Energy 300 (2024) 131546

Contents lists available at ScienceDirect

Energy

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

A novel hybrid model based on Empirical Mode Decomposition and Echo State Network for wind power forecasting

Ugur Yuzgec [a,*], Emrah Dokur [b], Mehmet Balci [c]

[a] Bilecik Seyh Edebali University, Department of Computer Engineering, Bilecik, 11210, Turkiye
[b] Bilecik Seyh Edebali University, Department of Electrical Electronics Engineering, Bilecik, 11210, Turkiye
[c] Bilecik Seyh Edebali University, Information Technology Department, Bilecik, 11210, Turkiye

ARTICLE INFO

Keywords:
Echo State Network
Empirical Mode Decomposition
Forecasting
Wind power

ABSTRACT

Accurate wind power forecasting is essential for (i) the management of wind energy, (ii) increasing the integration of generated power into the electrical grid, and (iii) enhancing maintenance efficiency. This paper proposes a novel hybrid model for short-term (1-hour ahead) wind power forecasting. The model integrates the echo-state network (ESN) architecture with empirical mode decomposition (EMD) to improve forecast accuracy. Unlike existing approaches that use separate models for each decomposed signal after EMD method, proposed model uses a single ESN structure to predict wind power by incorporating all decomposed signals and their past values. The main advantage of this architecture is that it eliminates the need to train multiple models, thereby streamlining the forecasting process. To evaluate the performance of the proposed model, one year of data from the West of Duddon Sands, Barrow, and Horns Power offshore wind farms is utilized. Firstly, the classical ESN model and the EMD-ESN hybrid model are compared for the three datasets. Then, a comprehensive evaluation is performed by comparing the results of the proposed model with commonly used standalone and hybrid forecasting models such as Multi-Layer Perceptron (MLP), Adaptive-Neuro Fuzzy Inference System (ANFIS), Long Short-Term Memory (LSTM), Bidirectional LSTM (BiLSTM), EMD-LSTM, EMD-BiLSTM, Variational Mode Decomposition based ESN (VMD-ESN), and Wavelet Decomposition based ESN (WD-ESN). The results show that the EMD-ESN hybrid model outperforms implemented models in predicting wind power in all three datasets. Furthermore, the proposed EMD-ESN model for an onshore wind turbine power data in Germany is compared with the SOTA models, such as Transformer, Informer, Autoformer, and Graph Patch Informer (GPI), in the literature. This study highlights the superior predictive capabilities of proposed model, making it a valuable tool for enhancing the accuracy of wind power forecasts, thereby contributing to the reliable integration of wind energy into the power grid.

1. Introduction

Wind energy has emerged as the primary driver for decarbonizing global energy production, playing a significant role in achieving net-zero targets worldwide. Global Wind Energy Council (GWEC), 2023 reported that the global wind power market saw a noteworthy expansion in its new capacity, reaching 77.6 GW in 2022 [1]. This surge contributed to a cumulative installed capacity of 906 GW, reflecting a remarkable 9% increase when compared to the figures from the preceding year. Projections indicate an annual addition of approximately 136 GW of wind power capacity over the next five years, marking a substantial compound annual growth rate of 15%. Due to increased capacity efficiency and advancements throughout the entire lifecycle of offshore wind processes, this technology is regarded as essential for addressing carbon mitigation requirements and enhancing its competitiveness [2]. There are expectations for a substantial 55% reduction in the levelized cost of electricity from offshore wind by the year 2030. The EU has hosted the majority of the world's offshore wind capacity and has subsequently elevated its offshore wind power capacity goal to reach 60 GW by the year 2030 [3].

Management of wind energy generation efficiently and dependably presents a significant challenge due to its inherently volatile nature. The precision of wind power forecasting is vital, representing a critical requirement for energy producers and grid operators alike [4,5]. The complex interplay of ever-changing wind patterns, topographical influences, and environmental factors necessitates precise prediction models that can adapt to these dynamic conditions. The seamless incorporation of wind energy into the power grid and the

* Corresponding author.
E-mail addresses: ugur.yuzgec@bilecik.edu.tr (U. Yuzgec), emrah.dokur@bilecik.edu.tr (E. Dokur), mehmet.balci@bilecik.edu.tr (M. Balci).

https://doi.org/10.1016/j.energy.2024.131546
Received 10 November 2023; Received in revised form 12 April 2024; Accepted 3 May 2024
Available online 7 May 2024
0360-5442/

List of Acronyms

ANFIS	Adaptive-Neuro Fuzzy Inference System
ANN	Artificial Neural Networks
ARIMA	Autoregressive Integrated Moving Average
Bi-LSTM	Bi-Directional Long Short-Term Memory
CEEMD	Complete Ensemble Empirical Mode Decomposition
CEEMDAN	Complete Ensemble Empirical Mode Decomposition with Adaptive Noise
CNN	Convolutional Neural Network
DL	Deep Learning
EMD	Empirical Mode Decomposition
EEMD	Ensemble Empirical Mode Decomposition
EMD-LSTM	Empirical Mode Decomposition-Long Short-Term Memory
EMD-BiLSTM	Empirical Mode Decomposition-Bi-Directional Long Short-Term Memory
ESN	Echo-State Network
EU	European Union
GPI	Graph Patch Informer
GRU	Gate Recurrent Unit
GWEC	Global Wind Energy Council
IMF	Intrinsic Mode Functions
LSTM	Long Short-Term Memory
LSTNet	Long- and Short-term Time-series network
LTSF	Long-term Time Series Forecasting
MAE	Mean Absolute Error
MLP	Multi-Layer Perceptron
MSE	Mean Squared Error
NWP	Numerical Weather Prediction
PSF	Pattern Sequence-based Forecasting
RMSE	Root Mean Squared Error
RNN	Recurrent Neural Network
SOTA	State-of-the-Art
Std	Standard Deviation
TCN	Temporal Convolutional Network
VMD	Variational Mode Decomposition
VMD-ESN	Variational Mode Decomposition based Echo-State Network
VMD-ESN-STO	Variational Mode Decomposition based Echo-State Network with Subseries to The Orginal Series
WD-ESN	Wavelet Decomposition based Echo-State Network

maximization of its contribution to a more environmentally friendly and sustainable energy landscape hinge on our ability to harness the accuracy of wind power forecasting [6]. Historically, numerous forecasting models have been originally devised and implemented in the context of onshore wind installations [7]. Nonetheless, the applicability of conventional forecasting methods to offshore settings necessitates significant enhancements due to a multitude of factors. In particular, offshore wind environments are characterized by greater wind speed continuity, stronger winds, and less turbulent conditions. Additionally, the availability of offshore wind speed observations is comparatively scarcer when compared to onshore locations. Furthermore, the influence of coastal effects must also be taken into account [8,9]. In order to facilitate the successful integration of large-scale offshore wind energy into the power system while upholding reliability, it is imperative to attain a comprehensive understanding of offshore wind speed characteristics and their associated properties.

Various approaches and methods have been developed to estimate wind power, given the value and importance of wind energy. In the literature, numerous studies provide multiple techniques and approaches for wind energy prediction. These research studies include a wide range of methods, from traditional approaches to machine learning and artificial neural networks [10]. Air flow modeling [11,12], numerical weather prediction (NWP) [13] and statistical approaches [14] have been frequently employed for wind speed and power forecasting. Furthermore, in recent years, Deep Learning (DL) methods, particularly Recurrent Neural Network (RNN) [15,16] and Convolutional Neural Networks (CNN) [17,18], have significant potential to improve wind forecasting. Most of these studies focus on improving the methods and modeling to increase forecasting accuracy and optimize energy production. However, the estimation results of current models in the wind energy sector indicate certain limitations and potential improvements.

Decomposition methods are often employed as a preprocessing in the literature, to enhance the effectiveness of models employed for wind energy forecasting. Decomposition-based models belong to a category of hybrid models that initially break down the wind speed/power time series into more stable subseries. Subsequently, individual forecasting models are constructed for each of these subseries. The attributes of each subseries are significantly influenced by both the decomposition length and the level of decomposition [19]. Therefore, it is imperative to meticulously ascertain these factors. Presently, in existing decomposition-based models, the determination of the decomposition length and level typically relies on empirical approaches. Enhancing the efficacy of decomposition-based models in real-world forecasting scenarios necessitates the development of methodologies for accurately selecting the appropriate decomposition length and level. In the majority of current decomposition-based models, the ultimate predictive outcomes commonly involve aggregating the forecasting results of each subseries directly. Nevertheless, as the predictability of individual subseries can vary, the distribution of prediction residuals may differ across them. A potential approach to enhance the effectiveness of decomposition-based models involves combining the prediction results from each subseries using an appropriate weighting model. Such an analysis approach dissects intricate data sets into simpler components or features to comprehend the underlying data structure, reveal concealed patterns, and concentrate on specific components. Decomposition methods facilitate the processing, visualization and analysis of data. They can significantly enhance the performance of wind energy forecast accuracy by converting wind data into an easily understandable and practical set of components [19]. The following are the most widely used decomposition techniques in signal processing: Empirical Mode Decomposition (EMD), Ensemble Empirical Mode Decomposition (EEMD), Variational Mode Decomposition (VMD), Wavelet Packet Transform (WPT), Complete Ensemble Empirical Mode Decomposition (CEEMD), and Complete Ensemble Empirical Mode Decomposition with Adaptive Noise (CEEMDAN).

Hybrid models, which combine decomposition techniques with other forecasting methods, are becoming increasingly important in wind power or speed forecasting. For example, Liu et al.'s (2012) [20] investigation introduces a hybrid EMD-ANN model for forecasting wind speeds, merging Empirical Mode Decomposition (EMD) with Artificial Neural Networks (ANNs). Another notable example is the research by Bokde and colleagues (2020) [21], who present two hybrid intelligent models, EEMD-PSF and EEMD-PSF-ARIMA, which demonstrate a significant improvement in the accuracy of wind speed and power prediction through the integration of Ensemble Empirical Mode Decomposition (EEMD) with Pattern Sequence-based Forecasting (PSF) and Autoregressive Integrated Moving Average (ARIMA) models. In addition, Duan et al. (2018) [22] have used a hybrid forecasting model based on a long short-term memory (LSTM) neural network combined with an improved variational mode decomposition to improve wind power forecasting. In the study by Yagang Zhang et al. (2022) [23], CEEMDAN, a decomposition algorithm, is selected to process wind speed sequences into IMFs, followed by further decomposition using SVD, optimizing Elman's parameters using PSO, and using ARIMA for prediction. These studies demonstrate a growing trend in the literature to develop hybrid models that exploit the benefits of decomposition techniques and other modeling approaches, resulting in better and more reliable wind energy forecasts. The results from these studies also confirm the effectiveness of decomposition-based hybrid approaches in improving forecast accuracy, outperforming stand-alone models.

