Energy Conversion and Management 251 (2022) 114983

Contents lists available at ScienceDirect

Energy Conversion and Management

journal homepage: www.elsevier.com/locate/enconman

A novel ensemble model for long-term forecasting of wind and hydro power generation

Priyanka Malhan *, Monika Mittal

National Institute of Technology, Kurukshetra, India

ARTICLE INFO

Keywords:
Renewable power generation forecasting
Univariate time series prediction
Ensemble model
Long term time-horizon
Diligent Search Algorithm

ABSTRACT

Power generation scenario modelling has become an integral part of long-term planning in power system due to high penetration of variable renewable energy. It requires accurate estimates of power generation from different resources to find cost-optimal mix of generations. Predicting the generation of weather dependent renewables in long-term is not feasible but an adaptive long-term forecasting model based on univariate time-series analysis can provide the solution. Therefore, an effort has been made through this paper to provide accurate medium to long-term forecasts (a week-ahead to a year-ahead) for wind and hydro power generation using a novel ensemble forecasting model. The proposed model is devised in three phases; Phase-I develops a hybrid model using ARIMA (Auto Regressive Integrated Moving Average) and Bi-LSTM (Bidirectional Long Short Term Memory) predictions. Phase-II integrates the forecasts of seasonal and off-season generation periods obtained via a Diligent Search Algorithm (DSA). DSA is an innovative algorithm, designed to identify the hidden seasonalities that are responsible for the intermittent behaviour of wind and hydro power generation time-series. Finally, Phase-III facilitates amalgamation of prediction results of Phase-I and Phase-II to build the proposed forecasting model. Results show that MAE (Mean Absolute Error) for wind and hydro power are 1.97% to 5.52% and 2.3% to 6.42% while RMSE (Root Mean Square Error) varies from 2.79% to 7.8% and 2.63% to 8.4% respectively in a week-ahead to a year-ahead scenarios. Since this model is specifically designed for a year-ahead forecasting scenario, its performance can become unstable beyond this time horizon.

1. Introduction

Nowadays surplus amount of electricity is generated from renewable energy that offers advantages of being clean, inexhaustible, eco-friendly and cost-effective source of power generation. This alternative power source has experienced tremendous growth in past few years across the globe. As per 2020 report of REN21, installed power capacity of renewables has surpassed (in 2019) all past records with capacity addition of more than 200 GW. Solar PV and Wind power are the leading producers among all the renewables in many countries such as China, United States, Germany, India, Japan and Spain [1]. As of 31st March 2021, total installed capacity of renewables in India is 94.43 GW, comprising 39.24 GW of wind power, 40 GW of solar power, 4.8 GW from small hydro power plants and rest from other renewable sources. However, the total share of hydro power in the country is 51 GW including both small and large hydro power plants [2].

Although renewable energy offers bundle of benefits over the conventional sources of power such as coal, oil and natural gas, large scale integration of renewable energy affects the reliability and stability of the existing power system. Power generation from renewable sources is predominantly dependent upon weather and climatic conditions, which varies from year to year or season to season. In case of wind and hydro power, fluctuating wind speeds and monsoon variabilities have a direct impact on their power generation [1]. This intermittent and unpredictable behavior of renewables poses challenges at various stages of power system planning, management and operations. For instance, optimum generation scheduling and coordination of load demand and generation require accurate estimates of available power from different sources. There, it becomes difficult to provide the estimates due to the uncertainties associated with power production from these resources. Additional reserve capacity is thus employed to serve the power deficit, thereby increasing the overall generation cost [3,4]. Therefore, such unavoidable circumstances need intelligent and effective power forecasting models to enhance stability of the existing power systems and to reduce the additional costs and complexities. In the literature, there exist several efficient machine learning based approaches for solar, wind

* Corresponding author.
E-mail addresses: priyanka_61900032@nitkkr.ac.in (P. Malhan), monika_mittal@nitkkr.ac.in (M. Mittal).

https://doi.org/10.1016/j.enconman.2021.114983
Received 11 September 2021; Received in revised form 1 November 2021; Accepted 2 November 2021
Available online 15 November 2021
0196-8904/

Table 1
Considered medium to long-term forecasting horizons.

Scenario	Forecasting time-horizon	Forecasting Period
1	W – Weekly (1 week ahead)	Medium-term
2	M – Monthly (28 days)	Medium-term
3	Q – Quarter-yearly (28*4 days)	Long-term
4	HY – Half-yearly (181 days)	Long-term
5	Y – Yearly (365 days)	Long-term

and hydropower forecasting [5–7]. The number of research studies in wind and solar power forecasting is more than those for hydro and other renewables. For instance, solar and wind constitutes about 40% of the total research studies each as compared to about 2.3% studies related to hydro power [5].

Different research studies on renewable power forecasting can be categorized based on: a) Forecasting time-horizon, and b) adopted approaches [8]. Adopted power forecasting models can be grouped further into four classes with respect to forecasting time horizon [6,9–13]:

- Very-short term forecasting: Few seconds to 1 h
- Short-term forecasting: 30 min to 6 h/ several hours to 1 day
- Medium-term forecasting: 6 h to 1 day/ several days to weeks
- Long-term forecasting: 1 day to 1 week/ Weeks to a month

The proposed methods for wind power forecasting can be broadly

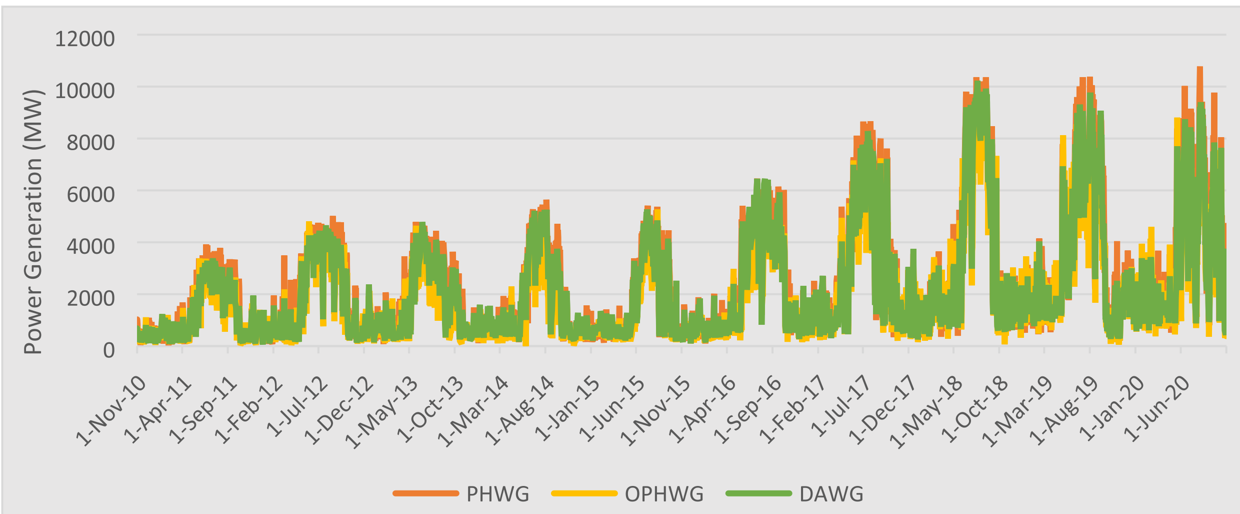

Fig. 1. Wind power generation scenarios (MW) from 1st November 2010 to 31st October 2020.

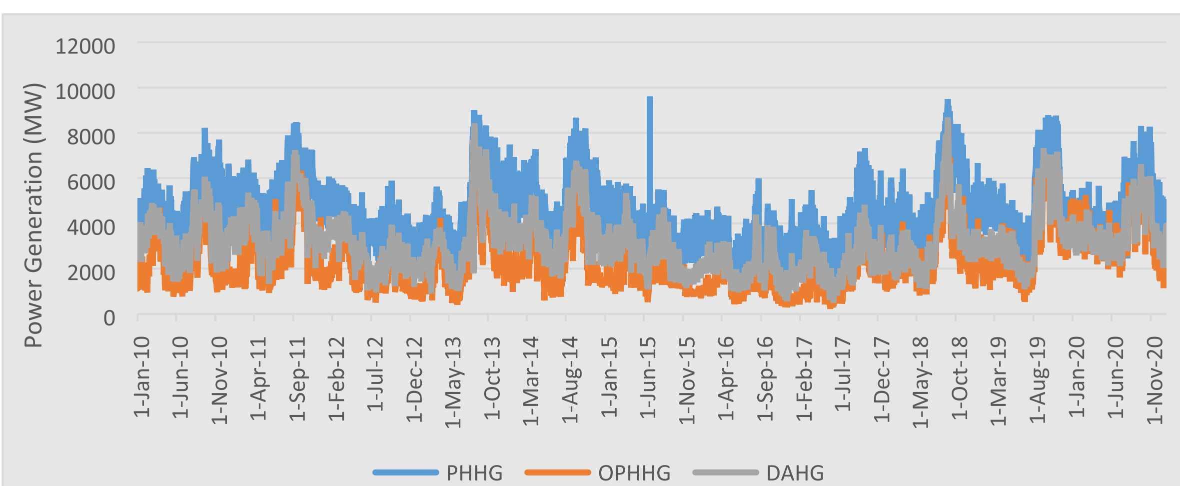

Fig. 2. Hydro power generation scenarios (MW) from 1st January 2010 to 31st December 2020.

Table 2
Statistical values of different wind power datasets.

Dataset Samples	DAWG (MW)				PHWG (MW)				OPHWG (MW)			
	Mean	SD	Max	Min	Mean	SD	Max	Min	Mean	SD	Max	Min
All	2199.2	2091.2	10,228	45	2422.5	2259.4	10,778	27	2017	1901	9675	0
Training	1970.1	1957	10,228	45	2178.1	2102.3	10,359	27	1748.8	1747.4	9675	0
Testing	3169.2	2480.4	9771	284	3496.1	2712	10,778	167	3034.4	2232.3	9100	56
Validation	2012	1328.1	9124	509	2088.7	1386.8	10,459	537	2204.2	1431	8872	451

Table 3
Statistical values of different hydro power datasets.

Dataset Samples	DAHG (MW)				PHHG (MW)				OPHHG (MW)			
	Mean	SD	Max	Min	Mean	SD	Max	Min	Mean	SD	Max	Min
All	3191.7	1277.3	8683	485	4504	1517.6	9609	717	2333	1309.2	8423	227
Training	3081.6	1274.2	8683	485	4458.4	1530.7	9609	717	2094.2	1220	8423	227
Testing	3682.2	1286.8	7418	1098	4785.1	1494.2	8779	1423	3295	1336.7	6980	530
Validation	3602	1203.6	7418	1098	4674.1	1455.8	8779	1423	3223	1245.8	6980	530

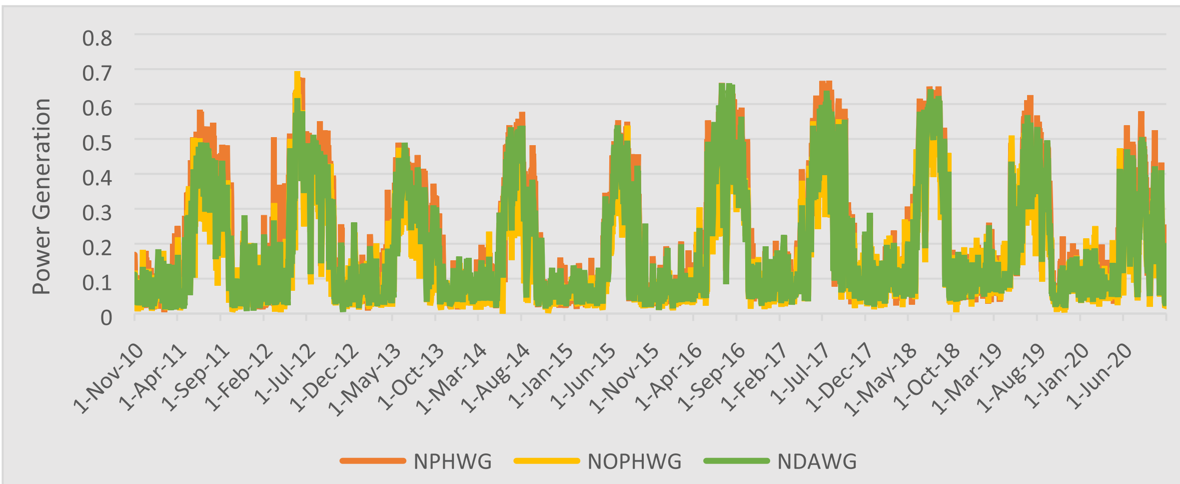

Fig. 3. Normalized wind power generation time series.

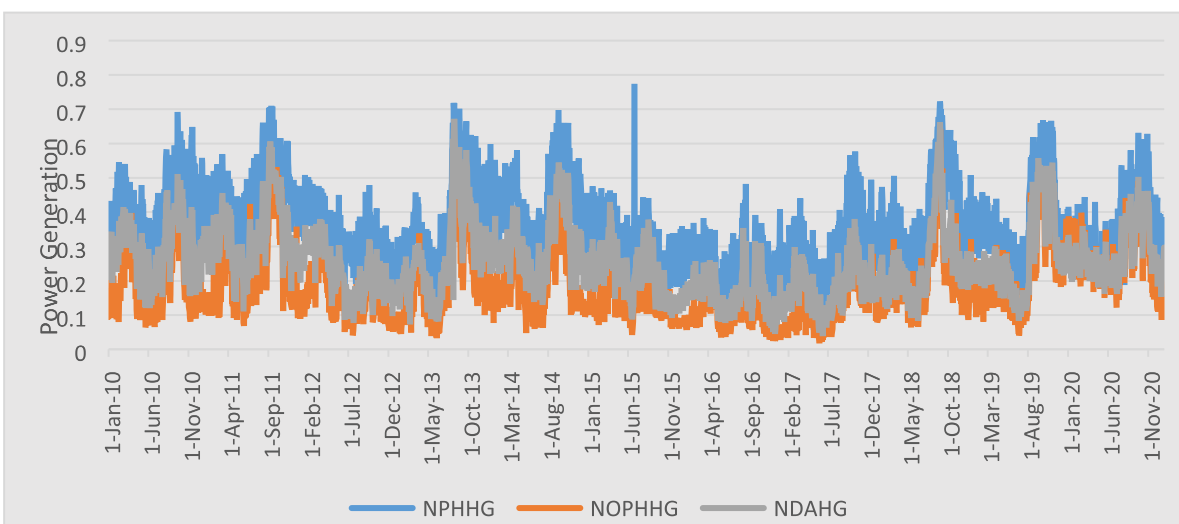

Fig. 4. Normalized hydro power generation time series.

classified into the following categories: Physical, Statistical, Artificial Intelligence, and Hybrid models [4,6,8–13]. Physical models rely on NWP (Numerical Weather Prediction) data for wind power forecasting. The meteorological data containing wind speed, wind direction, temperature, pressure etc. is processed to predict the wind speed. The predicted wind speed is thereby utilized to develop wind power forecasts on the basis of wind turbine power curve. Even though this method could provide accurate forecasts for accurate NWP data, it was found unsuitable for wind power forecasting because it is computationally inefficient. However, statistical models are way simpler, easier to implement and efficient than physical models. These models aim at extracting the essential information from the historical data which is then used to develop wind power predictions using various mathematical relations. Moreover, these can be sub classified as: a) Time Series models and b) Neural Network based models [6,10,14], and [64]. These time series models and neural network models are the earliest and most effective techniques in the field of wind power forecasting. These two powerful approaches have been widely adopted by various researchers for more than a decade [19,21,22,43,51], and [64–66].

Time series models can be further of two types: Multivariate time series and Univariate time series models. Multivariate time series model uses several variables like wind speed, wind direction, temperature, pressure, humidity etc., to forecast wind power generation. On the contrary, univariate time series models are single variable models that

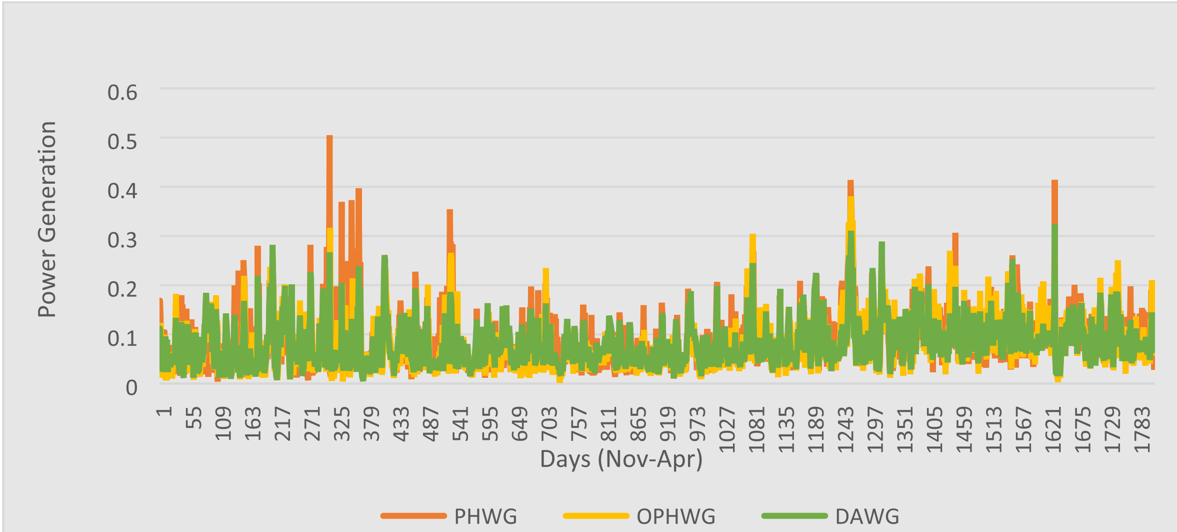

Fig. 5. Off-season (Nov-Apr) generation of wind power.

