İSTANBUL UNIVERSITY CERRAHPAŞA

ORIGINAL ARTICLE
Electrica 2024; 24(2): 346-356

A Combined Model Based on Secondary Decomposition and Long Short-Term Memory Networks for Enhancing Wind Power Forecast

Mehmet Balcı[1], Uğur Yüzgeç[2], Emrah Dokur[3]

[1]Department of Information Technology, Bilecik Şeyh Edebali University, Bilecik, Türkiye
[2]Department of Computer Engineering, Bilecik Şeyh Edebali University, Bilecik, Türkiye
[3]Department of Electrical Electronics Engineering, Bilecik Şeyh Edebali University, Bilecik, Türkiye

Cite this article as: M. Balcı, U. Yüzgeç and E. Dokur, "A combined model based on secondary decomposition and LSTM networks for enhancing wind power forecast," *Electrica*, 24(2), 346-356, 2024.

ABSTRACT

Accurately predicting the potential wind power generation is of paramount importance in advancing the contribution of wind energy within the overall energy production landscape. To reduce dependence on fossil fuels, there is an urgent need to accelerate the integration of renewable energy sources, such as wind power. Moreover, ensuring a stable equilibrium between energy supply and demand hinges upon a profound understanding of the anticipated energy generation capacity. This paper presents a short-term forecasting model using data from the West of Duddon Sands, Barrow, and Horns Power sites. In pursuit of this goal, we have meticulously developed hybrid prediction models based on long short-term memory (LSTM) and bi-directional LSTM (Bi-LSTM) architectures. These models entail an initial data decomposition stage followed by the prediction phase. While some models solely incorporate the empirical mode decomposition (EMD) method for decomposition, others combine EMD with wavelet decomposition (WD) and swarm decomposition (SWD) for a more comprehensive approach. This investigation encompasses a range of models, including EMD–LSTM, EMD–WD–LSTM, EMD–SWD–LSTM, Bi-LSTM, EMD–Bi-LSTM, EMD–WD–Bi-LSTM, and EMD–SWD–Bi-LSTM. After a meticulous analysis of the outcomes generated by each model, a consistent trend emerges: the EMD–SWD–LSTM model consistently yields elevated R^2 values, signifying a heightened level of predictive accuracy and success.

***Index Terms*—**Decomposition, deep learning, hybrid models, long short-term memory (LSTM), wind forecasting

I. INTRODUCTION

Advancements in industry and technology are amplifying energy consumption, while the escalating reliance on polluting and finite fossil fuels persists as the primary energy source. Consequently, environmentally conscious and renewable energy sources are experiencing a surge in popularity. Among these, wind energy has emerged as a rapidly expanding force in the realm of electricity generation from renewable sources in recent years. The wind is abundant, has no pollutant emissions, and is a clean source that is environmentally friendly [1], [2]. However, for wind to become a reliable source of energy, it needs to be predicted with a high degree of accuracy in advance [3]. The variability, intermittency, and randomness of wind make the forecasting process challenging [4], [5]. Therefore, wind can sometimes not be a reliable source. Accurate prediction of wind speed and power is crucial to improve its utilization and reliability [6]. Many forecasting methods have been developed in the literature with the aim of accurately predicting wind and improving its utilization. Prediction models include physical methods, statistical models, artificial intelligence models, and hybrid models [7].

Physical models take into account various characteristics such as humidity, air temperature, and the geographical structure of the region during the forecast process. These models are commonly known as Numerical Weather Prediction (NWP) models, namely Regional Atmospheric Modeling System (aka RAMS), High-Resolution Limited Area Model (aka HIRLAM), Aire Limitée Adaptation Dynamique Développement InterNational (aka ALADIN), and Weather Research and Forecasting [8]. Several noteworthy studies have been conducted on these models. For instance, Chen et al. [9] employed NWP for modeling wind speed prediction within an Arctic wind field. Additionally, Dong et al. [10] employed the Extreme Gradient Boosting (XGBoost) method to make predictions across seven distinct regions in China, augmenting their approach with NWP-based bias

Corresponding author:
Mehmet Balcı

E-mail:
mehmet.balci@bilecik.edu.tr

Received: September 19, 2023

Revision Requested: November 21, 2023

Last Revision Received: November 28, 2023

Accepted: January 12, 2024

Publication Date: March 18, 2024

DOI: 10.5152/electrica.2024.23138

correction techniques. Liu et al. have developed the NWPWS method by customizing NWP for wind prediction [11].

In statistical wind forecasting methods, predictions are generated by employing mathematical and statistical techniques on historical wind data. The autoregressive integrated moving average (ARIMA) and autoregressive moving average (ARMA) models are the most commonly used statistical models [12]. Studies utilizing these models can be found in the existing literature. A few notable examples include the work of Liu et al., who enhanced the ARIMA-based Seasonal Auto Regressive Integrated Moving Average (SARIMA) method for wind forecasting in coastal/offshore areas of Scotland [13]. Another study by Liu et al. introduced the repeated WT-based ARIMA (RWT-ARIMA) model, an extension of the ARIMA and Wavelet Transform-ARIMA (WT-ARIMA) models [14].

With the advancements in the field of computer science, artificial intelligence models are also being developed. Artificial intelligence models can be utilized in various domains, including wind prediction. A brief summary of the literature on artificial intelligence models in the field of wind forecasting is as follows: Qureshi et al. utilized the gated recurrent unit (GRU) artificial intelligence method to forecast wind energy by utilizing wind data from the Jhimpir-Pakistan region [15]. Yang et al. employed a predictive model developed by integrating the kernel-based extreme learning machine and multi-objective optimization methods [16]. Rodriguez et al. employed artificial neural networks (ANN) to perform a 10-minute ahead wind prediction [17].

In hybrid prediction models, multiple methods are used together, including techniques such as decomposition, normalization, and prediction methods. Nahid et al. attempted to forecast wind speed using the empirical mode decomposition (EMD)–convolutional long-short term memory (CLSTM) hybrid model, which was constructed by combining EMD, convolutional neural network (CNN), and long short-term memory (LSTM) [18]. Mansoor et al. performed power analysis and prediction employing a hybrid model known as Modified White Shark Optimization (MWSO)–Radial Basis Function Neural Network (RBFNN), which synergistically integrates the MWSO and RBFNN models [19]. Bommidi et al. made forward predictions ranging from 5 minutes to 48 hours using their developed hybrid model, which combines improved complete ensemble EMD with adaptive noise, transformer network, and novel kernel mean square error (MSE) loss function [20]. Zhao et al. conducted forward predictions of 15-minute wind power using a hybrid model that employed variable mode decomposition for decomposition and combined CNN with GRU [21].