Building on the positive outcomes achieved by integrating decomposition techniques and other modeling approaches into hybrid models, our focus turns to the essential component of wind forecasting: the Echo State Network (ESN) model. In addition to wind forecasting, ESN models have been used effectively in areas such as time series prediction of financial markets [24], robotics applications [25,26], fault diagnosis [27], and speech emotion recognition [28]. Comprising an input layer, a reservoir layer, and an output layer, ESN models are characterized by their rapid and efficient recurrent neural network structure. The neurons within the reservoir layer are sparsely connected, and the weights of the input and reservoir layers are predetermined. Linear regression allows for the output weights to be trainable. Several studies have utilized the ESN model for wind power and speed prediction. Chitsazan et al. (2019) introduced two novel methods for wind speed and direction forecasting based on the internal states of ESN [29]. Tian (2021) [30] introduced a hybrid approach that employs an enhanced particle swarm optimization technique to fine-tune the parameters of Echo State Networks (ESN) for the purpose of wind speed forecasting. Wang and colleagues (2019) also employed the ESN model with wavelet transform for wind power prediction, highlighting its superiority compared to traditional methods and thus providing reliable forecasts for wind energy producers [31]. Hu et al. (2023) presented a hybrid forecasting model named VMD-ESN-STO, which combines rolling Variational Mode Decomposition (VMD), Echo State Networks (ESN), and the subseries to the original series (STO) structure [32]. The model outperforms six comparative models across four wind speed datasets, demonstrating its effectiveness in wind speed forecasting. These examples highlight the adaptability and significant potential of ESN in enhancing wind power and speed projection. This makes it an essential resource for the renewable energy sector.

In this study, we aim to address the critical requirement for accurate short-term wind power forecasting, which is essential for the seamless integration of wind energy into the grid. Our approach involves introducing a novel hybrid model, combining Echo-State Networks (ESN) with Empirical Mode Decomposition (EMD), in order to considerably improve forecast accuracy. The rationale for our study originates from the inherent unpredictability of wind power, which necessitates dependable prediction techniques to maximize energy generation and grid functionality. In contrast to conventional methodologies, our proposed approach simplifies the procedure by removing the requirement for training multiple models for each decomposed signal, underscoring the efficiency and feasibility of our methodology. After undergoing thorough evaluation against established models, our study highlights the superior predictive capabilities of the EMD-ESN hybrid model, positioning it as a valuable tool for the renewable energy sector, contributing to sustainable energy production and grid reliability. The contributions made to the literature in this study are given below:

- This paper presents a novel hybrid model that combines Echo-State Networks (ESN) and Empirical Mode Decomposition (EMD) to enhance short-term wind power forecasting.
- The forecasting process is streamlined by using one ESN structure to predict wind power. This incorporates all decomposed signals and their past values, thereby eliminating the necessity of training multiple models.
- Via a comprehensive evaluation, we compare the results of the proposed model with commonly employed models in the literature, including Multi-Layer Perceptron (MLP), Adaptive-Neuro Fuzzy Inference System (ANFIS), Long Short-Term Memory (LSTM), Bidirectional Long Short-Term Memory (BiLSTM), Empirical Mode Decomposition based Long Short-Term Memory (EMD-LSTM), Empirical Mode Decomposition based Bidirectional Long Short-Term Memory (EMD-BiLSTM), Variational Mode Decomposition based Echo State Network (VMD-ESN), and Wavelet Decomposition based Echo State Network (WD-ESN).
- The superior predictive performance of the EMD-ESN hybrid model across all three datasets demonstrates its potential for accurately forecasting wind power.
- To evaluate the proposed EMD-ESN model with SOTA models, a public dataset from an onshore wind farm was utilized. The SOTA models includes LSTM, Gate Recurrent Unit (GRU), Temporal Convolutional Network (TCN), Long-term Time Series Forecasting (LTSF)-Linear (DLinear), Long- and Short-term Time-series network (LSTNet), Transformer, Informer, Autoformer, and Graph Patch Informer (GPI).

- The model's significance lies in its contribution to reliable integration of wind energy into the power grid, which supports sustainable and efficient energy generation.

The remaining parts of this study can be summarized as follows: Section 2 describes the methods used, such as EMD and Echo State Network Model. Furthermore, it describes how the hybrid model based on EMD and ESN is constructed and the datasets used. Performance evaluation metrics are also presented in this section. Section 3 presents the results obtained and the discussion of these results. Finally, a summary of the study is presented in the last section.

2. Methods and datasets

This section details a new approach for short-term wind power forecasting developed in this study. Our methodology exploits the synergy between Empirical Mode Decomposition (EMD) and Echo-State-Network (ESN) modeling. Unlike conventional methods that utilize distinct models for each decomposed signal, our technique combines all decomposed signals and their past data into a singular ESN model. The subsequent subsections present a comprehensive account of the EMD-ESN hybrid model, information about the datasets, and performance evaluation criteria employed in our investigation.

2.1. Empirical mode decomposition

Empirical Mode Decomposition (EMD) is a signal processing technique with notable attention in recent years. It is capable of efficiently analyzing and decomposing complex, non-linear, and non-stationary time series data. Dr. Norden E. Huang developed EMD method in the 1990s [33] for use in various fields such as meteorology, finance, and biomedical signal processing. At its essence, EMD is a method driven by data that aims to adaptively break down a given time series into a limited number of intrinsic mode functions (IMFs). These IMFs represent the fundamental oscillatory modes that are present in the data, and their derivation is based on the local properties of the data instead of predetermined basis functions or assumptions about the stationarity of the data [34].

The Empirical Mode Decomposition (EMD) algorithm comprises several essential procedures. Initially, the sifting process identifies the local extrema in the time series and generates upper and lower envelopes to acquire the first Intrinsic Mode Function (IMF) that represents the highest frequency oscillations. The extraction of IMF then proceeds by repeating the process iteratively, which involves subtracting the present IMF from the original signal, thereby obtaining subsequent IMFs, each signifying oscillatory components that vary in timescale. Finally, the residue is obtained. The residual signal, following the extraction of all IMFs, depicts the extended trend or low-frequency characteristics of the data [35]. Where $x(t)$ represents a particular time series, the EMD calculation steps are defined as follows.

Step 1: All local extrema within the signal, denoted as $x(t)$, are initially identified. These local minimum and maximum values are then employed for interpolation, resulting in the generation of respective upper and lower envelopes ($U_x(t)$ and $L_x(t)$) using cubic splines.

Step 2: The average envelope value $m(t)$ and the detailed component $d(t)$ are computed according to Eqs. (1) and (2), correspondingly.

$$m(t) = \frac{U_x(t) + L_x(t)}{2} \tag{1}$$

$$d(t) = x(t) - m(t) \tag{2}$$

Step 3: Until d(t) transforms into an IMF, the process proceeds based on the following criteria:

$$\sum_{t=1}^{T} \frac{\left[d_j(t) - d_{j-1}(t)\right]^2}{\left(d_{j-1}(t)\right)^2} \leq \delta \tag{3}$$

where T represents the signal length and j denotes the number of iterative calculations. The usual value for δ is typically set between 0.2 and 0.3.

Step 4: Continue iterating through steps one to three until all intrinsic mode functions (IMFs) and the detailed signal have been acquired. Ultimately, the original time series $x(t)$ may be decomposed as given below:

$$x(t) = c_i(t) + R_n(t) \tag{4}$$

where, $c_i(t)$ and $R_n(t)$ stand for IMF signals (i=[1 n]) and residual signal. An IMF within EMD captures signals that have significantly different scales, or it can represent a signal of the same scale appearing in different components.

The adaptability of EMD to capture both high-frequency and low-frequency components in a data-driven manner renders it an ideal preprocessing technique for wind power forecasting. By decomposing wind speed and power time series data into IMFs, EMD permits concentration on pertinent forecast-relevant components while removing noise and undesirable variations. In this study, we utilize EMD as a vital constituent of our combined wind power prediction model, showcasing its efficacy in identifying the underlying patterns in wind data and improving the precision of short-term wind power forecasts.

2.2. Echo state network model

The Echo State Network (ESN) model represents a capable and adaptable architecture for a Recurrent Neural Network (RNN). The model has gained significant attention in recent years due to its capacity for processing complicated temporal data. ESNs are particularly proficient in handling tasks such as time series prediction [36], pattern recognition [37], and signal processing [38].

At the heart of an ESN is a dynamic reservoir of interconnected nodes. These activations collectively capture temporal dependencies and complex patterns in input data. What differentiates ESNs from traditional RNNs is the fixed, random connectivity of these nodes, known as the "reservoir", which simplifies the challenges of training recurrent networks. ESNs are recognized for their streamlined architecture, effortless training, and exceptional performance across an extensive range of applications [39].

In Fig. 1, the basic architecture of the ESN model is shown. The ESN model comprises three main layers: an input layer, a dynamic reservoir, and an output layer. The input layer's role is to receive input signals, which stimulate the network's activation. The dynamic reservoir contains many sparsely connected neurons and replaces the hidden layer in ESN. This reservoir processes information. Finally, the output layer generates the network's output signal. Looking at the ESN architectural structure in Fig. 1, it seems that the input layer is assumed to possess H input units, the reservoir contains N reservoir neurons, and the output layer comprises L neurons. Here, $u(k) = \left[u_1(k), u_2(k), \ldots, u_H(k)\right]^T$ denotes input vector, $x(k) = \left[x_1(k), x_2(k), \ldots, x_N(k)\right]^T$ represents state vector of dynamic reservoir, and $y(k) = \left[y_1(k), y_2(k), \ldots, y_L(k)\right]^T$ is the output vector. The weight matrix among the internal reservoir units is represented by W. The input weight matrix (W_{in}) indicates the connectivity between the input layer and the reservoir neurons.

