodologies

As stated in section ELM-Based Methodologies, ELM is usually coupled with optimization approaches to perform better and result in reliable models. For example, weights and biases of ELM were trained by a newly devolved crisscross optimization (CSRO) algorithm for multistep wind speed forecasting (Yin et al., 2017). ELM was also integrated with PSO, GA, DE, and the forecasting results were compared. The CSO-ELM model was superior and had a higher ability in capturing the chaotic and non-linear behaviors of wind. In Zhang et al. (2017), ELM was trained for mean half-hour wind speed forecasting and incorporated its training process using a hybrid backtracking search algorithm (HBSA). Real-valued backtracking search algorithm (RBSA) was used to estimate the optimal weights and biases of the ELM, and the binary backtracking search algorithm (BBSA) was exploited as a feature selection after applying the PCA to choose the suitable input candidate features. It is essential to mention that, besides optimizing the inputs of the ELM and its structure, the authors also implemented optimized variational mode decomposition (OVMD) of the wind data as an additional denoising step. The HBSA-OVMD-ELM approach demonstrated a satisfactory execution compared to standalone ELM, SVM, HBSA-OVMDSVM, and other different combinations of the optimization algorithms and the ML approaches. This means that the advantageous computational speed of ELM was lost, and the balance between both speed and accuracy remains

a trade-off decision that engineers need to consider when employing such hybridized models. Another BSA application was used by Sun et al. (2019b) for tuning the parameters of different ML approaches for multistep wind power forecasting. The time series data were decomposed and fed to three ML forecasting engines (namely, ELM. LSSVM, and WNN). The BSA algorithm tuned the parameters of these ML models. Afterward, the optimal weights for combining these forecasters were also estimated by using the BSA to produce a reliable, highly accurate forecasting model.

Swarm-Based Optimization for Tuning ML Model Parameters

Inspired by the natural biological swarm movements, swarm-based optimization systems consist of agents cooperating with each other locally. By following simple rules, these agents search for an optimal solution among a set of possible solutions in a particular search space (Beni and Wang, 1993). Swarm-based metaheuristics have been utilized for optimizing the performances and outcomes in different applications in engineering, medicine, military, and economy (Martens et al., 2011).

Swarm Optimization Algorithms and ANN-Based Forecasting Methodologies

Aiming to improve the accuracy of a multistep wind energy prediction, Du et al. (2019) hybridized a WNN with different swarm optimization algorithms, including multi-objective moth-flame optimization (MOMFO) algorithm, multi-objective water cycle algorithm (MOWCA), multi-objective multi-verse optimization (MOMVO), and multi-objective whale optimization algorithm (MOWOA). The authors tested these four adopted algorithms for training a WNN using four data sets and compared their performance according to different evaluation metrics. According to the comparison results, the MOMFO surpassed the other algorithms and improved the precision and stability of prediction which other researchers sometimes ignore. The authors also claimed that this reliable forecasting approach could accurately predict in case of its utilization in different fields.

Swarm Optimization Algorithms and SVM-Based Forecasting Methodologies

Li et al. (2020) proposed a hybrid optimization approach for tuning the parameters of unimodal and multimodal functions to exploit both evolutionary and swarm optimization advantages. This hybrid optimization was performed by incorporating the differential evolutionary (DE) optimization into the dragonfly optimization (DA) algorithm. To validate the robustness of this hybrid optimization approach, five different kernel functions were tuned and tested in this study. Further validation was achieved by estimating the parameters of the SVM short-term wind power forecaster. The reported results showed that a DE optimization scheme positively boosted and supported the optimal search of the swarm-based optimization approach (i.e., the DA) compared to standalone DA and other optimization algorithms.

Another hybridization of the swarm optimization algorithm was proposed by Vinothkumar and Deeba (2020), where PSO was integrated into ant lion optimizer (ALO). In this work, to validate the robustness of this hybrid optimization technique, the authors used this technique to tune the kernel function of wavelet SVM and the parameters of LSTM-RNN for the wind speed prediction. The testing results showed that the proposed metaheuristic approach provided models approaching a minimal MAE% with a very reasonable computational time. The incorporation of the two techniques together was successful in finding the solution by various appropriate positional updates.

Swarm Optimization Algorithms and ELM-Based Forecasting Methodologies

To determine the most suitable swarm optimization algorithm for estimating the optimal parameters of ELM-forecasting PV power generation, a study conducted by Kumar et al. (2018) compared the performance of three ELM networks optimized by PSO, accelerate-PSO (APSO), and craziness-PSO (CRPSO). The APSO-ELM model quickly boosted the performance of ELM and provided considerably reliable, accurate forecasting results. Liu et al. (2019) compared the chicken swarm optimizer (CSO) and an improved chicken swarm optimizer (ICSO) for PV power short-term forecasting and concluded that ICSO is more efficient for optimizing the weights and biases of ELM and can outperform not only the CSO-ELM but also the benchmark approaches such as SVM.

METAHEURISTIC OPTIMIZATION OF THE NETWORK PARAMETERS OF THE ML SYSTEMS

The hyperparameters of ML networks are the variables that are set to construct a network structure. Tuning these parameters is essential since they can directly affect the performance of a training algorithm, which will eventually have a crucial control on the precision of the prediction model that is being learned and trained (Hutter et al., 2015). These parameters generally obtain the structures of networks (i.e., the numbers of units in the hidden layer, and the type of the activation functions) and weights and biases of the initializing schemes based on the selected activation function (Hutter et al., 2015). Unlike the learning-related parameters, the network's structure hyperparameter tuning is mainly favorable to be achieved through grid search, random search, and Bayesian optimization (Hutter et al., 2015).

The applications of metaheuristics for tuning of hyperparameters are not notably found in the literature studies. Only a few scholars reported their application of metaheuristics for tuning of the hyperparameters of networks. For example, for wind power forecasting systems, Jursa and Rohrig (2008) conducted a study to validate the importance of tuning the number of hidden neurons of a wind power ANN forecaster through the swarm and evolutionary optimization algorithms. This study defines the structure of an ANN prediction system by applying PSO and differential evolution algorithms (DEAs) through an automated selection approach. The proposed models

were tested for predicting the wind power of 10 different wind power stations in Germany. The reported results illustrated that the proposed automated approach through the two optimization approaches reduced the prediction error for most power stations compared to the manually tuned ANN forecasters. The PSO tuning approach enhanced the prediction by 9.6% and the DE approach by 6.8. Another application route of this ML metaheuristics integration can be found by Jahangir et al. (2020), the GA was employed to determine the optimal configuration of a stacked denoising autoencoder that was used as a pre-wind speed forecasting approach to denoise the data before processing them into forecasting network.