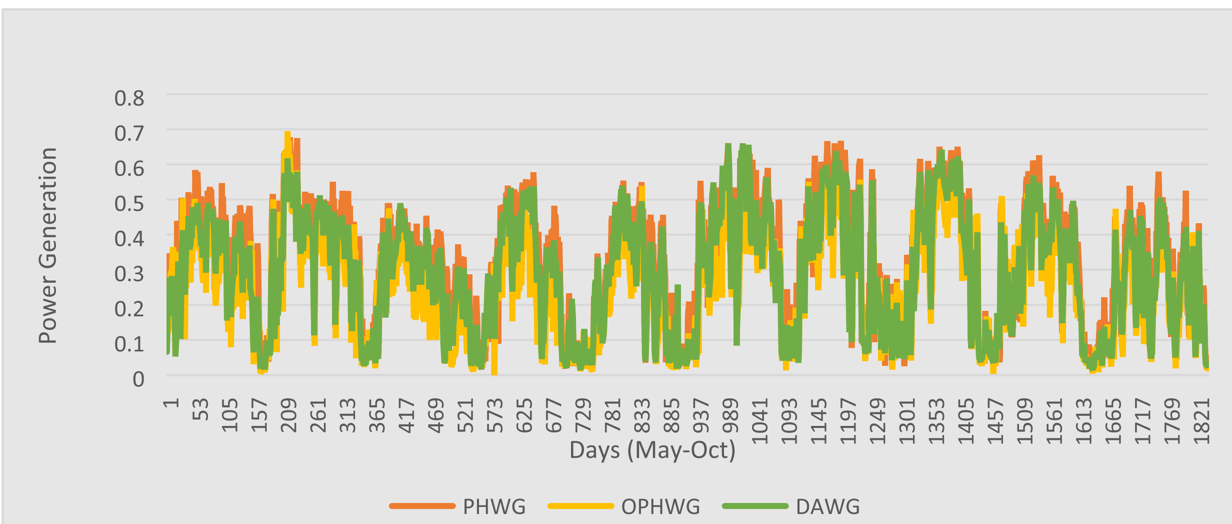

Fig. 6. Seasonal (May-Oct) generation of wind power.

forecast the future values of wind power based on lagged values of past wind power data [8]. AR (Auto-Regressive), ARMA (Auto-Regressive Moving Average), ARIMA (Auto-Regressive Integrated Moving Average) and ARX (Auto-Regressive with exogenous variable) are the most widely used statistical models by researchers [10]. These models provide accurate predictions for short-term forecasting but as forecasting time-horizon increases, the accuracy of statistical models falls sharply. In [15], an ARIMA based model namely, D-TARX (Decomposed-Threshold Auto Regressive with External Inputs), was applied for short to medium term wind power forecasting. The performance of the proposed model was evaluated on the basis of MAE (Mean Absolute Error), RMSE (Root Mean Square Error) and MAPE (Mean Absolute Percentage Error). MAPE for 48-hour ahead prediction for 16 days was reported to be 47%. Likewise, [16] employed ARMA and ARIMA techniques to forecast wind speed and wind power for medium-term and long-term applications. The time period over which the predictions were made ranged from 10-min to 1 day ahead. MAPE for wind power increased from 10% to 72% with the increasing prediction time-scale. Researchers have successfully exploited time-series models such as AR, ARX and VAR to develop multi-fold model for short-term wind power forecasting [6,10,17–20]. Neural Networks are considered as one of the most robust and versatile forecasting techniques for various applications. The strength of these models is they can easily depict the non-linear relationship between the input features and output target and are independent of mathematical model complexities. Existing studies used more than 50 different forms of ANN (Artificial Neural Network) based models. For instance, traditional ANNs used for wind power forecasting include RBFNN (Radial Basis Function Neural Network), WNN (Wavelet Neural Network), BPNN (Back-Propagation Neural Network), ELM (Extreme Learning Machine) etc. [17,21,22,27], and [50]. In [51], forecasting performances of neural network models RBFNN and BPNN were compared and it was established that the former is more efficient than the latter. Some other studies on wind and hydro power forecasting, that have used these two models for comparison, also confirmed that RBFNN gives better results in forecasting than BPNN [26,52]. Besides, ELM based models were found to be superior in short-term forecasting of wind power than BPNN [27,50].

In hybrid prediction approach, models are designed by combining different forecasting methods to take the advantage of merits of each technique, thereby improving the overall accuracy of the developed hybrid model [13,23–25]. Combination of ANN and fuzzy logic system is one of the most popular hybrid approach for wind power forecasting as

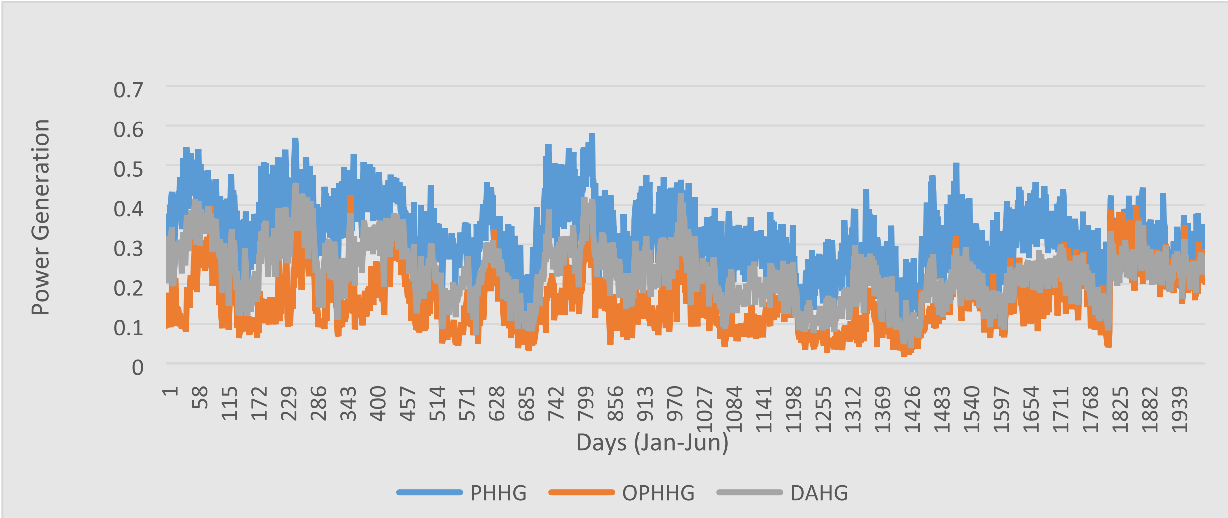

Fig. 7. Off-season (Jan-Jun) generation of hydro power.

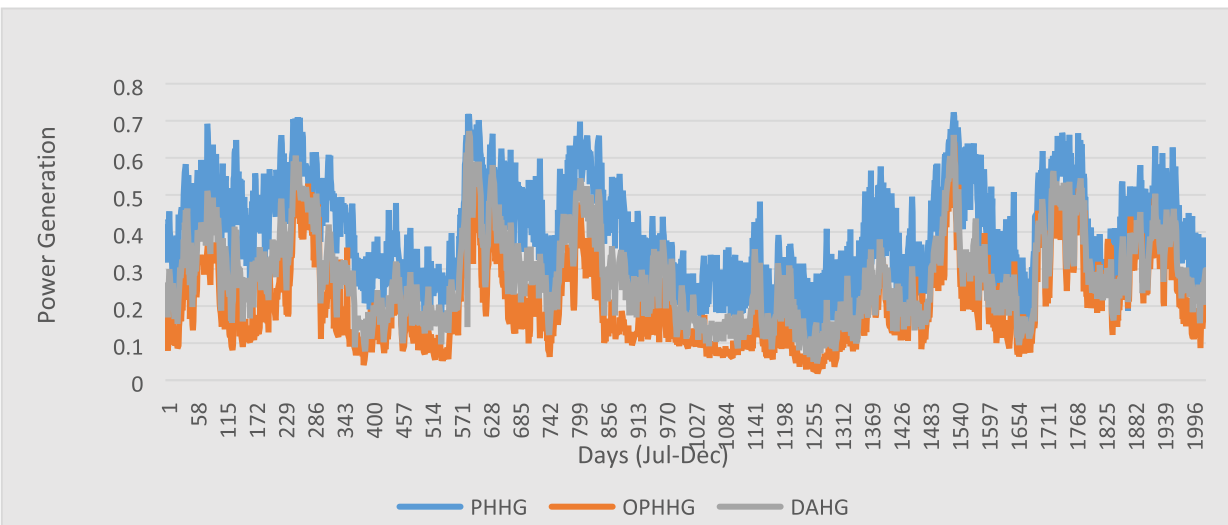

Fig. 8. Seasonal (Jul-Dec) generation of hydro power.

reported in the literature [12,26]. Ensemble modelling is a more advanced form of hybrid forecasting methods. It is derived from statistics and machine learning in which multiple learning algorithms and different training datasets are utilized to develop a diverse predictive model [8,11,27,28]. Apart from aforementioned techniques, deep learning based neural networks such as LSTM (Long-Short-Term Memory), RNN (Recurrent Neural Network), CNN (Convolutional Neural Network) and RBM (Restricted Boltzmann Machine) are also disseminating in the field of renewable power forecasting [25,29–33]. CNN, LSTM, CNN-RBFNN, hybrid CNN-LSTM, Bi-LSTM are some of the deep learning models that were successfully utilized for short-term wind speed and wind power forecasting [8,25,30–32,52,53]. CNN is used majorly for wind speed/power forecasting among these models. Identifying appropriate input features and output target value, that is to be predicted, is another perspective of understanding the state-of-art of various wind power forecasting methodologies. According to [6], more than 70% of the studies have considered multiple input features such as wind speed, wind direction, temperature, pressure, relative humidity, air density, and wind turbine dynamics to forecast the output wind power. While the rest considered only past wind power data as the input feature. LSTM, ARIMA, hybrid ARIMA-LSTM are widely used techniques for wind speed/power forecasting from the perspective of time-series forecasting [54–58]. Hybrid LSTM-ARIMA provides better forecasts than both ARIMA and LSTM [58].

Though the number of existing research studies related to hydro power forecasting is comparatively low as perceived from the review papers [5,34], and [35], both deep learning and time-series forecasting models are found to have application in hydro power. For instance in [59], ELM based hybrid model was developed to predict hydro power generation for 3 days to 21 days-ahead time horizon. Performance of the developed model was compared with that of ELM, LSTM, RBFNN, and BPNN. MAE of 5.75% to 10.81% and RMSE of 9% to 15% over selected time horizons verified the accuracy of this model as compared to other methods. It was also observed that LSTM model outperformed RBFNN, BPNN, and simple ELM. Hybrid ARIMA-LSTM model was employed to obtain improved forecast for hydrological time-series data in [60]. Apart from this, physical models employing NWP data and statistical based models were also used to predict hydro power generation in [67] and [68]. In [36], an optimal reservoir based operation model was introduced to fully utilize the long, medium and short term inflow forecasts. It used multiple linear regression and ANN models to obtain the proposed hybrid model that could provide best results over a time period of

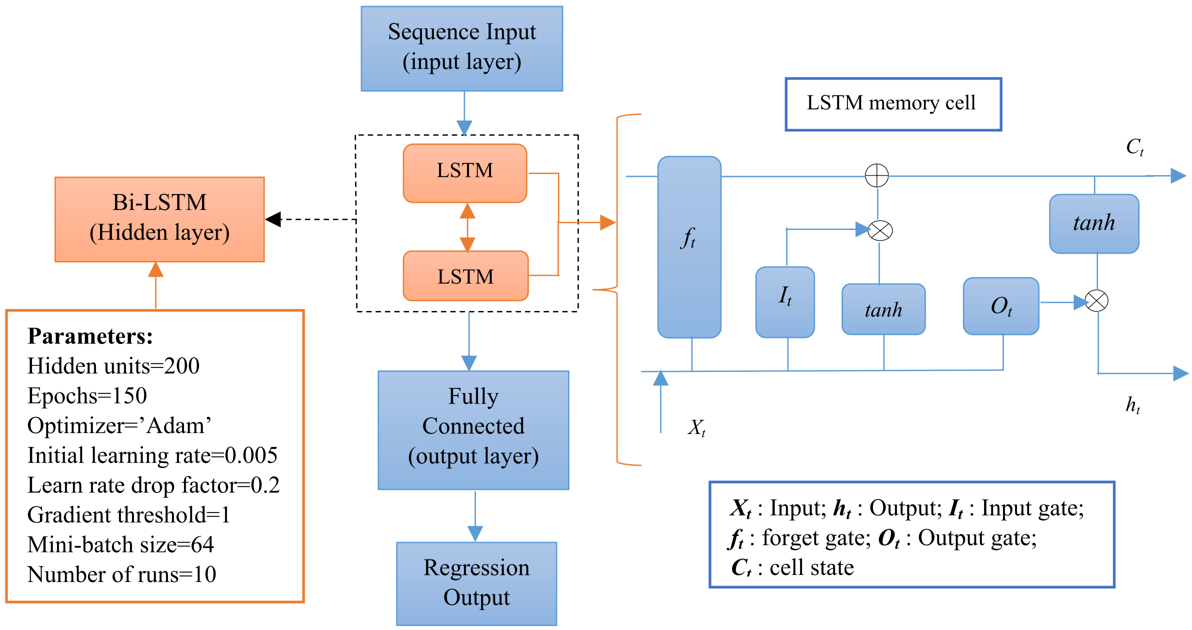

Fig. 9. Architecture of Bi-LSTM model with details of LSTM memory cell structure.

only 10 days. In [37], water level of a reservoir was forecasted for developing a sustainable hydropower generation strategy using different machine learning algorithms. There, Bayesian Linear Regression outperformed other models with 97% accuracy in predicting a week ahead scenario. In [38], a promising technique was presented for hydro power generation forecasting, employing metaheuristic optimization and adaptive neuro-fuzzy inference system, providing an accuracy of 79%.

The following challenges still exist in wind and hydro power forecasting as observed from extensive literature survey:

- There exists only a few studies on long-term forecasting, which is essential to estimate the extent of flexibility and various permutations & combinations of power generation required for large scale integration of renewables with the grid [39].
- In the above scenario modeling, extrapolation and interpolation based techniques are usually adopted for future prediction of the share of different generation resources [40]. But majority of such existing techniques aren't reliable in capturing the future variability of renewables efficiently in long-term.
- Majority of the existing studies developed multivariate predictive models that are dependent on weather data to carry out their generation forecasts. Use of such models need large data that is difficult to obtain in general and also involve more computational efforts. However, univariate predictive models can reduce the computational burden by eliminating the need of weather data for power forecasting [8,41].
- Different datasets from different locations such as Spain, Canada, USA, China, Turkey, Taiwan, France, Brazil, and India etc. have been considered in the recent studies for performance evaluation of the proposed forecasting models. However, there doesn't exist much literature related to wind power forecasting on Indian data despite of its fourth position in wind energy with 39 GW of installed power capacity [42,43].
- Diversity of the forecasting models has been validated on datasets from different locations in various existing studies. There doesn't exist a single model approach for prediction of two different renewable sources and is still a major challenge in the field of renewable power generation forecasting.

Most of the renewables like solar, wind and hydro primarily depend upon weather for power generation at any instant of time in a year. Therefore, many authors have used multivariate forecasting approach based on meteorological data to make accurate predictions for both wind and hydro power. Meteorological data parameters can be predicted accurately for short-term period but their accuracy decreases beyond a week-ahead forecasts. So, it is not reliable to involve weather data as input feature for long-term forecasting such as a year-ahead scenario for renewable power generation. Hence, univariate time-series forecasting approach has been used in this paper to develop a novel ensemble based long-term forecasting model for renewable power generation. It is observed from the extensive literature survey that statistical model viz. ARIMA and deep learning models viz. LSTM/Bi-LSTM exhibit an excellent performance for various time-series forecasting applications over longer time horizons. Therefore, these two models have been considered in this paper to design the proposed forecasting model. Also, performance of the proposed model is compared with other efficient machine learning algorithms such as RBFNN, ELM and LSTM.

It is a well-known fact that power generation from renewable sources exhibit an evident cyclic pattern which repeats itself every year. This periodic repetition is the result of seasonal variations. Along with seasonal variations the renewables also have some irregular periods of fluctuating generation. Thus, this paper focuses on capturing these seasonal effects that lie within steady and unsteady behavior of wind and hydro power time-series to effectively predict their generation in future. The major contributions of this paper are:

a) A novel ensemble model based on ARIMA and Bi-LSTM has been designed for forecasting wind and hydro power generation. This univariate time series model is suitable for other power generation sources also that manifest seasonal variations.
b) The proposed model provides accurate forecasts for medium to long-term prediction horizon. Various forecasting horizons considered in this paper are shown in Table 1.
c) An intelligent algorithm viz., DSA (Diligent Search Algorithm), has been proposed to be used for the first time, to identify hidden seasonality periods present in the considered time series to enhance quality of the proposed model.
d) The performance of the model is evaluated using the most recent wind and hydro power generation datasets from India.

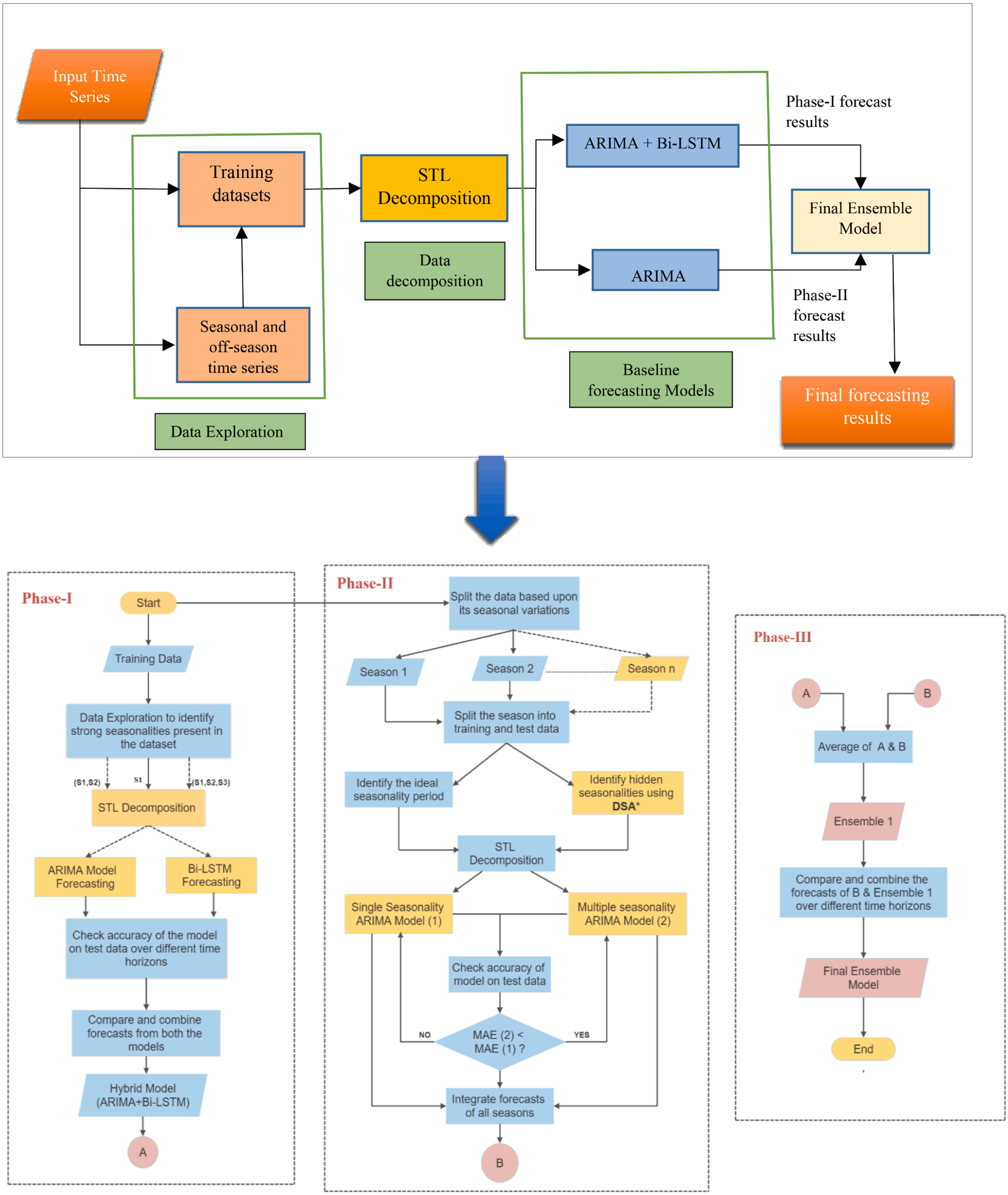

Fig. 10. General architecture and detailed flowchart of the proposed ensemble forecasting model.