This paper provides a comprehensive comparison of a hybrid model that combines deep learning with secondary decomposition for short-term wind power forecasting. The analysis involves integrating decomposition methods into LSTM-based forecasting models. The inclusion of decomposition techniques like EMD, wavelet decomposition (WD), and swarm decomposition (SWD) enhances model performance by mitigating nonlinearity. Moreover, the study explores secondary decomposition models that combine these techniques, yielding more accurate estimations. Notably, the overall model's effectiveness is notably enhanced by improving the prediction accuracy of high-frequency subcomponents in secondary decomposition. To evaluate model performance, comprehensive analyses are conducted, employing offshore wind power data from the Duddon Sands, Barrow, and Horns Power regions, and comparing results against established performance metrics. In this respect, the main contributions of this study can be summarized as follows:

- Short-term wind power forecasting frame based on a multiple signal processing strategy is successfully designed. The original data is decomposed using a robust preprocessing approach including EMD and SWD. The advantages of the EMD and SWD data processing strategies are hybridized to maximize the data preprocessing performance and provide convenience for high-accuracy prediction. Multiple decomposition approaches provide a unique perspective for analyzing and modeling data dynamics and complexities, resulting in a better understanding of the data.
- To improve the forecasting performance, the advantages of the secondary decomposition data processing strategies are hybridized to maximize performance with LSTM. Long short-term memory performs on offshore wind datasets to exploit its capability of learning features from nonlinear historical data. The proposed model is assessed on multivariate offshore wind farm data from three different regions.
- The robustness of the proposed scheme is assessed through variants of LSTM on all wind farms in terms of error measures. Several performance metrics including MSE, MAE, RMSE, and R^2 are used to quantify the forecasting accuracy and demonstrate the superior performance of the proposed models. By identifying significant variables in the offshore wind power data and subseries obtained from the multiple decomposition, the proposed forecasting model provides compelling analysis for wind power forecasting, catering to the needs of decision-makers.

The remainder of the paper is organized as follows: Methodology is presented in Section II. The comparative results of the models are given in Section 3. Section 4 presents conclusion remarks.

II. METHODOLOGY

A. Long-Short Term Memory (LSTM)

This model is a type of recurrent neural network (RNN) architecture in which values are memorized at random intervals. The stored values are not modified as the learned data progresses. RNNs allow forward and backward connections between neurons. Considering the unknown dimensions and time delays between events in time series classification, processing, and prediction, LSTM is a highly suitable model. LSTM networks consist of LSTM units and recurrent components. Due to the recurrent nature of the components, LSTM units can remember long or short time periods. As a result, the stored values remain unchanged through iterations and are trained over time through backpropagation. The mathematical formulas of the LSTM model are provided below [22]:

$$f_t = \sigma\left(W_f x_t + U_f h_{t-1} + b_f\right) \tag{1}$$

$$i_t = \sigma\left(W_i x_t + U_i h_{t-1} + b_i\right) \tag{2}$$

$$\tilde{c}_t = \tanh\left(W_c x_t + U_c h_{t-1} + b_c\right) \tag{3}$$

$$c_t = c_{t-1} \square f_t + i_t \square \tilde{c}_t \tag{4}$$

$$o_t = \sigma\left(W_o x_t + U_o h_{t-1} + b_o\right) \tag{5}$$

$$h_t = o_t \quad \tanh\left(c_t\right) \tag{6}$$

where W_f, W_i, W_c, W_o, U_f, U_i, U_c, U_o denote the weights of the input and recurrent connections, the subscript *f* is the forget gate, the subscript *i* is the input gate, the subscript *c* is the memory cell, and the subscript *o* is the output gate, $\odot$ is element-wise product operator, *b* represents the bias vector, x_t is the input of the LSTM, σ is Sigma function. f_t is the activation vector of the forget gate, i_t is the activation vector of the input/update gate, $\tilde{c}_t$ is activation vector of the cell input, c_t is the cell state vector, o_t is the activation vector of the output gate, and h_t is the hidden state vector.

This process continues by iterating through the steps described above. In addition, the weight (*W*) and bias (*b*) parameters used in the above equations are determined by the model in such a way as to minimize the difference between the LSTM outputs and the real training data. This ensures that the model optimally adjusts these parameters to achieve the best fit between predicted and actual results during the training process.

B. Bi-Directional Long Short-Term Memory

Bi-directional (Bi)-LSTM is composed of two independent artificial neural networks. In this structure, the networks possess information both in the forward and backward directions. Inputs are processed in two distinct ways: from the past to the future and from the future to the past. What sets it apart from unidirectional LSTM is that the information coming from the future is preserved and merged with two hidden layers. This allows the retention of information from both the past and the future at any given time. A diagram illustrating the general operation of Bi-LSTM is provided in Fig. 1.

C. Empirical Mode Decomposition

Empirical mode decomposition is a powerful signal processing technique that plays a significant role in enhancing the accuracy of wind power forecasting models. Empirical mode decomposition has found widespread application in various fields due to its ability to adaptively decompose nonstationary and nonlinear data into intrinsic mode functions (IMF). Empirical mode decomposition operates on the principle of sifting through a signal iteratively to extract its oscillatory components at different scales. These extracted IMFs capture the essential temporal patterns present in the signal. Each IMF represents a narrowband component that, when combined, reconstructs the original signal. The decomposition is performed in such a way that each IMF satisfies two main conditions: it has equal numbers of zero crossings and extrema, ensuring its adaptability to various data characteristics. In the formula [23], $X(t)$ represents the input signal, h_1 denotes the first component, and m_1 corresponds to the average of the upper and lower envelopes determined by the cubic-spline interpolation of local maxima and minima.

$$h_1(t) = X(t) - m_1(t) \tag{7}$$

In the second step, h_1 data are obtained, with m_1 being the average of the upper and lower envelopes of h_1 components.

$$h_{11}(t) = h_1(t) - m_{11}(t) \tag{8}$$

These first two steps are repeated k times until h_{1k} becomes an IMF.

$$h_{1k}(t) = h_{1(k-1)}(t) - m_{1k}(t) \tag{9}$$

Afterward, the first IMF $c_1 = h_{1k}$ is determined from the data containing the shortest period component of the signal. c_1 is separated from the rest of the data.

$$r_1(t) = X(t) - c_1(t) \tag{10}$$

These operations are repeated up to n times.

$$r_1 - c_2 = r_2, \ldots, r_{n-1} - c_n = r_n \tag{11}$$

D. Wavelet Decomposition

Wavelet decomposition operates by convolving a given signal with a family of wavelet functions, each of which has a specific frequency and time localization. This convolutive process generates a set of coefficients that represent the signal's behavior at different scales and positions. By iteratively applying this convolution and downsampling, the original signal is divided into a series of low-frequency and high-frequency components, known as approximation and detail coefficients, respectively. In the domain of wind power forecasting, WD is employed as a preprocessing technique to unveil underlying temporal patterns that might be masked in the original data. By capturing both short-term variations and longer-term trends, wavelet-based decomposition provides a comprehensive representation of the data's oscillatory behavior.

Wavelet decomposition can decompose a time-domain signal into different frequency groups. In WD, various functions indicating certain mathematical properties, such as a window length corresponding to a fixed number of periods, are used. A fundamental requirement in WD is that the basic functions should be within the bounds of the Hilbert space with a mean of zero.