COMPARATIVE DISCUSSION FOR ML AND METAHEURISTIC METHODOLOGIES

ML techniques (ANN, RNN, SVM, and ELM) have been successfully utilized for renewable power forecasting. In many references, and according to statistical evaluation metrics such as MAE, MSE, RMSE, and R, those techniques confirmed their ability. They surpassed various traditional forecasting approaches, especially for short-term and medium-term forecasting.

Using simple structures, ANN can capture the non-linear and chaotic features in data and generate reliable, accurate predictions, especially for short-term and medium-term forecasting horizons (Pedro and Coimbra, 2012). BPFF-ANN is a robust ANN; it is known for its ability to map non-linear patterns usually found in solar and wind power. Nevertheless, this type sometimes fails to tolerate oscillations and can easily fall in the local minima (Sözen et al., 2004). In addition, it suffers from a low convergence rate (Ding et al., 2011). On the other hand, the RBF-ANN is usually introduced for renewable power forecasting problems because it is faster in learning, and it is not computationally expensive compared to the regular BP-ANN (Rana et al., 2016).

Nevertheless, for the different ANN types, several parameters, either related to the training process or the network structure, directly affect the reliability of models (Vinothkumar and Deeba, 2020). Tuning these parameters requires an integration of different optimization algorithms that are considered time consuming in some cases. Besides, sufficiently large historical data are needed to train the networks.

ANN-based models based on different time horizons and approaches to renewable power forecasting in the recent literature studies are summarized in **Table 1**.

RNNs are special types of ANN that can preserve and utilize the features from previous time steps, making them able to learn to attain the temporal relations between the data (Yona et al., 2013). Although RNNs can generate accurate forecasting models, short-memory problems associated with them cause immature training issues. GRU and LSTM are special nodes introduced to overcome the RNN drawbacks; these nodes process data in different mathematical activation functions to benefit from the attributes of the previous time steps with longer memory terms. They actively confirmed their superiority for

time series forecasting with moderately short training times. However, this recursive mechanism in all types of RNN results in error accumulation, which causes exploding gradient concerns that affect the training process of networks (Niu et al., 2020). **Table 2** summarizes some studies that mitigated these issues, particularly for renewable power forecasting, and provided consistent, accurate results.

SVM approaches are also powerful ML techniques that are well-known for their global approximation abilities. They can simplify complex mathematical computations, and unlike the ANN, they can learn patterns with a moderately small size of data sets with a little dependence on prior knowledge (Voyant et al., 2017). Nevertheless, their performance highly depends on the kernel function parameters, which require the incorporation of optimization algorithms for tuning and training (Liu et al., 2019). Their prediction stability diminishes for longer forecasting horizons when the training dataset is extensive (Hutter et al., 2015). Furthermore, the overfitting issue also comes with the SVM training process, necessitating different resolutions during a training process (Akhter et al., 2019).

The system optimization requirement also appears when utilizing the ELM tools to estimate the appropriate weights and biases (Malvoni and Hatziargyriou, 2019). Although the convergence is quickly achieved for ELM training, this convergence could be premature in some cases. Therefore, the created model fails to be generalized, and the precision of forecasting becomes insufficient in some cases. In fact, this encouraged consideration of DL concepts with ELM approaches (Mellit et al., 2009). **Table 3** summarizes some recent papers utilizing SVM and ELM tools for wind power and solar power forecasting.

The hybridized ML approaches with metaheuristic algorithms are recommended solutions to increase the reliability of ML models and resolve their limitations. The metaheuristic approaches are used to tune the parameters of an ML model and/or the structure of networks. Incorporating these metaheuristics aims to achieve adequate convergence, resulting in higher prediction accuracy than standalone ML methodologies.

Based on our investigations in this paper, the hybridized ML tools with metaheuristics resulted in predictions with high accuracies represented in the evaluation metrics such as MSE, MAPE, MSE, R/R^2, SAMPE, and RMSE. Population-based metaheuristics are usually favorable to be combined with the ML approaches because they are known for the ability to determine the optimal global variables for different types of objective functions. Evolutionary and swarm-based optimization techniques are subsets of the population-based optimizers preferable by scholars for optimizing the parameters of ML models. In addition, hybrid combinations of the evolutionary and swarm optimizers can also robustly determine near-optimal values surpassing standalone optimizers. Although the hybridization resulted in reducing the prediction error and improved the reliability of ML networks, the exaction time of these models is noticeably higher than the time needed for the individual ones and requires robust computation machines in some cases. **Table 4** summarizes some findings from the studies

TABLE 1 | Summary of artificial neural network (ANN) methods.

References	Forecasting target and horizon	ANN method	Important conclusions	Accuracy of the proposed model
Nielson et al. (2020)	Monthly wind power	FFBP-ANN	- The atmospheric effects on wind power curved could be explained by incorporating wind density into the input's features.	Improved MAE% by 59% compared to IEC*
Chen et al. (2019)	Hourly wind speed	FFBP-ANN ADALINE -NN RBF-ANN	- The time-series input size dramatically affects the performance of the ANN approaches. - Different learning rates resulted in considerably different interpretations.	RBF-ANN: MAE = 1.112 m/s
Grassi and Vecchio (2010)	Monthly generate wind power.	Two-layer FFNN with sigmoid and Tanh activation functions	- Considering maintenance hours as an input has a significant impact on increasing the forecasting reliability.	Training MAE = 0.0109 MWh
Liu et al. (2017)	48 h-ahead of wind power	ANFIS combining BPNN, LSSVM, RBANN	- Utilization of PCC for input selection contributes to improving accuracy. - Combining three ANN approaches as inputs to ANFIS ensures accurate results throughout the four seasons.	Spring: MAPE = 11.76% Fall: MAPE = 6.7%
Abuella and Chowdhury (2015a)	Hourly solar power	FF-ANN	- Input data pre-processing and clustering improve results. - Elimination of night hours results in better predictions.	Testing R^2 = 0.9665
O'Leary and Kubby (2017)	Hourly solar power	ANN	- Masking the ANN input data into specific categories according to the error rate was why the highly reliable drawn-out forecasting results.	MAE = 5.99 W
Ozoegwu (2019)	Monthly mean daily global solar radiation	Hybridization of non-linear autoregressive + ANN	- Forecasting results by the proposed model are satisfactory for longer-term forecasting horizons (up to 2 years-ahead) under different climate conditions.	R^2 = 0.92
Yagli et al. (2019)	Hourly global horizontal irradiance (GHI)	68 ML and statistical approaches	- This study can offer researchers' guidance in using the appropriate forecasting approach based on the climate conditions for GHI furcating.	N/A
Ghimire et al. (2019)	Daily global horizontal irradiance	FFBP-ANN	- Nearest component analysis incorporates in choosing the appropriate features that result in better predictions. - Integrating optimization techniques is recommended and would increase the reliability of ANN and ELM forecasting approaches.	R = 0.967

IEC, The International Electrotechnical Commission.

integrating metaheuristics to ML models for different renewable power forecasting targets and horizons.