The model proposed in this paper captures the effects of different seasons occurring in a year on renewable power generation for designing the long-term forecasting ensemble. Therefore, accurate forecasts can be made for up to a year-ahead scenario only as its prediction accuracy would be severely affected for longer time scales beyond a year. This paper is organized as follows. Details of the data and preliminary studies on the characteristics of the selected wind and hydro power datasets are presented in Section 2. Section 3 elaborates the adopted framework & methodology to develop the proposed ensemble forecasting model. It also describes the performance evaluation criteria used in this study to assess accuracy of the model. Section 4 presents the obtained prediction results along with associated errors using both testing datasets and validation datasets. It also provides detailed discussion on comparison with other existing methods and models. Finally,

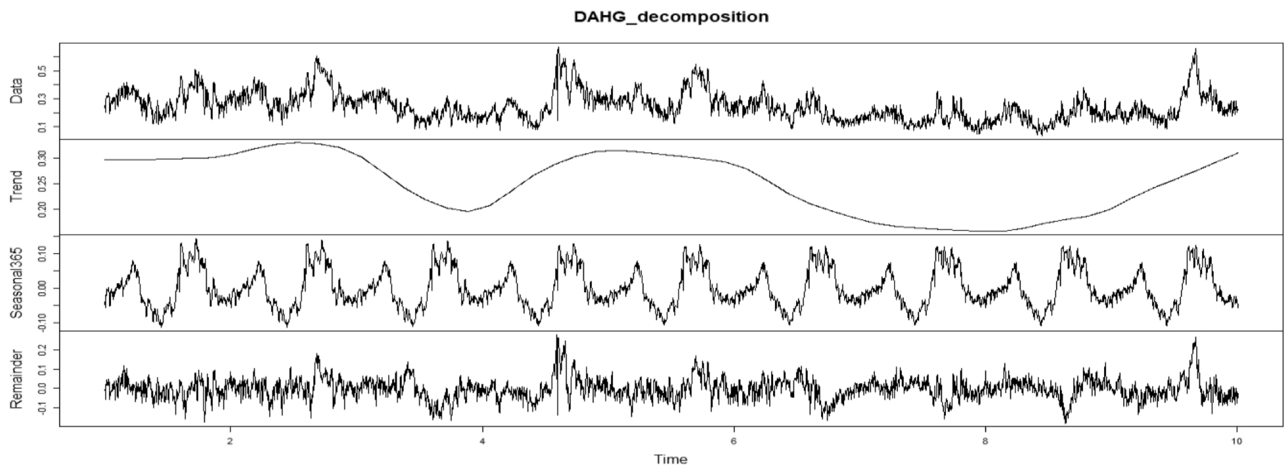

Fig. 11. STL decomposition of DAHG series with single seasonality period.

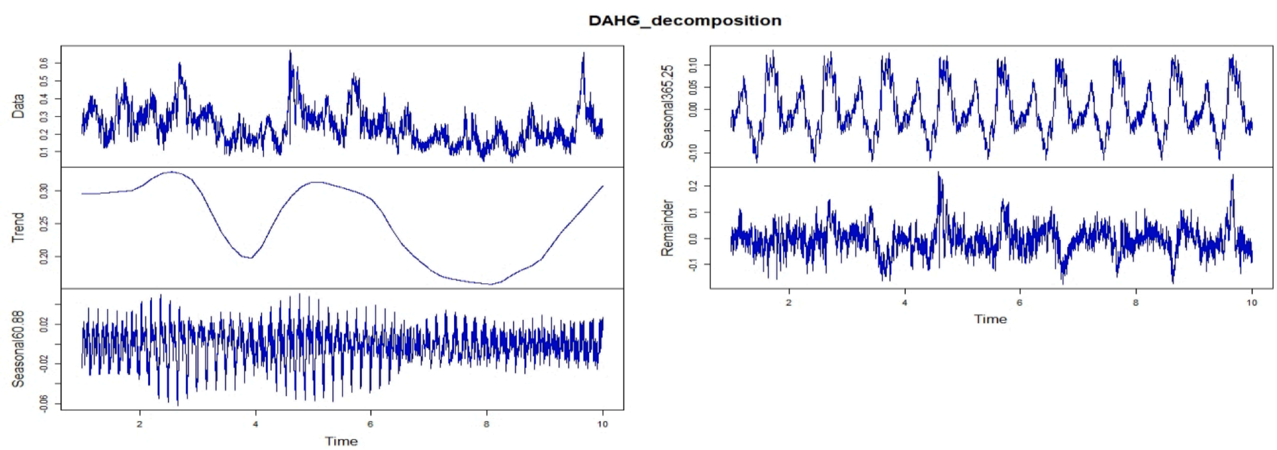

Fig. 12. STL decomposition of DAHG series with two seasonality periods.

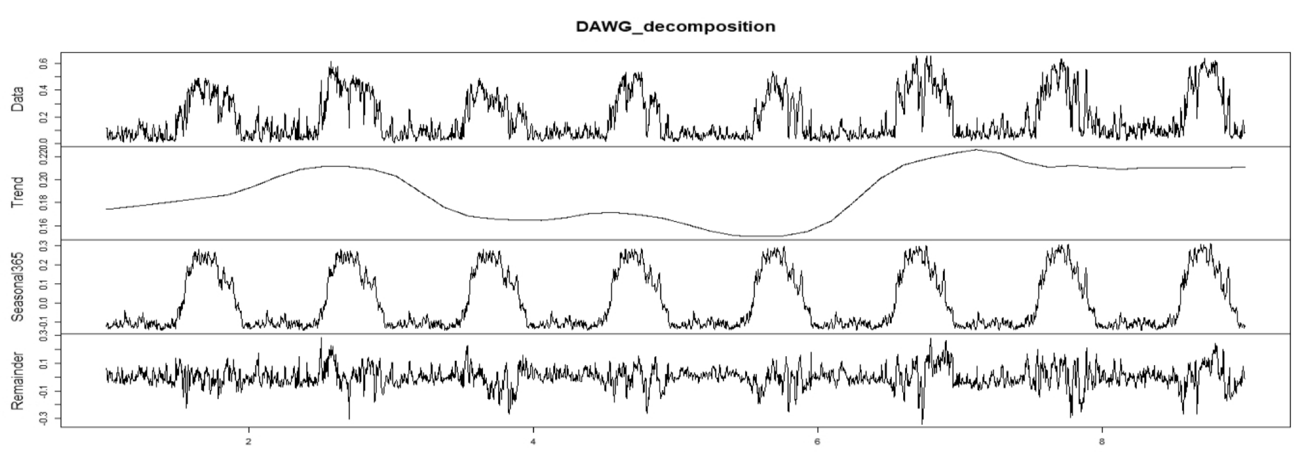

Fig. 13. STL decomposition of DAWG series with single seasonality period.

Section 5 concludes the paper summarizing different aspects of the benefits of the proposed model.

2. Data description and preliminary analysis

2.1. Data description

In this paper, six wind and hydro power generation datasets viz., DAWG (Daily Average Wind Generation), PHWG (Peak Hour Wind

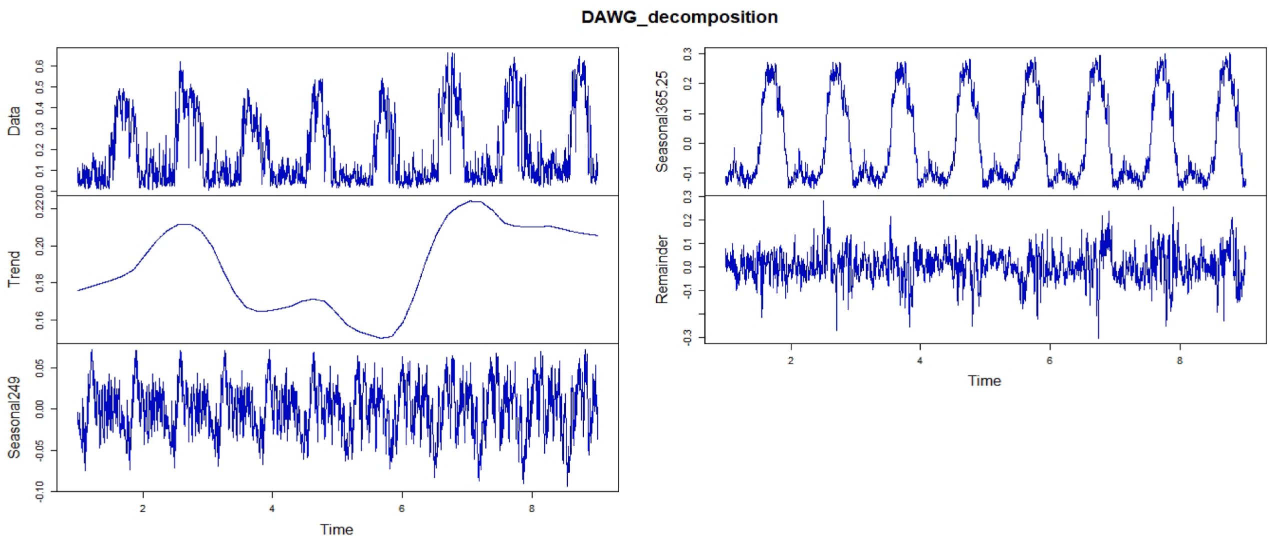

Fig. 14. STL decomposition of DAWG series with two seasonality periods.

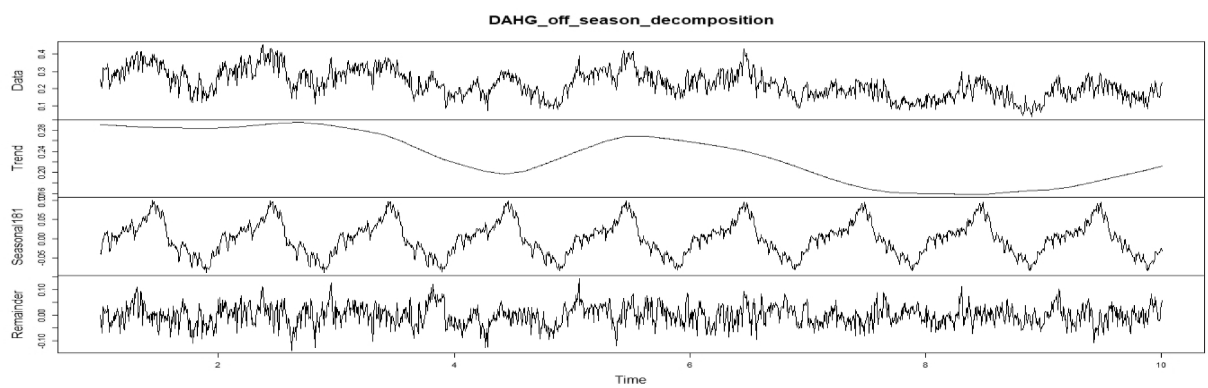

Fig. 15. STL decomposition of off-season generation period of DAHG.

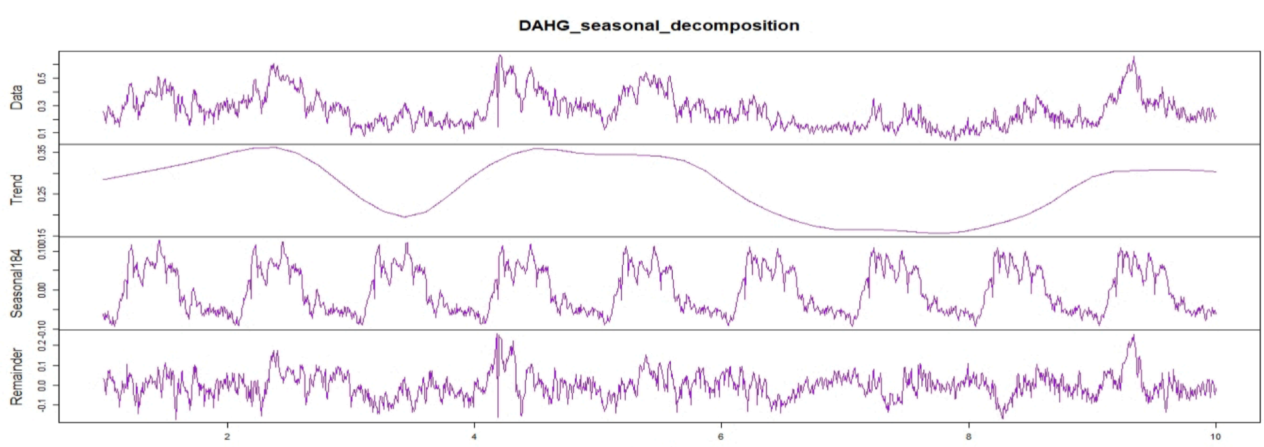

Fig. 16. STL decomposition of seasonal generation period of DAHG.

Generation), OPHWG (Off-Peak Hour Wind Generation), DAHG (Daily Average Hydro Generation), PHHG (Peak Hour Hydro Generation), and OPHHG (Off-Peak Hour Hydro Generation) are considered to develop the proposed forecasting model. Daily basis data of eleven years (i.e. from 2010 to 2020) is collected from SRLDC (Southern Regional Load Despatch Centre) of India that handles power generation of total five states [44]. All the six datasets are time-series data, therefore, power generation data from 2010 to 2018 are used for training and those from 2019 to 2020 are treated as testing datasets. Predictive power and diversity of the designed forecasting model is further validated using an unknown validation datasets comprising of power generation data for five months of the year 2021 in this study.

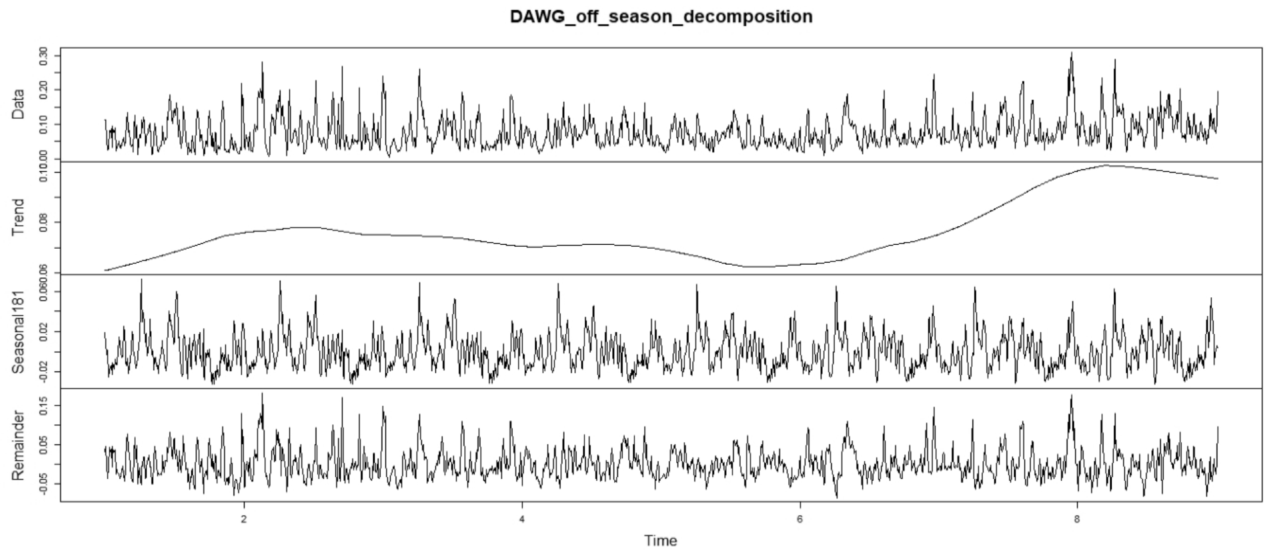

Fig. 17. STL decomposition of off-season generation period of DAWG.

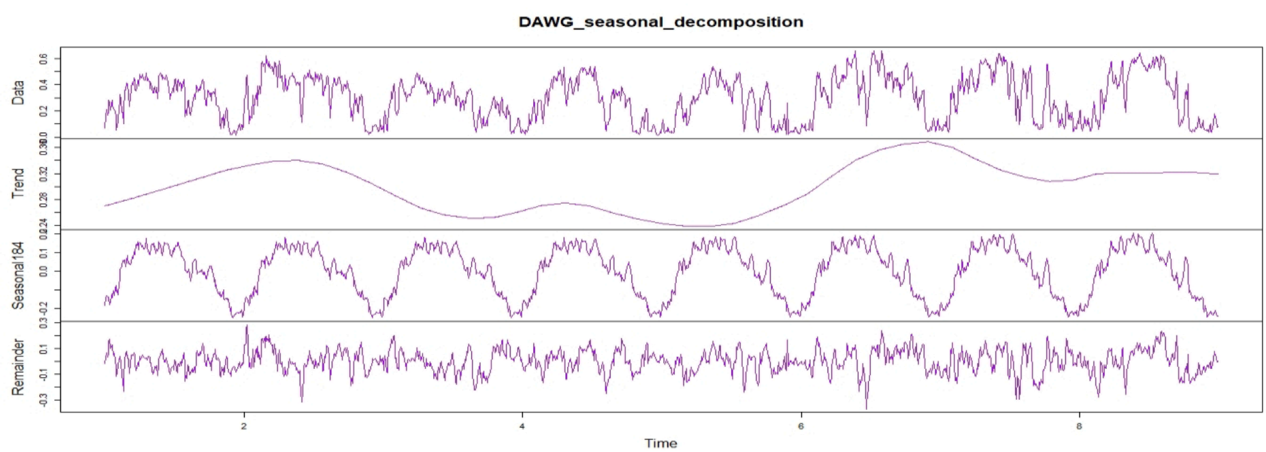

Fig. 18. STL decomposition of seasonal generation period of DAWG.

2.2. Preliminary data analysis

Since each renewable power source exhibits different characteristics, a preliminary analysis of its power generation time series is essential to get an insight about the seasonal variations before actually designing the forecasting model. The actual power generation (MW) scenarios of wind and hydro power plants are shown in Figs. 1 and 2, respectively. It can be quickly observed from Figs. 1 and 2 that power generation from both the renewables are cyclic in nature that repeat itself every year.

The descriptive statistics of all the considered datasets is shown in Table 2 & Table 3. Higher values of standard deviation specifically in wind power datasets indicate that the generation is highly variable and fluctuating. Also, the maximum and minimum generation values are far apart. The maximum value of wind power generation and standard deviation as observed in existing studies didn't exceed 3200 MW and 800 MW respectively [8,11,18,24,25,27]. On the contrary, values of standard deviation and mean for wind power datasets considered in this study are very close to each other. Hence, it is challenging to develop an accurate power forecasting model for these two highly intermittent and different renewable sources. In order to build a forecasting model that guarantees optimum performance for longer prediction time-horizon, it is crucial to process the original high-resolution data in its original form. Besides, it is also prudent to normalize the abruptly varying data on to a common scale without losing any information.

In Fig. 1, it can be noticed that wind power generation rises rapidly after 2016. Diagnostic analysis of the data showed that this continuous year-on-year increase in power generation is due to addition in installed capacity of the plants every year. In case of hydro power plants, the capacity addition is at smaller scales (i.e. small and micro hydro), and hence, the trend remains almost constant as shown in Fig. 2. Thus, all the power generation datasets are normalized to bring them at the same scale for forecasting using following equation:

$$NormalizedPower = \frac{ActualPowerGeneration(MW)}{TotalInstalledCapacity(MW)} \tag{1}$$

Next, the periodic patterns reflected in the normalized wind and hydro power graphs as shown in Figs. 3 and 4 are due to seasonal variations. To examine the effect of different seasons on power generation, these are further divided into two categories: a) Seasonal Power Generation and b) Off-season Power Generation as shown in Figs. 5–8. Seasonal generation period is from May to October and July to December for wind and hydro power respectively. Similarly, Off-season generation period for wind is from November to April and for hydro it is from January to June. In other words, seasonal and off-season time intervals are the periods of high and low power generation, respectively.

