Energy 310 (2024) 133208

Contents lists available at ScienceDirect

Energy

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

Enhanced offshore wind resource assessment using hybrid data fusion and numerical models

Basem Elshafei *, Atanas Popov, Donald Giddings

Faculty of Engineering, University of Nottingham, NG7 2RD, Nottingham, UK

ARTICLE INFO

Keywords:
Gaussian process regression
Temporal data fusion
Wind resource assessment
Data pre-processing

ABSTRACT

Wind resource assessments are crucial for pre-construction planning of wind farms, especially offshore. This study proposes a novel hybrid model integrating Complete Ensemble Empirical Mode Decomposition with Adaptive Noise (CEEMDAN) and Empirical Wavelet Transform (EWT) for enhanced wind speed forecasting. This secondary decomposition reduces forecasting complexity by processing high-frequency signals. A Bidirectional Long Short-Term Memory (BiLSTM) neural network optimized with the Grey Wolf Optimizer (GWO) is then employed for forecasting. The model's accuracy is evaluated using simulated wind speeds along the coast of Denmark, combined with lidar measurements through data fusion. This approach demonstrates significant improvements in prediction accuracy, highlighting its potential for offshore wind resource assessment.

1. Introduction

With rapid economic development and an increase in standards of living, human demand for energy has been significantly increasing. Additionally, the use of fossil fuels such as coal, oil, and natural gas has resulted in huge amounts of greenhouse gases in the atmosphere, damaging the environment and raising the global temperature, leading to global warming. Alternately, investments in renewable energy have increased significantly and have received more attention lately. Wind energy has the second largest potential after solar energy and has received extensive global attention. Wind capacity in 2023 is 906 GW with 136 GW per annum of new installations predicted according to the Global Wind Report [1].

Wind energy generation is an important component of the smart grid and plays a significant role in the supply and management of electricity, but its stochastic nature has a great influence on the stability and safety of grid-integrated wind power systems [2]. To solve the resulting scheduling, management and optimization challenges, accurate forecasting of future wind availability is of paramount importance [3]. Forecasting wind energy contributes to solving other issues in the wind energy sector such as reducing operating costs and enhancing competitiveness. Therefore, wind forecasting is a key technology in integrating wind energy into existing multiscale electricity grids. There are two overall methods of wind forecasting: wind power can be directly predicted from historic wind power data (which is not usually available when the forecasts are used to assess the wind resources of a new location), or wind speed can be predicted and converted to wind power forecasts based on wind power curves. Here, the focus is on wind speed predictions to provide wind resource assessments for potential wind farm locations based on available wind speed data [4].

The past decade has seen the development of numerous approaches to forecasting wind speed. Forecasting time horizons range from very short-term predictions (a few seconds to 30 min ahead) to long-term predictions (1 day to 1 week, or more ahead). Forecasting models include physical models, statistical models, machine learning-based models, and hybrid models. Numerical weather predictions, such as Weather Research Forecasting (WRF), use a physical approach that takes into consideration meteorological factors such as air pressure, humidity, and temperature. Physical approaches have the most accurate performance when forecasting wind speeds in the long-term time scale. Statistical models, on the other hand, include Auto-Regressive Moving Average (ARMA) and its variant models, which focus on characterizing time series data. Statistical models have superior performance in the very short-term time range [5].

Latterly, machine learning employing AI-based models has emerged in wind speed forecasting [6]. Different types of neural networks such as backpropagation [7], long-short-term [8], and convolutional [9] neural networks have demonstrated superior capability in capturing and dealing with non-linearity in datasets. Several studies showed that machine learning-based models performed better than statistical models [10]. Following the success of machine learning models in wind speed forecasting, many studies suggested combining different algorithms to characterize different aspects of wind speed fluctuation.

* Corresponding author.
E-mail address: basem.elshafei3@nottingham.ac.uk (B. Elshafei).

https://doi.org/10.1016/j.energy.2024.133208
Received 5 March 2024; Received in revised form 11 July 2024; Accepted 16 September 2024
Available online 25 September 2024

For example, Convolutional Neural Network (CNN) is used to capture spatial features of a wind speed dataset, followed by a Long Short Term Memory (LSTM) neural network to generate temporal features. Together, a combination of both neural networks produces the hybrid forecasting model, CNN-LSTM, which captures spatio-temporal characteristics, making full use of space–time information in the data set [11].

Wind resource assessment is critical for the development and optimization of offshore wind farms. Traditional assessment methods often face challenges due to the non-stationary and non-linear nature of wind speed time series. These methods typically rely on single decomposition techniques, which are inadequate for accurately capturing high-frequency signal components. Furthermore, existing models frequently overlook the integration of high-fidelity lidar measurements, limiting their predictive capabilities.

The performance of forecasting approaches is significantly enhanced by pre-processing the data used, and then setting the model configuration using intelligent optimization algorithms that provide optimal parameters for the predictor model [12]. Wind data preprocessing considers wind speed time series as a signal that can be decomposed into subseries of different frequencies that are projected into forecasting models and then aggregated to produce a forecast from the existing numerical time series [13]. The de-noised signal improves the suitability of the signal for training forecasting models. Wavelet-based approaches can be used to decompose the original wind speed series into several frequency ranges. In contrast, other approaches can only produce all subseries at one frequency level. Wavelet-based approaches include algorithms such as Wavelet Decomposition (WD), Empirical Wavelet Transform (EWT), and Empirical Mode Decomposition (EMD). In addition to employing a primary signal decomposition algorithm, hybrid approaches were proposed to combine the merits of different decomposition algorithms [12]. A secondary-decomposition approach considers more than one signal decomposition algorithm and improves wind speed forecasting. After preprocessing, intelligent optimization algorithms are employed to provide optimal hyperparameters for the forecasting model. Model configuration hyper-parameters include, but are not limited to, weights, batch size, learning rate, and number of hidden layers. Examples of optimization algorithms include the Grey Wolf optimizer, the Multi-Objective Bat Algorithm, and Multi-Objective Multi-Universe Optimization [14–16].