On the other hand, tuning the hyperparameters related to the ML network structure is another challenge when using ML forecasters. This challenge is regularly tackled through search grid, random grid, and Bayesian optimization. It is, in some cases, a time-consuming process that some researchers prefer to depend on previous knowledge and experience tuning of these hyperparameters.

Finally, based on our investigations in this article, to improve the ML-based forecasting techniques, the following steps are usually recommended for more reliable, accurate ML forecasters:

1. Increasing the data set size is usually especially for ANN.
2. Preprocessing and analyzing the data to detect and filter the outliers and missing data are essential and affect the prediction results.
3. The presence of NWP as input features to the ML networks is crucial and improves the forecasting results.

4. Shorter-term forecasting horizons are preferable when using the ML techniques to ensure higher accuracies.
5. Hybridizing the ML models with optimization techniques improves the outcomes but might decelerate the training process in some cases; therefore, it remains a trade-off process that scholars need to consider based on the forecasting applications.
6. Hybridizing the ML approaches with metaheuristics improves the results of a multistep prediction.

Section Challenges and Future Directions will also highlight some vital challenges accompanied by utilizing ML forecasters for renewable power forecasting, to direct scholars to the problems that require higher focus and considerations in future studies.

CHALLENGES AND FUTURE DIRECTIONS

Even though many research studies are conducted on the ML forecasters of renewable power, some remaining significant questions and problems have not been efficiently tackled:

TABLE 2 | Summary of recurrent neural network (RNN) methodologies.

References	Forecasting targe and horizon	RNN method	Important conclusions	Accuracy of the proposed model
Syu et al. (2020)	Ultra-short-term wind speed Multi-step	GRU	- Number of previous steps as inputs to GRU is essential and affects forecasting results. - GRU requires less parameter tuning and shorter training time.	Step 1: MAE = 0.130 m/s Step 2: MAE = 0.222 m/s Step 3: MAE = 0.302 m/s
Yu et al. (2019)	Short-term wind power	LSTM- enhancement forget gate (LSTM-EFG)	- Proposed approach accelerates convergence. - Pre-processing of data represented in clustering and filtering noticeably improves prediction results.	Clustering increases the accuracy by 18.3%
Niu et al. (2020)	Wind power at different time horizons (MIMO) forecasting	Attention mechanism GRU	- The proposed model can continuously extract the spatial-temporal features in data, which boosts the forecasted power accuracy. - MIMO model offers stable forecasting results at different time horizons.	Step 1: MAPE = 4.35% Step 2: MAPE = 7.97% Step 3: MAPE = 11.54%
Aslam et al. (2019)	Long-term solar radiation forecasting (1-year interval)	Comparative study between RNN, LSTM, and GRU and other ML approaches.	- Reinforce the effectiveness of different RNN models in forecasting.	N/A
Rana et al. (2016)	Half-hourly PV power output	Elman RNN.	- Multivariate input model to Elman RRN is more reliable. - Elman RNN-based model can alternate the conventional persistence forecasting methodologies.	Univariate model: MAE = 90.95 kW
Hosseini et al. (2020)	Direct normal irradiance. At different horizons	GRU and LSTM (multivariate)	- Highlight the effectiveness of incorporating wind speed and cloud coverage as inputs to the forecasting networks. - GRU is adequate, with no superiority of LSTM.	The multivariate model outperforms the univariate model by 34.42%
Hossain and Mahmood (2020)	Short-term PV power. (6-12-24 h ahead)	LSTM	- Using k-means clustering to create numerical synthetic approximations of the sky type and incorporating them as inputs to the LSTM dramatically improved the forecasting results.	MAE Fall for 24 h: 0.36 MW

1. Minimal studies have been conducted for regional wind power or solar power forecasting; most studies consider single locations or stations. Regional electrical grid optimal scheduling and managing would be achieved by constructing a model that forecasts the solar or wind power for multiple locations in a specific region. Hence, constructing a precise regional wind power or solar power forecasting model is one of the critical problems to be tackled in the future.
2. Probabilistic prediction of wind energy and solar energy is not adequately considered in the literature studies. These predictions can quantify the changes in the resources of renewable energies. This could improve the scheduling of the electricity networks based on the estimated odd operating conditions. Therefore, focusing on probabilistic forecasting of renewables is a future key direction for researchers.
3. While the one-step-ahead forecasting has been extensively studied and tested, the multistep ahead forecasting proposed models remain a complex task that is not considered adequately in the literature studies and needs to become more encountered by researchers.
4. Currently, most of the published studies do not look at the problem of renewable power forecasting through a core structure of the ML model; the mathematical correlations between the input features and the renewable power prediction targets are not fully systematically disclosed and explained. Moreover, the input attributes that majorly affect the forecasting behaviors and precision are not entirely unambiguously indicated. In other words, the appropriate mathematical way to describe a renewable power forecasting model needs to be seen by scholars in the future.
5. From a forecasting horizon perspective, it was investigated that the proposed ML methodologies in the literature studies mainly focused on very short-term and short-term forecasting. Although these time horizons of forecasting have various vital applications related to maintaining the stability of the microgrid, medium-term and long-term forecasting horizons are also essential for studying the economic feasibility of the renewable power integration to the electricity sector. Thus, a higher focus on longer-term forecasting is expected and needed and could improve the incorporation of renewables into the electricity networks.

CONCLUSION

Precise forecasting of the renewables generation will maintain a stable production of electrical grids, avoid power waste and

TABLE 3 | Summary of support vector machine (SVM) and extreme learning machine (ELM) models.