E. Swarm Decomposition

Swarm decomposition is a method developed to analyze non-stationary signals. The concept of SWD was introduced by Apostolidis et al. [24], and it has since been adopted across various fields of

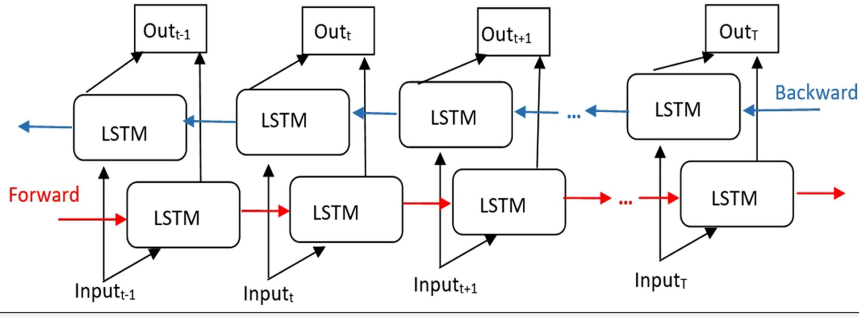

Fig. 1. Diagram depicting the general operation of Bi-LSTM. Bi, bidirectional; LSTM, long short-term memory.

study, encompassing fault analysis [25], [26] and renewable energy [27], [28].

This approach is primarily based on swarm filtering (SWF) [28]. The method involves the foraging behavior of a swarm. There are two interaction forces in the swarm for successful foraging: driving force and cohesion. The driving force is denoted as $F_{dr}(n, i)$;

$$F_{dr}\left(n,i\right)=P_{prey}\left(n\right)-P_i\left(n-1\right) \tag{12}$$

In the above equation, i and n represent the number of members and the number of steps, respectively. The location information of the prey is denoted by P_{prey}. For all members of the swarm, the alerted cohesion force is denoted as $F_{coh}(n, i)$;

$$F_{Coh,i}^{n}=\frac{1}{M-1}\sum_{j=1,j\neq i}^{M} f\left(P_i\left[n-1\right]-P_j\left(n-1\right)\right) \tag{13}$$

$$f\left(d\right)=-sgn\left(d\right).\ln\left(\frac{\left|d\right|}{d_{cr}}\right) \tag{14}$$

The above equations are represented in terms of the sign and logarithmic functions, sgn(.) and ln(.), respectively. The distances between members and the critical distance are denoted as d and d_{cr}, respectively. M represents the number of swarms, and in order for the swarm to track its prey, the position and velocity information are updated at each step as follows:

$$V_i\left[n\right]=V_i\left[n-1\right]+\delta.\left(F_{Dr,i}^{n}+F_{Coh,i}^{n}\right) \tag{15}$$

$$P_i\left[n\right]=P_i\left[n-1\right]+\delta.\left(V_i\left[n\right]\right) \tag{16}$$

The parameter δ, as shown in the equations, is used to control the flexibility of the swarm. The output of SWF is determined by the following equation.

TABLE I. INFORMATION ABOUT THE DATASETS

Turbine Name	Standard Deviation	Turbine Number	Power
West of Duddon Sands	72.0836	108	388.8 MW
Barrow	28.1497	30	90.0 MW
Horns Power	51.1955	49	160 MW

$$y\left[n\right]=\beta.\sum_{i=1}^{M}P_i\left[n\right] \tag{17}$$

In the above equation, β is the scale parameter that influences the members of the swarm.

III. EXPERIMENTAL RESULTS

In this study, a desktop computer equipped with an Intel i5-7500 processor and 16 GB RAM was used as the hardware. Matlab 2020b was employed for data analysis and prediction processes. The maximum number of epochs for the models was set to 50, with a minimum batch size of 16, and the Adam optimizer was selected for the training process. We used well-known metrics such as MSE, root MSE (RMSE), mean absolute error (MAE), and R^2 to compare the models.

A. Datasets

The forecasting models were trained and tested on wind power data from three different regions of real offshore wind farms. The datasets consist of 1-year records recorded at 1-hour intervals. These farms include the West of Duddon Sands and Barrow wind farms located between England and Ireland, and the Horns Power wind farm situated near the coasts of Denmark in the North Sea. Information regarding the datasets is summarized in Table I.

B. Prediction Results and Analysis

This section presents the results of the predictions made within the scope of the study, both for the training and testing phases, displayed in tabular form. Additionally, the test results of all models for each region are provided in graphical format.

C. Results for West of Duddon Sands Dataset

Table II and Table III display the forecasting results of the models based on LSTM and Bi-LSTM for the West of Duddon Sands dataset, respectively. Here, the hybrid models have been further categorized into EMD–WD and EMD–SWD, incorporating a secondary decomposition. Here, in the results in the table, the best metric values are indicated in bold.

Upon examining the values presented in the aforementioned tables, it becomes evident that the EMD–SWD–LSTM model emerges as the optimal forecasting model during the training phase. Conversely, the testing phase yields superior performance from the EMD–SWD–Bi-LSTM model. The R^2 values corresponding to the LSTM and Bi-LSTM models are illustrated graphically in Fig. 2.

Fig. 3 presents the forecasting test results of the implemented models for West of Duddon Sands. It is obtained that the EMD–SWD

TABLE II. THE FORECASTING RESULTS OF LSTM-BASED MODELS FOR THE WEST OF DUDDON SANDS REGION

	LSTM		EMD–LSTM		EMD–WD–LSTM		EMD–SWD–LSTM	
Metrics	**Training**	**Test**	**Training**	**Test**	**Training**	**Test**	**Training**	**Test**
MSE	384.23	381.38	118.57	189.56	313.33	342.16	**95.874**	**113.95**
RMSE	19.602	19.529	10.889	13.768	17.701	18.498	**9.7915**	**10.675**
MAE	13.951	14.207	7.938	10.872	10.814	11.565	**7.3311**	**8.384**
R^2	0.92774	0.91559	0.9777	0.95805	0.95126	0.9388	**0.98197**	**0.97478**

EMD, empirical mode decomposition; LSTM, long short-term memory; MAE, mean absolute error; MSE, mean square error; RMSE, root mean square error; SWD, swarm decomposition; WD, wavelet decomposition.

TABLE III. THE FORECASTING RESULTS OF BI-LSTM-BASED MODELS FOR THE WEST OF DUDDON SANDS REGION

	Bi-LSTM		EMD–Bi-LSTM		EMD–WD–Bi-LSTM		EMD–SWD–Bi-LSTM	
Metrics	**Training**	**Test**	**Training**	**Test**	**Training**	**Test**	**Training**	**Test**
MSE	362.21	349.84	144.97	136.22	327.79	342.58	**101.71**	**99.336**
RMSE	19.032	18.704	12.04	11.671	18.105	18.509	**10.085**	**9.9668**
MAE	11.962	11.145	9.134	8.8001	11.657	11.503	**7.4792**	**7.6334**
R^2	0.93188	0.92257	0.97274	0.96985	0.94901	0.93873	**0.98087**	**0.97801**

Bi, bidirectional; EMD, empirical mode decomposition; LSTM, long short-term memory; MAE, mean absolute error; MSE, mean square error; RMSE, root mean square error; SWD, swarm decomposition; WD, wavelet decomposition.