The output weight matrix is represented by W_{out}. The matrix W_{back} represents the feedback weight relationships linking the output layer to the reservoir neurons. The update on the status of ESN reservoirs and outputs can be expressed as follows:

$$x(k+1) = f(W_{in}u(k+1) + W x(k) + W_{back} y(k)) \tag{5}$$

$$y(k+1) = g(W_{out} x(k+1)) \tag{6}$$

where f stands for the activation functions of the neurons in the reservoir and g denotes the activation functions of the neurons in

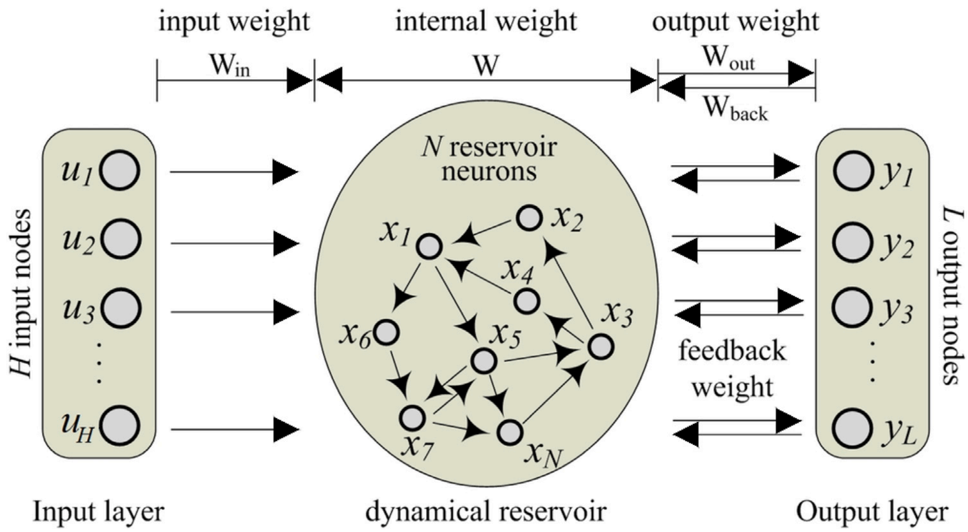


Fig. 1. ESN model architecture [40].

the output layer. Throughout the training process, the input weight matrix W_{in}, internal weight matrix W, and feedback weight matrix W_{back} remain static, maintaining their initialized values, while only the output weight matrix W_{out} is subject to adaptation through simple linear regression, guided by the training dataset. This approach streamlines the training process by maintaining the original random initialization of weight matrices, specifically those representing the reservoir's internal dynamics, and keeping them constant during training. Therefore, the primary focus of modeling an ESN concerns the computation of the output weight matrix. The ESN model was trained using the ridge regression method. The steps of the ESN training algorithm are given below:

1. Parameters Initialization: Set the reservoir's size (N), leakage rate (α), spectral radius ρ, W_{in}, W, W_{back}, and washout time step (WOT).
2. Reservoir States Update: Calculate the new reservoir states $x(k+1)$ with the following formula:

$$x(k+1) = (1-\alpha)x(k) + \alpha \cdot f(W_{in}u(k+1) + W x(k) + W_{back} y(k)) \quad (7)$$

   Here, the parameter named leakage rate, denoted by α, controls the extent to which the previous state is retained.
3. Output Weight Matrix Elements Calculation: Calculate output weights W_{out} with Eq. (8) using the reservoir state vector (X) and the expected output vector (Y) at moments greater than or equal to WOT.

$$W_{out} = Y^T X^T \left(X X^T + \lambda I\right)^{-1} \quad (8)$$

   Here, I denotes the identity matrix, and λ represents the regularization parameter (used as $1e-8$).

2.3. Hybrid model based on Empirical Mode Decomposition (EMD) and Echo State Network (ESN)

This subsection presents a framework of the new hybrid model for wind power forecasting. Our model, which combines Empirical Mode Decomposition (EMD) and Echo State Network (ESN), provides a reliable and efficient approach to address the intricate and ever-changing nature of wind power generation. The combination of these efficient techniques aims to maximize the benefits of both approaches, ultimately leading to improved accuracy and dependability in short-term wind power predictions. Throughout this section, we explore the structure and essential elements of the proposed innovative composite model. This model adopts EMD's signal decomposition and ESN's modeling to produce precise wind power forecasts. Firstly, the EMD provides a way to refine the wind data. It is clear that more stable and linear signals must be obtained to improve the forecasting performance. The EMD tackles the challenges effectively due to the non-linearity of the signal. It also preserves the main characteristics of the original signal. Secondly, the proposed ESN model includes a novel input approach process. The forecasting process is streamlined by using one ESN structure to predict offshore wind power. Fig. 2 illustrates the structure of the proposed hybrid EMD-ESN model. The steps of working the proposed EMD-ESN model are indicated with numbers in the figure. The steps outlining the forecast process for the EMD-ESN hybrid model are provided below:

1. Data collection: The real wind power data extending over one-year periods were collected from three distinct offshore wind farms. The data was gathered at hourly intervals, covering intervals of time ranging from December 4, 2015, to March 31, 2020, for West of Duddon Sands, December 31, 2014, to April 30, 2020, for Horns Power, and May 1, 2019, to April 30, 2020, for Barrow. Each dataset consists of 8760 hourly observations, rendering a comprehensive outlook of the wind power performance at these sites.
2. Wind power data decomposition: The hourly wind power data obtained from offshore wind farms is decomposed into more basic and inherent components referred to as empirical modes via the implementation of the EMD method.
3. Data normalization: Normalization of data is performed on each empirical mode signal that undergoes the EMD technique to achieve consistent values within a range of 0 to 1.
4. Data preparation: For each normalized empirical mode signal, time series were created for use in both the input and output of the ESN model. To obtain these time series, the sliding window technique was used to generate four input series ($x(t)$, $x(t-1)$, $x(t-2)$, and $x(t-3)$) and one output series ($x(t+1)$).
5. Setting the dataset for training and testing: 70% of the prepared time series of all normalized empirical mode signals were used in the training of the ESN model and the rest in its testing process.
6. Training ESN model: To train the ESN model, the model input is IMF signal data obtaining by EMD preprocessing method, while wind power data as the model output. All of the IMF signals are utilized at the inputs of the model. At iterations t, (t − 1), (t − 2), and (t − 3), each IMF signal serves as an input for the wind power forecast at the (t + 1)th iteration. The number of model entries is the product of the number of decomposed

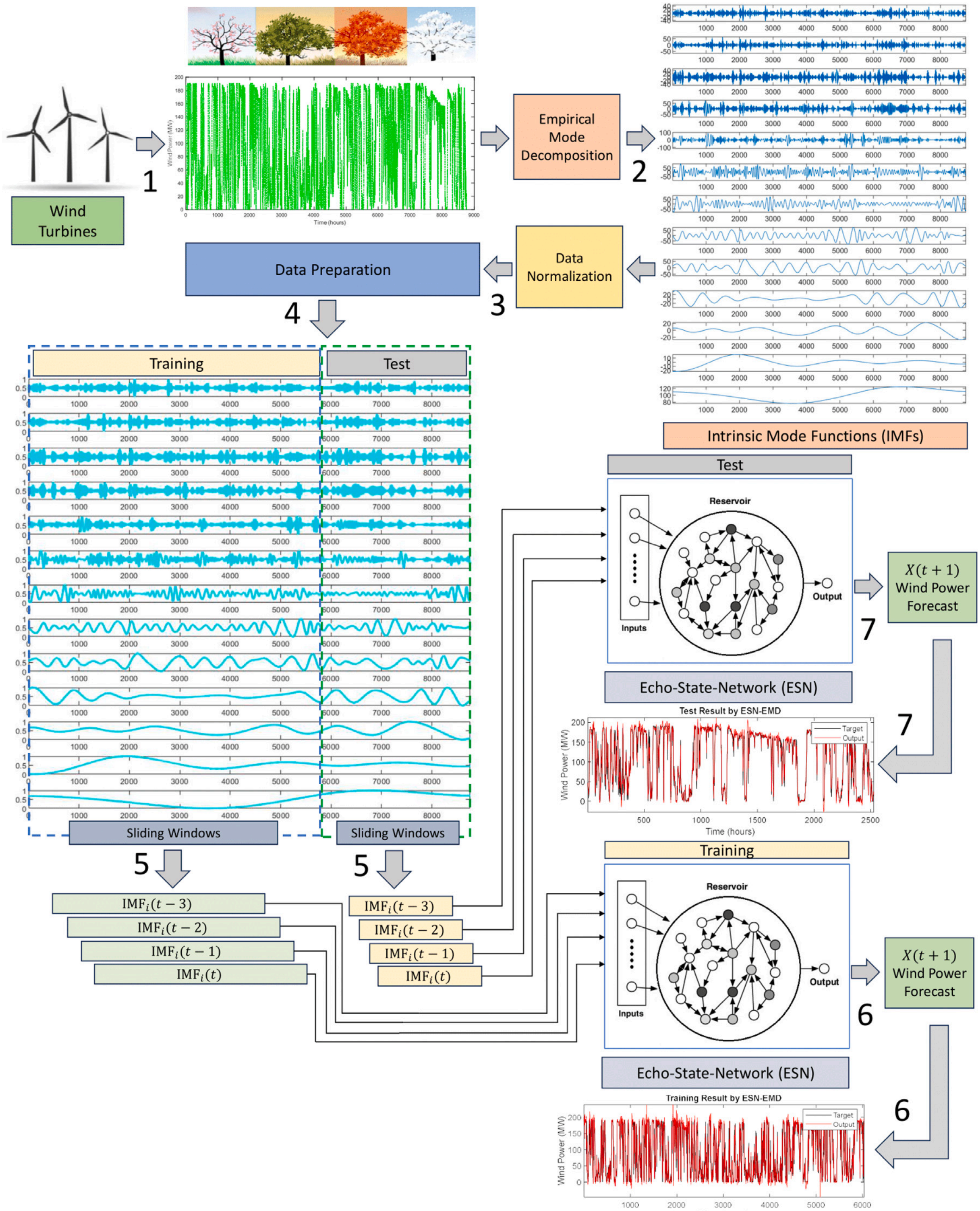


Fig. 2. Framework of hybrid EMD-ESN model.

signals and one more of the past time step. In our study, the number of signals decomposed by EMD is 13 and the number of past time steps is 3, so the total number of ESN model inputs is 52. Here, ESN model is employed due to its excellent nonlinear fitting ability and high training efficiency.

7. Test of the ESN model: At this stage of the experiment, the ESN model, which has been trained beforehand, is tested with 30% of the data set that has never been used during the training process. Consequently, the model's reaction to a previously unknown data set is established. RMSE, MSE, MAE and R2 metrics were utilized to assess the model's performance.