References	Forecasting targe and horizon	SVM/ELM method	Important conclusions	Accuracy of the proposed model
He and Xu (2019)	Ultra-short term wind speed	SVM with wavelet + polynomial hybrid kernel function	- The SVM model's combined kernel function can outperform the single regular functions and improve the interpolation of the model's extrapolation ability. - Pre-processing of historical data and clustering also contribute to increased accuracy.	The hybrid function reduced the mean error by 3.94 %.
Tabari et al. (2012)	Short-term wind power	SVM with a linear kernel function.	- SVM models tackle the issue of local optimality that appears in other ML networks.	DBSCAN clustering reduced MAE by 54%.
Quej et al. (2017)	Daily global solar irradiance	SVM RBF-kernel function	- SVM with kernel function can tolerate noisy input data in humid areas and result in an accurate model with reasonable computational time. It outperforms the ANFIS and ANN models.	$RMSE = 3.089$ $MJm^{-2}d^{-1}$
Ahmad et al. (2020)	PV module and power generation parameters	Polynomial kernel SVM RBF kernel SVM	- Detailed SVM-based modeling study can offer helpful guidance for solar power-related problems.	Poly kernel: total RMSE $= 5.1674$ W BBFkernel: total RMSE $= 2.342$W
Hamamy and Omar (2019)	Solar irradiance at different time horizons.	LSSVM with RBF kernel	- LSSVM models perform better for short-term forecasting; the longer the forecasting term, the lower the predictions' accuracy.	Not given
Li et al. (2019)	Wind power	ELM with kernel mean P-power error loss function	- The ELM model with the special loss function can outperform the classical BP-ANN; nevertheless, the fast convergence of ELM is lost when the proposed loss function is introduced to ELM. Hybridizing ELM with optimization algorithms to estimate optimal weights and biases is effective and recommended.	$MAE = 255.4860$kW
Hossain et al. (2017)	Hourly and daily PV solar power	Classical ELM	- ELM can perform robustly for long-term forecasting with short training time.	Testing $R^2 = 0.8936$

intermittency, and respond to the environmental concerns regarding pollution and global warming because it will contribute to integrate environmentally friendly resources into the electricity sector. The conclusions drawn out from this paper illustrate that the ML techniques for wind power and solar power forecasting can outperform the traditional statistical tools, especially if the optimization algorithms accompany them.

First, a general review of renewable energy forecasting using mathematical prediction approaches was conducted. These reviews show that the persistence models to date are considered adequate and reliable for very short-term forecasting. However, researchers lost interest in the physical methodologies because they are computationally expensive and require complex mathematical operational computations. On the other hand, the statistical approaches are favorable in many forecasting cases as they are simple and do not require a massive preprocessing and filtering of data. ARIMA is a robust example of the statistical forecasting tools that showed a powerful performance for short and medium forecasting horizons (up to 2 days).

After that, detailed ML forecasting methodologies (ANN, RNN, SVM, and ELM) were reviewed and analyzed. These methodologies consider both the historical data and climate conditions for forecasting. It was concluded that the ANN models are adequate for non-linear systems and result in reliable predictions; however, their training process can fall

in local minima and/or overfitting and requires an enormous size of training input data. In addition, their ideal weights, biases, and structures (including the number of layers and nodes) require tuning and optimization, which might be time consuming in some cases. RNN, a special type of ANN, also showed a powerful performance because of its ability to preserve and capture the information from the previous time steps. Although this unique structure contributed to increasing accuracy, it can result in the buildup error issues that produce gradient vanishing problems and cause a failure to train the network and update its weights and biases.

Our survey also investigated that the SVM network can solve some of the issues that appear in the other ML methodologies by producing generalized, reliable models with reduced mathematical complexities. However, the overfitting problem also complements the SVM applications in some cases. In addition, a suitable associated kernel function in the SVM and optimizing their parameters remain the issue that needs various experiments and studies. Finally, the ELM tool that is known for its extremely fast convergence was also considered in our survey for renewables forecasting. The results showed that it is suitable for simple models only because it fails to capture enough features and learn adequately. It also requires either optimizing its initial parameters or extending it to become a deep network with multiple layers.

TABLE 4 | Summary of machine learning- (ML-) metaheuristic methods.

References	Forecasting target and horizon	ML method-metaheuristic method	Important conclusions	Accuracy of the proposed model
Jafarian-Namin et al. (2019)	Monthly wind power generation	ANN-GA ANN-PSO	- Hybridization with optimization algorithms increases accuracy - GA is superior to PSO in this case.	ANN-GA: R^2 = of 0.9212 GA reduced RMSE by 3.9% PSO reduced RMSE by 3.07%
Pedro and Coimbra (2012)	1 and 2 h-ahead solar power output	ANN-GA	- GA-ANN surpassed ARIMA and KNN approaches in improving the accuracy of forecasting	1 h ahead R^2 = 0.96 2 h ahead R^2 = 0.93 RMSE improved by; 32.3% for 1 h 35.1% for 2 h
Flores et al. (2019)	Day-ahead wind speed	NN-DE	- NN-DE approaches resulted in the highest accuracy predictions compared to MLP-ANN	SMAPE = 33.552 m/s
Salcedo-Sanz et al. (2011)	Short-term wind speed	SVM-EP SVM-PSO	- Proposed models surpassed the multi-layer perceptron forecasting approach	EP-SVMr: MAE = 1.7921 m/s PSO-SVMr:MAE = 1.7823 m/s
Tian et al. (2020)	Short-term wind speed	LSSVM-BSOA	- LSSVM-BSOA is superior to LSSVM-GA, LSSVM-PSO, and other standalone approaches.	MAE = 0.1113 m/s
Wang et al. (2015)	Short-term wind speed and multi-step ahead	SVM-COA SVM-PSO	- Implementation of the COA-SVR model is advanced to that of the PSO-SVR and GA-SVR methods,	1 step ahead: MAE = 0.6836 m/s 2 steps ahead: MAE = 1.0051 m/s
Yin et al. (2017)	Wind power forecasting and multi-step ahead	ELM-CSO	- CSO can address the premature convergence of ELM and surpass other algorithms.	1 step ahead: MAE=0.104 m/s 2 steps ahead: MAE = 0.157 m/s 3 steps ahead: MAE = 0.186 m/s
Zhang et al. (2017)	Wind speed a mean half-hour	HBSA-ELM	- Compared to other approaches, the proposed model has robust performance in capturing the non-linear attributes of wind speed	MAE = 0.372 m/s
Du et al. (2019)	Short term wind power multi-step ahead	WNN-MOMFO WNN-MOWCA WNN-MOMVO WNN- MOWOA	- Integration of the multi-objective optimizers improves the predictions compared to GNN and WNN	1 step ahead: MAPE = 5.0661% 2 steps ahead: MAPE = 7.7877% 3 steps ahead: MAPE = 10.6968%
Li et al. (2020)	Short-term wind power	SVM-DA-DEA	- The model showed a better performance compared to other models such as ANN	R^2 = 0.9791 for winter dataset R^2 = 0.9544 for autumn dataset.
Vinothkumar and Deeba (2020)	Short-term wind speed	SVM-ALO-PSO LSTM-ALO-PSO	- The effectiveness of the proposed optimizer integrated into the ML forecasters is proven comparing to other benchmarking forecasting approaches	ALO-PSO MAE = 0.0027 m/s ALO-LSTM = 0.0126 m/s
Kumar et al. (2018)	PV power generation	ELM-PSO ELM-APSO ELM-CRPSO	- ELM model surpassed BP-ANN - The hybridized ELM with different PSO approaches reduced the prediction error.	ELM: MAPE = 2.9417% PSO-ELM: MAPE = 2.7736% CRPSO-ELM: MAPE = 2.2207% APSO-ELM: MAPE = 1.440%
Liu et al. (2019)	Short-term PV power outputs	ICSO-ELM	- Improving the CSO considerably enhanced the forecasting ability of the ELM	MAPE = 5.54%