In this study, we concentrate on the development of assessment for wind resources, which involves a detailed evaluation of the potential wind resources in a specific geographic area. This assessment mainly revolves around accurately predicting wind speeds at various locations and heights to determine the optimal placement of wind turbines. Our approach involves collecting meteorological data from multiple sources, including wind speed and direction, and analysing historical wind patterns and trends. Factors such as diurnal variations are considered in our assessment. Wind resource assessment plays an essential role in site selection and turbine positioning, which are vital for optimizing energy generation.

This study addresses some of the aforementioned limitations by proposing a hybrid model that combines secondary decomposition techniques with data fusion. Specifically, the integration of Complete Ensemble Empirical Mode Decomposition with Adaptive Noise (CEEMDAN) and Empirical Wavelet Transform (EWT) effectively handles high-frequency signals, reducing the complexity of wind speed forecasting. The decomposed modes are predicted using a Bidirectional Long Short-Term Memory (BiLSTM) neural network optimized by the Grey Wolf Optimizer (GWO). This novel approach not only enhances the accuracy of wind speed predictions but also leverages the strengths of both numerical simulations and lidar measurements through data fusion.

The innovative contributions of this study include: 1. The application of a secondary decomposition model integrating CEEMDAN and EWT to improve the handling of high-frequency wind speed components. 2. The use of a BiLSTM neural network optimized with GWO to accurately predict decomposed modes. 3. The fusion of lidar measurements with numerical simulations to enhance overall prediction accuracy.

This paper demonstrates the potential of the proposed hybrid model in improving offshore wind resource assessments, providing a more reliable foundation for pre-construction planning and optimization.

The innovative aspect of this study lies in the integration of CEEMDAN and EWT for secondary decomposition, which effectively handles high-frequency signal components that traditional single decomposition methods fail to address. This hybrid approach, coupled with the application of a BiLSTM neural network optimized by the Grey Wolf Optimizer (GWO), significantly enhances the accuracy of wind speed forecasts. Moreover, the fusion of lidar measurements with numerical simulations further refines the prediction accuracy, demonstrating a novel method for offshore wind resource assessments.

Current wind energy resource assessment methods often struggle with non-stationary and non-linear characteristics of wind speed signals. Traditional single decomposition techniques are inadequate for accurately forecasting high-frequency signals. Existing models also fail to fully leverage high-fidelity lidar measurements, limiting their predictive capabilities. This study addresses these limitations by proposing a hybrid model that combines secondary decomposition techniques with data fusion, providing a more robust and accurate assessment of wind resources.

1.1. Motivation and contribution of this work

The literature on hybrid models for predicting wind speed contains more than 250 articles testing different combinations of approaches, where the majority provide more accurate forecasts and enhanced performance relative to previous non-hybrid work. Fig. 1 shows the standard stages of building a wind speed forecasting process. However, the majority of studies conducted numerical experiments on a single source of data generated by numerical simulations for its wide availability and relatively low cost. Data generated using simulations is considered of relatively low fidelity and despite achieving high accuracy in predicting future trends of the numerical wind forecast, results are rarely compared to other sources of data such as lidar observations or mast measurements. Therefore, this work investigates a hybrid approach that employed Complete Ensemble Empirical Mode Decomposition With Adaptive Noise (CEEMDAN) and EWT algorithms for data processing, Bidirectional Long Short Term Memory (BiLSTM) neural networks for the prediction model, and Grey Wolf Optimizer (GWO) to optimize the weights of the BiLSTM networks using WRF simulation data. The prediction results are then compared to lidar observations to accurately assess the performance of the model when compared to a high-fidelity wind speed dataset. The main objective of the work is to demonstrate the importance and added value of data fusion between high-fidelity data generated using lidars and low-fidelity data from simulations for the assessment of wind resources.

The approach of this work is to optimize the hybrid combination of time series data prediction methods available in the literature based solely on numerical simulation data. This is compared to the method of data fusion, using both numerical simulation data and empirical data, and both results at unobserved periods are compared with the empirical data. The data processing methods are described in this section. In this study, we emphasize that our focus is on the assessment of wind resources, which involves the accurate prediction of wind speeds at various locations and heights within a geographic area. This information is crucial for the optimal positioning and placement of wind turbines, which directly impacts potential energy generation from the site. Although wind power generation is a separate and specialized field dedicated to harnessing the kinetic energy of the wind and converting

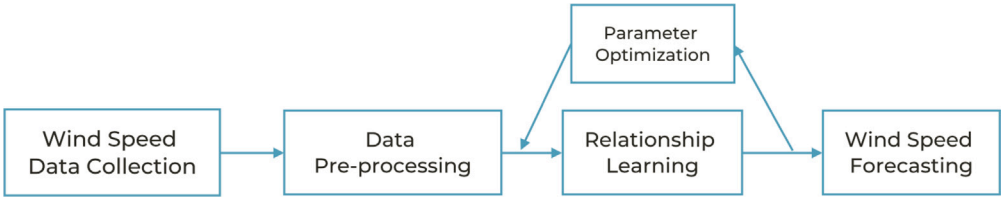

Fig. 1. Stages of building a forecasting model, which includes collection of data, pre-processing the signals, training a forecasting model with optimized parameters from optimization algorithm.

it into electricity, our work primarily concerns the assessment of wind resources.

The innovative aspect of this study lies in the integration of CEEMDAN and EWT for secondary decomposition, effectively handling high-frequency signal components. This hybrid approach, coupled with the application of a BiLSTM neural network optimized by the Grey Wolf Optimizer (GWO), significantly enhances the accuracy of wind speed forecasts. Moreover, the fusion of lidar measurements with numerical simulations further refines the prediction accuracy, demonstrating a novel method for offshore wind resource assessments.

This study begins with an introduction to the importance of wind resource assessment for the wind energy industry, followed by a review of the existing literature on the topic in Section 1. We then present the two forecasting models in Section 2; first, the proposed numerical simulations only secondary decomposition approach for wind speed forecasting, which combines CEEMDAN and EWT preprocessing approaches with the BiLSTM recurrent neural network and GWO optimization algorithm; second, the data fusion approach, which combines multi-fidelity Gaussian regression with NARX using both numerical simulations and measured observation data. In Section 3, we present the case description showing the generation of the data and a comparison of the metrics used to assess wind speed forecasting. Section 4 presents the results in detail showing a comparison of the numerical simulation model and the data fusion techniques, and discusses potential applications of the findings for the wind energy industry. Finally, Section 5 presents the conclusion of the study, demonstrating the challenges, and concludes with a summary of the key findings and suggestions for future research in this area.