2.4. Datasets

This study used datasets comprising one year of real wind power data from three diverse offshore wind farms to assess the effectiveness of the proposed hybrid model. The wind farms mentioned are West of Duddon Sands and Barrow, situated between England and Ireland, and Horns Power, located near the coast of Denmark in the North Sea. The capacity of West of Duddon Sands is 388.8 MW with an 80.0-m hub height, and it utilizes 108 Siemens SWT-3.6-120 turbines. The offshore wind farm, Barrow, has been operational from midnight on December 4th, 2015 until 9 p.m., on March 31st, 2020. It boasts a capacity of

90.0 MW and a hub height of 75.0 m, utilizing 30 Vestas V90-3000 offshore turbines. Similarly, Horns Power, with a capacity of 160 MW and a hub height of 70.0 m, employs 49 Vestas V80-2000 turbines. Horns Power has been in operation since midnight on December 31st, 2014 until 10 p.m. on April 30th, 2020. The data sets, comprising hourly measurements, were collected over one-year periods from each diverse location. Each of the three data sets comprises 8760 observations gathered hourly between 02.12.2018 at 21:00 and 31.03.2020 at 21:00 for West of Duddon Sands, between 01.05.2019 at 23:00 and 30.04.2020 at 22:00 for Horns Power, and between 22.12.2018 at 21:00 and 31.03.2020 at 21:00 for Barrow. In the Horns Power dataset, data recorded between the hours of 6603 and 6801 during a one-year segment was registered as zero due to multiple technical issues with the wind turbines at that time. Therefore, we removed this anomalous, insignificant data from this dataset.

2.5. Performance metrics

When evaluating different models for estimating short-term wind power, the study utilized four key performance metrics to thoroughly assess their effectiveness. These metrics included Mean Squared Error (MSE), which calculates the average of squared errors between predicted and observed values, offering insight into overall prediction accuracy. Root Mean Squared Error (RMSE), however, is a valuable indicator since it takes the square root of MSE, rendering it interpretable in the same unit as the original data and thus providing a more intuitive comprehension of prediction errors. Mean Absolute Error (MAE) calculated the average absolute differences between predicted and actual values, providing a durable indicator of model performance. Lastly, the models' capabilities in short-term wind power estimation were comprehensively assessed using R-squared (R^2), a widely recognized coefficient of determination, which measured the proportion of variance in the wind power data explained by the models. These metrics together provided a comprehensive evaluation of the models' performance. Calculation formulas for all metrics are provided in the equations presented below:

$$MSE = \frac{\sum_{i=1}^{N}(y_i - \tilde{y}_i)^2}{N} \tag{9}$$

$$RMSE = \sqrt{\frac{\sum_{i=1}^{N}(y_i - \tilde{y}_i)^2}{N}} \tag{10}$$

$$MAE = \frac{1}{N}\sum_{i=1}^{N}|y_i - \tilde{y}_i| \tag{11}$$

$$R^2 = 1 - \frac{\sum_{i=1}^{N}(y_i - \tilde{y}_i)^2}{\sum_{i=1}^{N}(y_i - \bar{y}_i)^2} \tag{12}$$

where y_i and $\tilde{y}_i$ represent the actual and predicted values, respectively. N denotes the number of samples.

3. Results and discussion

This section presents the experimental results of the proposed EMD-ESN hybrid model. Firstly, we acquired comparative outcomes of both the original ESN model and the hybrid EMD-ESN model for three different offshore wind power datasets. Next, an in-depth analysis was conducted using popular models such as Multi-Layer Perceptron (MLP), Adaptive-Neuro Fuzzy Inference System (ANFIS), Long Short-Term Memory (LSTM), Bidirectional Long Short-Term Memory (BiLSTM), Empirical Mode Decomposition based Long Short-Term Memory (EMD-LSTM), and Empirical Mode Decomposition based Bidirectional Long Short-Term Memory (EMD-BiLSTM), frequently employed in prediction processes in the literature. Finally, a comparative study with State-of-the-Art (SOTA) models for different horizon values for a onshore wind turbine power data in Germany is presented.

3.1. Experiment 1: Comparison of EMD-ESN and other models for offshore wind farms

The EMD-ESN hybrid model was implemented and assessed through MATLAB on a personal computer that features an Intel Core i7-8550U Microprocessor, clocked at 1.80 GHz, a NVIDIA GeForce MX150 graphics card, and 20 GB of RAM. For the ESN model, washout was set to 100, reservoir size 50, leakage rate 0.3, and spectral radius 0.5. Furthermore, the activation function in the reservoir was set to the tanh function, while the outlet layer was set to the linear function. Initially, we compared the performance of the ESN model and our hybrid proposal. For both models, the training and testing phases were carried out 10 times. In Fig. 3, the best training and testing results obtained for the Barrow dataset are shown comparatively. Here, the blue color represents the 1-h wind power forecast generated by the proposed EMD-ESN model, the red color represents the 1-h wind power forecast produced by the ESN model, and the black color shows the actual 1-h wind power signal. Table 1 summarizes the statistical results obtained from the training process of the original ESN and the proposed hybrid EMD-ESN model the for Barrow dataset. This table includes the error metrics (MSE, RMSE, and MAE) and R^2 values for both models. Similarly, Table 2 displays the statistical test outcomes of both models using the same dataset. The best metric values are highlighted in bold in both tables for the Barrow dataset.

The training and testing results of the models demonstrate the superiority of the proposed EMD-ESN model over the classical ESN model for this dataset. The primary reason for this is due to the EMD decomposition method utilized before the model implementation. Nonetheless, the literature indicates the common practice of employing separate models for each decomposed signal, thus leading to a substantial increase in training time for the hybrid model. Thanks to the EMD-ESN model architecture proposed in this study, training running times that can compete with the ESN model have been attained. For example, the training time of the proposed new EMD-ESN model architecture (0.1956 s) is only two times longer than that of the ESN model (0.0987 s). However, in the classical architecture where one ESN model is used for each IMF signal, this training time is up to ten times longer.

In the second analysis, the performance of both models was evaluated using the West of Duddon Sands dataset. The best training and testing outcomes achieved for the West of Duddon Sands dataset are demonstrated in Fig. 4. Additionally, Tables 3 and 4 provide statistical results of both models after being run 10 times during the training and testing phases for the West of Duddon Sands dataset. Here, the analysis indicates that the proposed hybrid EMD-ESN model outperforms the ESN model, similar to the Barrow dataset.

The mean R^2 values of the hybrid model were 98.08% (93.58% for ESN) and 97.52% (92.69% for ESN) for training and testing, respectively. During the training phase, it is evident that the ESN model outperforms the hybrid model in terms of running times. However, the results indicate that the hybrid model's running times are competitive with those of the ESN model. Statistical results are obtained using the third dataset, Horns Power dataset, after comparing the original ESN model and the proposed hybrid EMD-ESN model. The best training and test plots, acquired by running both models for 10 times, are presented in Fig. 5.

Table 5 presents the statistical outcomes for the error metrics and running times of both models achieved through training. In Table 6, the comparative test performances of the ESN and EMD-ESN models are depicted. As in the previous tables, the best results are shown in bold.

The error metrics in both the training and test results demonstrate that the proposed hybrid EMD-ESN model is more effective than the classical ESN model when predicting wind power for this dataset. R^2 results demonstrate the same success of the EMD-ESN model in the training and testing phases. As the running time is as crucial as prediction performance in training, the analysis of running times for

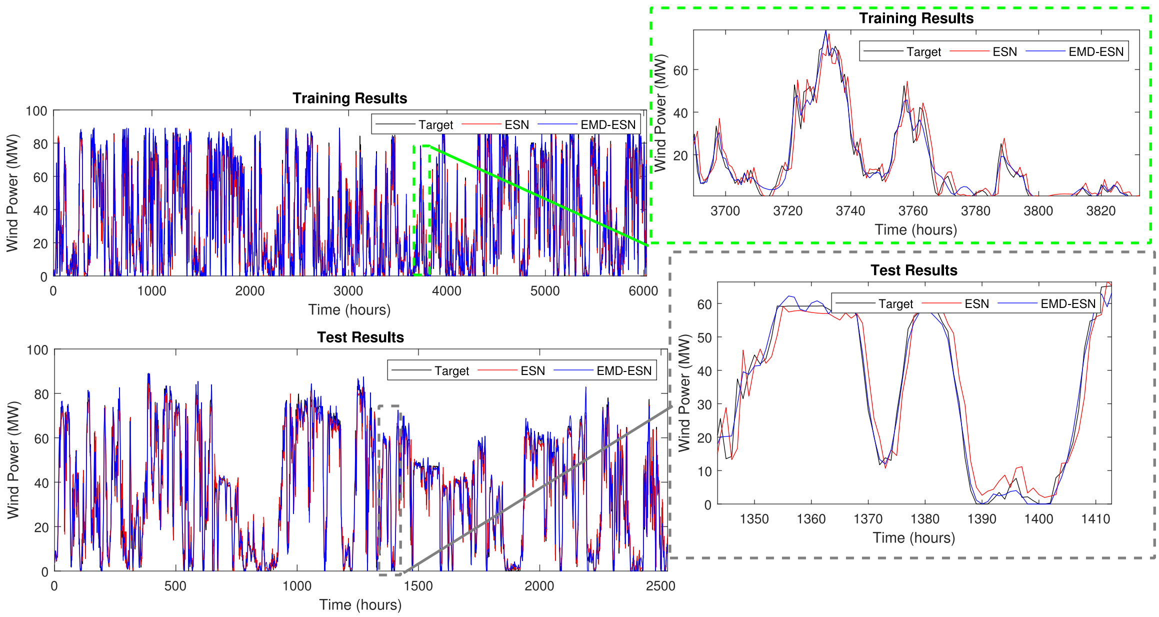


Fig. 3. The training and test results of the proposed EMD-ESN hybrid model for Barrow dataset.

Table 1
Training results of the proposed EMD-ESN and ESN models for Barrow dataset.

Metric	Best		Worst		Median		Mean		Std	
	ESN	EMD-ESN	ESN	EMD-ESN	ESN	EMD-ESN	ESN	EMD-ESN	ESN	EMD-ESN
MSE	72.693	**18.633**	73.128	**18.817**	72.807	**18.713**	72.814	**18.718**	1.239E−01	**6.528E−02**
RMSE	8.526	**4.317**	8.552	**4.338**	8.533	**4.326**	8.533	**4.326**	**7.256E−03**	7.543E−03
MAE	5.600	**3.038**	5.636	**3.057**	5.620	**3.048**	5.620	**3.048**	1.010E−02	**5.652E−03**
R2	0.9149	**0.9782**	0.9144	**0.9780**	0.9147	**0.9781**	0.9147	**0.9781**	1.451E−04	**7.644E−05**
Run time	**0.0245**	0.0299	**0.0294**	0.0339	**0.0252**	0.0304	**0.0258**	0.0311	1.642E−03	**1.485E−03**

Table 2
Test results of the proposed EMD-ESN and ESN models for Barrow dataset.