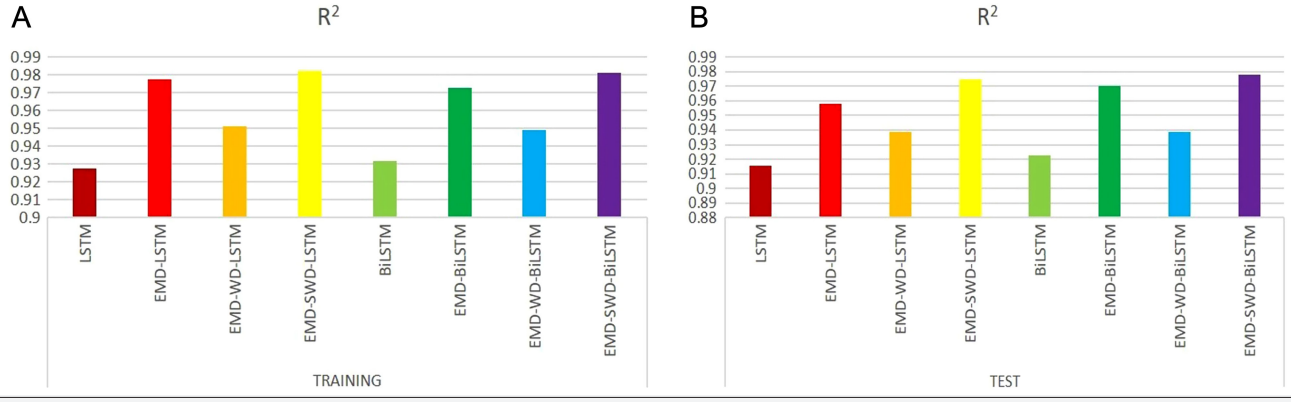

Fig. 2. Results of (a) training and (b) testing phases of hybrid deep learning models for West of Duddon Sands dataset. Bi, bidirectional; EMD, empirical mode decomposition; LSTM, long short-term memory; MAE, mean absolute error; MSE, mean square error; RMSE, root mean square error; SWD, swarm decomposition; WD, wavelet decomposition.

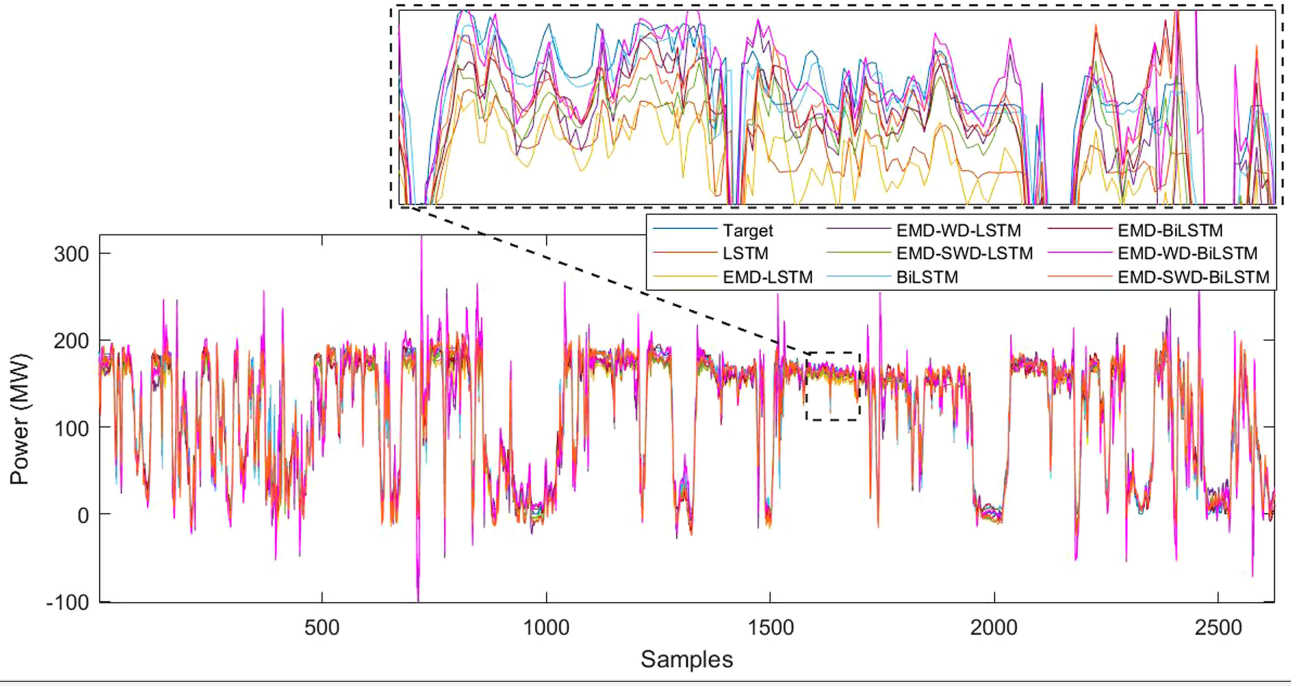

Fig. 3. Results of hybrid deep learning models for West of Duddon Sands. Bi, bidirectional; EMD, empirical mode decomposition; LSTM, long short-term memory; MAE, mean absolute error; MSE, mean square error; RMSE, root mean square error; SWD, swarm decomposition; WD, wavelet decomposition.

TABLE IV. THE FORECASTING RESULTS OF LSTM-BASED MODELS FOR THE BARROW REGION

	LSTM		EMD–LSTM		EMD–WD–LSTM		EMD–SWD–LSTM	
Metrics	**Training**	**Test**	**Training**	**Test**	**Training**	**Test**	**Training**	**Test**
MSE	85.584	77.555	19.553	19.988	**11.686**	**18.836**	17.53	19.245
RMSE	9.2512	8.8065	4.4219	4.4708	**3.4185**	**4.34**	4.1868	4.3869
MAE	5.8682	5.6011	3.2652	**3.3104**	**2.5464**	3.5442	3.1402	3.3108
R^2	0.89963	0.87972	0.97707	0.969	**0.98606**	**0.9707**	0.97944	0.97015

EMD, empirical mode decomposition; LSTM, long short-term memory; MAE, mean absolute error; MSE, mean square error; RMSE, root mean square error; SWD, swarm decomposition; WD, wavelet decomposition.

based hybrid proposed model results give the best fit to the real data, while the standalone models' results display the most discrepancies. Upon examination of the plots, it becomes evident that the EMD–SWD–Bi-LSTM model exhibits the highest degree of success.