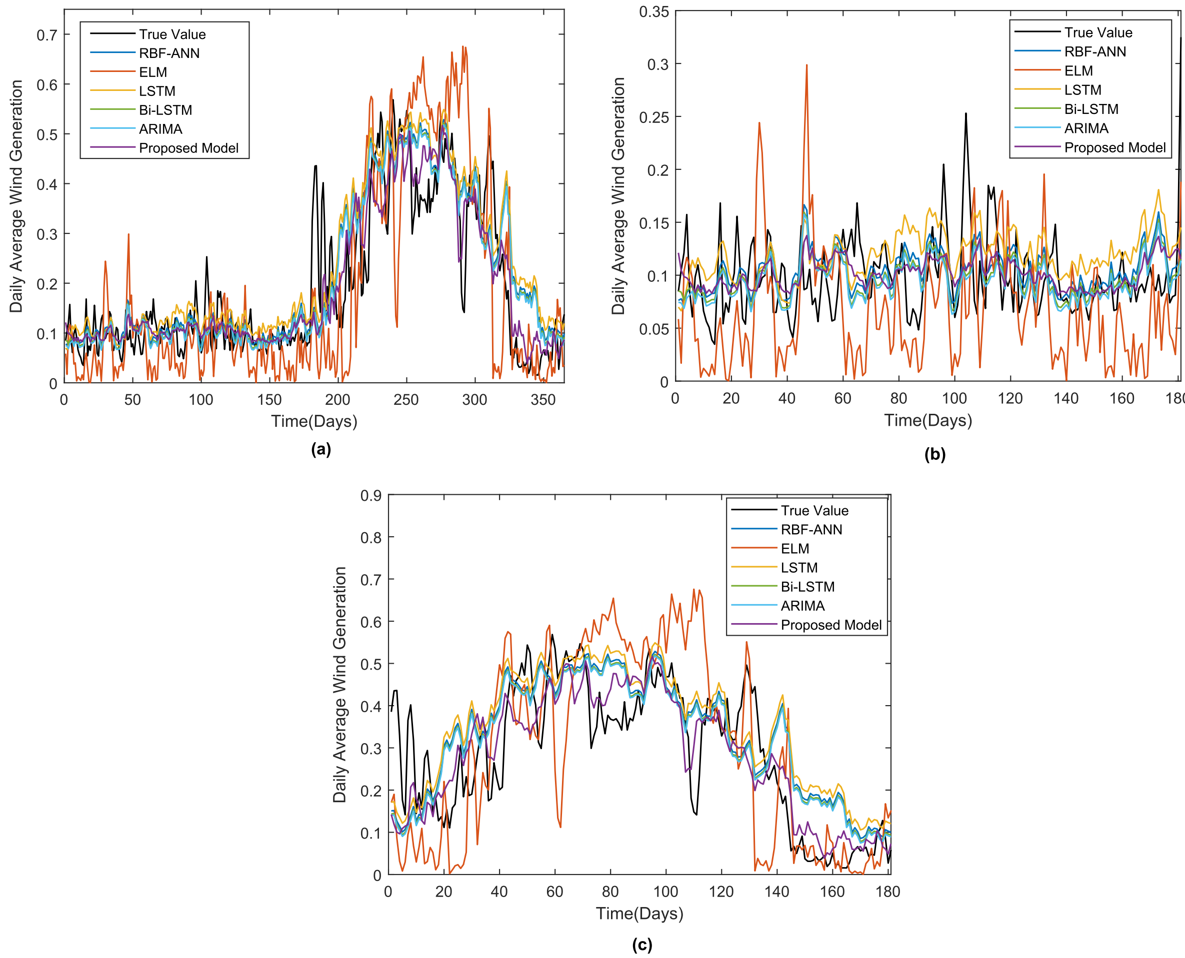

Fig. 19. DAWG testing dataset forecasts for **a)** A year ahead scenario; **b)** Off-season generation; **c)** Seasonal generation.

3. Methodology

The accuracy and robustness of ensemble models are ensured by the diversity of its sub-models [8]. In this paper, ARIMA and Bi-LSTM are selected as the baseline models for time-series forecasting. The statistical models have a great history in providing accurate forecasts for time-series data as already stated in Section 1. Hence, the most widely used ARIMA model is adopted to meticulously capture the persistent behavior of wind and hydro power generation. Deep learning model Bi-LSTM (Bidirectional Long Short-Term Memory) is well known for its ability to learn long-term dependencies between time steps of the time-series in both the directions.

3.1. ARIMA

The design of ARIMA model is based upon the concept of detecting autocorrelation between the past data points. The lags and lagged forecast errors are then used to develop the parameters of the fitted model which in turn can be utilized to predict the future values. AR (autoregressive) model, MA (moving average) model and ARMA (autoregressive + moving average) model all together forms ARIMA (autoregressive integrated moving average) model.

3.1.1. MA model

The moving average model creates a set of random variables using the past values that is quite similar to the random walk with zero mean and constant variance of the white noise. These models are expressed as MA (q), where q is the order of the process. The order decides the number of back steps to see along the white noise sequence for taking the average over the time series. The model can be written as

$$X_t = Z_t + \theta_1 Z_{t-1} + \theta_2 Z_{t-2} + \cdots + \theta_q Z_{t-q} \tag{2}$$

Where, Z_t is normally distributed noise and θ represents the weights such that

$$\theta_1 > \theta_2 > \cdots > \theta_q$$

X_t is the observed time series at time t.

3.1.2. AR model

The autoregressive model defines the current position in a time series as a function of its previous position and a noise variable as

$$X_t = Z_t + \varnothing_1 X_{t-1} + \varnothing_2 X_{t-2} + \cdots + \varnothing_p X_{t-p} \tag{3}$$

This model is denoted by AR (p), where p is the order of the process and $\varphi_1,\ldots,\varphi_p$ are the parameters of the model to be estimated. This expression can be written in terms of back-shift operator B as

$$\varnothing(B) X_t = Z_t \tag{4}$$

Where,

$$\varnothing(B) = 1 - \varnothing_1 B - \cdots - \varnothing_p B^p$$

and the individual terms are defined as $X_{t-1} = BX_t, X_{t-2} = B^2 X_t, \cdots X_{t-p} = B^p X_t$

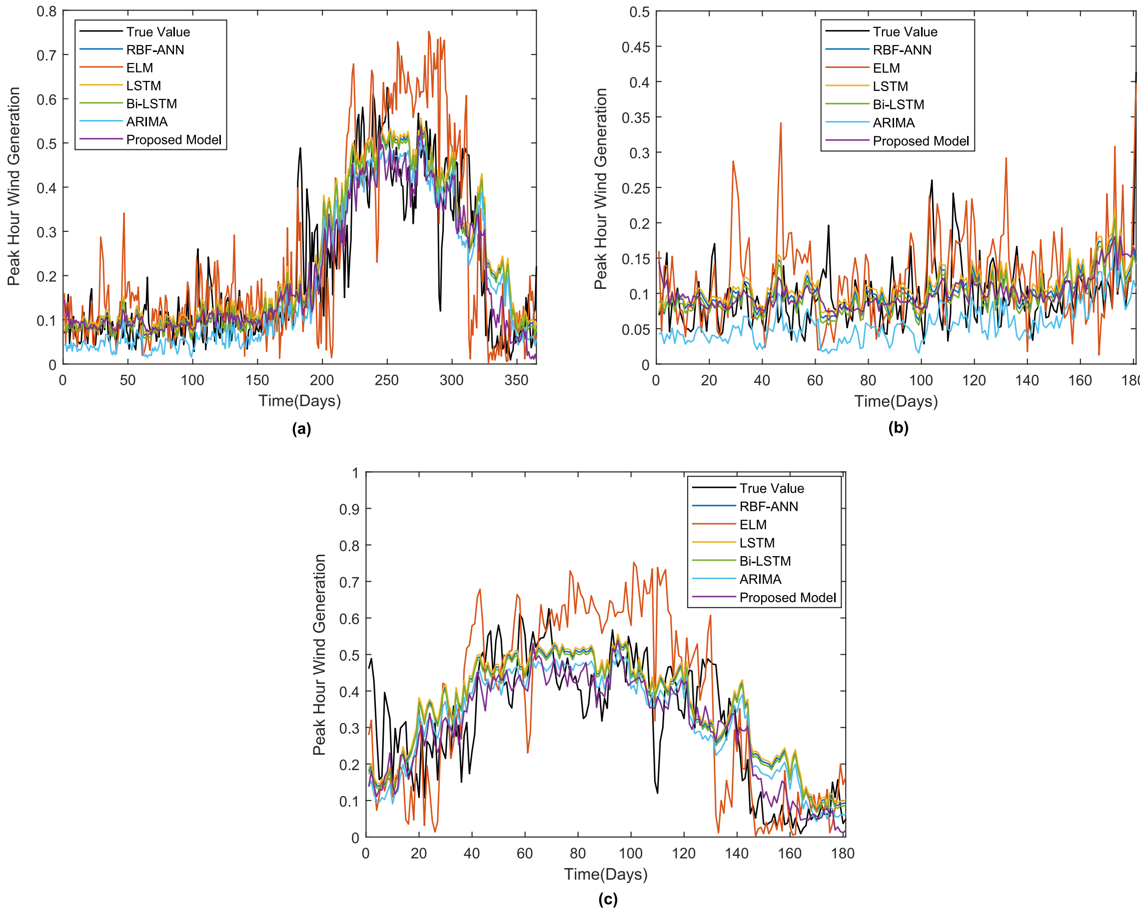

Fig. 20. PHWG testing dataset forecasts for **a)** A year ahead scenario; **b)** Off-season generation; **c)** Seasonal generation.

3.1.3. ARMA model

ARMA (p, q) model with p autoregressive terms and q moving average terms provides more flexibility in fitting the time series. This model can be expressed as

$$X_t = Noise + ARterm + MAterm \tag{5}$$

$$X_t = Z_t + \varnothing_1 X_{t-1} + \cdots + \varnothing_p X_{t-p} + \theta_1 Z_{t-1} + \cdots + \theta_q Z_{t-q} \tag{6}$$

In terms of back-shift operator B as

$$\varnothing_p(B) X_t = \theta_q(B) Z_t \tag{7}$$

3.1.4. ARIMA model

AR, MA, and ARMA models are statistical models which are designed to deal with the stationary time series. The real-life datasets are non-stationary in nature as they might contain systematic change in trend. In order to apply these statistical forecasting models it is necessary to de-trend or difference the given time series to convert it into a stationary time series. The difference operator (∇) modifies ARMA into ARIMA model. The autoregressive integrated moving average of time series X_t of order (p, d, q) thus can be represented as

$$\nabla^d X_t = Z_t + \varnothing_1 X_{t-1} + \cdots + \varnothing_p X_{t-p} + \theta_1 Z_{t-1} + \cdots + \theta_q Z_{t-q} \tag{8}$$

Or

$$\varnothing(B) \nabla^d X_t = \theta(B) Z_t \tag{9}$$

Where, $\nabla^d X_t = (1 - B)^d X_t$, and d = order of differencing.

ARIMA modelling process involves basic steps viz. model identification, parameter estimation and diagnostic checking and forecasting. These models are also referred as Box-Jenkins models [10].

3.2. Bi-LSTM

Bi-LSTM model is a combination of two LSTMs. A LSTM is a type of recurrent neural network (RNN) which provides more flexibility in controlling the output as compared to that provided by RNN. These are the most suitable deep learning models to make predictions for time-series data. The network of LSTM comprises a processor to identify only the useful information from the input. The hidden units of an ordinary recurrent network are replaced by the recurrently connected cells in LSTM. Input layer gate, forget-gate and output layer gate are the three doors in a cell [31]. Forget-gate removes the irrelevant information from the memory of the network using a sigmoid function. Input gate handles the new information added in the memory with the help of its 'sigmoid' and 'tanh' layer. These two layers further decide which values to be updated and added in the memory. And, the output gate checks whether the existing value in the cell contributes to the output or not [45].

Bi-LSTM network consists of two LSTMs so as one LSTM takes the input in forward direction and the other in backward direction. In this way more information is fed into the learning algorithm which improves the overall efficiency of the model. Bi-LSTM enables additional training by processing the data twice, and hence, provides better prediction results as compared to traditional LSTM [45]. The general architecture of the Bi-LSTM model is shown in Fig. 9 along with details of a single LSTM architecture for ready reference. Parameters used to train the Bi-LSTM network in the proposed model are also shown in this figure. Sequence Input layer and Bi-LSTM layer are the two main components of this network. Sequence Input layer inputs the time-series data into the network and Bi-LSTM layer learns their long-term dependencies.

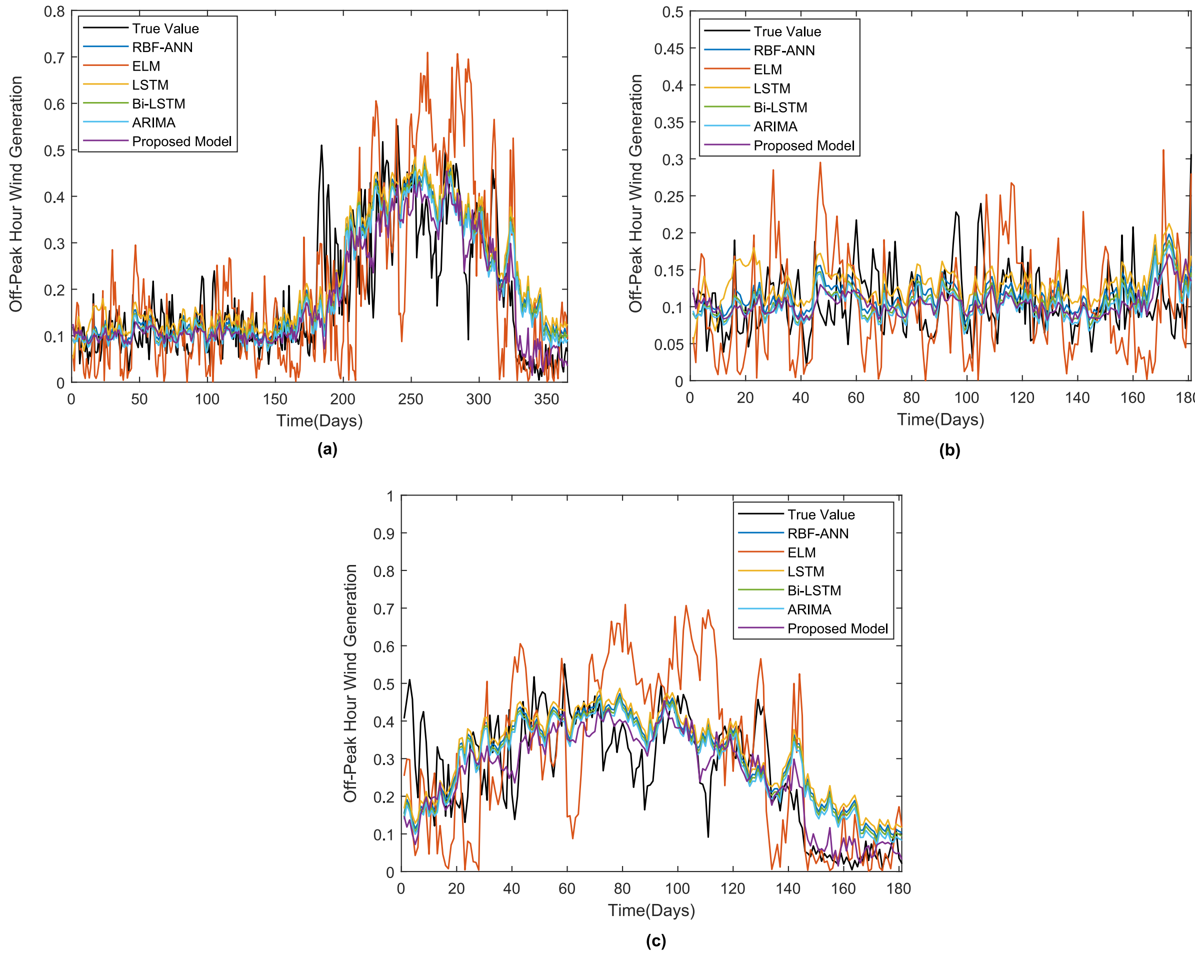

Fig. 21. OPHWG testing dataset forecasts for **a)** A year ahead scenario; **b)** Off-season generation; **c)** Seasonal generation.

3.3. Performance evaluation

The accuracy of the proposed ensemble model is evaluated using three performance metrics. Mean Absolute Error (MAE), Root Mean Square Error (RMSE), and Mean Arctangent Absolute Percentage Error (MAAPE). MAE and RMSE are the most commonly used criterions for measuring accuracy of the forecasts of not only wind energy but also of many other renewables [5,8,27]. Though Mean Absolute Percentage Error (MAPE) is an important unit less measure of accuracy but it generates infinite or undefined values when the actual values are zero or close to zero. Hence, as an alternative solution to this problem a new metric MAAPE is currently being used as reported in various studies for different applications [8,46]. It offers several advantages over MAPE like it is more robust, more effective, less biased and simple to calculate. It can be intuitively used as an absolute percentage error. Smaller MAAPE means better accuracy and its value ranges from 0 to π/2.

Mathematically MAE, RMSE and MAAPE are expressed as:

$$MAE = \frac{1}{n}\sum_{j=1}^{n}\left(\left|X_j - Y_j\right|\right) \tag{10}$$

$$RMSE = \sqrt{\frac{1}{n}\sum_{j=1}^{n}\left(X_j - Y_j\right)^2} \tag{11}$$

$$MAAPE = \frac{1}{n}\sum_{j=1}^{n}\arctan\left(\frac{\left|X_j - Y_j\right|}{X_j}\right) \tag{12}$$

Where X_j, Y_j and n are actual value, forecasted value and total number of observations of power generation in the testing datasets.

3.4. Data decomposition

Data pre-processing has always been an important part of machine learning based forecasting methods. The pre-processing strategies reportedly used by various researchers for wind energy forecasting involve decomposition, feature selection, feature extraction, de-noising, residual error modelling, outlier detection and filter based correction. Out of these seven strategies, decomposition is the most commonly used data pre-processing technique due to its ability of reducing the forecasting error by more than 50% [47]. Further, data decomposition is of several types: EMD (Empirical Mode Decomposition), EEMD (Ensemble Empirical Mode Decomposition), CEEMD (Complementary Ensemble Empirical Mode Decomposition), VMD (Variational Mode Decomposition), SAM (Seasonal Adjust Method), ASD (Atomic Sparse Decomposition) etc. [48].