Metric	Best		Worst		Median		Mean		Std	
	ESN	EMD-ESN	ESN	EMD-ESN	ESN	EMD-ESN	ESN	EMD-ESN	ESN	EMD-ESN
MSE	64.177	**15.458**	65.102	**21.103**	64.661	**16.788**	64.635	**17.141**	**3.674E−01**	1.600E+00
RMSE	8.011	**3.932**	8.069	**4.594**	8.041	**4.097**	8.040	**4.136**	**2.285E−02**	1.877E−01
MAE	5.173	**2.833**	5.263	**3.515**	5.229	**3.006**	5.222	**3.034**	**2.697E−02**	1.989E−01
R2	0.8995	**0.9758**	0.8980	**0.9670**	0.8987	**0.9737**	0.8988	**0.9732**	**5.755E−04**	2.506E−03

Table 3
Training results of the proposed EMD-ESN and ESN models for West of Duddon Sands dataset.

Metric	Best		Worst		Median		Mean		Std	
	ESN	EMD-ESN	ESN	EMD-ESN	ESN	EMD-ESN	ESN	EMD-ESN	ESN	EMD-ESN
MSE	340.300	**101.400**	342.110	**102.820**	341.130	**102.320**	341.150	**102.280**	8.043E−01	**3.872E−01**
RMSE	18.447	**10.070**	18.496	**10.140**	18.470	**10.115**	18.470	**10.113**	2.177E−02	**1.916E−02**
MAE	11.513	**7.052**	11.562	**7.078**	11.539	**7.063**	11.541	**7.064**	1.635E−02	**8.277E−03**
R2	0.9360	**0.9809**	0.9357	**0.9807**	0.9358	**0.9808**	0.9358	**0.9808**	1.513E−04	**7.282E−05**
Run time	**0.0258**	0.0299	**0.0562**	0.0847	0.0439	**0.0328**	**0.0426**	0.0436	**1.143E−02**	2.005E−02

Table 4
Test results of the proposed EMD-ESN and ESN models for West of Duddon Sands dataset.

Metric	Best		Worst		Median		Mean		Std	
	ESN	EMD-ESN	ESN	EMD-ESN	ESN	EMD-ESN	ESN	EMD-ESN	ESN	EMD-ESN
MSE	333.570	**99.584**	336.730	**158.190**	334.780	**107.620**	334.980	**113.840**	**1.211E+00**	1.911E+01
RMSE	18.264	**9.979**	18.350	**12.577**	18.297	**10.374**	18.302	**10.639**	**3.307E−02**	8.524E−01
MAE	11.259	**7.013**	11.454	**8.957**	11.334	**7.452**	11.340	**7.568**	**6.634E−02**	6.105E−01
R2	0.9272	**0.9783**	0.9265	**0.9655**	0.9269	**0.9765**	0.9269	**0.9752**	**2.643E−04**	4.173E−03

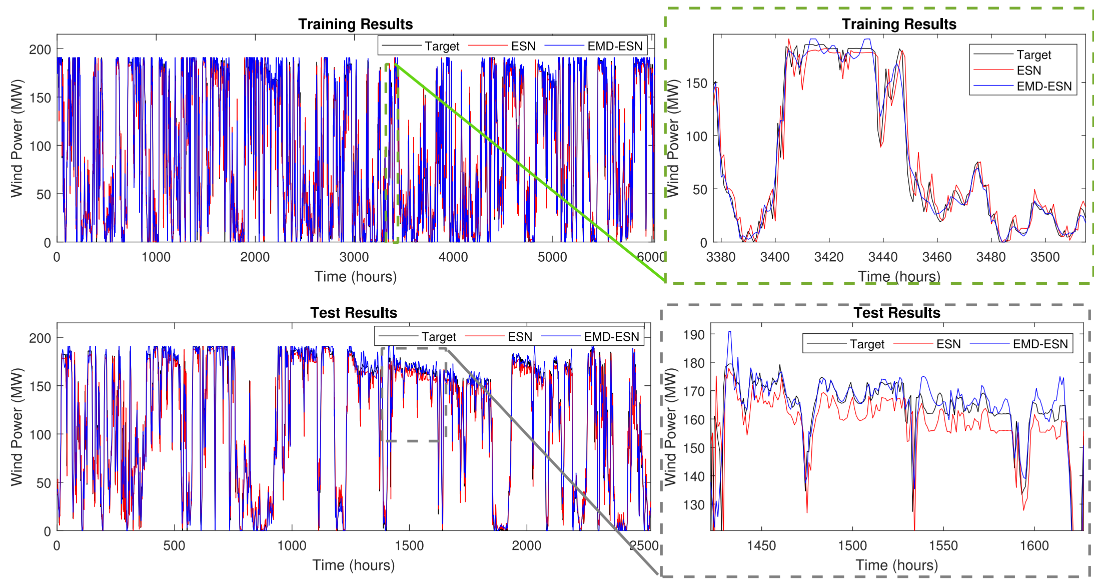


Fig. 4. The training and test results of the proposed EMD-ESN hybrid model for West of Duddon Sands dataset.

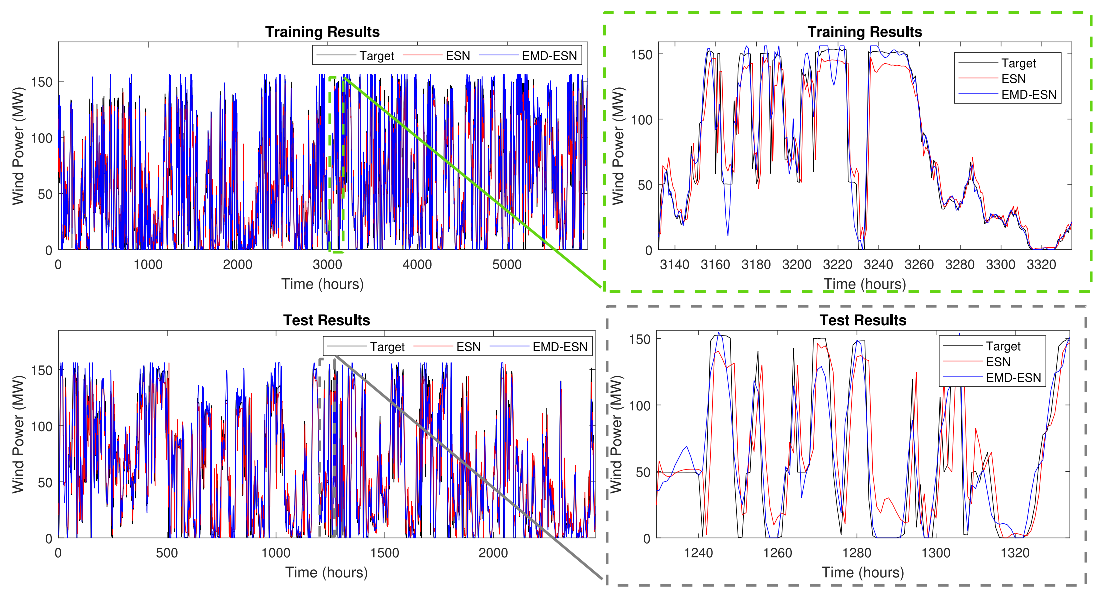


Fig. 5. The training and test results of the proposed EMD-ESN hybrid model for Horns Power dataset.

Table 5
Training results of the proposed EMD-ESN and ESN models for Horns Power dataset.

Metric	Best		Worst		Median		Mean		Std	
	ESN	EMD-ESN	ESN	EMD-ESN	ESN	EMD-ESN	ESN	EMD-ESN	ESN	EMD-ESN
MSE	384.950	**124.730**	388.970	**126.290**	387.120	**125.530**	386.990	**125.450**	1.163E+00	**4.524E−01**
RMSE	19.620	**11.168**	19.722	**11.238**	19.675	**11.204**	19.672	**11.200**	2.957E−02	**2.020E−02**
MAE	12.483	**7.577**	12.588	**7.625**	12.516	**7.610**	12.518	**7.609**	3.063E−02	**1.437E−02**
R2	0.8514	**0.9519**	0.8498	**0.9513**	0.8506	**0.9515**	0.8506	**0.9516**	4.491E−04	**1.747E−04**
Run time	**0.0261**	0.0296	**0.0293**	0.0617	**0.0266**	0.0453	**0.0269**	0.0454	**1.076E−03**	1.597E−02

Table 6
Test results of the proposed EMD-ESN and ESN models for Horns Power dataset.

Metric	Best		Worst		Median		Mean		Std	
	ESN	EMD-ESN	ESN	EMD-ESN	ESN	EMD-ESN	ESN	EMD-ESN	ESN	EMD-ESN
MSE	577.700	**239.800**	**586.770**	758.320	584.400	**285.810**	583.940	**332.430**	**2.524E+00**	1.549E+02
RMSE	24.035	**15.486**	**24.223**	27.538	24.174	**16.901**	24.165	**17.914**	**5.230E−02**	3.578E+00
MAE	15.032	**10.614**	**15.304**	22.507	15.148	**12.570**	15.159	**13.464**	**7.382E−02**	3.399E+00
R2	0.7759	**0.9070**	**0.7724**	0.7058	0.7733	**0.8891**	0.7735	**0.8710**	**9.790E−04**	6.011E−02

both models in this phase was also conducted. The same outcome was achieved with this dataset as with the other two datasets. The EMD-ESN model has a longer training time than the ESN model due to the computational load caused by EMD preprocessing. However, when compared to other decomposed hybrid models in previous studies, the singular ESN model architecture employed in this research displays a favorable run-time for the training process.

To summarize, it is understood that the proposed EMD-ESN model predicts wind power more successfully than the original ESN model for all three data sets. According to the average values of the statistical MSE error metrics presented in the tables, the success of the proposed model is numerically demonstrated. For the Barrow dataset, the mean MSE achieved during ESN model training decreased by 74.293% with the implementation of the proposed EMD-ESN model. The test results showed a 73.480% improvement in MSE performance. In terms of the West of Duddon Sands dataset, the EMD-ESN model was able to reduce the average MSE metric value during training by 70.019% and during testing by 66.016%. For the last dataset, Horns Power, our hybrid model achieved a 67.583% error reduction in training and a 43.071% error reduction in testing. By integrating the ESN architecture with empirical mode decomposition (EMD), the model effectively decomposes the intricate dynamics of wind behavior into distinct IMFs, each representing a specific timescale or frequency component.