D. Results for Barrow Dataset

In this subsection, for the Barrow dataset, we presented the comparison results of all hybrid deep learning-based forecasting models used in this study. The error performance values of the LSTM, EMD–LSTM, EMD–WD–LSTM, and EMD–SWD–LSTM models for predictions in the Barrow region are shown in Table IV, while the error performance values for predictions using the Bi-LSTM, EMD–Bi-LSTM, EMD–WD–Bi-LSTM, and EMD–SWD–Bi-LSTM models are shown in Table V. Here, within the table results, the superior metric values are highlighted in bold.

Upon examining the results in these tables, it is observed that the EMD–WD–LSTM model was the most successful during the training phase, while the EMD–SWD–Bi-LSTM model performed best during the testing phase. R^2 values for the models using LSTM and Bi-LSTM are shown in the graphs in Fig. 4.

TABLE V. THE FORECASTING RESULTS OF BI-LSTM-BASED MODELS FOR THE BARROW REGION.

	Bi-LSTM		EMD–Bi-LSTM		EMD–WD–Bi-LSTM		EMD–SWD–Bi-LSTM	
Metrics	**Training**	**Test**	**Training**	**Test**	**Training**	**Test**	**Training**	**Test**
MSE	79.715	69.033	**17.634**	15.658	19.86	21.79	18.111	**14.05**
RMSE	8.9284	8.3086	**4.1993**	3.9571	4.4565	4.6679	4.2557	**3.7484**
MAE	6.3167	5.8006	**3.0517**	2.9132	3.5261	3.4785	3.156	**2.7601**
R^2	0.90651	0.89294	**0.97932**	0.97572	0.97631	0.96611	0.97876	**0.97821**

Bi, bidirectional; EMD, empirical mode decomposition; LSTM, long short-term memory; MAE, mean absolute error; MSE, mean square error; RMSE, root mean square error; SWD, swarm decomposition; WD, wavelet decomposition.

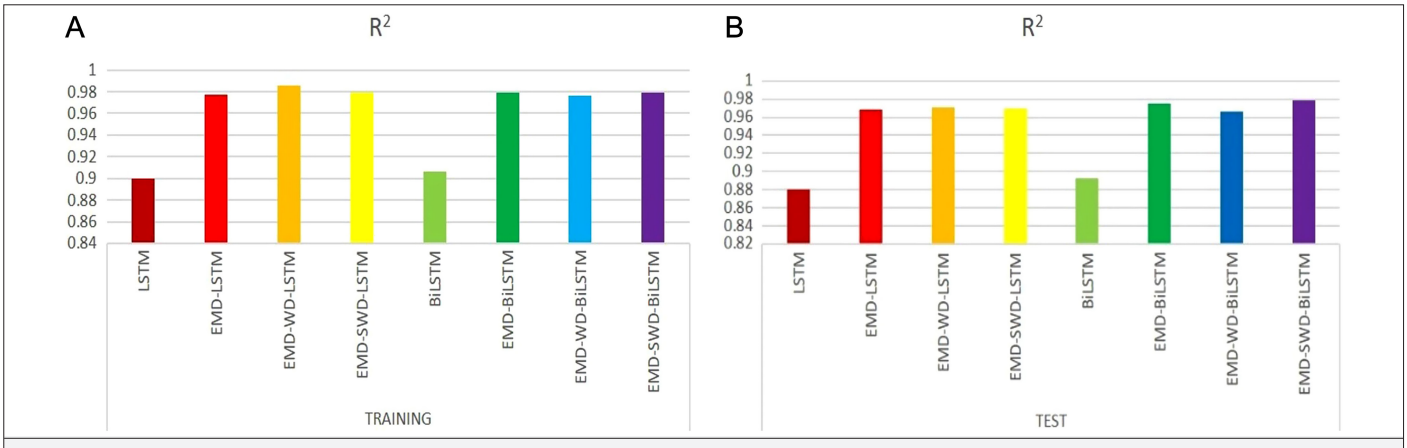

Fig. 4. Results of (a) training and (b) testing phases of hybrid deep learning models for Barrow dataset. Bi, bidirectional; EMD, empirical mode decomposition; LSTM, long short-term memory; MAE, mean absolute error; MSE, mean square error; RMSE, root mean square error; SWD, swarm decomposition; WD, wavelet decomposition.

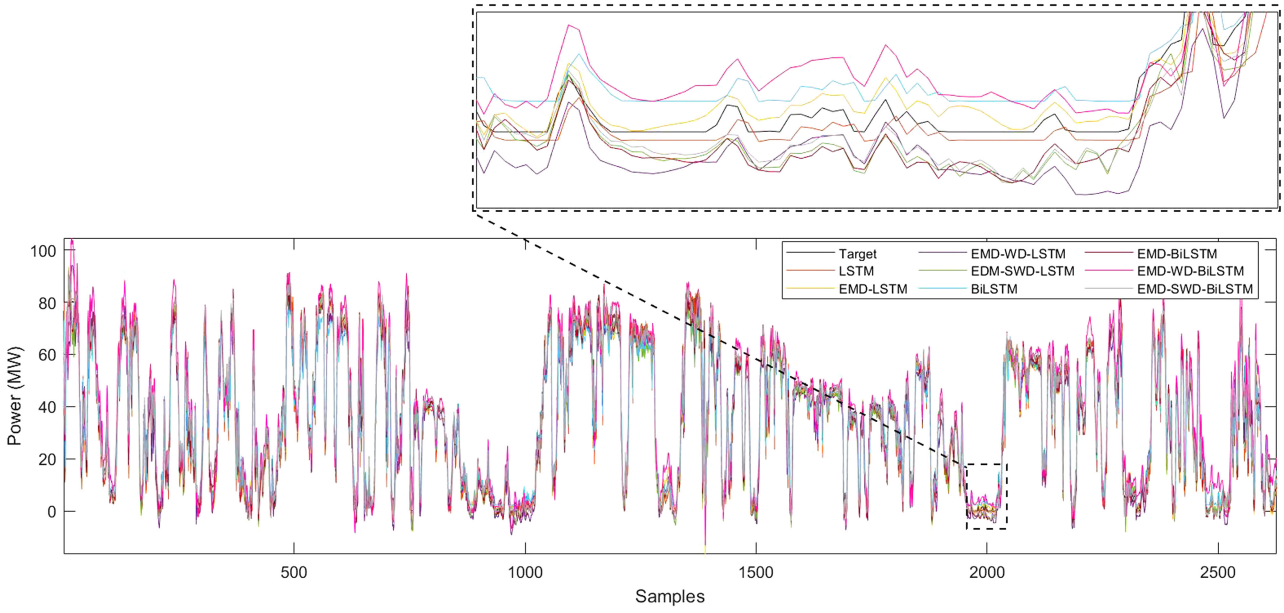

Fig. 5. Results for Barrow LSTM and Bi-LSTM-based models. Bi, bidirectional; EMD, empirical mode decomposition; LSTM, long short-term memory; MAE, mean absolute error; MSE, mean square error; RMSE, root mean square error; SWD, swarm decomposition; WD, wavelet decomposition.