2. Methodology

Wind speed time series exhibit non-stationary and non-linear characteristics, necessitating advanced preprocessing for accurate forecasting. This study introduces a secondary decomposition model integrating CEEMDAN and EWT, effectively handling high-frequency signal components. The decomposed modes are forecasted using a BiLSTM neural network optimized by the GWO algorithm, enhancing prediction accuracy and reducing complexity.

Section 2.1 presents the demonstration of the general implementation of the proposed hybrid model that uses only numerical data from WRF simulations. Fluctuations in the time series are reduced by CEEMDAN prepossessing and supported by EWT to further reduce the noise and fluctuations in the decomposed signal. The forecasting model is trained by a recurrent neural network BiLSTM, a special type of LSTM neural network that provides both past and future data. The parameters of the Artificial Neural Network (ANN) forecasting model are tuned using the Grey Wolf Optimizer. Section 2.2 presents an alternative data fusion model of wind speed data from both numerical and empirical data that is used to verify the hybrid model using only numerical data.

2.1. Hybrid technique for numerical data: CEEMDAN + EWT + BiLSTM + GWO

The hybrid model is based on hybrid CEEMDAN+EWT double signal decomposition and the BiLSTM-GWO forecasting model. The flowchart of the proposed forecasting model is shown in Fig. 2. In stage one, Complete Ensemble Empirical Mode Decomposition With Adaptive Noise (CEEMDAN) is used to decompose the original wind speed time series into a set of Intrinsic Modal Functions (IMF)s each with a different frequency. The first IMF generated, IMF1, is the most fluctuating and disorderly signal amongst all outputs and hence is not suitable for forecasting and would reduce the prediction accuracy. Therefore, another signal decomposition iteration aimed at reducing the fluctuations and noise in that signal is essential to reduce the forecast difficulty. Empirical Wavelet Transform (EWT) can decompose high-frequency signals such as IMF1 into more steady components. Hence, EWT is employed to further decompose the decomposed wind speed signal before feeding signals into the neural networks. In the following step, predictions on the decomposed signals are generated, including the modes generated by EWT and IMFs generated by CEEMDAN, excluding IMF1, using Bidirectional Long Short Term Memory (BiLSTM) optimized by Grey Wolf Optimizer (GWO). In the last stage, all the forecast results from the modes and IMFs were aggregated to obtain the final forecast time series.

The data fusion approach integrates lidar measurements, which provide high-fidelity wind speed data, with numerical simulations. This fusion leverages the strengths of both data sources, improving overall prediction accuracy. Lidar data, known for its precision, enhances the robustness of numerical simulations, which cover broader temporal and spatial scales. The fusion process involves aligning the datasets temporally and spatially, followed by applying the hybrid model to predict wind speeds more accurately.

2.1.1. Data pre-processing

Signal processing approaches address the entire dataset indiscriminately and hence are usually used for two purposes, data decomposition and data denoising. Combined pre-processing algorithms have shown high efficiency for wind speed forecasting. In most cases, the data time series is decomposed into different sub-series with various frequencies. Then, the higher frequency sub-series is further decomposed using a different algorithm to catch in-depth trends in the dataset. The predictability of the decomposed sub-series is found to be stronger than that of the original series as it contains less noise, hence will generate more accurate forecasts of each level of the original dataset [17]. Results from a study that combined Wavelet Packet Decomposition (WPD) with Fast Ensemble Empirical Mode Decomposition (FEEMD) has shown that forecasts from hybrid pre-process data is far more accurate than a single signal processing approach [18]. Therefore, further studies investigated combining different approaches and several studies have developed secondary decomposition algorithms with accurate results, some of the developed algorithms are summarized in Table 1. For the current work, CEEMDAN will generate different frequency intrinsic mode functions and the highest frequency level is later further decomposed using EWT to reduce forecasting complexity.

While our approach aims to be comparable to existing models, it introduces several novel elements. First, we employ a hybrid decomposition algorithm that combines CEEMDAN and EWT for preprocessing wind speed signals. This preprocessing step generates data that is less sensitive and more suitable for the prediction model. Furthermore, we integrate the Grey Wolf Optimizer, which dynamically updates our

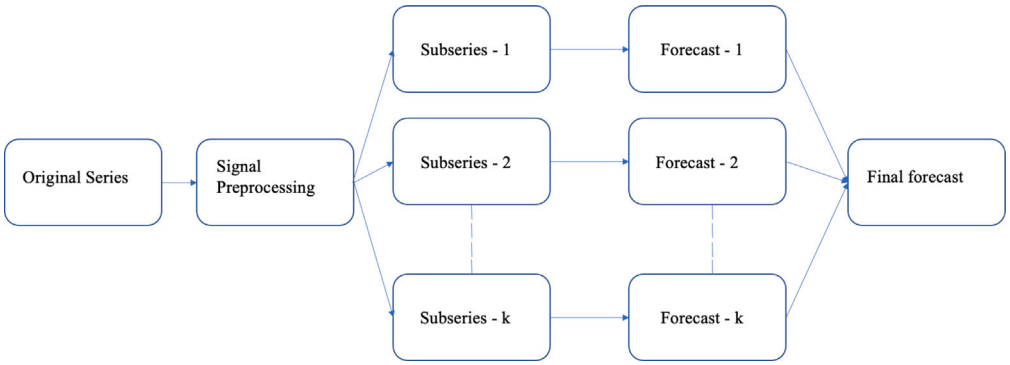

Fig. 2. Flowchart of the proposed forecasting model.

Table 1
Existing Secondary decomposition algorithms.

Article	Secondary decomposition method
Mi et al.	Wavelet Packet Decomposition (WPD) +EMD
Wu and Xiao	EWT + Singular Spectrum Analysis (SSA)
Moreno ET al.	Variational Mode Decomposition (VMD) + SSA
Peng et al.	CEEMDAN + VMD
Lie et al.	EEMD + WPD

prediction neural network at each iteration, enhancing the adaptability and performance of the model. These elements distinguish our method and contribute to the broader field of wind resource assessment.