Unlike traditional approaches that rely on separate models for individual IMFs, our model ingeniously incorporates all IMFs and their historical data into a single ESN framework. This comprehensive methodology enables the model to dynamically adjust to the evolving nature of wind data, efficiently discerning the correlation between short-term fluctuations and longer-term trends. The model additionally derives from the echo-state network's inherent capacity for managing nonlinear relationships and complex data, enabling it to detect nuanced, intricate connections within the wind data.

Numerous wind speed and power estimation models exist in the literature. In our research, we evaluated the performance of the proposed hybrid EMD-ESN model against commonly used models. These models investigated include Multi-Layer Perceptron (MLP), Adaptive-Neuro Fuzzy Inference System (ANFIS), Long Short-Term Memory (LSTM), Bidirectional Long Short-Term Memory (BiLSTM), Empirical Mode Decomposition based Long Short-Term Memory (EMD-LSTM), Empirical Mode Decomposition based Bidirectional Long Short-Term Memory (EMD-BiLSTM), Variational Mode Decomposition based Echo State Network (VMD-ESN), and Wavelet Decomposition based Echo State Network (WD-ESN). In this comparison study, we added the best result values of the original ESN model to the tables to analyze the results better. The MLP model is a traditional model comprising of three layers with nodes arranged in the format 4-8-1. The ANFIS model is a model with four inputs, one output, and five layers with each input consisting of two Gaussian membership functions. Two layers, each with 100 units, were utilized for both the LSTM and BiLSTM models. The models were optimized using the Adam optimizer, with a minimum batch size of 16 and a maximum of 50 training epochs. The dropout value was 0.5. In the model structure utilizing EMD decomposition methodology (EMD-LSTM and EMD-BiLSTM), an individual LSTM/BiLSTM model is employed for every decomposed signal. For the two different decomposition methods (VMD and WD) integrated into the ESN model, the model framework proposed in this study is preferred, i.e., instead of using one ESN prediction model for each decomposed signal, a single ESN model is used for all signals and their historical values. Thus, the EMD, VMD and WD decomposition methods are fairly compared within the same ESN model framework without emphasizing the advantage of the proposed model.

In Table 7, we present the comparison results of the proposed EMD-ESN model alongside other models for the Barrow dataset. For West of Duddon Sands dataset, the EMD-ESN model and the other forecasting models were compared and the obtained results are summarized in Table 8. Comparative outcomes for the final dataset, Horns Power, are presented in Table 9. All tables contain four metric values for both the training and testing phases of the models. These are MSE, RMSE, MAE, and R^2. The best metric values obtained between the models are indicated in bold in the tables.

Examining the table results for the Barrow dataset, it is evident that the EMD-BiLSTM model was the best for training, and the EMD-ESN model was the best for testing. Our model ranked second for the training phase, just after the EMD-BiLSTM model. In both phases, the EMD-BiLSTM model and the proposed EMD-ESN model displayed very similar results for this specific dataset. The findings from the West of Duddon Sands dataset indicate that the EMD-ESN model performs the best during both training and testing stages. The VMD-ESN and EMD-BiLSTM models, which are based on the primary decomposition method, ranked second and third, respectively, for this particular dataset. As anticipated, the primary decomposition methods, such as EMD, VMD, reveal significant components within the wind dataset and thus improving the performance of the models. However, in the comparison of EMD-BiLSTM and the proposed EMD-ESN models, the fact that our proposed model and architecture comes to the fore reveals the advantages of the method used.

Nevertheless, from the comparison of VMD-ESN and the proposed model using the same model framework, it is clear that the EMD decomposition method improves the results by 21.55% in training and 24.41% in testing in terms of the MSE metric compared to the VMD decomposition method. Based on the model training and testing outcomes for the Horns Power dataset, it is evident that the EMD-ESN model is the best performing model in both the training and testing phases. Here, in the training results, the EMD-BiLSTM model achieved the closest results to the metric results of the proposed EMD-ESN model. The main difference between the two models is their framework: our proposed hybrid framework uses an ESN model for all decomposed signals and their past values. The test results indicate that the metric results of the VMD-ESN model and the proposed EMD-ESN model are quite close.

Fig. 6 shows the R^2 results of the EMD-ESN and the other models for training and test stages. As shown in the sub-figures, the suggested hybrid EMD-ESN model is the most suitable model for both training and testing, except for the training outcomes of the Barrow dataset.

Among the three datasets, the EMD-ESN model achieved the best R^2 value for the West of Duddon Sands dataset (98.09% for training and 97.83% for testing). In the Barrow data set, the R^2 value of the model was found to be 97.82% for training and 97.58% for testing. The lowest performance of the model (R^2 metric) was achieved on the Horns Power dataset (95.19% for training and 90.70% for testing). This is because the data sets used show different characteristics.

3.2. Experiment 2: Comparison of EMD-ESN and SOTA models for onshore wind farm

In this section, the proposed EMD-ESN model is compared with the most recent SOTA models in the literature for a German onshore wind turbine power data. These SOTA models are LSTM, Gate Recurrent Unit (GRU) [41], Temporal Convolutional Network (TCN) [42], Long-term Time Series Forecasting (LTSF)-Linear (DLinear) [43], Long- and Short-term Time-series network (LSTNet) [44], Transformer [45], Informer [46], Autoformer [47], and Graph Patch Informer (GPI) [48].

The dataset used for comparison consists of 15 min of power data from a terrestrial wind farm in Germany for the period 01.01.2011 to 30.12.2021 [49]. From this dataset, only the wind power data from 1 January 2020 to 19 June 2020 were taken and the proposed EMD-ESN model results were found for different horizon values (3-6-12 steps). 70% of the data set is reserved for training and the rest for testing. Twice the prediction horizon values constitute the look-back window size, in short, the number of inputs of the proposed model. For instance, a prediction horizon of 3 means that the model will predict 3 steps ahead. The model has 6 inputs and a look-back window size of 6.

Table 7
Comparison results of the proposed EMD-ESN and other models for Barrow dataset.

Metrics	LSTM		BiLSTM		MLP		ANFIS		ESN	
	Training	Test	Training	Test	Training	Test	Training	Test	Training	Test
MSE	85.584	77.555	79.715	69.033	75.323	66.689	83.991	76.158	72.693	64.177
RMSE	9.2512	8.8065	8.9284	8.3086	8.6789	8.1663	9.1647	8.7269	8.5260	8.0111
MAE	5.8682	5.6011	6.3167	5.8006	5.8548	5.2974	6.3036	5.8011	5.5998	5.1728
R2	0.8996	0.8797	0.9065	0.8929	0.9117	0.8966	0.9015	0.8819	0.9149	0.8995
	EMD-LSTM		EMD-BiLSTM		VMD-ESN		WD-ESN		EMD-ESN	
	Training	Test	Training	Test	Training	Test	Training	Test	Training	Test
MSE	25.422	22.108	**17.634**	15.658	26.797	23.950	33.340	56.206	18.633	**15.458**
RMSE	5.0420	4.7019	**4.1993**	3.9571	5.1766	4.8939	5.774	7.4971	4.3165	**3.9316**
MAE	3.7225	3.3991	**3.0517**	2.9132	3.8822	3.6506	3.6223	5.2118	3.0380	**2.8334**
R2	0.9702	0.9657	**0.9793**	0.9757	0.9686	0.9625	0.9610	0.9120	0.9782	**0.9758**

Table 8
Comparison results of the proposed EMD-ESN and other models for West of Duddon Sands dataset.

Metrics	LSTM		BiLSTM		MLP		ANFIS		ESN	
	Training	Test	Training	Test	Training	Test	Training	Test	Training	Test
MSE	384.23	381.38	362.21	349.84	361.25	345.47	426.13	413.55	340.30	333.57
RMSE	19.602	19.529	19.032	18.704	19.007	18.587	20.643	20.336	18.447	18.264
MAE	13.951	14.207	11.962	11.145	12.360	11.540	13.301	12.464	11.513	11.259
R2	0.9277	0.9156	0.9319	0.9226	0.9321	0.9235	0.9199	0.9085	0.9360	0.9272
	EMD-LSTM		EMD-BiLSTM		VMD-ESN		WD-ESN		EMD-ESN	
	Training	Test	Training	Test	Training	Test	Training	Test	Training	Test
MSE	162.85	169.59	144.97	136.22	129.26	131.75	169.58	756.59	**101.40**	**99.584**
RMSE	12.761	13.023	12.040	11.671	11.369	11.478	13.022	27.506	**10.070**	**9.9792**
MAE	9.5021	9.9876	9.1340	8.8001	8.5003	8.3069	8.1357	25.160	**7.0519**	**7.0126**
R2	0.9694	0.9625	0.9727	0.9699	0.9757	0.9712	0.9681	0.8348	**0.9809**	**0.9783**

Table 9
Comparison results of the proposed EMD-ESN and other models for Horns Power dataset.

Metrics	LSTM		BiLSTM		MLP		ANFIS		ESN	
	Training	Test	Training	Test	Training	Test	Training	Test	Training	Test
MSE	416.42	541.45	430.73	568.02	423.16	553.63	435.64	550.86	384.95	577.70
RMSE	20.406	23.269	20.754	23.833	20.571	23.529	20.872	23.470	19.620	24.035
MAE	14.221	15.497	12.405	13.472	13.979	15.356	14.067	15.007	12.483	15.032
R2	0.8403	0.7940	0.8348	0.7839	0.8377	0.7894	0.8329	0.7904	0.8514	0.7759
	EMD-LSTM		EMD-BiLSTM		VMD-ESN		WD-ESN		EMD-ESN	
	Training	Test	Training	Test	Training	Test	Training	Test	Training	Test
MSE	168.76	785.19	125.86	770.13	167.45	240.14	190.50	459.27	**124.73**	**239.80**
RMSE	12.991	28.021	11.219	27.751	12.940	15.496	13.802	21.431	**11.168**	**15.486**
MAE	9.0148	15.910	7.9517	15.016	8.7534	10.807	8.6392	16.108	**7.5774**	**10.614**
R2	0.9353	0.7013	0.9517	0.7070	0.9354	0.9068	0.9265	0.8218	**0.9519**	**0.9070**

Fig. 7 shows the EMD-ESM model's wind power forecasting test results for 3, 6, and 12 horizon steps. As expected, the best forecasting curve is obtained for the lowest horizon value of 3, and it is seen that the prediction curves of the model deteriorate as the prediction horizon value increases. Table 10 summarizes the results of the proposed EMD-ESN and SOTA models for three different horizon values. Here, the wind power prediction results of SOTA models are taken from the study of Liu and Fu [48].