Afterward, to explore the effect of hybridizing the ML forecasting approaches with optimization techniques, we reviewed some hybridized systems in the literature studies to show how this hybridizing could boost the performance of ML approaches and ensure reliable, accurate forecasting. Finally, a comprehensive comparison between the reviewed models was conducted, and some crucial drawn-out observations highlighting the challenges and future trends related to the review are exemplified.

AUTHOR CONTRIBUTIONS

AA, QZ, and AE: conceptualization. HA, AA, QZ, and AE: methodology, validation, formal analysis, and investigation. HA and AA: software, data curation, and writing—original draft preparation. QZ and AE: resources, writing—review and editing, supervision, project administration, and funding acquisition. All authors have read and agreed to the published version of the manuscript.

REFERENCES

Abuella, M., and Chowdhury, B. (2015a). "Solar power forecasting using artificial neural networks," in *2015 North American Power Symposium, (NAPS)* (Charlotte, NC), 1–5. doi: 10.1109/NAPS.2015.7335176

Abuella, M., and Chowdhury, B. (2015b). "Solar power probabilistic forecasting by using multiple linear regression analysis," in *Conference Proceedings - IEEE SOUTHEASTCON* (Florida, FL), 1–5. doi: 10.1109/SECON.2015.7132869

Ahmad, A., Jin, Y., Zhu, C., Javed, I., and Akram, M. W. (2020). Support vector machine based prediction of photovoltaic module and power station parameters. *Int. J. Green Energy* 17, 219–232. doi: 10.1080/15435075.2020.1722131

Ahmed, A., and Khalid, M. (2019). A review on the selected applications of forecasting models in renewable power systems. *Renew. Sustain. Energy Rev.* 100, 9–21. doi: 10.1016/j.rser.2018.09.046

Akhter, M. N., Mekhilef, S., Mokhlis, H., and Shah, N. M. (2019). Review on forecasting photovoltaic power generation based on machine learning and metaheuristic techniques. *IET Renew. Power Generation* 13, 1009–1023. doi: 10.1049/iet-rpg.2018.5649

Alba, E. (2005). *Parallel Metaheuristics: A New Class of Algorithms*. Wiley. doi: 10.1002/0471739383

Aslam, M., Lee, J. M., Kim, H. S., Lee, S. J., and Hong, S. (2019). Deep learning models for long-term solar radiation forecasting considering microgrid installation: a comparative study. *Energies* 13:147. doi: 10.3390/en13010147

Atique, S., Noureen, S., Roy, V., Subburaj, V., Bayne, S., and MacFie, J. (2019). "Forecasting of total daily solar energy generation using ARIMA: a case study," in *2019 IEEE 9th Annual Computing and Communication Workshop and Conference, CCWC* (Las Vegas, NV), 114–119. doi: 10.1109/CCWC.2019.8666481

Bacher, P., Madsen, H., and Nielsen, H. A. (2009). Online short-term solar power forecasting. *Solar Energy* 83, 1772–1783. doi: 10.1016/j.solener.2009.05.016

Bashir, Z., and El-Hawary, M. E. (2000). "Short term load forecasting by using wavelet neural networks," in *Canadian Conference on Electrical and Computer Engineering*, Vol. 1 (Halifax, NS), 163–166. doi: 10.1109/CCECE.2000.849691

Beni, G., and Wang, J. (1993). "Swarm intelligence in cellular robotic systems," in *Robots and Biological Systems: Towards a New Bionics?* (Berlin; Heidelberg), 703–712. doi: 10.1007/978-3-642-58069-7_38

Bianchini, M., Maggini, M., and Jain, L. C. (2013). *Handbook on Neural Information Processing*. Berlin; Heidelberg, 29–65. doi: 10.1007/978-3-642-36657-4

Cai, H., Jia, X., Feng, J., Li, W., Hsu, Y. M., and Lee, J. (2020). Gaussian Process Regression for numerical wind speed prediction enhancement. *Renew. Energy* 146, 2112–2123. doi: 10.1016/j.renene.2019.08.018

Chakraborty, S., Senjyu, T., Saber, A. Y., Yona, A., and Funabashi, T. (2012). A fuzzy binary clustered particle swarm optimization strategy for thermal unit commitment problem with wind power integration. *IEEJ Trans. Electrical Electronic Eng.* 7, 478–486. doi: 10.1002/tee.21761

Chen, K. S., Lin, K. P., Yan, J. X., and Hsieh, W. L. (2019). Renewable power output forecasting using least-squares support vector regression and google data. *Sustainability* 11:3009. doi: 10.3390/su11113009

Cohoon, J., Kairo, J., and Lienig, J. (2003). "Evolutionary algorithms for the physical design of VLSI circuits," in *Advances in Evolutionary Computing*, eds A. Ghosh and S. Tsutsui (Berlin; Heidelberg: Springer), 683–711. doi: 10.1007/978-3-642-18965-4_27

Ding, M., Wang, L., and Bi, R. (2011). An ANN-based approach for forecasting the power output of photovoltaic system. *Proc. Environ. Sci.* 11, 1308–1315. doi: 10.1016/j.proenv.2011.12.196

Du, P., Wang, J., Yang, W., and Niu, T. (2019). A novel hybrid model for short-term wind power forecasting. *Appl. Soft Comp. J.* 80, 93–106. doi: 10.1016/j.asoc.2019.03.035

Erdem, E., and Shi, J. (2011). ARMA based approaches for forecasting the tuple of wind speed and direction. *Appl. Energy* 88, 1405–1414. doi: 10.1016/j.apenergy.2010.10.031