Complete Ensemble Empirical Mode Decomposition with Adaptive Noise (CEEMDAN):

In [19], Huang et al. developed the empirical mode decomposition technique, an adaptive data analysis method that deals with both non-linearity and non-stationarity in signals. EMD aims to decompose complex signals into a finite number of Intrinsic Modal Functions (IMF) based on the signal's local characteristics. The IMFs comply with two conditions: (a) the entire data set contains several extremes and zero crossings that are either equal or differ by one unit, and (b) the mean value of the envelope defined by the local minima and local maxima is zero at any point of the defined overall envelope of the local sample. The EMD process of the original signal can be defined as follows:

$$n(t) = \sum_{k=1}^{m} imf_{m(t)} + r_{m(t)} \tag{1}$$

where n(t) represents the non-linearity or non-stationarity in the signal, $imf_m(t)$ is the mth IMF of the signal, and $r_m(t)$ is the residual.

To overcome the drawback of mode mixing in EMD, a noise-assisted analysis of the data, EEMD, was proposed by Wu and Huang [20]. In Ensemble Empirical Mode Decomposition (EEMD), the true IMF components are identified based on the mean of an ensemble of trials. The signal is reconstructed with white noise incorporated, providing a uniform reference scale, and the EMD process is applied. The process is repeated in several trials to produce artificially noisy versions of the source signal. The white noise is finally cancelled due to averaging of the ensemble. However, as each trial adds white noise to the decomposition result, the added white Gaussian noise signal after a finite number of iterations results in a reconstruction error that may not be eliminated, and the accuracy of the forecasts will be affected. It is possible to decrease the reconstruction error by increasing the number of iterations at a high computational cost. Contrarily, a complete EEMD with adaptive noise, CEEMDAN, was developed by Torres et al. [21] providing three main advantages: (a) a noise coefficient value to control the noise level at each decomposition, (b) complete and noise-free reconstruction of the signal, and (3) fewer trials are required compared to other approaches. The decomposition process of CEEMDAN is as follows:

Step 1: Some noise-added series is generated for the EEMD method:

$$n^i(t) = n(t) + p_0 w^i(t) \tag{2}$$

where n(t) denotes the original signal, $w^i(t)(t = 1, \ldots, I)$ denotes different white Gaussian noise with $N(0,1)$ and p_0 is a noise coefficient which controls the signal-to-noise ratio.

Step 2: Decompose each of the generated noise using EMD to get the corresponding first modes $IMF_1^i(t)$. Then calculate the first mode of CEEMDAN by averaging all modes:

$$\overline{IMF_1^i(t)} = \frac{1}{I}\sum_{i=1}^{I} IMF_1^i(t) \tag{3}$$

Step 3: Calculate residual $r_1(t) = n(t) - \overline{IMF_1^i(t)}$ and decompose the noise-added residual $r_1(t) + p_1 E_1(w^i(t))$ to obtain the second mode:

$$\overline{IMF_2^i(t)} = \frac{1}{I}\sum_{i=1}^{I} E_1(r_1(t) + p_1 E_1(w^i(t))) \tag{4}$$

where $E_1(\bullet)$ is a function to produce the first mode of EMD.

Step 4: The process is repeated to obtain the subsequent modes until the residual component does not have at least two extreme values.

Empirical Wavelet Transform (EWT):

EWT is a three-model decomposition algorithm used in forecasting. Several studies demonstrated that it could be used to achieve good forecasting results for non-stationary time series such as wind speed series [22]. EWT can extract meaningful information from the series by designing appropriate wavelet filter banks. In our work, preprocessing of the WRF time series generated by ERA-5 can adaptively represent the processed signal by generating the adaptive wavelet, and then decomposing the signal into a finite number of modes as per previous literature [22]. The algorithm is based on identifying and extracting the different intrinsic modes of a time series by relying on robust pre-processing for peak detection and then performing spectrum segmentation based on detecting maxima to construct a corresponding wavelet filter bank.

The process can be divided into five steps [22]:

1. Extending the signal.
2. Fourier transforms.
3. Extracting boundaries.
4. Building a filter bank.
5. Extracting the subband.

The original wind speed signal had considerable high-frequency fluctuations. The three-level decomposition attained by the EWT algorithm describes the wind speed series in a meaningful way. Three uncorrelated filter modes are extracted from the wind speed series and a residual is also obtained from the extraction.

2.1.2. Recurrent Neural Network (RNN)

Recurrent Neural Network (RNN), is a member of the neural network family, it solves the long-term dependence of time series using a circular internal outline feedback connection which enables the use of past information. It provides a reliable and robust solution for processing time series data with variable periods. RNNs have a major drawback, where the gradient tends to disappear for a long time series and the neural network starts suffering from short-term memory [23], see Fig. 4.

Long-short Term Memory (LSTM)

In 1998, Hochreiter et al. [24] proposed the concept of LSTMs to overcome the drawback in RNNs. As a special variant of RNNs, LSTM proposes a gating mechanism, which gives the traditional RNN the ability to store or forget temporal state information. A classic LSTM is composed of a cell, an input gate, an output gate, and a forget gate. The cell stores information over arbitrary time intervals, while the memory block in the network consists of the gates, which regulate the information flow into and out of the cell, as indicated in Fig. 4. The forget gate filters the information for retention from the past memory cell, the input gate determines the part of the information to be updated, and the output gate decides the information to be exported from the memory block. LSTM uses additional long-term information to enhance the accuracy of time series predictions, and it is widely used to obtain deterministic or probabilistic wind speed forecasts. For example in [25], LSTM generated deterministic forecasts by producing prediction intervals using a beta distribution tuned for the counterpart forecasting error, and in [26], an LSTM-based model to generated day-ahead hourly wind speed forecasts by using a multi-scale network that integrated information for each temporal scale. A detailed implementation of the corresponding gates is as follows:

Input gate $i_t = \sigma(w_i \cdot [x_t, h_{t-1} + b_i])$

Output Gate $o_t = \sigma(w_o \cdot [x_t, h_{t-1}] + b_o)$

Forget Gate $f_t = \sigma(w_f \cdot [x_t, h_{t-1}] + b_f)$

Cell state $c_t = c_{t-1} * f_t + (tanh(w_c \cdot [x_t, h_{t-1}] + b_c)) * i_{t-1}$