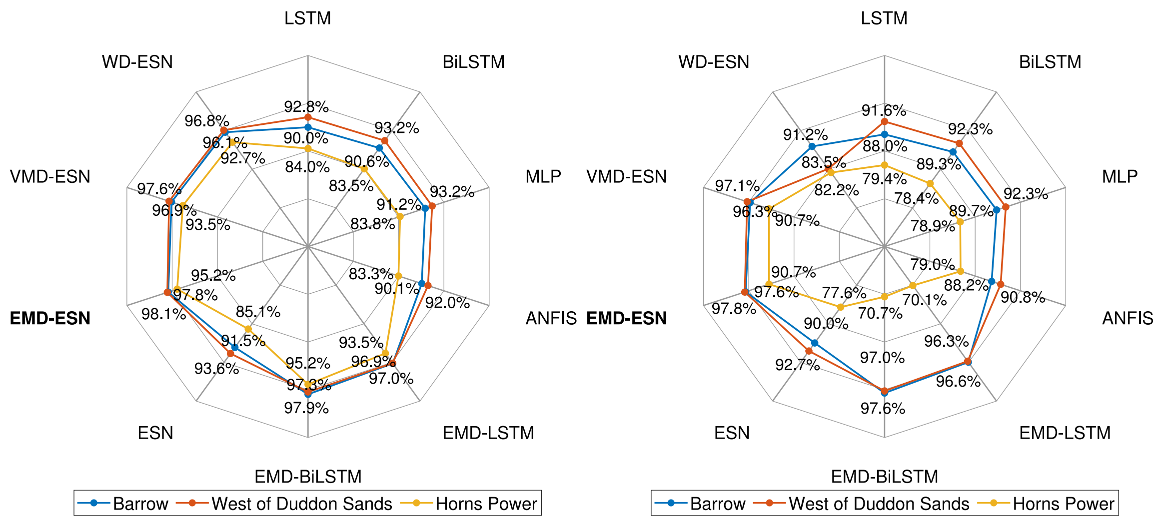
The table consists of two error metrics (MSE and MAE) for each model used in the wind power forecasting 3, 6 and 12 steps ahead. The best model values are shown in bold in the table. Based on the MSE and MAE values, the EMD-ESN hybrid model proposed in this study gives the best results for all horizon values. The main reasons for the success of the model are the EMD decomposition method and the ESN architecture proposed in this study.

4. Conclusion

In this study, a hybrid forecasting model for short-term wind power prediction is proposed by integrating the echo-state network (ESN) architecture with empirical mode decomposition (EMD). The approach diverges from conventional methods that use separate models for individual decomposed signals after EMD. The model utilizes a unified ESN structure to forecast wind power by considering all decomposed signals and their historical values. This innovative architecture simplifies the forecasting process by eliminating the necessity of training multiple models.

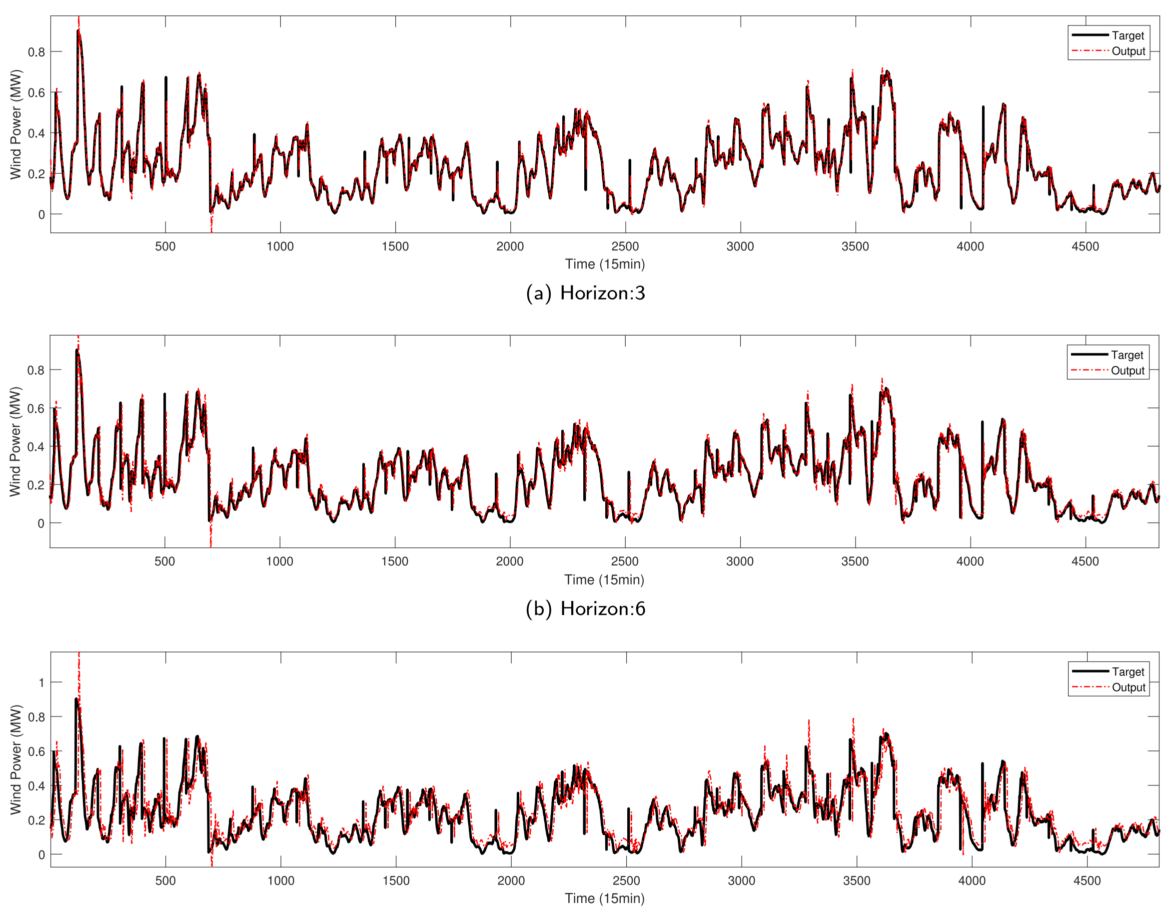
To evaluate the efficacy of our newly proposed model, we conducted widespread experiments using one year's worth of data obtained from offshore wind farms situated in the West of Duddon Sands, Barrow, and Horns Power. We initially compared the classical ESN model with the EMD-ESN hybrid model across these three data sets. Subsequently, we conducted a thorough analysis by comparing our model against commonly used models in the literature, such as Multi-Layer Perceptron (MLP), Adaptive-Neuro Fuzzy Inference System (ANFIS), Long Short-Term Memory (LSTM), Bidirectional Long Short-Term Memory (BiLSTM), Empirical Mode Decomposition based Long Short-Term Memory (EMD-LSTM), and Empirical Mode Decomposition based Bidirectional Long Short-Term Memory (EMD-BiLSTM).

When the results of the comparison of the proposed hybrid EMD-ESN model with independent single models (LSTM, BiLSTM, MLP, ANFIS, and ESN) are examined, the MSE metric value decreased by



(a) Training (b) Test

Fig. 6. R^2 results of EMD-ESN and other models for offshore datasets.



(a) Horizon:3

(b) Horizon:6

(c) Horizon:12

Fig. 7. EMD-ESN model's test results for 3-6-12-step wind power forecasting.

76.55% in the training phase and 78.14% in the testing phase for the Barrow dataset. Similarly, the West of Duddon Sands dataset showed an improvement in the MSE metric of 72.95% in the training phase and 72.70% in the testing phase. For the Horns Power dataset, the proposed model reduced the MSE metric value by 70.17% during training and 57.05% during testing. When comparing the four different hybrid models used in this study (EMD-LSTM, EMD-BiLSTM, VMD-ESN, and WD-ESN) with the proposed EMD-ESN model, it was found that the proposed model for the Barrow dataset reduced the MSE metric value by 27.77% for training and 47.57% for testing. Similarly, for the West of Duddon Sands dataset, the average reduction in the MSE metric was 33.14% for training and 66.64% for testing. The MSE metrics for the

Table 10
Comparison results of the proposed EMD-ESN and SOTA models for onshore wind power dataset.

Model	Horizon	3	6	12
LSTM	MSE	0.020	0.031	0.059
	MAE	0.074	0.102	0.159
GRU	MSE	0.020	0.031	0.059
	MAE	0.075	0.100	0.155
TCN	MSE	0.018	0.029	0.054
	MAE	0.069	0.095	0.144
DLinear	MSE	0.041	0.038	0.073
	MAE	0.131	0.124	0.184
LSTNet	MSE	0.018	0.030	0.058
	MAE	0.068	0.098	0.152
Transformer	MSE	0.015	0.024	0.059
	MAE	0.061	0.092	0.165
Informer	MSE	0.017	0.027	0.057
	MAE	0.066	0.099	0.157
Autoformer	MSE	0.024	0.031	0.076
	MAE	0.096	0.102	0.181
GPI	MSE	0.013	0.020	0.042
	MAE	0.052	0.080	0.133
EMD-ESN	MSE	**0.00188**	**0.00383**	**0.00868**
	MAE	**0.02182**	**0.03711**	**0.06504**

Horns Power dataset decreased by 23.55% during the training phase and 57.46% during the test phase. The MSE metrics demonstrate that the proposed EMD-ESN hybrid model outperforms the single models, as expected. This is the contribution of the primary decomposition (EMD) method. The superiority of the proposed model over the LSTM and BiLSTM hybrid models using the EMD decomposition method is due to the ESN model architecture proposed here, which utilizes all IMF signals. The success achieved in countering the ESN hybrid model using various decomposition methods, such as VMD and WD, can be attributed to the EMD decomposition method.

In addition to the experimental studies with the offshore wind dataset, the proposed hybrid model is tested for a publicly available onshore wind dataset with the latest popular prediction SOTA models such as Transformer, Informer, Autoformer and Graph Patch Informer. The proposed EMD-ESN hybrid model is the most successful among other SOTA models in wind power forecasting for different forecast horizons. Our research conclusively verifies the superior forecasting efficiency of the EMD-ESN hybrid model in both onshore and offshore datasets, surpassing alternative models. This study emphasizes the predictive capabilities of our proposed model, positioning it as a valuable tool for improving the accuracy of wind power predictions. Such advancements are crucial for the smooth integration of wind energy into the power grid, enabling a reliable and sustainable energy supply.