Ezzat, A. A., Jun, M., Ding, Y., and Member, S. (2018). Spatio-temporal asymmetry of local wind fields and its impact on short-term wind forecasting. *Transactions Sustain. Energy X* 9, 1437–1447. doi: 10.1109/TSTE.2018.2789685

Fentis, A., Bahatti, L., Tabaa, M., and Mestari, M. (2019). Short-term nonlinear autoregressive photovoltaic power forecasting using statistical learning approaches and in-situ observations. *Int. J. Energy Environ. Eng.* 10, 189–206. doi: 10.1007/s40095-018-0293-5

Flores, J. J., Cedeño González, J. R., Rodríguez, H., Graff, M., Lopez-Farias, R., and Calderon, F. (2019). Soft computing methods with phase space reconstruction for wind speed forecasting—a performance comparison. *Energies* 12, 1–19. doi: 10.3390/en12183545

Ghimire, S., Deo, R. C., Downs, N. J., and Raj, N. (2019). Global solar radiation prediction by ANN integrated with European Centre for medium range weather forecast fields in solar rich cities of Queensland Australia. *J. Clean. Prod.* 216, 288–310. doi: 10.1016/j.jclepro.2019.01.158

Ghofrani, M., and Alolayan, M. (2018). "Time series and renewable energy forecasting," in *Time Series Analysis and Applications*, 78–92. doi: 10.5772/intechopen.70845

Giebel, G., Brownsword, R., Kariniotakis, G., Denhard, M., and Draxl, C. (2011). *The State-Of-The-Art in Short-Term Prediction of Wind Power*. doi: 10.13140/RG.2.1.2581.4485

Giebel, G., Kariniotakis, G., and Brownsword, R. (2018). "The state-of-the-art in short term prediction of wind power from a danish perspective," in *4th International Workshop on Large Scale Integration of Wind Power and Transmission Networks for Offshore Wind Farms* (Billund). Available online at: https://hal-mines-paristech.archives-ouvertes.fr/hal-00529986

Global energy transformation: A roadmap to 2050 (2019). Available online at: https://www.irena.org/publications/2019/Apr/Global-energy-transformation-A-roadmap-to-2050-2019Edition (accessed August 17, 2020).

Gomes, P., and Castro, R. (2012). Wind speed and wind power forecasting using statistical models: AutoRegressive Moving Average (ARMA) and Artificial Neural Networks (ANN). *Int. J. Sustain. Energy Dev.* 1, 41–50. doi: 10.20533/ijsed.2046.3707.2012.0007

Grassi, G., and Vecchio, P. (2010). Wind energy prediction using a two-hidden layer neural network. *Commun. Nonlinear Sci. Numer. Simul.* 15, 2262–2266. doi: 10.1016/j.cnsns.2009.10.005

Hamamy, F., and Omar, A. M. (2019). Least square support vector machine technique for short term solar irradiance forecasting. *AIP Conf. Proc.* 39, 535–576. doi: 10.1063/1.5118141

He, J., and Xu, J. (2019). Ultra-short-term wind speed forecasting based on support vector machine with combined kernel function and similar data. *Eurasip J. Wireless Commun. Network.* 2019, 1–7. doi: 10.1186/s13638-019-1559-1

Hong, Y., Lian, C., and Rioflorido, P. P. (2019). A hybrid deep learning-based neural network for 24-h ahead wind power forecasting. *Appl. Energy* 250, 530–539. doi: 10.1016/j.apenergy.2019.05.044

Hossain, M., Mekhilef, S., Danesh, M., Olatomiwa, L., and Shamshirband, S. (2017). Application of extreme learning machine for short term output power forecasting of three grid-connected PV systems. *J. Clean. Prod.* 167, 395–405. doi: 10.1016/j.jclepro.2017.08.081

Hossain, M. S., and Mahmood, H. (2020). Short-term photovoltaic power forecasting using an LSTM neural network and synthetic weather forecast. *IEEE Access* 8, 172524–172533. doi: 10.1109/ACCESS.2020.3024901

Hosseini, M., Katragadda, S., Wojtkiewicz, J., Gottumukkala, R., Maida, A., and Chambers, T. L. (2020). Direct normal irradiance forecasting using multivariate gated recurrent units. *Energies* 13, 1–15. doi: 10.3390/en13153914

Huang, X., Maier, A., Hornegger, J., and Suykens, J. A. K. (1999). Indefinite kernels in least squares support vector machines and principal component analysis. *Appl. Comput. Harmon. Anal.* 43, 162–172. doi: 10.1016/j.acha.2016.09.001

Hutter, F., Lücke, J., and Schmidt-Thieme, L. (2015). Beyond Manual Tuning of Hyperparameters. *Kunstliche Intelligenz* 29, 329–337. doi: 10.1007/s13218-015-0381-0

Jafarian-Namin, S., Goli, A., Qolipour, M., Mostafaeipour, A., and Golmohammadi, A. M. (2019). Forecasting the wind power generation using Box–Jenkins and hybrid artificial intelligence: a case study. *Int. J. Energy Sector Manag.* 13, 1038–1062. doi: 10.1108/IJESM-06-2018-0002

Jahangir, H., Aliakbar, M., Alhameli, F., Mazouz, A., Ahmadian, A., and Elkamel, A. (2020). Short-term wind speed forecasting framework based on stacked denoising auto-encoders with rough ANN. *Sustain. Energy Technol. Assessments* 38:100601. doi: 10.1016/j.seta.2019.100601

Jiang, Y., Huang, G., Peng, X., Li, Y., and Yang, Q. (2018). Journal of wind engineering & industrial aerodynamics a novel wind speed prediction method :

hybrid of correlation-aided DWT, LSSVM and GARCH. *J. Wind Eng. Industrial Aerodynamics* 174, 28–38. doi: 10.1016/j.jweia.2017.12.019

Jursa, R., and Rohrig, K. (2008). Short-term wind power forecasting using evolutionary algorithms for the automated specification of artificial intelligence models. *Int. J. Forecast* 24, 694–709. doi: 10.1016/j.ijforecast.2008.08.007

Kavasseri, R. G., and Seetharaman, K. (2009). Day-ahead wind speed forecasting using f -ARIMA models. *Renew. Energy* 34, 1388–1393. doi: 10.1016/j.renene.2008.09.006

Keshtegar, B., Mert, C., and Kisi, O. (2018). Comparison of four heuristic regression techniques in solar radiation modeling: kriging method vs RSM, MARS and M5 model tree. *Renew. Sustain. Energy Rev.* 81, 330–341. doi: 10.1016/j.rser.2017.07.054