Hidden state $h_t = tanh(c_{t-1}) * o_t$

where c represents a cell state, $h(t-1)$ is the hidden state, W represents the weight matrices, b represents the biases, which are not time-dependent, additionally, the activation function for the three gates is the hyperbolic tangent function tanh(.), and the activation function of the state update is the logistic sigmoid function sigma(.), which are defined as follows:

$$\sigma(x) = \frac{1}{(1+e^{-x})} \tag{5}$$

$$tanh(x) = \frac{(e^{x} - e^{-x})}{(e^{x} + e^{-x})} \tag{6}$$

The above equations are all computed for a one-time step, implying that this set of equations is recomputed for every time step. Moreover, as weight and biases are time-independent, they remain constant in every iteration for the set of equations in each time step.

Bidirectional Long-short term memory (BiLSTM):

Bidirectional Long Short Term Memory (BiLSTM) processes sequential data in both forward and backward directions so that forecasting is not only dependent on past information but also future data in the time series. Two separate hidden layers, produce an aggregated output. The structure of a BiLSTM neural network is shown in Fig. 5. However, BiLSTM is relatively new and advanced and its competitiveness has not been comprehensively demonstrated in the wind speed predictions sector. In [28], the performance of an LSTM was compared to a BiLSTM, with different epochs and unit values, and the results showed that BiLSTM performs better due to the additional future information passed to the network. In [29], an ensemble of BiLSTM models as base predictors generated more accurate predictions than other ensembles of deep neural networks. In [30], BiLSTM was used to extract features from a time series, which were then fed into another BiLSTM to obtain wind speed predictions. The results showed 39% improvements compared with traditional models. Finally, in [31] different BiLSTM models were utilized to predict wind speed time series in different sub-clusters produced by k-means clustering, a machine-learning approach that groups similar data points into k-numbered clusters and makes predictions based on information from all points in the cluster.

2.1.3. Grey Wolf Optimizer (GWO):

GWO is an optimization approach that aids in deciding weights for the forecasting models as part of CEEMDAN. The algorithm is based on a swarm intelligence-based computation technique, where the leave-one-out strategy was developed to integrate the individual models [32].

GWO is a Meta-heuristic optimization algorithm, which simulates the hunting behaviour of grey wolves in nature. Four types of wolves, alpha (α), beta (β), delta (δ), and omega (ω) are hired to imitate leadership hierarchy. The nature followed is that omega wolves follow the optimization process which is guided by the other three wolves. Four main steps followed by grey wolves during hunting, namely, encircling prey, hunting, attacking prey, and searching for prey are key steps implemented during the process. The algorithm is categorized as follows:

Social hierarchy:

To mathematically simulate this process, the first three wolves are assigned to the first three best solutions as a, b, and c, respectively, while the remaining solutions – wolves – are assigned to as w. due to the nature of GWO, solutions a, b, and c are employed to guide the optimization process and w will follow.

Encircling prey:

The encircling behaviour of wolves can be expressed as:

$$\overrightarrow{D} = |\overrightarrow{C} \cdot \overrightarrow{X_p}(t) - \overrightarrow{X}(t)| \tag{7}$$

$$\overrightarrow{X}(t+1) = \overrightarrow{X_p}(t) - \overrightarrow{A} \cdot \overrightarrow{D} \tag{8}$$

where t represents the current iteration, $\overrightarrow{A}$ and $\overrightarrow{C}$ are coefficient vectors, $\overrightarrow{X_p}(t)$ is the position vector of the prey, and $\overrightarrow{X}(t)$ represents the position vector of a grey wolf vector, They are calculated as follows:

$$\overrightarrow{A} = 2\overrightarrow{e} \cdot \overrightarrow{r_1} - \overrightarrow{e} \tag{9}$$

$$\overrightarrow{C} = 2\overrightarrow{r_2} \tag{10}$$

where $\vec{e}$ components are linearly decreased from 2 to 0 throughout iterations and r_1 and r_2 are random vectors in [0,1].

Hunting:

Grey wolves can identify the position and location of the prey and then encircle it. To mathematically simulate this process of hunting the prey, the previous statement that wolves a, b, and c are the optimum solutions and will have a better understanding of the positions and locations of prey is followed. Subsequently, the fittest three solutions obtained are assigned and the remaining search agents, including w wolves, update their positions accordingly. This is mathematically represented as follows:

$$\overrightarrow{D_a} = |\overrightarrow{C_1} \cdot \overrightarrow{X_a} - \overrightarrow{X}|, \overrightarrow{D_b} = |\overrightarrow{C_2} \cdot \overrightarrow{X_b} - \overrightarrow{X}|, \overrightarrow{D_c} = |\overrightarrow{C_3} \cdot \overrightarrow{X_c} - \overrightarrow{X}| \tag{11}$$

$$\overrightarrow{X_1} = \overrightarrow{X_a} - \overrightarrow{A_1}(\overrightarrow{D_a}), \overrightarrow{X_2} = \overrightarrow{X_b} - \overrightarrow{A_2}(\overrightarrow{D_b}), \overrightarrow{X_3} = \overrightarrow{X_c} - \overrightarrow{A_3}(\overrightarrow{D_c}) \tag{12}$$

$$\overrightarrow{X}(t+1) = \frac{\overrightarrow{X_1} + \overrightarrow{X_2} + \overrightarrow{X_3}}{3} \tag{13}$$

where $\overrightarrow{A_1}$, $\overrightarrow{A_1}$, $\overrightarrow{A_1}$ are random vectors, and $\overrightarrow{X_a}$, $\overrightarrow{X_b}$ and $\overrightarrow{X_c}$ represent the positions of and wolves, respectively.