In this study, the proposed time-series forecasting model doesn't employ any exogenous variables and thus, feature selection and extraction become redundant. There is a possibility that the peculiarities associated with present hydro and wind power generation may also prevail in future generation scenarios. So, de-noising or data cleaning is not used here to keep the generation data in its original form. Data decomposition is the only processing technique that is used in this paper and outlier detection has been used to remove the visible outliers in peak hydro power generation (Fig. 2). As stated in [48], characteristics of wind or any other renewable power time-series vary according to seasonal, climatic and topographical conditions. Therefore, forecasting model that is developed for a group of datasets may not be applicable to

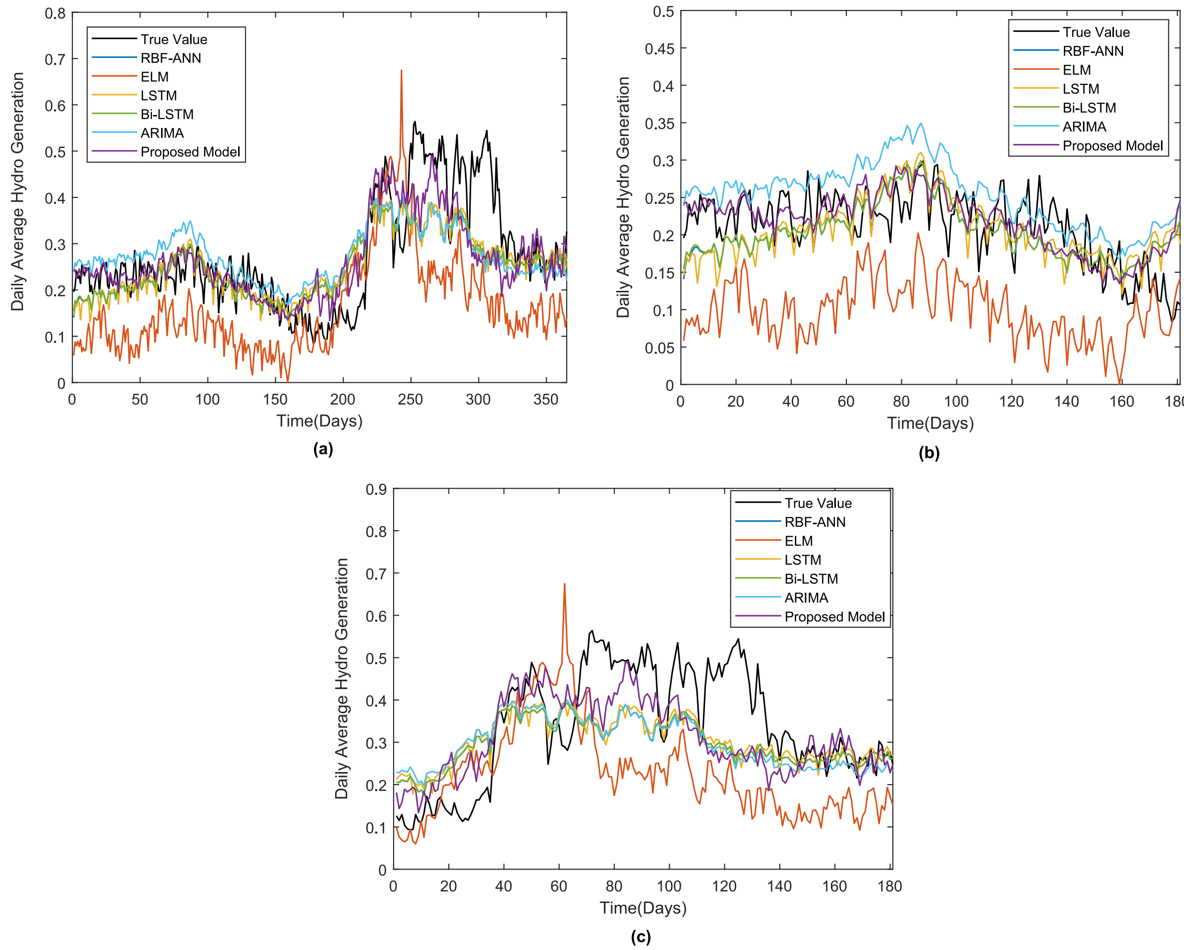

Fig. 22. DAHG testing dataset forecasts for **a)** A year ahead scenario; **b)** Off-season generation; **c)** Seasonal generation.

other datasets belonging to completely different climates and topographies. Hence, this paper mainly focuses on designing a model that can rigorously capture the effect of different seasons on different types of power generation. To design such a model for long-term forecasting application, it becomes essential to first extract the seasonality reflected over longer durations like on yearly/monthly/quarterly basis in the considered time-series. Therefore, STL (Seasonal and Trend decomposition using Loess) decomposition method is used to enhance the effectiveness of the proposed forecasting model. Additionally, since ARIMA alone can't detect non-linear variations of the time-series, hence STL decomposition can help in estimating them. Here, SARIMA (Seasonal ARIMA) model has not been used since they are suitable for the time-series with only single seasonality component. Though EMD decomposition has also been used quite frequently for different forecasting applications but there doesn't exist any study that specifically compares STL and EMD based wind and hydro power forecasting models which could prove superiority of one over the other. However, few other studies, such as [61–63], have established that STL based prediction models yield better results than EMD and EEMD based models especially for long-term prediction. In [61], ANN model combined with STL decomposition performed better than both EMD and EEMD in long-term prediction of water discharge. Similarly, STL-LSTM produced better results than EMD-LSTM in short-term forecasting of metro ridership in [63]. Moreover, STL decomposition offers the advantages of comparable speed and precision with that provided by EMD-LSTM [62].

STL decomposition disintegrates the original time-series into three main components: a) trend, b) seasonal component and c) remainder. It is a versatile method for decomposing a time-series and has several advantages like capability to handle multiple seasonalities such as weekly, monthly or any user defined seasonality simultaneously. The unusual occasional events, if present in the time-series data, don't affect the smooth trend cycle and seasonal components. Instead they are included in the remainder.

3.5. Framework of proposed model

The process of designing the proposed ensemble forecasting model is completed in three phases. The architecture and flowchart of the whole process is shown in Fig. 10. The proposed ensemble model utilizes two different variants viz., 'hybrid ARIMA + Bi-LSTM' and 'ARIMA with DSA' to produce the final forecasts. The former provides improved forecasts in comparison to both the baseline models, while the latter takes care of all prominent seasonal effects manifesting in the renewable power generation time-series. These variants or models are developed in Phase-I and Phase-II, respectively to handle different forms of the training datasets. Training datasets employed to train each model include original training datasets and subsets of the original training datasets. These subsets correspond to seasonal and off-seasonal wind and hydro power generation time-series as shown in Figs. 5-8. In Phase-I, original training datasets are first decomposed using STL decomposition method and then ARIMA and Bi-LSTM are used to generate independent forecasts which are then combined to form 'hybrid ARIMA + Bi-LSTM'. In Phase-II, all the subsets of training datasets are considered as individual time-series that are split further into training and testing datasets. These new training datasets are used to train ARIMA with DSA after STL decomposition. Finally, forecasts of all the seasons are integrated at the end of Phase-II. In Phase-III, mean output of both the models are used to form final ensemble using an additional step of their

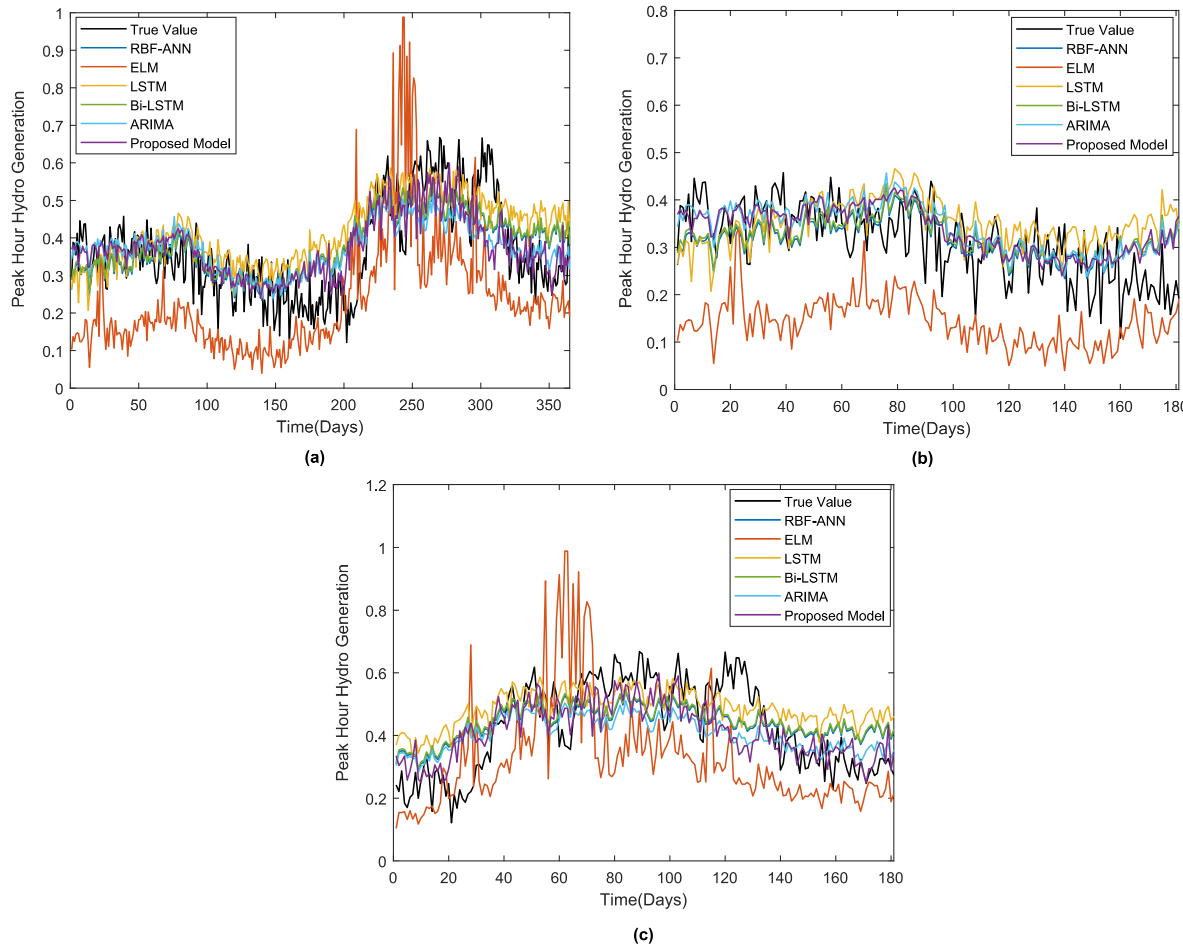

Fig. 23. PHHG testing dataset forecasts for **a)** A year ahead scenario; **b)** Off-season generation; **c)** Seasonal generation.

mutual comparison to avoid any case of over fitting.

3.5.1. Phase-I

This is the first part of the process and it uses the aforementioned baseline models to develop a hybrid (ARIMA + Bi-LSTM) model as per the following steps;

Step –I: Each of the normalized wind and hydro power generation time series, shown in Figs. 3 and 4, is split into training and testing datasets as discussed in Section 2.1.

Step –II: The training datasets are investigated to identify the visual strong seasonalities (S1, S2, S3,….,Sn) present in each power generation time-series. In order to avoid over-fitting of the data, initially only two such seasonalities are considered for constructing the model. One accounts for the annual effect and the other for weekly or monthly or quarterly or any other, whichever is certain and persistent in the respective power generation. For example: S1 = 365.25, S2 = 249 for DAWG and S1 = 365.25, S2 = 60.875 for DAHG (Figs. 3 and 4).

Step –III: After successfully identifying the seasonalities, groups {say Group1: S1, Group 2: (S1, S2), Group n: (S1, S2,…,Sn)} are formed to analyze individual and cumulative effects of discovered seasonalities on power generation. Then the training datasets are decomposed into trend, seasonal and remainder components using STL decomposition according to the groups. The decomposed series for DAWG and DAHG are shown in Figs. 11–14.

Step –IV: All the decomposed time-series are fed as input to ARIMA and Bi-LSTM learning algorithms. Both the models are utilized to predict the seasonally adjusted time-series and combine the predictions with the seasonal components at the end to obtain the final forecasts.

Step –V: The performance of the forecasts obtained from ARIMA as well as Bi-LSTM model is evaluated on the basis of the accuracy measures described in Section 3.3. The final predictions of both the models are compared over different time scales beginning from a week to a year ahead. Those with minimum error on testing datasets are selected as the optimum prediction values.

Step –VI: Finally, accurate forecasts of both the models over different time-horizons are combined to develop an optimum hybrid model.

3.5.2. Phase-II

Even though effect of dominating seasonalities on power generation has been apprehended in Phase-I, it can be observed from Figs. 11–14 that the remainder is not completely random and still holds some essential information about power generation at different instant of time. Further, Fig. 11 & Fig. 12 show that original data and remainder almost follow the same pattern. This indicates that the remainder consists of plenty of significant information that need to be extracted. Similarly, in Fig. 13. Fig. 14 it can be clearly seen that although there is no definite period in the remainder that repeats over time but frequent ups and downs do have a cyclic sequence. Thus, Phase-II is designed to assess the seasonal and off-season generation independently for mitigating this issue and to exploit the additional information contained in the power generation datasets as per the steps given below:

Step –I: Different normalized datasets are divided into different seasons (season 1,…., season n) based upon their inherent seasonal variations. Both hydro and wind have perceptible seasonal and off-season power generation as discussed in Section 2.2. Therefore, each type of power generation is divided into two different periods of generation i.e., high generation period and low generation period.

Step –II: Both high and low generation periods are split into training

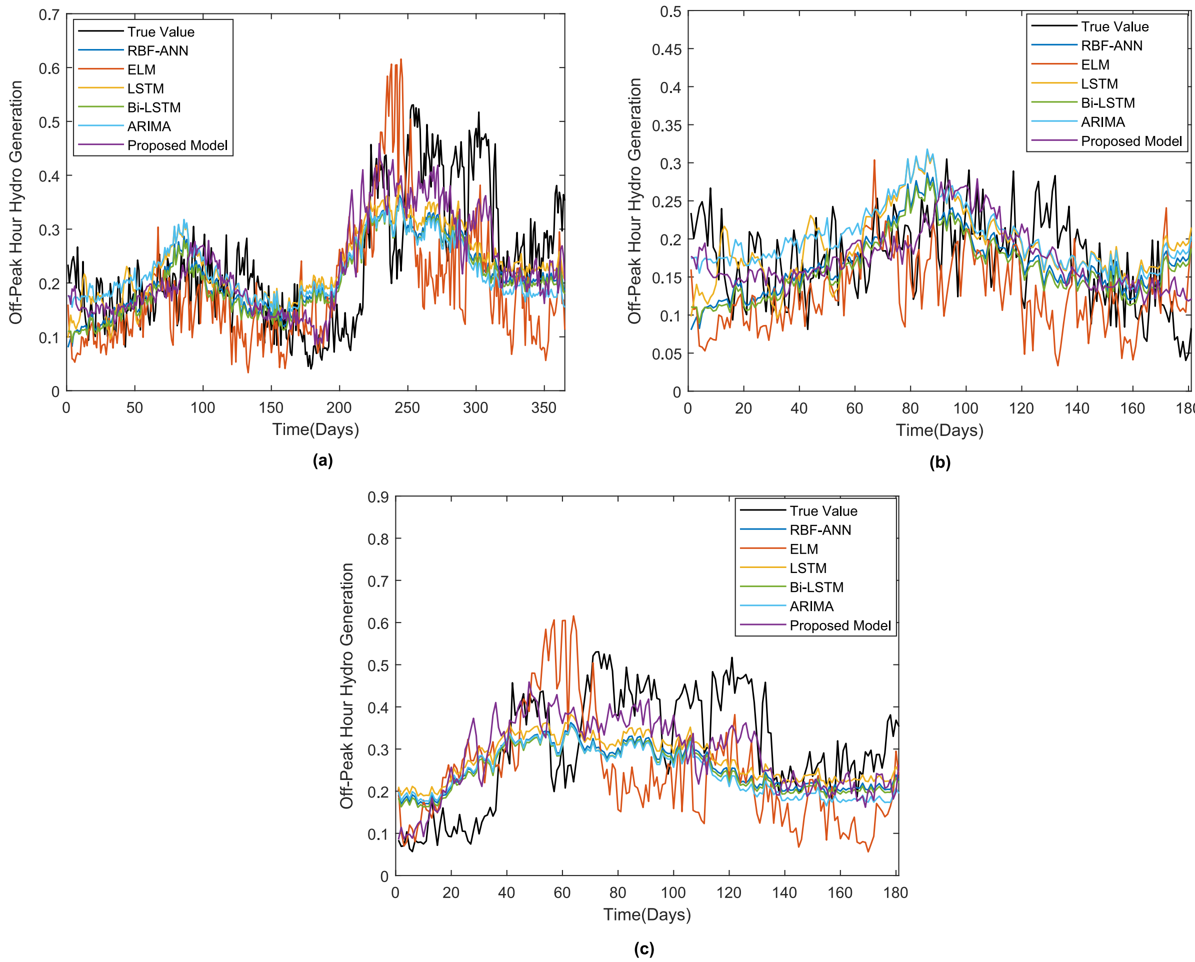

Fig. 24. PHHG testing dataset forecasts for **a)** A year ahead scenario; **b)** Off-season generation; **c)** Seasonal generation.

and testing datasets for all wind and hydro power generation.

Step –III: The training datasets are used to identify the ideal seasonality period of each season. There is an identical seasonality present in each of the two seasons. For instance, seasonal generation of both wind and hydro (Figs. 6 and 8) shows repeatability after every 184 days (6 months) while off-season generation patterns (Figs. 5 and 7) repeat after approximately 181 days (6 months). Hence, ideal seasonality for these two seasons becomes 184 and 181 days, respectively.

Step –IV: Besides ideal seasonality, seasonal generation series exhibit some irregular fluctuations. These fluctuations don't follow any definite pattern and are quirky in nature. In order to analyze the effect of these fluctuations on power generation and to improve the quality of the forecasts, the concept of Diligent Search Algorithm (DSA) is introduced here for the first time. This algorithm is solely designed to identify the hidden seasonalities present in the time-series which consequently address the instant of time at which these fluctuations occur.

Step –V: Once both ideal and hidden seasonalities are determined, the training datasets are decomposed into trend, seasonal components and remainder, using STL decomposition, firstly, as per ideal seasonality and then cumulatively afterwards. The decomposed DAHG and DAWG series are shown in Figs. 15–18.

Step –VI: The decomposed time-series are fed as input to ARIMA model '1' and ARIMA model '2'. Model '1' is the single seasonality model that handles only ideal seasonality while model '2' is multiple seasonality model that takes into account all identified seasonalities. Both the models predict the seasonally adjusted time-series and then re-seasonalizes the predictions to obtain the final forecasts.

Step –VII: The forecasts of both the models are evaluated on the basis of mean absolute error calculated using testing datasets. The model which shows least error is considered as the most appropriate model for the selected season. It is also observed that model '1' is well suited for off-season generation data while model '2' efficiently recognized fluctuations in seasonal generation.

Step –VIII: Finally, the optimum predictions of both the seasons are combined to produce a year ahead forecast for all hydro and wind power generation datasets.

3.5.3. DSA algorithm

In time-series analysis, correlation between current and preceding values decreases as the time span increases. This implies that power generation of current year would have a strong relationship with immediately preceding year generation as compared to that from earlier years. Hence, the effect of recent years' generation would be more on predictions of the future generation. Keeping this in view, the testing datasets are tuned to identify the hidden seasonalities that contain the information related to occasional or irregular events associated with renewable power generation. The pseudo code of DSA algorithm is shown below:

Table 4
Comparison of MAE values obtained by the proposed and existing models on testing datasets.