Fig. 5 illustrates the outcome of model predictions during the testing phase. Upon closer examination of the graphs, it becomes evident that the EMD–SWD–Bi-LSTM hybrid model exhibits the most accurate and effective prediction curve compared to other models.

E. Results for Horns Power Dataset

In this subsection, the wind power prediction results of all models for the Horns Power dataset are evaluated. The error performance values of the LSTM-based hybrid models for forecasts in the West of Duddon Sands region are presented in Table VI, while the error performance values found for forecasts using Bi-LSTM-based hybrid models are summarized in Table VII. Here, in the table results, the best metric values are highlighted in bold.

As can be seen from these tables, the best hybrid model for the training phase is the EMD–SWD–LSTM model. For the testing phase, the EMD–SWD–LSTM model showed the best performance, except for the MAE value where the LSTM model performed best. Additionally, the R^2 values for LSTM and Bi-LSTM-based hybrid models are presented in the graphs in Fig. 6.

Fig. 7 presents the graphs illustrating the wind power prediction results generated by each hybrid model during the testing phase. Examination of both the graphs and accompanying tables reveals that the LSTM model for Horns Power achieved noteworthy performance metrics, with RMSE, MSE, and MAE values of 23.833, 568.02, and 13.472, respectively. These findings lead to the conclusion that the LSTM model outperformed other models in terms of predictive accuracy. In contrast to other datasets, the efficacy of this secondary decomposition hybrid model diminished for the present dataset due to extended intervals of zero crossing points when power generation is absent. Specifically, this unique scenario was intentionally included in the analysis without applying data filtering or reference point selection. The prolonged duration of zero crossing points leads to adverse transitions in the separated signals, thereby reasoning to the degradation of model performance within the hybrid models.

IV. CONCLUSION

In this study, LSTM and Bi-LSTM-based hybrid models were developed for wind predictions. During the development of the hybrid

TABLE VI. THE FORECASTING RESULTS OF LSTM-BASED MODELS FOR THE HORNS POWER REGION.

	LSTM		EMD–LSTM		EMD–WD–LSTM		EMD–SWD–LSTM	
Metrics	**Training**	**Test**	**Training**	**Test**	**Training**	**Test**	**Training**	**Test**
MSE	416.42	**541.45**	115.94	654.87	528.41	25609	**98.163**	558.56
RMSE	20.406	**23.269**	10.767	25.59	22.987	160.03	**9.9077**	23.634
MAE	14.221	15.497	7.5895	14.346	14.031	50.659	**6.8884**	**13.083**
R^2	0.84025	**0.79402**	0.95552	0.75087	0.88104	0.48722	**0.96234**	0.78751

EMD, empirical mode decomposition; LSTM, long short-term memory; MAE, mean absolute error; MSE, mean square error; RMSE, root mean square error; SWD, swarm decomposition; WD, wavelet decomposition.

TABLE VII. THE FORECASTING RESULTS OF BI-LSTM-BASED MODELS FOR THE HORNS POWER REGION.

	Bi-LSTM		EMD–Bi-LSTM		EMD–WD–Bi-LSTM		EMD–SWD–Bi-LSTM	
Metrics	**Training**	**Test**	**Training**	**Test**	**Training**	**Test**	**Training**	**Test**
MSE	430.73	**568.02**	125.86	770.13	494.26	22200	**98.504**	568.96
RMSE	20.754	**23.833**	11.219	27.751	22.232	149	**9.9249**	23.853
MAE	12.405	13.472	7.9517	15	13.24	47.182	**6.913**	**13.089**
R^2	0.83476	**0.78391**	0.95172	0.70702	0.88872	0.55549	**0.96221**	0.78355

Bi, bidirectional; EMD, empirical mode decomposition; LSTM, long short-term memory; MAE, mean absolute error; MSE, mean square error; RMSE, root mean square error; SWD, swarm decomposition; WD, wavelet decomposition.

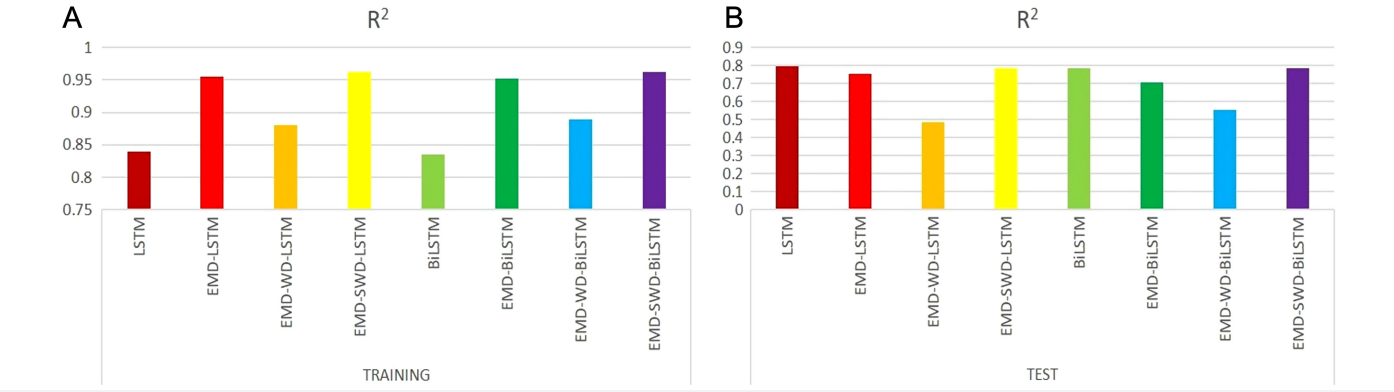

Fig. 6. Results of (a) training and (b) testing phases of hybrid deep learning models for Horns power dataset. Bi, bidirectional; EMD, empirical mode decomposition; LSTM, long short-term memory; MAE, mean absolute error; MSE, mean square error; RMSE, root mean square error; SWD, swarm decomposition; WD, wavelet decomposition.