Attacking prey:

The wolves attack when the movement of the prey stops. To approach the prey in the mathematical simulation, the value of $\vec{e}$ is reduced and the fluctuation range of $\overrightarrow{A}$ is also decreased with $\overrightarrow{A}$. That is by decreasing $\vec{e}$ from 2 to 0, the random value of $\overrightarrow{A}$ is changed in

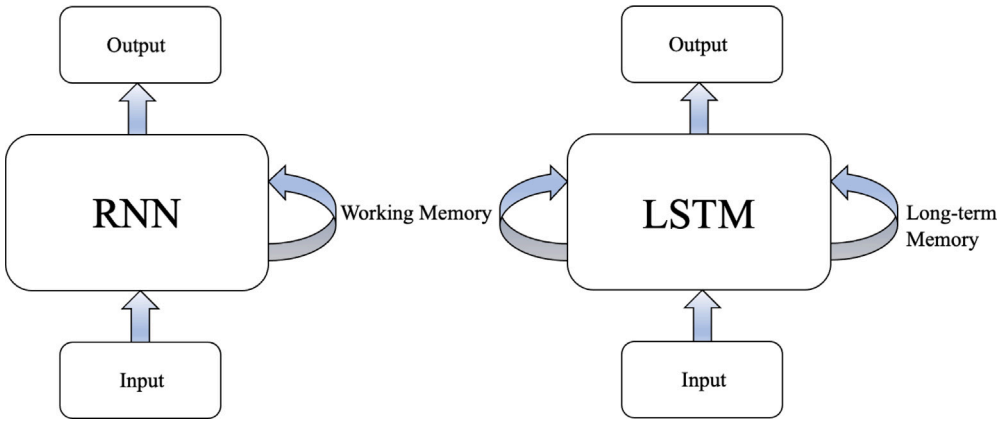

Fig. 3. General architecture of RNN and LSTM. In RNN, the network maintains information over time in the working memory. However, LSTM networks add a long-term memory for temporal, which allows them to solve the vanishing gradient problem in a regular RNN.

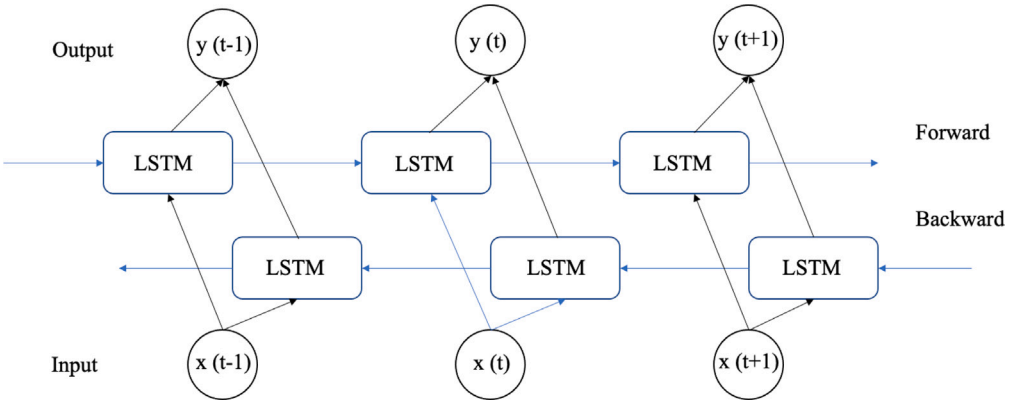

Fig. 4. Flowchart of the architecture of BiLSTM, where input is processed in both directions for better understanding and analysis of the time series.

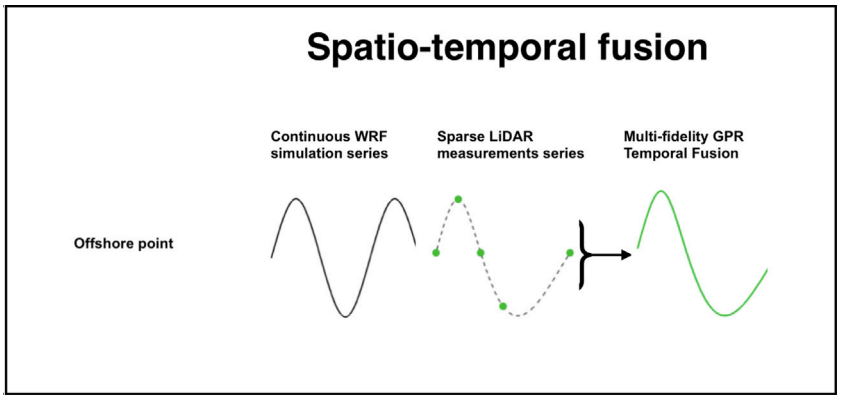

Fig. 5. Flow chart for temporal data fusion, where low-fidelity numerical simulations (WRF) are merged with high-fidelity measurements (Lidars) using multi-fidelity Gaussian process regression [27].

the interval [-e, e] during the iterations. The next position of the search agent could be anywhere between the current position and the position of the prey, given the value of is in [−1, 1].

The Grey Wolf Optimizer (GWO) was chosen due to its effectiveness in optimizing complex neural network structures like BiLSTM. GWO mimics the leadership hierarchy and hunting mechanism of grey wolves in nature, providing a balance between exploration and exploitation in the optimization process. The parameter selection involved defining the search space for key BiLSTM parameters, including the number of layers, number of neurons per layer, and learning rate. Extensive trials were conducted to determine the optimal parameter values, which significantly enhanced the model's performance. Comparative analysis with other optimization algorithms validated the superiority of GWO in this context.

2.2. Procedure for numerical data fusion with measurements data

Temporal data fusion of low and high-fidelity data from simulations and measurements at a given location is performed using deep multi-fidelity-GPR. As illustrated in Fig. 3, the low-fidelity results (from numerical simulations) are assumed to be available across a continuous time domain T as in Section 2.1 High-fidelity results (from the lidar measurements) are available in a reduced time domain T_{re} where T_{re} is

a subset of T and can be discontinuous. Thereafter, the objective is to combine the low and high-fidelity results to reach a data fusion on the full-time domain T [27].

2.2.1. Pre-processing

The pre-processing approach for this model is Empirical Wavelet Transform (EWT), which is covered at the end of 2.1.1 high-frequency fluctuations in it. Uncorrelated filter modes are extracted from the wind speed series and a residual is obtained from the extraction. The reconstructed wind speed series shows a significant decrease in high-frequency fluctuations and is used as input to the data fusion model.