GenerationType	ForecastingModel	MAE (% of IC) W	M	Q	HY	Y
DAWG	RBF-ANN	2.97	3.09	3.29	3.09	6.35
	ELM	3.44	4.49	5.56	5.01	9.36
	LSTM	3.56	3.75	3.63	3.76	7.02
	Bi-LSTM	2.7	3.05	3.23	2.95	6.2
	ARIMA	3.23	3.18	3.26	2.94	6.16
	Proposed	**2.44**	**2.92**	**2.95**	**2.72**	**4.92**
PHWG	RBF-ANN	2.53	2.79	3.37	3.52	6.75
	ELM	4.3	2.89	5.13	5.2	9.43
	LSTM	**2.49**	2.88	3.51	3.74	6.95
	Bi-LSTM	2.78	**2.72**	3.32	3.4	6.6
	ARIMA	5.3	4.51	4.86	4.55	7
	Proposed	3.16	3.07	**3.2**	**3.23**	**5.52**
OPHWG	RBF-ANN	2.92	3.33	3.86	3.82	6.68
	ELM	4	4.72	6.01	5.88	10.33
	LSTM	3.95	5.56	4.65	4.57	7.31
	Bi-LSTM	2.92	3.06	3.78	3.7	6.5
	ARIMA	2.85	2.96	3.76	3.68	6.34
	Proposed	**1.97**	**2.83**	**3.6**	**3.45**	**5.15**
DAHG	RBF-ANN	3.68	3.44	4.66	5.31	6.38
	ELM	14.5	12.37	11.67	11.31	13.24
	LSTM	6.23	5.4	3.45	3.64	6.21
	Bi-LSTM	4.72	4.5	3.4	3.65	6.37
	ARIMA	3.28	3.16	5.04	4.75	7.28
	Proposed	**2.3**	**1.76**	**2.3**	**2.82**	**5.47**
PHHG	RBF-ANN	6.6	6.06	5.35	5.83	7.68
	ELM	23.8	20.15	18.73	17.04	16.87
	LSTM	8.25	6.29	5.95	7.04	8.74
	Bi-LSTM	5.6	5.53	5.2	5.81	7.72
	ARIMA	4.17	4	4.84	5.5	7.54
	Proposed	**3.84**	**3.62**	**4.28**	**5.05**	**6.42**
OPHHG	RBF-ANN	12.78	6.86	4.6	4.9	7.62
	ELM	13.65	8.71	5.9	6.39	10.46
	LSTM	10.28	6.12	5.31	5.45	7.6
	Bi-LSTM	11.73	6.8	4.52	4.86	7.72
	ARIMA	5.26	4.72	5.3	5.33	8.41
	Proposed	**4.64**	**3.87**	**3.65**	**3.86**	**6.17**

Algorithm:. (*DSA*)

Objective function: min $\{f = MAE = \frac{1}{n}\sum_{J=1}^{n}|x_j - y_j|\}$
Testing dataset: X $\{x_1, x_2, \cdots\cdots, x_n\}$
Forecast output: Y $\{y_1, y_2, \cdots\cdots, y_n\}$
Parameters:
$\boldsymbol{n}$ = length of testing dataset
$\boldsymbol{Z_d}$ = Decomposed Training data
Model '1': Single seasonality ARIMA
Model '2': Multiple seasonality ARIMA
new_k: Hidden seasonality

1 /* Initial seasonality = ideal seasonality period. */
2 /* Apply Model '1' on $\boldsymbol{Z_d}$ with initial seasonality. */
3 /* Calculate Objective function value $\boldsymbol{f}$. */
4 /* Initialize error = $\boldsymbol{f}$. */
5 /* Define **N** as ones vector of size n. */
6 **FOR** k in (2:n) **DO**
7 /* Create multi-time series object of training data with seasonal periods= (k, n). */
8 /* Apply Model '2' on $\boldsymbol{Z_d}$ to produce **Y**. */
9 /* Update value of $\boldsymbol{f}$. */
10 **IF** $\boldsymbol{f}$ < error **DO**
11 N(k) = $\boldsymbol{f}$.
12 **END IF**
13 **END FOR**
14 error = min (N)
15 /* new_k = index of 'error' in N. */
16 /* Update the seasonal periods with (k, new_k, n). */
17 /* Repeat the procedure from code line 5 to 14. */
18 **IF** error = $\boldsymbol{f}$ **DO**
19 **STOP**
20 **ELSE DO**
21 /* Go to code line 15. */
22 **END IF**

This algorithm is applicable for all the seasons as discussed in Step –I of Phase –II.

3.5.4. Phase-III

In this last phase of the process both Phase-I and Phase-II forecast results are scrutinized to develop an optimum and robust ensemble model. 'Ensemble 1' is produced by taking the average of hybrid model obtained from Phase-I (denoted by 'A') and Phase-II model (denoted by 'B'). Here, Model 'A' acts as the benchmark model for Model 'B' so that the forecasts of 'B' are not biased with respect to the identified hidden seasonalities. On the other hand, it is also important to maintain the optimum value of the forecasts as obtained from 'B'. Therefore, power generation forecasts of 'Ensemble 1' are further compared with that obtained from 'B' over different time horizons as detailed in Table 1. Hence, final ensemble model is prepared after integrating the forecasts of 'Ensemble 1' and 'B'.

Thus, the proposed ensemble model has used two different types of models or variants to produce accurate forecasts over long-term. The first model is trained on the entire training datasets to extract annual effects from power generation time-series while the second model deals with both the seasonal and off-seasonal power generation individually. In general, complexity of an ensemble model increases with the use of multiple models and varying training datasets which in turn increases the computational cost and training time. But in this paper, complexity is kept at minimum by using ARIMA and Bi-LSTM only and by employing same training datasets used in different ways for training them as explained in Section 3.5 above. However, the proposed method would entail more training time with the increase in the number of seasons during Phase-II and with increased prediction horizon as Bi-LSTM reaches the equilibrium state much slower than ARIMA and

Table 5
Comparison of RMSE values obtained by the proposed and existing models on testing datasets.

Generation Type	ForecastingModel	RMSE (% of IC) W	M	Q	HY	Y
DAWG	RBF-ANN	3.75	3.74	4.15	4.07	8.94
	ELM	4.19	5.02	6.92	6.27	13.13
	LSTM	4.14	4.31	4.44	4.58	9.62
	Bi-LSTM	3.56	3.68	4.11	4	8.78
	ARIMA	4.02	3.8	4.15	4.03	8.72
	Proposed	**3.05**	**3.53**	**3.77**	**3.71**	**7.14**
PHWG	RBF-ANN	3.51	3.46	4.5	4.81	9.2
	ELM	5.07	3.5	7.15	7.1	13.11
	LSTM	**3.2**	3.46	4.58	4.95	9.37
	Bi-LSTM	3.51	**3.4**	4.48	4.75	9.07
	ARIMA	6.31	5.65	6.27	6.19	9.37
	Proposed	3.93	3.8	**4.31**	**4.52**	**7.81**
OPHWG	RBF-ANN	3.38	3.8	4.65	4.76	8.92
	ELM	4.78	5.5	7.58	7.38	14.34
	LSTM	4.76	6.36	5.5	5.5	9.45
	Bi-LSTM	3.23	3.54	4.63	4.68	8.75
	ARIMA	3.13	**3.46**	4.68	4.69	8.61
	Proposed	**2.79**	**3.46**	**4.5**	**4.5**	**7.34**
DAHG	RBF-ANN	5.54	5.17	4.12	4.45	8.62
	ELM	14.68	12.9	12.41	12.33	15.34
	LSTM	6.46	5.75	4.16	4.46	8.46
	Bi-LSTM	5.01	5.02	4.08	4.43	8.61
	ARIMA	3.72	3.82	6.03	5.87	9.44
	Proposed	**2.63**	**2.13**	**2.89**	**3.7**	**7.8**
PHHG	RBF-ANN	7.61	7.05	6.63	7.17	9.33
	ELM	24.29	21.37	19.94	18.56	19.44
	LSTM	8.93	7.47	7.44	8.79	10.74
	Bi-LSTM	6.77	6.55	6.5	7.22	9.41
	ARIMA	**4.67**	5.04	6.41	7.02	9.46
	Proposed	4.85	**4.68**	**5.53**	**6.41**	**8.48**
OPHHG	RBF-ANN	13.04	8.59	6.05	6.21	9.7
	ELM	14.33	10.45	7.57	8.02	13.29
	LSTM	10.5	7.15	6.58	6.68	9.46
	Bi-LSTM	11.97	8.42	5.94	6.14	9.87
	ARIMA	5.77	5.26	6.29	6.35	10.45
	Proposed	**4.87**	**4.54**	**4.5**	**4.7**	**8.2**

other machine learning methods. In order to evaluate the performance and accuracy of the proposed model in long-term forecasting, it is compared with the baseline models and some other efficient machine learning algorithms such as ELM, RBFNN, and LSTM etc. with appropriate parameters.

Classical ELM method with random initialization of weights, biases and 'radbas' activation function has been used for comparison. It has a single hidden layer with 200 neurons and one layer each for input and output. RBFNN also has one layer each for input and output but with single neuron. The default number of neurons for the hidden layer in RBFNN is usually specified in the simulation software, for instance in this paper, maximum number of epochs and neurons were set as 150 and 25, respectively. LSTM network has one input layer and one output layer followed by LSTM layer with 200 hidden units. Other parameters of LSTM were kept same as used for Bi-LSTM network (shown in Fig. 9) for the purpose of an exhaustive comparison.

4. Results and discussion

The proposed forecasting model is applied to various wind and hydro power generation datasets as mentioned in Section 2.1. This model deals with seasonal and off season generation in an efficient manner to improve the performance of the model. Therefore, if different seasons of a power generation could be determined properly then this model can be applied to any renewable energy.

A year ahead forecast for all the six types of wind and hydro power generation are shown in Figs. 19–24 (a). Figs. 19–24 (b) & (c) show the seasonal and off-season generation forecasts. It can be noticed here that the intermittent characteristics of seasonal generation period are efficiently computed by the proposed model specifically for wind power. Unlike seasonal wind generation, seasonal generation of hydro power exhibit periodicity with high variance from one cycle to another (Fig. 8). Thus, it was quite difficult to obtain accurate and precise forecasts over a long-time scale of one year for such generation patterns. However, the proposed model successfully managed to provide satisfactory prediction results for these inconsistent hydro power generations. The forecasting errors measured to substantiate the accuracy of the model for different types of generation over different time-horizons (from a week ahead to a year ahead) are tabulated in Tables 4-6. In order to justify the effectiveness of the proposed model, prediction results are compared with baseline models ARIMA & Bi-LSTM and other intelligent machine learning models such as RBF-ANN, ELM, and LSTM. The evaluation unit of MAE and RMSE is percentage of installed capacity (% of IC). In Tables 4-6, bold values denote the least prediction error with respect to different time-horizons and type of generation.

4.1. Performance of proposed model

In a week ahead forecast, minimum MAE is obtained for off-peak wind with a value of 1.97%. The maximum MAE is obtained for off-peak hydro with a value of 4.64%. The minimum value of RMSE and MAAPE is 2.63% and 0.104 for average hydro and peak hydro generation respectively. The maximum value of RMSE and MAAPE is obtained for off-peak hydro and peak wind with values of 4.87% and 0.36 respectively. For a month ahead prediction, the best response of MAE, RMSE and MAAPE are obtained for average hydro generation with values of 1.76%, 2.13% and 0.08 respectively. The maximum values of MAE, RMSE and MAAPE are recorded as 3.87%, 4.68% and 0.38 for off-peak hydro, peak hydro and peak wind generation. Range of MAE, RMSE and MAAPE are 2.3% to 4.28%, 2.89% to 5.53% and 0.1 to 0.36, respectively for quarterly forecasts. These minimum and maximum values of MAE and RMSE are obtained for average hydro and peak hydro generation while peak wind has maximum MAAPE. In case of half yearly forecasts, MAE, RMSE and MAAPE varies from 2.72% to 5%, 3.7% to 6.4% and 0.14 to 0.33. For a year ahead forecast, the best response of MAE, RMSE and MAAPE are recorded as 4.92%, 7.1% and 0.191 for average wind and peak hydro generation. The worst response of MAE, RMSE and MAAPE are 6.4%, 8.4% and 0.33 for peak hydro and peak wind generation. According to [6], RMSE for wind power forecasting must be within 10% of installed capacity and this value may vary as per the prescribed standards of different countries. For example, presently China is the leading country in wind energy where maximum acceptable limit of RMSE is 20% for short term wind power forecasting. After carefully analyzing the prediction results over all selected time-scales, it is found that MAE, RMSE and MAAPE varies from 1.7% to 6.4%, 2.13% to 8.4%, and 0.08 to 0.38 respectively. The most accurate forecasts are achieved for DAWG, DAHG, OPHWG, and PHHG out of total six datasets considered in this paper.

4.2. Comparison with other models

Accuracy of the proposed ensemble model in medium to long-term forecasts is compared with RBFNN, ELM, LSTM and baseline models ARIMA and Bi-LSTM as shown in Tables 4-6. This comparison not only

Table 6
Comparison of MAAPE values obtained by the proposed and existing models on testing datasets.

Generation Type	ForecastingModel	MAAPE				
		W	M	Q	HY	Y
DAWG	RBF-ANN	0.2632	0.3527	0.3339	0.3091	0.3786
	ELM	0.3483	0.4756	0.4921	0.4528	0.4738
	LSTM	0.3309	0.4254	0.3732	0.385	0.4281
	Bi-LSTM	0.2377	0.3401	0.3204	0.2869	0.364
	ARIMA	0.2802	0.3461	0.3189	0.2815	0.356
	Proposed	**0.2357**	**0.3391**	**0.3041**	**0.2759**	**0.3015**
PHWG	RBF-ANN	**0.2515**	0.3480	0.3769	0.3624	0.3903
	ELM	0.4738	0.3438	0.4785	0.4669	0.4620
	LSTM	0.2648	0.3680	0.3983	0.3890	0.4075
	Bi-LSTM	0.2869	**0.3296**	0.3631	0.3429	0.3763
	ARIMA	0.4678	0.4280	0.4372	0.391	0.3885
	Proposed	0.3606	0.3843	**0.3612**	**0.3378**	**0.3344**
OPHWG	RBF-ANN	0.3341	0.3958	0.3739	0.3617	0.4163
	ELM	0.3828	0.4747	0.4838	0.4757	0.5003
	LSTM	0.4078	0.5702	0.4487	0.4334	0.4627
	Bi-LSTM	0.3324	0.36	0.3544	0.3406	0.4
	ARIMA	0.325	0.3441	0.3454	0.331	0.3873
	Proposed	**0.2563**	**0.3374**	**0.3351**	**0.3146**	**0.3191**
DAHG	RBF-ANN	0.2299	0.1955	0.1485	0.1850	0.2328
	ELM	0.572	0.49	0.4548	0.4699	0.4378
	LSTM	0.2682	0.2295	0.1486	0.1836	0.2291
	Bi-LSTM	0.2038	0.1885	0.1468	0.1840	0.2323
	ARIMA	0.1511	0.1458	0.223	0.2434	0.2751
	Proposed	**0.1059**	**0.0815**	**0.1042**	**0.1489**	**0.2013**
PHHG	RBF-ANN	0.1661	0.1587	0.1547	0.2036	0.2288
	ELM	0.5672	0.4979	0.4750	0.4784	0.4142
	LSTM	0.2126	0.1644	0.1768	0.2501	0.2718
	Bi-LSTM	0.1387	0.1453	0.1522	0.2056	0.2323
	ARIMA	0.1170	0.1173	0.1502	0.2008	0.2190
	Proposed	**0.104**	**0.1044**	**0.1331**	**0.1869**	**0.1913**
OPHHG	RBF-ANN	0.513	0.3147	0.2387	0.2897	0.3239
	ELM	0.5361	0.4015	0.2860	0.3392	0.4002
	LSTM	0.4264	0.3318	0.2874	0.3307	0.3413
	Bi-LSTM	0.4776	0.3121	0.2322	0.2839	0.3226
	ARIMA	0.2225	0.2788	0.3022	0.3344	0.3630
	Proposed	**0.2012**	**0.2112**	**0.2058**	**0.2431**	**0.2715**

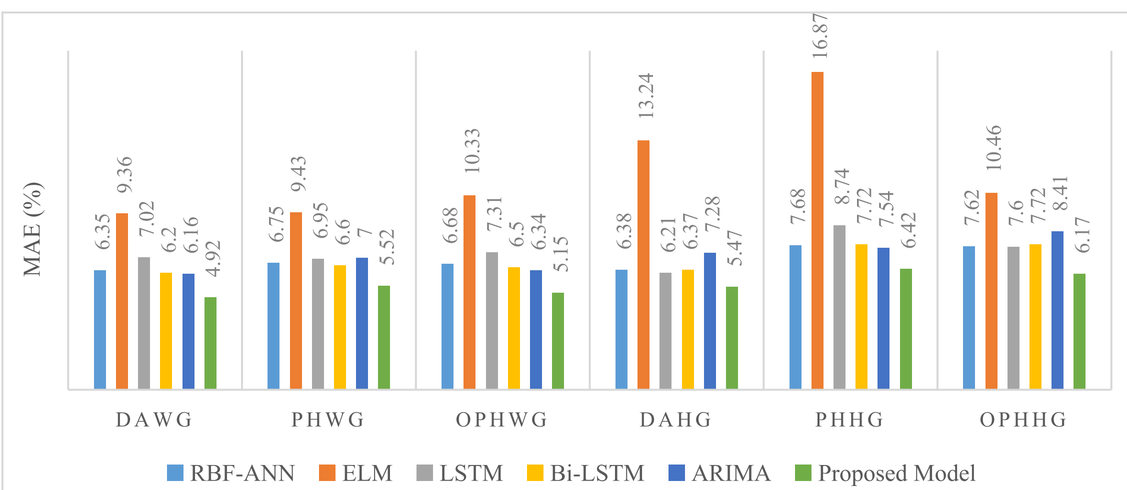

Fig. 25. MAE (%) values of all wind and hydro datasets in a year-ahead forecast.

illustrates the performance of proposed model over different time horizons but also gives an insight about the reason behind choosing ARIMA & Bi-LSTM for developing the proposed ensemble. ARIMA and Bi-LSTM models produce least error as compared to RBF-ANN, ELM, and LSTM when tested individually over a week-ahead to a year-ahead forecasting scenario for both wind and hydro power generations. All the evaluation metrics indicate that the proposed model outperformed both ARIMA & Bi-LSTM and other machine learning models in long-term forecasting of power generation. MAE, RMSE and MAAPE values of all the considered models on all the selected six datasets for a year-ahead forecast are shown in Figs. 25-27. Here, it can be observed that MAE is reduced by a maximum value of 4.44%, 3.91%, 5.18%, 7.77%, 10.45% and 4.29% for DAWG, PHWG, OPHWG, DAHG, PHHG, and OPHHG, respectively. Similarly, reduction in RMSE and MAAPE for these datasets vary from 5.09% to 10.96% and 0.128 to 0.236, respectively. Further, improvement percentages of baseline models ARIMA and Bi-LSTM are computed

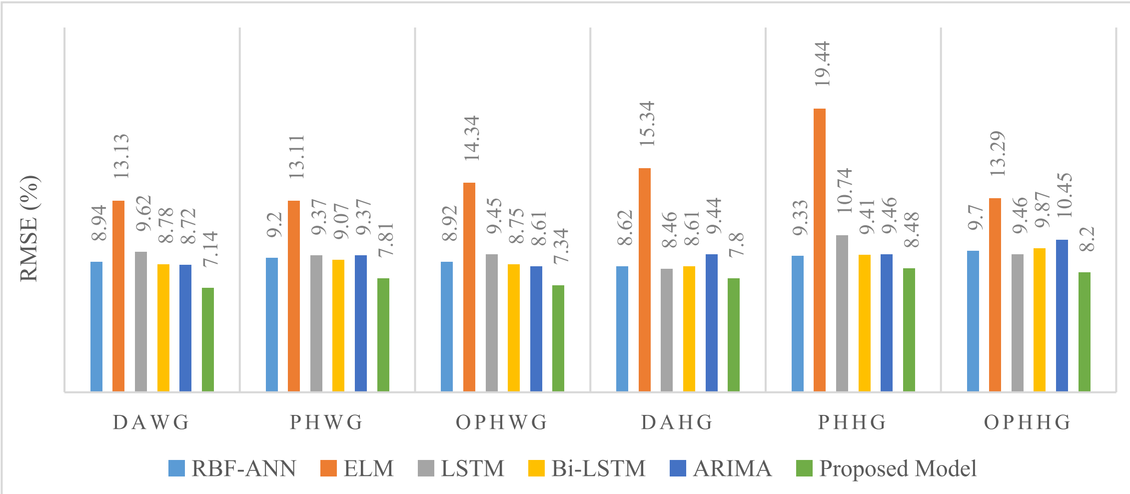

Fig. 26. RMSE (%) values of all wind and hydro datasets in a year-ahead forecast.