This study has limitations, but we are confident that our proposed EMD-ESN model architecture is effective. To further verify its general applicability, we recommend using more diverse data sets. It is important to note that while our model may not be the most appropriate for certain situations, we are confident in its overall performance. Further research is required to determine the effectiveness of alternative model architectures or approaches in varying weather conditions, geographical regions or scales. It is important to note that the data sets and conditions used in our study may not fully reflect real-world applications. Therefore, additional experimental studies and field tests are necessary to verify the applicability of the proposed model in the field. To achieve more comprehensive and reliable results in future studies, it is crucial to be aware of these limitations.

Further studies are necessary to investigate various model approaches and decomposition methods to enhance the accuracy and efficiency of wind power forecasting. The effectiveness of a novel model architecture combining EMD and ESN was demonstrated in this study. But, it is vital to explore the potential of other model combinations. For instance, new and diverse deep learning structures could be investigated to improve the accuracy of predictions. Furthermore, it would be worthwhile to investigate additional methods of decomposition and explore their possible combinations. Numerous feature engineering techniques could also be assessed in order to enhance data quality and optimize predictions. It may also prove beneficial to conduct comparative analyses of wind energy and power data from various geographical regions and to examine studies that take into account region-specific variations in model performance for future research. Further investigation of these methods and approaches has the potential to enhance the future reliability and integration of wind energy forecasting.

Funding

None.

CRediT authorship contribution statement

Ugur Yuzgec: Writing – review & editing, Writing – original draft, Software, Methodology, Conceptualization. **Emrah Dokur:** Writing – review & editing, Validation, Methodology. **Mehmet Balci:** Software, Data curation.

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.

Data availability

Data will be made available on request.

References

[1] Sun S, Du Z, Jin K, Li H, Wang S. Spatiotemporal wind power forecasting approach based on multi-factor extraction method and an indirect strategy. Appl Energy 2023;350:121749.

[2] Dokur E, Erdogan N, Salari ME, Karakuzu C, Murphy J. Offshore wind speed short-term forecasting based on a hybrid method: Swarm decomposition and meta-extreme learning machine. Energy 2022;248:123595.

[3] Deveci M, Cali U, Kucuksari S, Erdogan N. Interval type-2 fuzzy sets based multi-criteria decision-making model for offshore wind farm development in Ireland. Energy 2020;198:117317.

[4] Maradin D. Advantages and disadvantages of renewable energy sources utilization. Int J Energy Econ Policy 2021.

[5] Bórawski P, Bełdycka-Bórawska A, et al. Development of wind energy market in the European union. Renew Energy 2020;161:691–700.

[6] Li J, Wang G, Li Z, Yang S, Chong WT, Xiang X. A review on development of offshore wind energy conversion system. Int J Energy Res 2020;44(12):9283–97.

[7] Archer C, Simão H, Kempton W, Powell WB, Dvorak M. The challenge of integrating offshore wind power in the US electric grid. Part I: Wind forecast error. Renew Energy 2017;103:346–60.

[8] Deveci M, Özcan E, John R, Pamucar D, Karaman H. Offshore wind farm site selection using interval rough numbers based best-worst method and MARCOS. Appl Soft Comput 2021;109:107532.

[9] Lins DR, Guedes KS, Pitombeira-Neto AR, Rocha PAC, de Andrade CF. Comparison of the performance of different wind speed distribution models applied to onshore and offshore wind speed data in the northeast Brazil. Energy 2023;278:127787.

[10] Maldonado-Correa J, Solano J, Rojas-Moncayo M. Wind power forecasting: A systematic literature review. Wind Eng 2021;45(2):413–26.

[11] Magnusson M, Smedman A-S. Air flow behind wind turbines. J Wind Eng Ind Aerodyn 1999;80(1–2):169–89.

[12] Song M, Chen K, He Z, Zhang X. Wake flow model of wind turbine using particle simulation. Renew Energy 2012;41:185–90.

[13] Andrade JR, Bessa RJ. Improving renewable energy forecasting with a grid of numerical weather predictions. IEEE Trans Sustain Energy 2017;8(4):1571–80.

[14] Wadi M, Elmasry W. Statistical analysis of wind energy potential using different estimation methods for Weibull parameters: a case study. Electr Eng 2021;103(6):2573–94.
[15] Kisvari A, Lin Z, Liu X. Wind power forecasting–A data-driven method along with gated recurrent neural network. Renew Energy 2021;163:1895–909.
[16] Huang B, Liang Y, Qiu X. Wind power forecasting using attention-based recurrent neural networks: a comparative study. IEEE Access 2021;9:40432–44.
[17] Yildiz C, Acikgoz H, Korkmaz D, Budak U. An improved residual-based convolutional neural network for very short-term wind power forecasting. Energy Convers Manage 2021;228:113731.
[18] Putz D, Gumhalter M, Auer H. A novel approach to multi-horizon wind power forecasting based on deep neural architecture. Renew Energy 2021;178:494–505.
[19] Qian Z, Pei Y, Zareipour H, Chen N. A review and discussion of decomposition-based hybrid models for wind energy forecasting applications. Appl Energy 2019;235:939–53.
[20] Liu H, Chen C, Tian H-q, Li Y-f. A hybrid model for wind speed prediction using empirical mode decomposition and artificial neural networks. Renew. Energy 2012;48:545–56.
[21] Bokde N, Feijóo A, et al. The hybridization of ensemble empirical mode decomposition with forecasting models: Application of short-term wind speed and power modeling. Energies 2020;13(7):1666.
[22] Duan J, Wang P, et al. Short-term wind power forecasting using the hybrid model of improved variational mode decomposition and correntropy long short-term memory neural network. Energy 2021;214:118980.
[23] Zhang Y, Chen Y. Application of hybrid model based on CEEMDAN, SVD, PSO to wind energy prediction. Environ Sci Pollut Res 2022;1–14.
[24] Kim T, King BR. Time series prediction using deep echo state networks. Neural Comput Appl 2020;32:17769–87.
[25] Chatzis SP, Demiris Y. The copula echo state network. Pattern Recognit 2012;45(1):570–7.
[26] Niederhauser L, Prabhakar A, Reber D, Billard A. A predictive model for tactile force estimation using audio-tactile data. IEEE Robot Autom Lett 2023.
[27] Long J, Zhang S, Li C. Evolving deep echo state networks for intelligent fault diagnosis. IEEE Trans Ind Inf 2019;16(7):4928–37.
[28] Daneshfar F, Kabudian SJ. Speech emotion recognition system by quaternion nonlinear echo state network. 2021, arXiv preprint arXiv:2111.07234.
[29] Chitsazan MA, Fadali MS, Trzynadlowski AM. Wind speed and wind direction forecasting using echo state network with nonlinear functions. Renew Energy 2019;131:879–89.
[30] Tian Z. Approach for short-term wind power prediction via kernel principal component analysis and echo state network optimized by improved particle swarm optimization algorithm. Trans Inst Meas Control 2021;43(16):3647–62.
[31] Wang H, Lei Z, Liu Y, Peng J, Liu J. Echo state network based ensemble approach for wind power forecasting. Energy Convers Manage 2019;201:112188.
[32] Hu H, Wang L, Zhang D, Ling L. Rolling decomposition method in fusion with echo state network for wind speed forecasting. Renew Energy 2023;216:119101.
[33] Huang NE, et al. The empirical mode decomposition and the Hilbert spectrum for nonlinear and non-stationary time series analysis. Proc R Soc Lond. Ser A: Math, Phys Eng Sci 1998;454(1971):903–95.
[34] Bokde N, Feijóo A, Villanueva D, Kulat K. A review on hybrid empirical mode decomposition models for wind speed and wind power prediction. Energies 2019;12(2):254.
[35] Fan C, Ding C, Zheng J, Xiao L, Ai Z. Empirical mode decomposition based multi-objective deep belief network for short-term power load forecasting. Neurocomputing 2020;388:110–23.
[36] Hu H, Wang L, Lv S-X. Forecasting energy consumption and wind power generation using deep echo state network. Renew Energy 2020;154:598–613.
[37] Jamshidi MB, Daneshfar F. A hybrid echo state network for hypercomplex pattern recognition, classification, and big data analysis. In: 2022 12th international conference on computer and knowledge engineering. ICCKE, IEEE; 2022, p. 007–12.
[38] Mustaqeem K, El Saddik A, Alotaibi FS, Pham NT. AAD-Net: Advanced end-to-end signal processing system for human emotion detection & recognition using attention-based deep echo state network. Knowl-Based Syst 2023;270:110525.
[39] Jaeger H. The ''echo state'' approach to analysing and training recurrent neural networks-with an erratum note. Bonn, Ger: Ger Natl Res Cent Inf Technol GMD Tech Rep 2001;148(34):13.
[40] Li G, Li B-J, Yu X-G, Cheng C-T. Echo state network with Bayesian regularization for forecasting short-term power production of small hydropower plants. Energies 2015;8(10):12228–41.
[41] Gao Y, Glowacka D. Deep gate recurrent neural network. In: Asian conference on machine learning. PMLR; 2016, p. 350–65.
[42] Bai S, Kolter JZ, Koltun V. An empirical evaluation of generic convolutional and recurrent networks for sequence modeling. 2018, arXiv preprint arXiv:1803.01271.
[43] Zeng A, Chen M, Zhang L, Xu Q. Are transformers effective for time series forecasting? In: Proceedings of the AAAI conference on artificial intelligence. volume 37, 2023, p. 11121–8.
[44] Lai G, Chang W-C, Yang Y, Liu H. Modeling long-and short-term temporal patterns with deep neural networks. In: The 41st international ACM SIGIR conference on research & development in information retrieval. 2018, p. 95–104.
[45] Vaswani A, Shazeer N, Parmar N, Uszkoreit J, Jones L, Gomez AN, Kaiser Ł, Polosukhin I. Attention is all you need. Adv Neural Inf Process Syst 2017;30.
[46] Zhou H, Zhang S, Peng J, Zhang S, Li J, Xiong H, Zhang W. Informer: Beyond efficient transformer for long sequence time-series forecasting. In: Proceedings of the AAAI conference on artificial intelligence. volume 35, 2021, p. 11106–15.
[47] Wu H, Xu J, Wang J, Long M. Autoformer: Decomposition transformers with auto-correlation for long-term series forecasting. Adv Neural Inf Process Syst 2021;34:22419–30.
[48] Liu J, Fu Y. Renewable energy forecasting: A self-supervised learning-based transformer variant. Energy 2023;284:128730.
[49] Dataset P. Wind power. Kaggle, No date. URL: https://www.kaggle.com/datasets/l3llff/wind-power.