2.2.2. Temporal data fusion

This section introduces the algorithms for temporal data fusion by combining the low-fidelity continuous numerical time series and the high-fidelity intermittent measured one. The prototype of the Gaussian process regression is briefly introduced and then multi-fidelity GPR is presented. The use of different co-variance functions such as constant, linear, squared exponential, Matern kernel, and rational quadratic, defines the method of prediction for the Gaussian process. The model of merging two signals of different fidelities generated from different sources is shown in Fig. 6.

The data fusion approach integrates lidar measurements, which provide high-fidelity wind speed data, with numerical simulations. This fusion leverages the strengths of both data sources, improving overall prediction accuracy. Lidar data, known for its precision, enhances the robustness of numerical simulations, which cover broader temporal and spatial scales. The fusion process involves aligning the datasets temporally and spatially, followed by applying the hybrid model to predict wind speeds more accurately.

Gaussian Process Regression (GPR):

GPR is a non-parametric, stochastic process that follows the Bayesian approach for regression, working well on small data sets and having the ability to provide uncertainty measurements on predictions. Predictions are derived using a probability distribution over all possible values of a time series using prior functions w of training points f at observed points t, and targeted values f^* at unobserved points t^* are calculated from a predictive distribution, $p(f^*|t^*, f, t)$, by considering all possible predictions using their calculated posterior distribution [33]:

$$p\left(f^*|t^*, f, t\right) = \int p\left(f^*|t^*, w\right) p\left(w|f, t\right) dw. \tag{14}$$

To trace the integration process of Eq. (14), all terms of the equation are assumed Gaussian. The prior function defines the Gaussian distribution [33]:

$$f(t) \sim GP(m, k(t, t')), \tag{15}$$

where m is the mean function, which represents the trend of the function, and the covariance function (kernel), $k(t, t')$, represents the dependence of the structure, defined by the hyperparameters [34].

Multi-fidelity Gaussian Process Regression:

This section discusses advanced temporal data fusion using data with multiple fidelities to enhance the accuracy of prediction. The data sets are obtained using different techniques mathematically, the multi-fidelity technique considers the high-fidelity model as a function of two variables (t, s) and then uses the low-fidelity data as the s variable [33]:

$$f_h(t) = g(t, f_1(t)), \tag{16}$$

where in the present work $f_h(t)$ and $f_l(t)$ are the high-fidelity lidar measurements and low-fidelity WRF simulations, respectively. Such nonlinear auto-regressive Gaussian processes (NARGP) have been observed to produce highly accurate prediction when $f_h(t)$ is non-linearly dependent on $f_l(t)$, and GPR is then performed in a two-dimensional space.

To implement this, the co-kriging model is adopted, which uses multivariate functions concerning different levels of fidelities to reflect different accuracies. The additional data set is later introduced to the Gaussian distribution and the terms of the first data set (t, s) and the second data set (t', s') are added, while the mean function is zero, through [33]:

$$f(t) \sim GP(m, k((t, s), (t', s'))). \tag{17}$$

Merging of two or more sets that are approximately linearly dependent on scaling and shifting parameters is approached by Kennedy and O'Hagan [35]. However, due to the presence of nonlinear dependencies generally between the datasets, the quality of results degraded and this is a major issue for linear data fusion algorithms. To overcome and resolve the nonlinear dependencies, a space-dependent scaling factor $\rho(x)$ [33] or deep multi-fidelity GP [36] is introduced. However, the improvement brings further optimizations of additional hyperparameters. Here the NARGP algorithm, which is an implicit Automatic Relevance Determination (ARD) weighting, is employed in the extended space, parameterized by t and s, which produces a different scaling of the existing hyperparameters for each dimension in the kernel [37].

Additionally, the formulation can be extended through functions of the low-fidelity data set. The high-fidelity data can be further considered as a function of t, $f_l(t)$ and the derivatives of $f_l(t)$, since $f_l(t)$ has a similar trend to $f_h(t)$ [33]:

$$f_h(t) = g(t, f_l(t), f_l^1(t), \ldots, f_l^i(t)), \tag{18}$$

where $f_l^i(t)$ is the ith derivative of the low fidelity data.

3. Test case description

To test the methods, a case associated with the RUNE project is considered, which was a near-shore experiment conducted on the west coast of Denmark (see Fig. 6 (*b*)) [38]. The surrounding area is nearly flat coastal farmland and moving northwards from position 1 to 3, the sand embankment separating the North Sea and the grasslands transforms into cliffs covered by grass. In this work, the data from dual-Doppler scans performed nearly perpendicular to the coastline from about 5 km offshore to 2 km onshore are used. These scans are performed by synchronizing measurements from two scanning lidars, which are modified versions of WLS200S Leosphere units, one located at position 1 and the other at position 3. The dual-Doppler scans were performed at 50 m above mean sea level (amsl) during the period 2015-12-08 to 2016-02-17. Due to filtering of high noise/low signal strength and system availability, only 114 10-min samples of data are available at all the dual-Doppler positions shown as black markers (dots) in Fig. 6 (*a*). Further details concerning the experimental campaign and the instrumentation can be found in [38].

The experiments applied a one-step-ahead forecasting horizon, with plans to extend to multi-step forecasting in future work. The current setup ensures high accuracy for immediate predictions, critical for operational decision-making in wind farms.