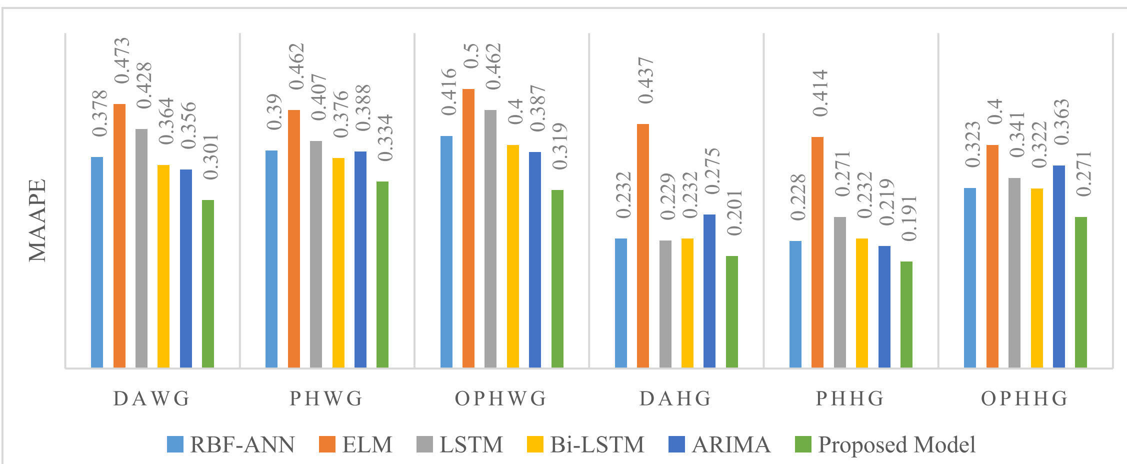

Fig. 27. MAAPE values of all wind and hydro datasets in a year-ahead forecast.

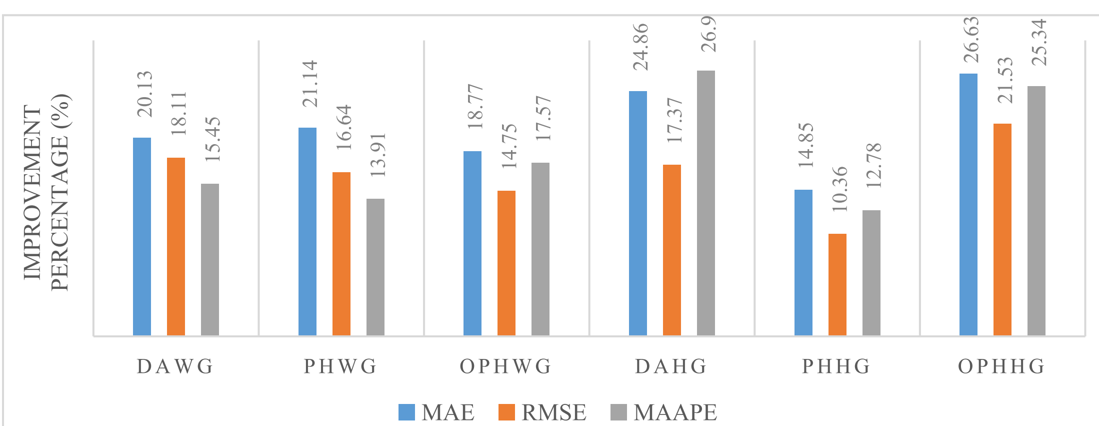

Fig. 28. Improvement percentages of ARIMA by proposed model for a year-ahead forecasts.

with respect to proposed model for a year-ahead forecasts as shown in Figs. 28 & 29. Highest improvements observed in MAE of ARIMA and Bi-LSTM are 26.63% and 20.77%, respectively. RMSE of ARIMA and Bi-LSTM are improved by 21.53% and 18.67%, respectively. And MAAPE of these two models are improved by 26.9% and 20.25%.

4.3. Comparison with existing studies

The performance of the proposed model is also compared with other published studies which have used same performance indicators with same evaluation units as considered in this paper (Table 7). In 2019,

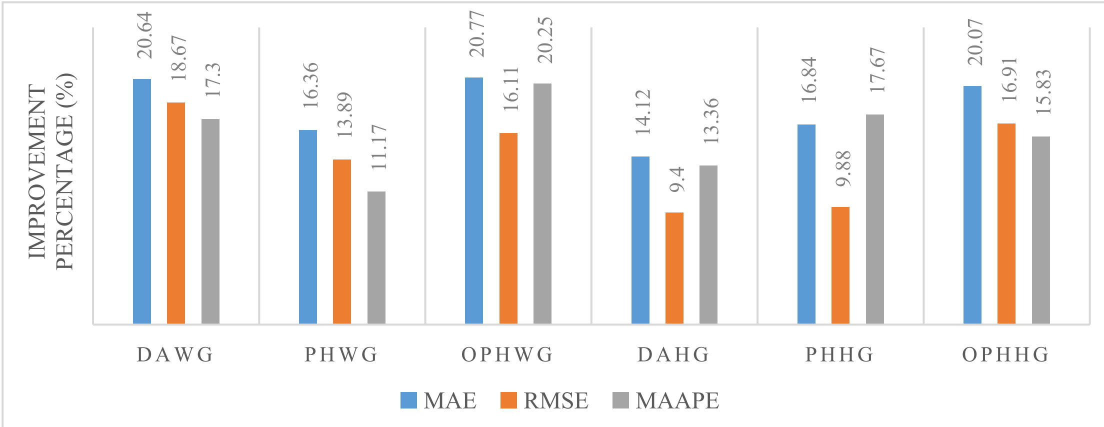

Fig. 29. Improvement percentages of Bi-LSTM by proposed model for a year-ahead forecasts.

Table 7
Comparison of proposed model forecasting errors with the existing studies.

Generation Type	Evaluation metric	Proposed Model		Existing Studies									
				1 h ahead to a day ahead forecasts [28]		540 to 3140 s forecasts [49]		3 steps-ahead (18 h) forecasts [8]		2 to 4 h ahead forecasts [50]		3 to 21 days ahead forecasts [59]	
		Min	Max	Min	Max	Min	Max	Min	Max	Min	Max	Min	Max
Wind Power	MAE	1.97% (W)	5.52% (Y)	6.63%	22.63%	–	–	–	–	–	–	–	–
	RMSE	2.79% (W)	7.81% (Y)	8.37%	27.8%	6.37%	31.29%	–	–	7%	13%	–	–
	MAAPE	0.235 (W)	0.384 (M)	–	–	–	–	0.356	0.507	–	–	–	–
Hydro Power	MAE	1.76% (M)	6.42% (Y)	–	–	–	–	–	–	–	–	5.75%	10.81%
	RMSE	2.13% (M)	8.48% (Y)	–	–	–	–	–	–	–	–	9%	15%
	MAAPE	0.081 (M)	0.271 (Y)	–	–	–	–	–	–	–	–	–	–

Zhang et al., [49] developed a hybrid wind power forecasting model employing LSTM with RMSE reported as 6.37%. In 2020, Sun et al., [28] introduced an ensemble probabilistic multi-step ahead wind power forecasting method in which range of MAE for six steps ahead forecast was 5.38% to 22.63% and RMSE was 8.37% to 27.8%. In 2020, [8] presented a multi-learner ensemble approach for medium term wind power forecasting using four wind power series. Minimum and maximum values of MAAPE for the four datasets varied from 0.35 to 0.5 in three steps ahead forecasting. Contrastingly, in present work, MAE, RMSE and MAAPE for wind power generation forecasts varies from 1.97% to 5.5%, 2.8% to 7.8% and 0.23 to 0.38 respectively. In [59], ELM based hybrid model was developed for small hydro power generation forecasting for a period of 3 to 21 days-ahead. MAE and RMSE were reported as 5.75% to 10.81% and 9% to 15%, respectively. While, the range of MAE and RMSE for hydro power forecasts in the proposed work is 1.76% to 6.42% and 2.13% to 8.48%. Hence, the existing studies developed different models for short-term to medium-term wind power forecasting while better prediction results are obtained in this paper for much longer time horizons.

4.4. Validation datasets results

Efficacy and versatility of the proposed ensemble model is further verified by using unknown validation datasets of wind and hydro power generation (Figs. 30 & 31). These datasets include wind power generation from November 2020 to April 2021 and hydro power generation from January 2021 to April 2021. The proposed model is used to forecast the power generation for these unknown generation datasets. Forecast errors thus obtained are shown in Table 8 & 9. In case of wind power generation the best response of MAE, RMSE and MAAPE is acquired for average wind generation with errors of 1.6%, 2% and 0.22 for a week ahead forecast. For hydro power generation, the best values of MAE, RMSE and MAAPE are 2.56%, 3.6% and 0.11 over weekly and monthly forecasts.

Generally, forecasting error increases by significant amount with the increasing time-horizon or time steps. However, this study don't reflect any major increment in MAE and RMSE values for different datasets from a week-ahead to half-yearly forecasts. In addition to this, MAAPE values also remain constant over the entire selected time-horizon. Therefore, the consistency achieved in keeping the prediction error less over longer time-horizon verifies the robustness of the proposed model. Hence, all the performance indicators confirm the aptness of the perceptive forecasting model for long-term forecasting of wind and hydro power generation.

5. Conclusion

Long-term forecasting of renewable power generation plays an important role in planning of generation mix scenarios. In this paper, a novel ensemble based univariate time-series model is presented for long-term forecasting of wind as well as hydro power generation. The strength of the proposed forecasting model lies in accurately detecting the seasonal effects of renewable power generation with the aid of a novel algorithm 'DSA'. The proposed model utilized ARIMA and Bi-

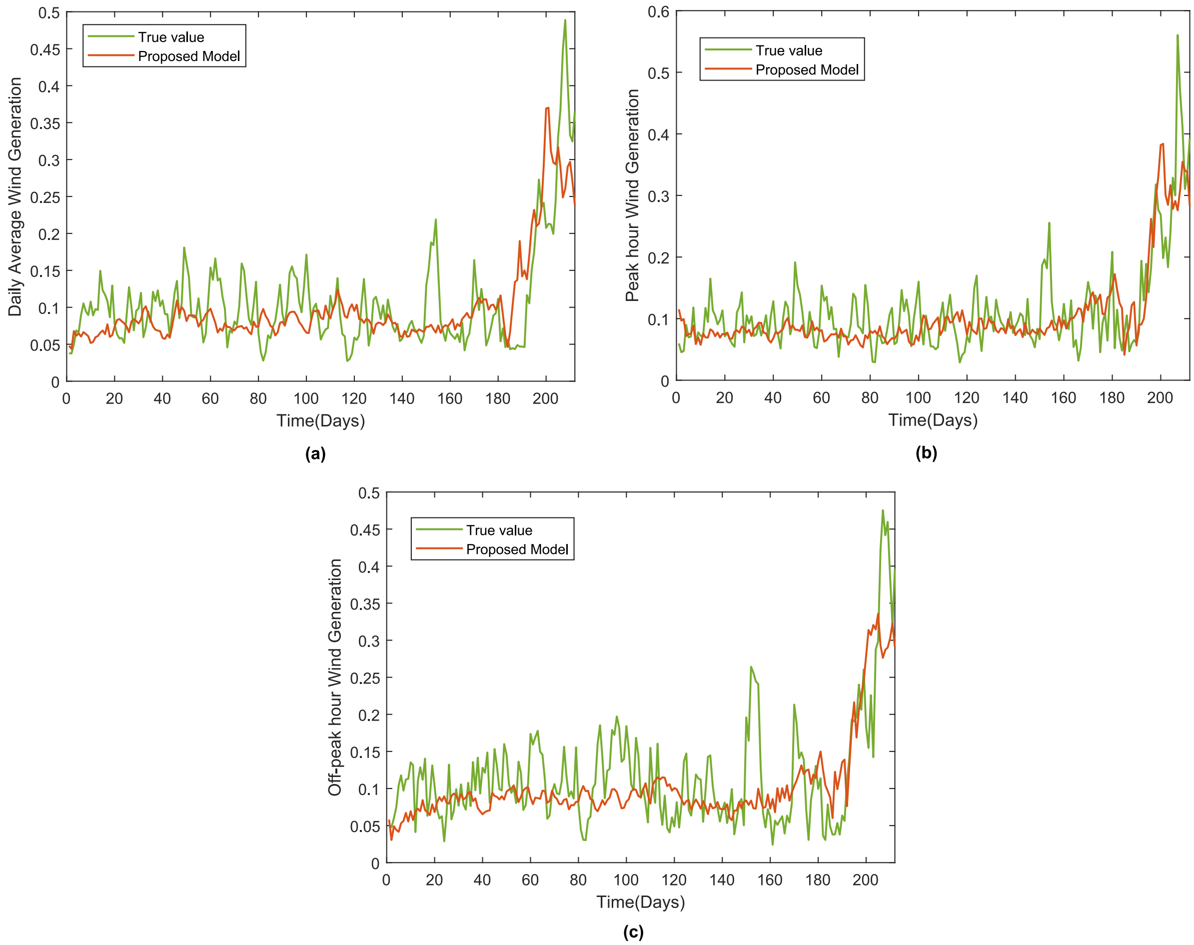

Fig. 30. Proposed Model forecasts for validation dataset of **a)** DAWG; **b)** PHWG; **c)** OPHWG.

LSTM to exploit the strongly persistent characteristics of the selected datasets and the search algorithm DSA helped in improving the overall performance by computing the hidden seasonality periods of the variable datasets. This model offers several advantages in terms of accuracy, uniformity and diversity:

- Accuracy: Both the testing datasets and validation datasets of wind and hydro power generation establish accuracy of the model. The obtained results show that the proposed model outperformed RBFNN, ELM, LSTM, ARIMA and Bi-LSTM in terms of accuracy. It provides accurate forecasts over long time horizon from a week-ahead to a year-ahead generation scenario. MAE, RMSE, and MAAPE values vary from 4.92% to 6.42%, 7.14% to 8.48%, and 0.1913 to 0.3344, respectively for a year-ahead forecast. For wind power, minimum to maximum value range of MAE, RMSE, and MAAPE are 1.97% to 5.52%, 2.79% to 7.8%, and 0.235 to 0.384, respectively. Similarly MAE, RMSE, and MAAPE are 1.76% to 6.42%, 2.13% to 8.48%, and 0.081 to 0.271, respectively for hydro power. MAE of DAWG, PHWG, OPHWG, DAHG, PHHG, and OPHHG in a year-ahead forecast are 4.92%, 5.52%, 5.15%, 5.47%, 6.42%, and 6.17%, respectively. RMSE of DAWG, PHWG, OPHWG, DAHG, PHHG, and OPHHG in a year-ahead forecast are 7.14%, 7.81%, 7.34%, 7.8%, 8.48%, and 8.2%, respectively. MAAPE of DAWG, PHWG, OPHWG, DAHG, PHHG, and OPHHG in a year-ahead forecast are 0.301, 0.334, 0.319, 0.201, 0.191, and 0.271, respectively. Maximum values of MAE and RMSE of wind in this paper are 17% and 23% less as compared to those reported in existing studies.
- Diversity: Total six datasets of wind and hydro power generation are considered in this study to validate the performance of the proposed model. The proposed forecasting model provides accurate results for two different renewable energy sources establishing its diversity. This quality of the model makes it adaptable to provide long term forecasts for other renewable energy sources that have identifiable and separable high and low power generation periods. In this way, the proposed univariate time-series model eliminates the need of different weather data for different renewable sources.
- Uniformity: Unlike other forecasting methods the prediction error doesn't increase continuously with increasing time scale in the proposed method. The change in prediction errors from one time step to another is steady and small. MAE and RMSE values increases slightly for each time step while the changes in MAAPE values are completely uniform.

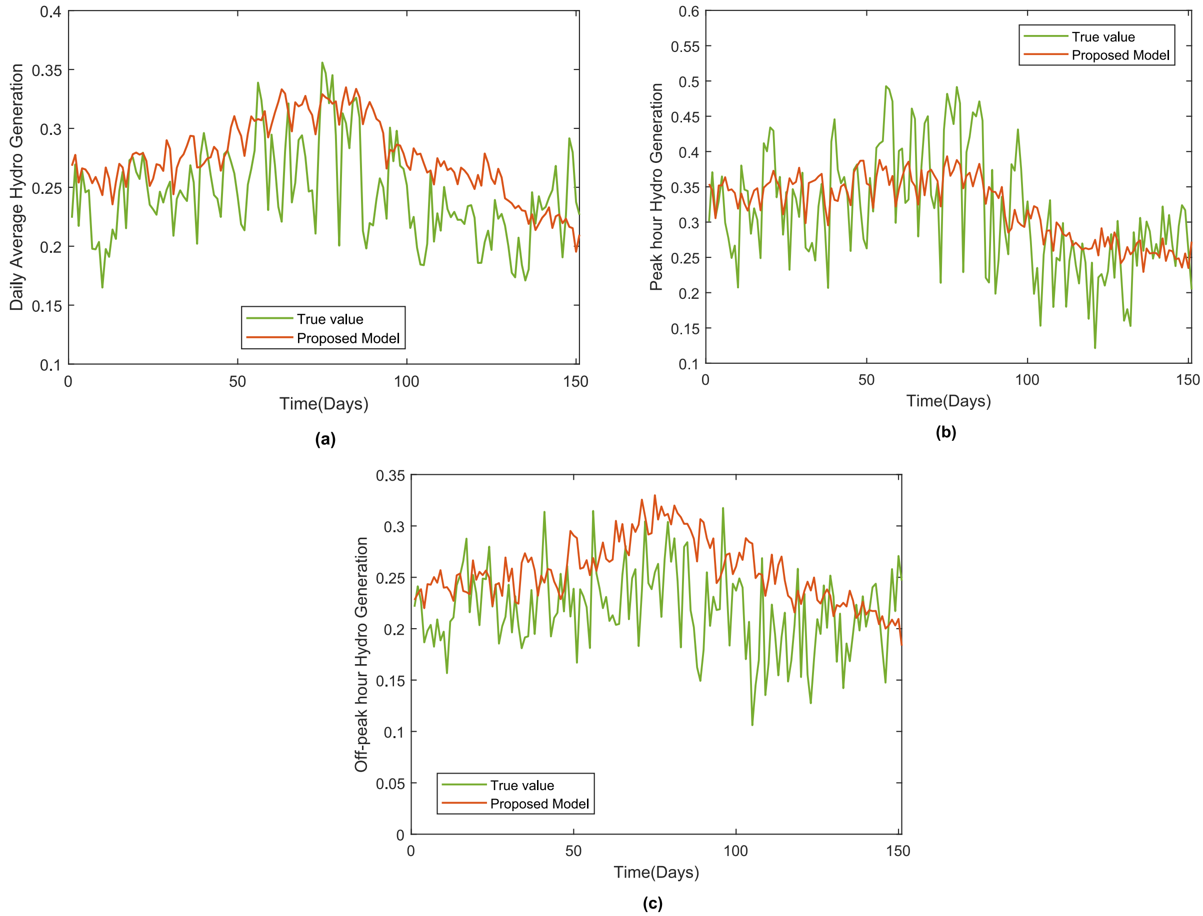

Fig. 31. Proposed Model forecasts for validation dataset of **a)** DAHG; **b)** PHHG; **c)** OPHHG.