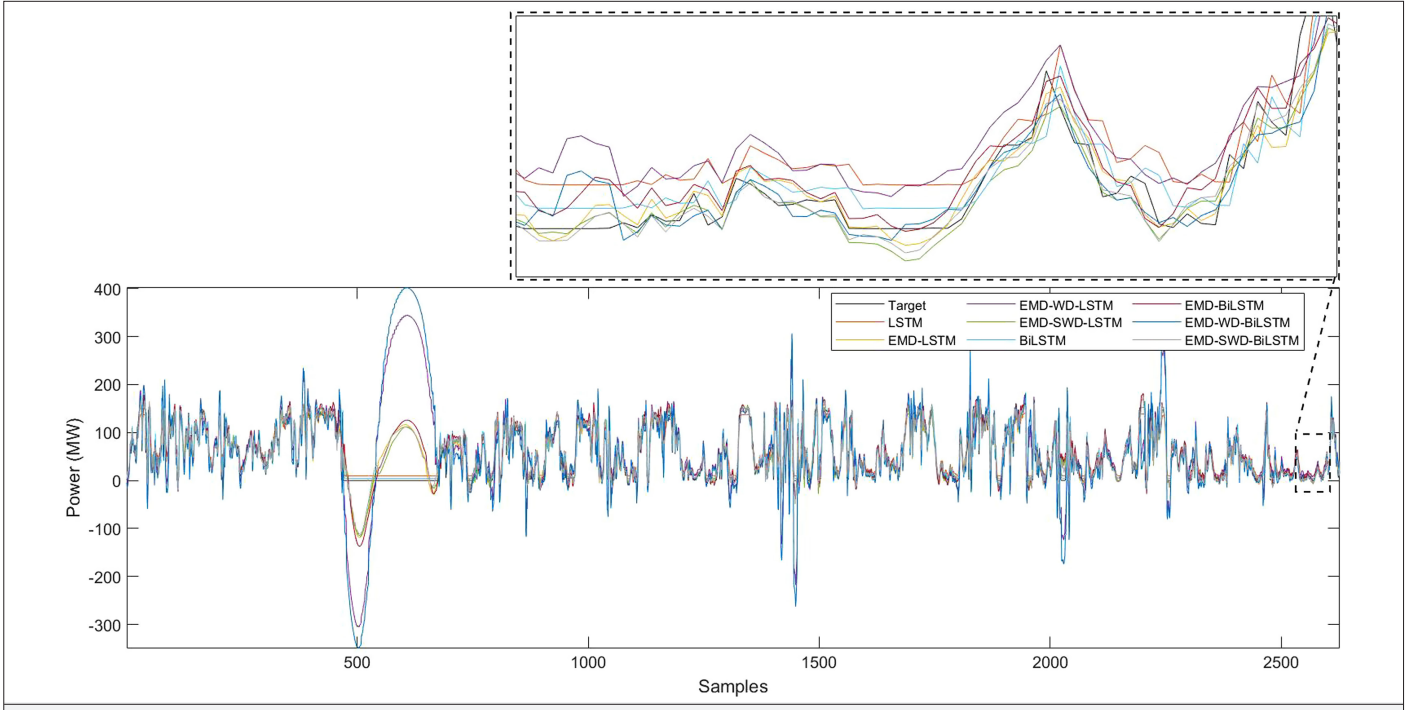

Fig. 7. Results for Horns power LSTM and Bi-LSTM-based models. Bi, bidirectional; EMD, empirical mode decomposition; LSTM, long short-term memory; MAE, mean absolute error; MSE, mean square error; RMSE, root mean square error; SWD, swarm decomposition; WD, wavelet decomposition.

models, the EMD, WD, and SWD decomposition methods were used. The EMD method served as the primary decomposition, while the WD and SWD methods were used as secondary decomposition methods. It was observed that the use of secondary decomposition methods generally improved the success rate of the models in making predictions.

The EMD–WD–LSTM model demonstrated superior performance during the training phase for the Barrow dataset, achieving an impressive MSE of 11.686 and an R^2 value of 0.98606. However, in the testing phase, the EMD–SWD–Bi-LSTM model claimed the top position with an MSE of 14.05 and an R^2 of 0.97821. Turning to the Horns Powers dataset, the EMD–SWD–LSTM model excelled during training, boasting MSE and R^2 values of 98.163 and 0.96234, respectively. Nevertheless, during the testing phase, the LSTM model emerged as the frontrunner, attaining the best results with an MSE of 541.45 and an R^2 of 0.79402. Finally, for the West of Duddon Sands dataset, the EMD–SWD–LSTM model showcased superior performance in the training phase, recording MSE and R^2 values of 95.874 and 0.98197. Conversely, in the testing phase, the EMD–SWD–Bi-LSTM model outperformed its counterparts with MSE and R^2 rates of 99.336 and 0.97801, respectively. These comprehensive findings underscore the nuanced dynamics of model performance across diverse datasets and phases. The difference between LSTM and Bi-LSTM-based models in the successful models was generally minimal. Thus, it can be concluded that LSTM and Bi-LSTM-based models exhibit similar success rates.

In conclusion, the results of wind power predictions for the three datasets show that hybrid models including secondary decomposition methods generally produce more accurate forecasts. However, as the number of methods used in hybrid models increases, the running times required for the model to perform the predictions also increase. For future research, the exploration of incorporating meta-heuristic approaches to fine-tune LSTM parameters could be considered.

Peer-review: Externally peer-reviewed.

Author Contributions: Concept – M.B., U.Y.; Design – M.B., U.Y., E.D.; Supervision – U.Y., E.D.; Data Collection and/or Processing – M.B., E.D.; Analysis and/or Interpretation – M.B., U.Y., E.D.; Literature Review – M.B., E.D.; Writing – M.B., U.Y.; Critical Review – U.Y., E.D.

Declaration of Interests: The authors have no conflict of interest to declare.

Funding: The authors declared that this study has received no financial support.