Kisvari, A., Lin, Z., and Liu, X. (2021). Wind power forecasting – a data-driven method along with gated recurrent neural network. *Renew. Energy* 163, 1895–1909. doi: 10.1016/j.renene.2020.10.119

Kumar, M., Majumder, I., and Nayak, N. (2018). Solar photovoltaic power forecasting using optimized modified extreme learning machine technique. *Eng. Sci. Technol. Int. J.* 21, 428-438. doi: 10.1016/j.jestch.2018.04.013

Lara-Fanego, V., J. A., Ruiz-Arias, D., and Pozo-Va'zquez, ⇑, F. J., Santos-Alamillos, J. (2012). Evaluation of the WRF model solar irradiance forecasts in Andalusia. *Solar Energy* 86, 2200–2217. doi: 10.1016/j.solener.2011.02.014

Lauret, P., David, M., and Pedro, H. T. C. (2017). Probabilistic solar forecasting using quantile regression models. *Energies* 10, 1–17. doi: 10.3390/en10101591

Lei, M., Shiyan, L., Chuanwen, J., Hongling, L., and Yan, Z. (2009). A review on the forecasting of wind speed and generated power. *Renew. Sustain. Energy Rev.* 13, 915–920. doi: 10.1016/j.rser.2008.02.002

Li, G., and Shi, J. (2010). On comparing three artificial neural networks for wind speed forecasting. *Appl. Energy* 87, 2313–2320. doi: 10.1016/j.apenergy.2009.12.013

Li, L. L., Zhao, X., Tseng, M. L., and Tan, R. R. (2020). Short-term wind power forecasting based on support vector machine with improved dragonfly algorithm. *J. Clean. Prod.* 242, 118447. doi: 10.1016/j.jclepro.2019.118447

Li, N., He, F., and Ma, W. (2019). Wind power prediction based on extreme learning machine with kernel mean p-power error loss. *Energies* 12, 1–19. doi: 10.3390/en12040673

Liu, J., Wang, X., and Lu, Y. (2017). A novel hybrid methodology for short-term wind power forecasting based on adaptive neuro-fuzzy inference system. *Renew. Energy* 103, 620–629. doi: 10.1016/j.renene.2016.10.074

Liu, Z., Li, L., Tseng, M., and Lim, M. K. (2019). Prediction short-term photovoltaic power using improved chicken swarm optimizer - Extreme learning machine model. *J. Clean. Prod.* 248:119272. doi: 10.1016/j.jclepro.2019.119272

Maldonado-Correa, J., Solano, J. C., and Rojas-Moncayo, M. (2019). Wind power forecasting: a systematic literature review. *Wind Engineering.* 2019, 1–14. doi: 10.1177/0309524X19891672

Malvoni, M., and Hatziargyriou, N. (2019). "One-day ahead PV power forecasts using 3D Wavelet Decomposition," in *SEST 2019 - 2nd International Conference on Smart Energy Systems and Technologies* (Porto), 1–6. doi: 10.1109/SEST.2019.8849007

Martens, D., Baesens, B., and Fawcett, T. (2011). Editorial survey: Swarm intelligence for data mining. *Mach. Learn.* 82, 1–42. doi: 10.1007/s10994-010-5216-5

Marugán, A. P., Márquez, F. P. G., Perez, J. M. P., and Ruiz-Hernández, D. (2018). A survey of artificial neural network in wind energy systems. *Appl. Energy* 228, 1822–1836. doi: 10.1016/j.apenergy.2018.07.084

Massidda, L., and Marrocu, M. (2017). Use of multilinear adaptive regression splines and numerical weather prediction to forecast the power output of a PV plant in Borkum, Germany. *Solar Energy* 146, 141–149. doi: 10.1016/j.solener.2017.02.007

Mellit, A., and Kalogirou, S. A. (2008). Artificial intelligence techniques for photovoltaic applications: a review. *Progress Energy Combustion Sci.* 34, 574–632. doi: 10.1016/j.pecs.2008.01.001

Mellit, A., Kalogirou, S. A., Hontoria, L., and Shaari, S. (2009). Artificial intelligence techniques for sizing photovoltaic systems: a review. *Renew. Sustain. Energy Rev.* 13, 406–419. doi: 10.1016/j.rser.2008.01.006

Murata, A., Ohtake, H., and Oozeki, T. (2018). Modeling of uncertainty of solar irradiance forecasts on numerical weather predictions with the estimation of multiple confidence intervals. *Renew. Energy* 117, 193–201. doi: 10.1016/j.renene.2017.10.043

Nielsen, T. S., Joensen, A., and Madsen, H. (1998). A new reference for wind power forecasting. *Wind Energy* 34, 29–34. doi: 10.1002/(SICI)1099-1824(199809)1:1<29::AID-WE10>3.0.CO;2-B

Nielson, J., Bhaganagar, K., Meka, R., and Alaeddini, A. (2020). Using atmospheric inputs for arti ficial neural networks to improve wind turbine power prediction. *Energy* 190:116273. doi: 10.1016/j.energy.2019.116273

Niu, Z., Yu, Z., Tang, W., Wu, Q., and Reformat, M. (2020). Wind power forecasting using attention-based gated recurrent unit network. *Energy* 196, 117081. doi: 10.1016/j.energy.2020.117081

O'Leary, D., and Kubby, J. (2017). Feature selection and ANN solar power prediction. *J. Renew. Energy* 2017, 1–7. doi: 10.1155/2017/2437387

Ozoegwu, C. G. (2019). Artificial neural network forecast of monthly mean daily global solar radiation of selected locations based on time series and month number. *J. Clean. Prod.* 216, 1–13. doi: 10.1016/j.jclepro.2019.01.096

Pasari, S., and Shah, A. (2020). *Time Series Auto-Regressive Integrated Moving Average Model for Renewable Energy Forecasting.* Pilani: Springer International Publishing. doi: 10.1007/978-3-030-44248-4_7

Pedro, H. T. C., and Coimbra, C. F. M. (2012). Assessment of forecasting techniques for solar power production with no exogenous inputs. *Solar Energy* 86, 2017–2028. doi: 10.1016/j.solener.2012.04.004

Quej, V. H., Almorox, J., Arnaldo, J. A., and Saito, L. (2017). ANFIS, SVM and ANN soft-computing techniques to estimate daily global solar radiation in a warm sub-humid environment. *J. Atmospheric Solar Terrestrial Phys.* 155, 62–70. doi: 10.1016/j.jastp.2017.02.002