Table 8
Proposed model forecasting errors on wind power generation validation datasets.

Generation Type	MAE (% of IC)				RMSE (% of IC)				MAAPE			
	W	M	Q	HY	W	M	Q	HY	W	M	Q	HY
DAWG	1.6	3.07	3.11	3.02	2.03	3.67	3.92	4	0.224	0.315	0.294	0.302
PHWG	3.08	2.62	2.85	3.12	3.71	3.35	3.71	4.25	0.43	0.28	0.287	0.308
OPHWG	3.36	3.76	3.86	4.02	4	4.3	4.8	5.23	0.362	0.387	0.325	0.354

Table 9
Proposed model forecasting errors on hydro power generation validation datasets.

Generation Type	MAE (% of IC)			RMSE (% of IC)			MAAPE		
	W	M	Q	W	M	Q	W	M	Q
DAHG	2.56	2.95	3.86	10.98	11.2	9.64	0.427	0.444	0.353
PHHG	3.52	4.57	5.83	4.29	5.58	7.02	0.114	0.152	0.185
OPHHG	3.01	2.88	4.83	3.69	3.67	6.03	0.151	0.138	0.227

6. Study limitations

The proposed forecasting model is based on univariate time series prediction that gives accurate forecasts up to a year-ahead scenario. Since this study considered renewable power generation characteristics that vary on annual and seasonal basis, the accuracy of the model can become unstable while producing forecasts for much longer time horizon say two to three year-ahead scenarios. Also, the model has a complex structure which can be further simplified in future without sacrificing the accuracy of the model.

CRediT authorship contribution statement

Priyanka Malhan: Conceptualization, Methodology, Software, Writing – original draft. **Monika Mittal:** Investigation, Supervision, Writing – review & editing.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

References

[1] REN21, *Renewables 2020 Global Status Report*. 2020. [Online]. Available: https://www.ren21.net/wp-content/uploads/2019/05/GSR2021_Full_Report.pdf.
[2] Central Eectricity Authority, "All India Installed Capacity (in MW) of Power Stations," *Cent. Eectricity Authority, Minist. power*, vol. 4, pp. 1–7, 2021. [Online]. Available: https://cea.nic.in/wp-content/uploads/installed/2021/03/installed_capacity.pdf.
[3] Foley AM, Leahy PG, Marvuglia A, McKeogh EJ. Current methods and advances in forecasting of wind power generation. Renew. Energy 2012;37(1):1–8. https://doi.org/10.1016/j.renene.2011.05.033.
[4] Wang H, Lei Z, Zhang X, Zhou B, Peng J. A review of deep learning for renewable energy forecasting. Energy Convers. Manag. 2019;198:111799. https://doi.org/10.1016/j.enconman.2019.111799.
[5] J. P. Lai, Y. M. Chang, C. H. Chen, and P. F. Pai, "A survey of machine learning models in renewable energy predictions," Appl. Sci., vol. 10, no. 17, 2020, doi: 10.3390/app10175975.
[6] Hanifi S, Liu X, Lin Z, Lotfian S. A Critical Review of Wind Power Forecasting Methods-Past, Present and Future. Energies 2020;13(15):1–24. https://doi.org/10.3390/en13153764.
[7] Zendehboudi A, Baseer MA, Saidur R. Application of support vector machine models for forecasting solar and wind energy resources: A review. J. Clean. Prod. 2018;199:272–85. https://doi.org/10.1016/j.jclepro.2018.07.164.
[8] Chen C, Liu H. Medium-term wind power forecasting based on multi-resolution multi-learner ensemble and adaptive model selection. Energy Convers. Manag. 2020;206:112492. https://doi.org/10.1016/j.enconman.2020.112492.
[9] Okumus I, Dinler A. Current status of wind energy forecasting and a hybrid method for hourly predictions. Energy Convers. Manag. 2016;123:362–71. https://doi.org/10.1016/j.enconman.2016.06.053.
[10] D. B. De Alencar, C. De Mattos Affonso, R. C. L. De Oliveira, J. L. M. Rodríguez, J. C. Leite, and J. C. R. Filho, "Different Models for Forecasting Wind Power Generation: Case Study," *Energies*, vol. 10, no. 12, 2017, doi: 10.3390/en10121976.
[11] da Silva RG, Ribeiro MHDM, Moreno SR, Mariani VC, dos S. Coelho L. A novel decomposition-ensemble learning framework for multi-step ahead wind energy forecasting. Energy 2021;216:119174. https://doi.org/10.1016/j.energy.2020.119174.
[12] Sharifian A, Ghadi MJ, Ghavidel S, Li L, Zhang J. A new method based on Type-2 fuzzy neural network for accurate wind power forecasting under uncertain data. Renew. Energy 2018;120:220–30. https://doi.org/10.1016/j.renene.2017.12.023.
[13] Li L-L, Zhao X, Tseng M-L, Tan RR. Short-term wind power forecasting based on support vector machine with improved dragonfly algorithm. J. Clean. Prod. 2020;242:118447. https://doi.org/10.1016/j.jclepro.2019.118447.
[14] S. S. Soman, H. Zareipour, O. Malik, and P. Mandal, "A review of wind power and wind speed forecasting methods with different time horizons," 2010, doi: 10.1109/NAPS.2010.5619586.
[15] Robles-Rodriguez CE, Dochain D. Decomposed Threshold ARMAX Models for short- to medium-term wind power forecasting. IFAC-PapersOnLine 2018;51(13):49–54. https://doi.org/10.1016/j.ifacol.2018.07.253.
[16] I. Colak, S. Sagiroglu, M. Yesilbudak, E. Kabalci, and H. Ibrahim Bulbul, "Multi-time series and-time scale modeling for wind speed and wind power forecasting part II: Medium-term and long-term applications," in *2015 International Conference on Renewable Energy Research and Applications, ICRERA 2015*, 2015, pp. 215–220, doi: 10.1109/ICRERA.2015.7418698.
[17] Zhao Y, Ye L, Li Z, Song X, Lang Y, Su J. A novel bidirectional mechanism based on time series model for wind power forecasting. Appl. Energy 2016;177:793–803. https://doi.org/10.1016/j.apenergy.2016.03.096.
[18] Croonenbroeck C, Ambach D. A selection of time series models for short- to medium-term wind power forecasting. J. Wind Eng. Ind. Aerodyn. 2015;136:201–10. https://doi.org/10.1016/j.jweia.2014.11.014.
[19] Durán MJ, Cros D, Riquelme J. Short-Term Wind Power Forecast Based on ARX Models. J. Energy Eng. Sep. 2007;133(3):172–80. https://doi.org/10.1061/(asce)0733-9402(2007)133:3(172).
[20] Gallego C, Pinson P, Madsen H, Costa A, Cuerva A. Influence of local wind speed and direction on wind power dynamics - Application to offshore very short-term forecasting. Appl. Energy 2011;88(11):4087–96. https://doi.org/10.1016/j.apenergy.2011.04.051.
[21] Mishra AK, Ramesh L. "Application of neural networks in wind power (generation). prediction" 2009. https://doi.org/10.1109/SUPERGEN.2009.5348160.
[22] J. P. S. Catalão, H. M. I. Pousinho, and V. M. F. Mendes, "An artificial neural network approach for short-term wind power forecasting in Portugal," Dec. 2009, doi: 10.1109/ISAP.2009.5352853.
[23] Yang W, Wang J, Lu H, Niu T, Du P. Hybrid wind energy forecasting and analysis system based on divide and conquer scheme: A case study in China. J. Clean. Prod. 2019;222:942–59. https://doi.org/10.1016/j.jclepro.2019.03.036.
[24] Du P, Wang J, Yang W, Niu T. A novel hybrid model for short-term wind power forecasting. Appl. Soft Comput. J. 2019;80:93–106. https://doi.org/10.1016/j.asoc.2019.03.035.
[25] Hong YY, Rioflorido CLPP. A hybrid deep learning-based neural network for 24-h ahead wind power forecasting. Appl. Energy 2019;250(April):530–9. https://doi.org/10.1016/j.apenergy.2019.05.044.
[26] Liu J, Wang X, Lu Y. A novel hybrid methodology for short-term wind power forecasting based on adaptive neuro-fuzzy inference system. Renew. Energy 2017;103:620–9. https://doi.org/10.1016/j.renene.2016.10.074.
[27] Y. Hao and C. Tian, "A novel two-stage forecasting model based on error factor and ensemble method for multi-step wind power forecasting," *Appl. Energy*, vol. 238, no. July 2018, pp. 368–383, 2019, doi: 10.1016/j.apenergy.2019.01.063.
[28] Sun M, Feng C, Zhang J. Multi-distribution ensemble probabilistic wind power forecasting. Renew. Energy 2020;148:135–49. https://doi.org/10.1016/j.renene.2019.11.145.
[29] Niu Z, Yu Z, Tang W, Wu Q, Reformat M. Wind power forecasting using attention-based gated recurrent unit network. Energy 2020;196:117081. https://doi.org/10.1016/j.energy.2020.117081.
[30] Wang H-zhi, Li G-qiang, Wang G-bin, Peng J-chun, Jiang H, Liu Y-tao. Deep learning based ensemble approach for probabilistic wind power forecasting. Appl. Energy 2017;188:56–70. https://doi.org/10.1016/j.apenergy.2016.11.111.
[31] Yu R, Gao J, Yu M, Lu W, Xu T, Zhao M, et al. LSTM-EFG for wind power forecasting based on sequential correlation features. Futur. Gener. Comput. Syst. 2019;93:33–42. https://doi.org/10.1016/j.future.2018.09.054.
[32] Yin H, Ou Z, Huang S, Meng A. A cascaded deep learning wind power prediction approach based on a two-layer of mode decomposition. Energy 2019;189:116316. https://doi.org/10.1016/j.energy.2019.116316.
[33] Demolli H, Dokuz AS, Ecemis A, Gokcek M. Wind power forecasting based on daily wind speed data using machine learning algorithms. Energy Convers. Manag. 2019;198:111823. https://doi.org/10.1016/j.enconman.2019.111823.
[34] A. Mosavi, M. Salimi, S. F. Ardabili, T. Rabczuk, S. Shamshirband, and A. R. Varkonyi-Koczy, "State of the art of machine learning models in energy systems, a systematic review," *Energies*, vol. 12, no. 7. MDPI AG, p. 1301, Apr. 04, 2019, doi: 10.3390/en12071301.
[35] J. F. Bermejo, J. F. G. Fernández, F. O. Polo, and A. C. Márquez, "A review of the use of artificial neural network models for energy and reliability prediction. A study of the solar PV, hydraulic and wind energy sources," *Applied Sciences (Switzerland)*, vol. 9, no. 9. MDPI AG, p. 1844, May 05, 2019, doi: 10.3390/app9091844.
[36] Zhang X, Peng Y, Xu W, Wang B. An Optimal Operation Model for Hydropower Stations Considering Inflow Forecasts with Different Lead-Times. Water Resour. Manag. Jan. 2019;33(1):173–88. https://doi.org/10.1007/s11269-018-2095-1.
[37] M. Sapitang, W. M. Ridwan, K. F. Kushiar, A. N. Ahmed, and A. El-Shafie, "Machine learning application in reservoir water level forecasting for sustainable hydropower generation strategy," *Sustain.*, vol. 12, no. 15, 2020, doi: 10.3390/su12156121.
[38] M. Dehghani *et al.*, "Prediction of hydropower generation using Grey wolf optimization adaptive neuro-fuzzy inference system," *Energies*, vol. 12, no. 2, 2019, doi: 10.3390/en12020289.
[39] Central Electricity Authority, "Flexible Operation of Thermal Power Plant for Integration of Renewable Generation," January, 2019. [Online]. Available: https://cea.nic.in/old/reports/others/thermal/trm/flexible_operation.pdf.
[40] Huber M, Dimkova D, Hamacher T. Integration of wind and solar power in Europe: Assessment of flexibility requirements. Energy May 2014;69:236–46. https://doi.org/10.1016/j.energy.2014.02.109.
[41] E. Sharma, "Energy forecasting based on predictive data mining techniques in smart energy grids," *Energy Informatics*, vol. 1, no. S1, 2018, doi: 10.1186/s42162-018-0048-9.
[42] M. N. Jyothi and P. V. R. Rao, "Very-short term wind power forecasting through Adaptive Wavelet Neural Network," Jul. 2016, doi: 10.1109/PESTSE.2016.7516513.
[43] M. Carolin Mabel and E. Fernandez, "Analysis of wind power generation and prediction using ANN: A case study," *Renew. Energy*, vol. 33, no. 5, pp. 986–992, May 2008, doi: 10.1016/j.renene.2007.06.013.
[44] Power System Operation Corporation Limited, "SRLDC-Daily-Reports,". (Accessed Jun. 22, 2021). [Online]. Available: https://www.srldc.in/Daily-Reports.
[45] S. Siami-Namini, N. Tavakoli, and A. S. Namin, "The Performance of LSTM and BiLSTM in Forecasting Time Series," *Proc. - 2019 IEEE Int. Conf. Big Data, Big Data 2019*, pp. 3285–3292, 2019, doi: 10.1109/BigData47090.2019.9005997.
[46] Kim S, Kim H. A new metric of absolute percentage error for intermittent demand forecasts. Int. J. Forecast. 2016;32(3):669–79. https://doi.org/10.1016/j.ijforecast.2015.12.003.
[47] Liu H, Chen C. Data processing strategies in wind energy forecasting models and applications: A comprehensive review. Appl. Energy 2019;249(May):392–408. https://doi.org/10.1016/j.apenergy.2019.04.188.
[48] Z. Qian, Y. Pei, H. Zareipour, and N. Chen, "A review and discussion of decomposition-based hybrid models for wind energy forecasting applications," *Appl. Energy*, vol. 235, no. April 2018, pp. 939–953, 2019, doi: 10.1016/j.apenergy.2018.10.080.
[49] Zhang J, Yan J, Infield D, Liu Y, Sang Lien F. Short-term forecasting and uncertainty analysis of wind turbine power based on long short-term memory network and Gaussian mixture model. Appl. Energy May 2019;241:229–44. https://doi.org/10.1016/j.apenergy.2019.03.044.
[50] Acikgoz H, Yildiz C, Sekkeli M. "An extreme learning machine based very short-term wind power forecasting method for complex terrain", *Energy Sources*. Part A Recover. Util. Environ. Eff. 2020;42(22):2715–30. https://doi.org/10.1080/15567036.2020.1755390.
[51] Z. Caiqing, Q. Ruonan, and Q. Zhiwen, "Comparing BP and RBF neural network for forecasting the resident consumer level by MATLAB," *Proc. 2008 Int. Conf. Comput. Electr. Eng. ICCEE 2008*, pp. 169–172, 2008, doi: 10.1109/ICCEE.2008.35.
[52] C. Yildiz, H. Acikgoz, D. Korkmaz, and U. Budak, "An improved residual-based convolutional neural network for very short-term wind power forecasting," *Energy Convers. Manag.*, vol. 228, no. November 2020, p. 113731, 2021, doi: 10.1016/j.enconman.2020.113731.

[53] Acikgoz H, Budak U, Korkmaz D, Yildiz C. WSFNet: An efficient wind speed forecasting model using channel attention-based densely connected convolutional neural network. Energy 2021;233:121121. https://doi.org/10.1016/j.energy.2021.121121.

[54] V. Bali, A. Kumar, and S. Gangwar, "A novel approach for wind speed forecasting using LSTM-ARIMA deep learning models," *Int. J. Agric. Environ. Inf. Syst.*, vol. 11, no. 3, pp. 13–30, 2020, doi: 10.4018/IJAEIS.2020070102.

[55] Shivani, K. S. Sandhu, and A. Ramachandran Nair, "A Comparative Study of ARIMA and RNN for Short Term Wind Speed Forecasting," *2019 10th Int. Conf. Comput. Commun. Netw. Technol. ICCCNT 2019*, pp. 1–7, 2019, doi: 10.1109/ICCCNT45670.2019.8944466.

[56] H. Liu, X. Mi, and Y. Li, "Smart multi-step deep learning model for wind speed forecasting based on variational mode decomposition, singular spectrum analysis, LSTM network and ELM," *Energy Convers. Manag.*, vol. 159, no. November 2017, pp. 54–64, 2018, doi: 10.1016/j.enconman.2018.01.010.

[57] Liu X, Lin Z, Feng Z. Short-term offshore wind speed forecast by seasonal ARIMA - A comparison against GRU and LSTM. Energy 2021;227:120492. https://doi.org/10.1016/j.energy.2021.120492.

[58] Liu M-D, Ding L, Bai Y-L. Application of hybrid model based on empirical mode decomposition, novel recurrent neural networks and the ARIMA to wind speed prediction. Energy Convers. Manag. 2021;233:113917. https://doi.org/10.1016/j.enconman.2021.113917.

[59] Yildiz C, Açikgöz H. "Forecasting diversion type hydropower plant generations using an artificial bee colony based extreme learning machine method", *Energy Sources*. Part B Econ. Plan. Policy 2021;16(2):216–34. https://doi.org/10.1080/15567249.2021.1872119.

[60] Guo Y, Xu Y-P, Xie J, Chen H, Si Y, Liu J. A weights combined model for middle and long-term streamflow forecasts and its value to hydropower maximization. J. Hydrol. 2021;602:126794. https://doi.org/10.1016/j.jhydrol.2021.126794.

[61] Li FF, Wang ZY, Zhao X, Xie E, Qiu J. Decomposition-ANN Methods for Long-Term Discharge Prediction Based on Fisher's Ordered Clustering with MESA. Water Resour. Manag. 2019;33(9):3095–110. https://doi.org/10.1007/s11269-019-02295-8.

[62] Yang S, Deng Z, Li X, Zheng C, Xi L, Zhuang J, et al. A novel hybrid model based on STL decomposition and one-dimensional convolutional neural networks with positional encoding for significant wave height forecast. Renew. Energy 2021;173:531–43. https://doi.org/10.1016/j.renene.2021.04.010.

[63] Chen D, Zhang J, Jiang S. Forecasting the Short-Term Metro Ridership with Seasonal and Trend Decomposition Using Loess and LSTM Neural Networks. IEEE Access 2020;8:91181–7. https://doi.org/10.1109/Access.628763910.1109/ACCESS.2020.2995044.

[64] Wang X, Guo P, Huang X. A review of wind power forecasting models. Energy Procedia 2011;12:770–8. https://doi.org/10.1016/j.egypro.2011.10.103.

[65] Zhao P, Wang J, Xia J, Dai Y, Sheng Y, Yue J. Performance evaluation and accuracy enhancement of a day-ahead wind power forecasting system in China. Renew. Energy 2012;43:234–41. https://doi.org/10.1016/j.renene.2011.11.051.

[66] Bhaskar K, Singh SN. AWNN-Assisted wind power forecasting using feed-forward neural network. IEEE Trans. Sustain. Energy 2012;3(2):306–15. https://doi.org/10.1109/TSTE.2011.2182215.

[67] Monteiro C, Ramirez-Rosado IJ, Fernandez-Jimenez LA. Short-term forecasting model for aggregated regional hydropower generation. Energy Convers. Manag. 2014;88:231–8. https://doi.org/10.1016/j.enconman.2014.08.017.

[68] Monteiro C, Ramirez-Rosado IJ, Fernandez-Jimenez LA. Short-term forecasting model for electric power production of small-hydro power plants. Renew. Energy 2013;50:387–94. https://doi.org/10.1016/j.renene.2012.06.061.