REFERENCES

1. Z. Shang, Y. Chen, Y. Chen, Z. Guo, and Y. Yang, "Decomposition-based wind speed forecasting model using causal convolutional network and attention mechanism," *Expert Syst. Appl.*, vol. 223, p. 119878, 2023. [CrossRef]
2. E. Tefera, M. Martínez-Ballesteros, A. Troncoso, and F. Martínez-Alvarez, "A new hybrid cnn-LSTM for wind power forecasting in ethiopia," in Int. Conf. Hybrid Artif. Intell. Syst. Berlin: Springer, 2023, pp. 207–218.
3. X. Liu, L. Zhang, J. Wang, Y. Zhou, and W. Gan, "A unified multi-step wind speed forecasting framework based on numerical weather prediction grids and wind farm monitoring data," *Renew. Energy*, vol. 211, pp. 948–963, 2023. [CrossRef]
4. X. Wang, X. Li, and J. Su, "Distribution drift-adaptive short-term wind speed forecasting," *Energy*, vol. 273, p. 127209, 2023. [CrossRef]
5. M. Galphade, V. Nikam, B. Banerjee, and A. W. Kiwelekar, "Comparative analysis of wind power forecasting using LSTM, bilstm, and gru," in Int. Conf. Frontiers Intell. Comput.: Theory Appl. Berlin: Springer, 2022, pp. 483–493.
6. K. Moharm, M. Eltahan, and E. Elsaadany, "Wind speed forecast using LSTM and bi-LSTM algorithms over gabal el-zayt wind farm,"in Int. Conf. Smart Grids Energy Syst. (SGES), Vol. 2020. IEEE Publications, 2020, pp. 922–927. [CrossRef]
7. Y. Hao, W. Yang, and K. Yin, "Novel wind speed forecasting model based on a deep learning combined strategy in urban energy systems," *Expert Syst. Appl.*, vol. 219, p. 119636, 2023. [CrossRef]
8. S. Tuy, H. S. Lee, and K. Chreng, "Integrated assessment of offshore wind power potential using weather research and forecast (wrf) downscaling with sentinel-1 satellite imagery, optimal sites, annual energy production and equivalent CO_2 reduction," *Renew. Sustain. Energy Rev.*, vol. 163, p. 112501, 2022. [CrossRef]
9. H. Chen, "A comprehensive statistical analysis for residuals of wind speed and direction from numerical weather prediction for wind energy," *Energy Rep.*, vol. 8, pp. 618–626, 2022. [CrossRef]
10. J. Dong, W. Zeng, L. Wu, J. Huang, T. Gaiser, and A. K. Srivastava, "Enhancing short-term forecasting of daily precipitation using numerical weather prediction bias correcting with xgboost in different regions of china," *Eng. Appl. Artif. Intell.*, vol. 117, p. 105579, 2023. [CrossRef]
11. C. Liu, X. Zhang, S. Mei, Q. Zhou, and H. Fan, "Series-wise attention network for wind power forecasting considering temporal lag of numerical weather prediction," *Appl. Energy*, vol. 336, p. 120815, 2023. [CrossRef]
12. M.-D. Liu, L. Ding, and Y.-L. Bai, "Application of hybrid model based on empirical mode decomposition, novel recurrent neural networks and the arima to wind speed prediction," *Energy Convers. Manag.-Ment*, vol. 233, p. 113917, 2021. [CrossRef]
13. X. Liu, Z. Lin, and Z. Feng, "Short-term offshore wind speed forecast by seasonal arima-a comparison against gru and lstm," *Energy*, vol. 227, p. 120492, 2021. [CrossRef]
14. S. Singh, S. N. Singh, and A. Mohapatra, "Repeated wavelet transform based arima model for very short-term wind speed forecasting," *Renew. Energy*, vol. 136, pp. 758–768, 2019. [CrossRef]
15. S. Qureshi, F. Shaikh, L. Kumar, F. Ali, M. Awais, and A. E. Gürel, "Short-term forecasting of wind power generation using artificial intelligence," *Environ. Chall.*, vol. 11, p. 100722, 2023. [CrossRef]
16. W. Yang, Z. Tian, and Y. Hao, "A novel ensemble model based on artificial intelligence and mixed-frequency techniques for wind speed forecasting," *Energy Convers. Manag.*, vol. 252, p. 115086, 2022. [CrossRef]
17. F. Rodríguez, A. M. Florez-Tapia, L. Fontán, and A. Galarza, "Very short-term wind power density forecasting through artificial neural networks for microgrid control," *Renew. Energy*, vol. 145, pp. 1517–1527, 2020. [CrossRef]
18. F. A. Nahid, W. Ongsakul, and N. M. Manjiparambil, "Short term multi-steps wind speed forecasting for carbon neutral microgrid by decomposition based hybrid model," *Energy Sustain. Dev.*, vol. 73, pp. 87–100, 2023. [CrossRef]
19. M. Mansoor, A. F. Mirza, M. Usman, and Q. Ling, "Hybrid forecasting models for wind-pv systems in diverse geographical locations: Performance and power potential analysis," *Energy Convers. Manag.*, vol. 287, p. 117080, 2023. [CrossRef]
20. B. S. Bommidi, K. Teeparthi, and V. Kosana, "Hybrid wind speed forecasting using iceemdan and transformer model with novel loss function," *Energy*, vol. 265, p. 126383, 2023. [CrossRef]
21. Z. Zhao *et al.*, "Hybrid vmd-cnn-gru-based model for short-term forecasting of wind power considering spatio-temporal features," *Eng. Appl. Artif. Intell.*, vol. 121, p. 105982, 2023. [CrossRef]
22. M. C. Yilmaz, and Z. Orman, "LSTM derin öğrenme yaklaşımı ile Covid-19 pandemi sürecinde twitter verilerinden duygu analizi," *Acta Infologica*, vol. 5, no. 2, pp. 359–372, 2019. [CrossRef]
23. Z. Guo, W. Zhao, H. Lu, and J. Wang, "Multi-step forecasting for wind speed using a modified emd-based artificial neural network model," *Renew. Energy*, vol. 37, no. 1, pp. 241–249, 2012. [CrossRef]
24. G. K. Apostolidis, and L. J. Hadjileontiadis, "Swarm decomposition: A novel signal analysis using swarm intelligence," *Signal Process.*, vol. 132, pp. 40–50, 2017. [CrossRef]
25. G. Vashishtha, S. Chauhan, M. Singh, and R. Kumar, "Bearing defect identification by swarm decomposition considering permutation entropy

measure and opposition-based slime mould algorithm," *Measurement*, vol. 178, p. 109389, 2021. [CrossRef]
26. C. Xiao, and J. Yu, "Adaptive swarm decomposition algorithm for compound fault diagnosis of rolling bearings," *IEEE Trans. Instrum. Meas.*, vol. 72, pp. 1–14, 2023. [CrossRef]
27. E. Dokur, "Swarm decomposition technique based hybrid model for very short-term solar pv power generation forecast," *Electron. Elektrotech.*, vol. 26, no. 3, 79–83, 2020. [CrossRef]
28. E. Dokur, N. Erdogan, M. E. Salari, C. Karakuzu, and J. Murphy, "Offshore wind speed short-term forecasting based on a hybrid method: Swarm decomposition and meta-extreme learning machine," *Energy*, vol. 248, p. 123595, 2022. [CrossRef]

Mehmet Balcı received the BS degree from the Electrical and Electronics Department, and the MS degree in the Computer Engineering Department, Bilecik Seyh Edebali University, Bilecik, Türkiye in 2012 and 2019, respectively. He earned his PhD in electronics and computer engineering from Bilecik Seyh Edebali University in 2019. His research interests include deep learning algorithms, artificial neural networks, and renewable energies. He has been an IT staff member at Bilecik Seyh Edebali University since 2015.

Uğur Yüzgeç received the BS degree from the Electronics and Communication Engineering Department, Yıldız Technical University, Istanbul, Türkiye, in 1995, and the MS and PhD degrees from the Electronics and Communication Engineering Department, Kocaeli University, Kocaeli, Türkiye, in 1999 and 2005, respectively. From 1998 to 2010, he was a research assistant with the Electronics and Communication Engineering Department, Kocaeli University. Since 2020, he has been a Professor with the Computer Engineering Department, Faculty of Engineering, Bilecik Seyh Edebali University, Türkiye. His research interest includes intelligent systems and control, fuzzy and neuro-fuzzy systems, meta-heuristic algorithms, unmanned aerial vehicles, and numeric techniques in optimization problems.

Emrah Dokur received the BSc degree in electrical and electronics engineering from Istanbul University, Türkiye, in 2010, the MSc degree in electrical engineering from Istanbul Technical University, Türkiye, in 2013, and the PhD degree in energy systems engineering from Bilecik Şeyh Edebali University. He was a Postdoctoral Researcher with the Marine and Renewable Energy Center (MaREI), University College Cork, Ireland. Since 2022, he has been an associate professor with the Electrical Electronics Engineering Department, Bilecik Şeyh Edebali University. His current research interests include high voltage engineering, renewable energy systems, and power system analysis. He received the Postdoctoral Fellowship Award from the Scientific and Technological Research Council of Türkiye, in 2020, to conduct research in Ireland.