Rana, M., Chandra, R., and Agelidis, V. G. (2016). "Cooperative neuro-evolutionary recurrent neural networks for solar power prediction," in *2016 IEEE Congress on Evolutionary Computation, CEC*, 4691–4698. doi: 10.1109/CEC.2016.7744389

Salcedo-Sanz, S., Ortiz-García, E. G., Pérez-Bellido, Á. M., Portilla-Figueras, A., and Prieto, L. (2011). Short term wind speed prediction based on evolutionary support vector regression algorithms. *Expert Syst. Appl.* 38, 4052–4057. doi: 10.1016/j.eswa.2010.09.067

Santhosh, M., and Venkaiah, C. (2019). Sustainable Energy, Grids and Networks Short-term wind speed forecasting approach using Ensemble Empirical Mode Decomposition and Deep Boltzmann Machine. *Sustain. Energy Grids Networks* 19:100242. doi: 10.1016/j.segan.2019.100242

Sindhu, V., and Nivedha, S. P. M. (2020). An empirical science research on bioinformatics in machine learning. *J. Mech. Continua Mathematical Sci.* 7, 86–94. doi: 10.26782/jmcms.spl.7/2020.02.00006

Sözen, A., Arcaklioglu, E., Özalp, M., and Kanit, E. G. (2004). Use of artificial neural networks for mapping of solar potential in Turkey. *Appl. Energy* 77, 273–286. doi: 10.1016/S0306-2619(03)00137-5

Su, H., Zio, E., Zhang, J., Xu, M., Li, X., and Zhang, Z. (2019). A hybrid hourly natural gas demand forecasting method based on the integration of wavelet transform and enhanced Deep-RNN model. *Energy* 178, 585–597. doi: 10.1016/j.energy.2019.04.167

Sun, S., Cao, Z., Zhu, H., and Zhao, J. (2019a). A survey of optimization methods from a machine learning perspective. *arXiv[Preprint].arXiv:1906.06821.*

Sun, S., Fu, J., and Li, A. (2019b). A compound wind power forecasting strategy based on clustering, two-stage decomposition, parameter optimization, and optimal combination of multiple machine learning approaches. *Energies* 12, 1–22. doi: 10.3390/en12183586

Syu, Y. D., Wang, J. C., Chou, C. Y., Lin, M. J., Liang, W. C., Wu, L. C., et al. (2020). "Ultra-short-term wind speed forecasting for wind power based on gated recurrent unit," in *2020 8th International Electrical Engineering Congress, iEECON* (Chiang Mai) 2020, 3–6. doi: 10.1109/iEECON48109.2020.229518

Tabari, H., Kisi, O., Ezani, A., and Hosseinzadeh Talaee, P. (2012). SVM, ANFIS, regression and climate based models for reference evapotranspiration modeling using limited climatic data in a semi-arid highland environment. *J. Hydrol.* 444–445, 78–89. doi: 10.1016/j.jhydrol.2012.04.007

Tian, Z., Ren, Y., and Wang, G. (2020). An application of backtracking search optimization–based least squares support vector machine for prediction of short–term wind speed. *Wind Engineering* 44, 266–281. doi: 10.1177/0309524X19849843

Vinothkumar, T., and Deeba, K. (2020). Hybrid wind speed prediction model based on recurrent long short-term memory neural network and support vector machine models. *Soft Comp.* 24, 5345–5355. doi: 10.1007/s00500-019-04292-w

Voyant, C., Notton, G., Kalogirou, S., Nivet, M., Paoli, C., Motte, F., et al. (2017). Machine learning methods for solar radiation forecasting : a review. *Renew. Energy* 105, 569–582. doi: 10.1016/j.renene.2016.12.095

Wang, G., Su, Y., and Shu, L. (2016). One-day-ahead daily power forecasting of photovoltaic systems based on partial functional linear regression models. *Renew. Energy* 96, 469–478. doi: 10.1016/j.renene.2016.04.089

Wang, H., Liu, Y., Zhou, B., Li, C., Cao, G., Voropai, N., et al. (2020). Taxonomy research of artificial intelligence for deterministic solar power forecasting. *Energy Convers. Manag.* 214:112909. doi: 10.1016/j.enconman.2020.112909

Wang, J., Zhou, Q., Jiang, H., and Hou, R. (2015). Short-term wind speed forecasting using support vector regression optimized by cuckoo optimization algorithm. *Math. Probl. Eng.* 11, 13–30. doi: 10.1155/2015/619178

Wang, S., and Chen, C. (2020). "Short-term wind power prediction based on DBSCAN clustering and support vector machine regression," in *2020 5th International Conference on Computer and Communication Systems, ICCCS* (Shanghai), 941–945. doi: 10.1109/ICCCS49078.2020.9118606

Yagli, G. M., Yang, D., and Srinivasan, D. (2019). Automatic hourly solar forecasting using machine learning models. *Renew. Sustain. Energy Rev.* 105, 487–498. doi: 10.1016/j.rser.2019.02.006

Yang, L., and Shami, A. (2020). On hyperparameter optimization of machine learning algorithms: theory and practice. *Neurocomputing* 415, 295–316. doi: 10.1016/j.neucom.2020.07.061

Yin, H., Dong, Z., Chen, Y., Ge, J., Lai, L. L., Vaccaro, A., et al. (2017). An effective secondary decomposition approach for wind power forecasting using extreme learning machine trained by crisscross optimization. *Energy Convers. Manag.* 150, 108–121. doi: 10.1016/j.enconman.2017.08.014

Yona, A., Senjyu, T., Funabashi, T., and Kim, C. H. (2013). Determination method of insolation prediction with fuzzy and applying neural network for long-term ahead PV power output correction. *IEEE Trans. Sustain. Energy* 4, 527–533. doi: 10.1109/TSTE.2013.22 46591

Yu, R., Gao, J., Yu, M., Lu, W., Xu, T., Zhao, M., et al. (2019). LSTM-EFG for wind power forecasting based on sequential correlation features. *Future Gener. Comput. Syst.* 93, 33–42. doi: 10.1016/j.future.2018. 09.054

Zerrahn, A., Schill, W. P., and Kemfert, C. (2018). On the economics of electrical storage for variable renewable energy sources. *Eur. Econ. Rev.* 108, 259–279. doi: 10.1016/j.euroecorev.2018.07.004

Zhang, C., Zhou, J., Li, C., Fu, W., and Peng, T. (2017). A compound structure of ELM based on feature selection and parameter optimization using hybrid backtracking search algorithm for wind speed forecasting. *Energy Conv. Manag.* 143, 360–376. doi: 10.1016/j.enconman.2017. 04.007

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.