Here is a numerical experiment, which was part of several numerical simulations performed using the WRF model v3.6 to supplement the measurements of RUNE [38]. This particular experiment was set up with 4 nested domains, with the outermost covering northwestern Europe, and a 2-km horizontal resolution innermost domain covering the west coast of Denmark. Spectral nudging to the ERA5 reanalysis is used in the upper model levels of the outermost domain. The simulation had 8 vertical levels within the first 100 m, and instantaneous output was produced every 10 min. The experiment also used the Mellor–Yamada–Janjic planetary boundary layer scheme, a sea surface temperature product from the Danish Meteorological Institute [39], and the CORINE land cover description. Fig. 7a shows an illustration of the low-fidelity WRF simulation data set of the most onshore point and Fig. 7b shows the high-fidelity lidar measurements for the same point. It is worth noting that data fusion models in the literature frequently combine

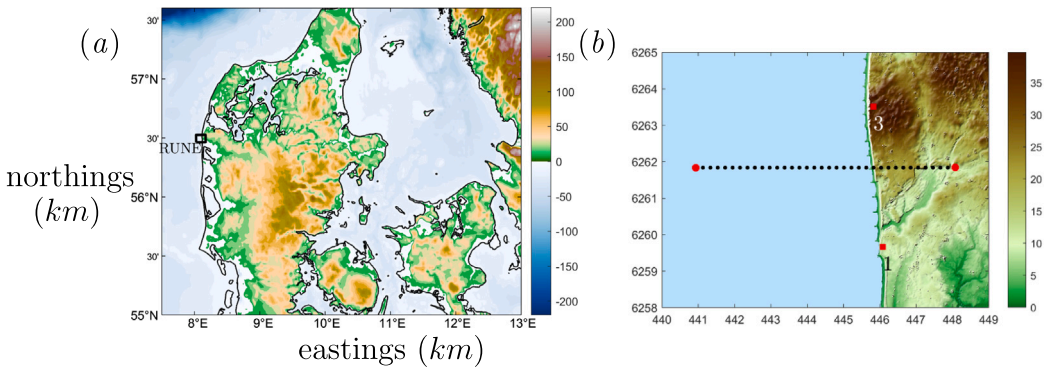

Fig. 6. (*a*) RUNE experimental area. Positions of the lidars are shown in red square markers and the dual-Doppler scans in black and red dot markers. The colour bar indicates the terrain elevation in meters above mean sea level. (*b*) The location of the RUNE experiment site (black rectangle) in Denmark.

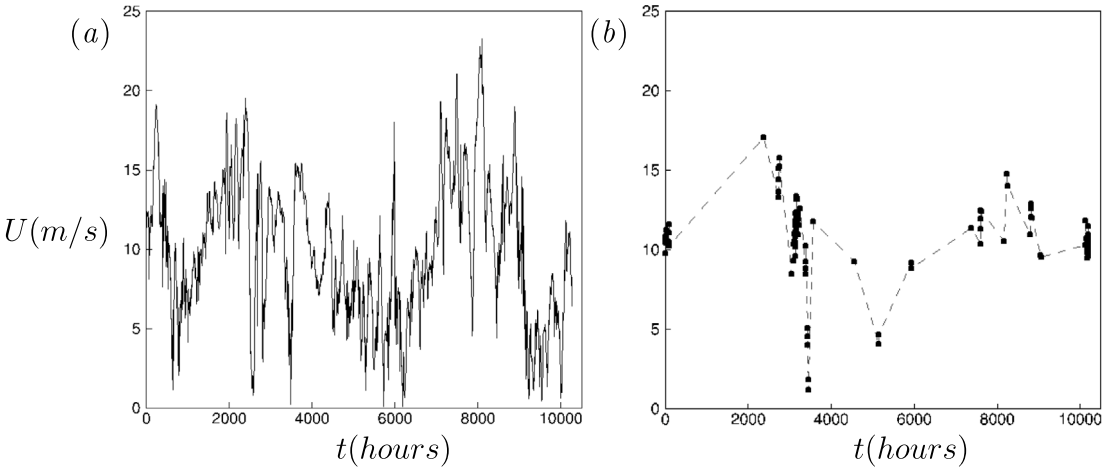

Fig. 7. (*a*) Low-fidelity data from numerical simulation (WRF) at the most onshore point.(*b*) High-fidelity data from the dual-Doppler lidar setup at the most onshore point.

continuous low-fidelity data sources, such as numerical WRF simulations, with limited high-fidelity data sources like LiDAR. In our case, LiDAR data accounts for 7% of the dataset, surpassing the percentages used in many existing data fusion models. This aligns with the primary aim of our approach, which is to reduce reliance on and acquisition of high-fidelity data sources.

3.1. Evaluation metrics

Many evaluation metrics are researched and applied to evaluate the effectiveness of different forecasting models. Although, no general standard for performance metrics is followed. The most common error metrics include mean absolute error (MAE), mean squared error (MSE), and root mean squared error (RMSE) results from multiple studies show that the optimum model indicated its dependence on the evaluation metric used. After investigating the equations for each metric and comparing the results from different studies, the abnormality is found to be associated with RMSE and MSE. This is because as models are improved absolute errors have reduced significantly to less than 1, i.e. 0.9, 0.6, and 0.4 m/s. In the equations for RMSE and MSE, the error is squared before it is aggregated to find the mean error, which means that the error is very much decreased for iterations where performance is high (error less than 1), and this is a property of both RMSE and MSE; they penalize both high and low errors more than other metrics. Hence, an iteration that generates forecasts with an error less than 1, when squared, will produce an error that is smaller compared to using other metrics. If the iteration generates forecasts with errors greater than 1, when squared, the error will increase, and the final mean error is significantly affected.

Table 2
Table of different metrics and their equations to calculate the forecasting error.

Metric	Definition	Equation
MAE	Mean absolute error for D forecasting results	$\frac{1}{D}\sum_{i=1}^{D}\lvert x_i - y_i\rvert$
MSE	Mean squared error for D forecasting results	$\frac{1}{D}\sum_{i=1}^{D}(x_i - y_i)^2$
RMSE	Root MSE for D forecasting results	$\sqrt{\frac{1}{D}\sum_{i=1}^{D}(x_i - y_i)^2}$

MAE, MSE, and RMSE are used to evaluate the average difference between the predicted result and the original data since the average of variance is required. The equations and definitions for each of the three metrics are given in Table 2, where x and y represent the actual and predicted values, respectively, and D is the sample size.

4. Results and discussion

In this work, a hybrid forecasting model is used to predict data from a WRF simulation dataset. IMFs from CEEMDAN are presented. The second signal decomposition of IMF1 using EWT then reduces the high-frequency noise. The generated signals are fed into an optimized BiLSTM algorithm optimized using GWO and predicted signals are aggregated and compared to the original dataset. Subsequently, the predicted signal is compared to the lidar observations and observed next to the results from the data fusion model, which employed data fusion of WRF simulations and Lidar measurements to investigate and assess the benefits of using data generated from different sources at different fidelities.

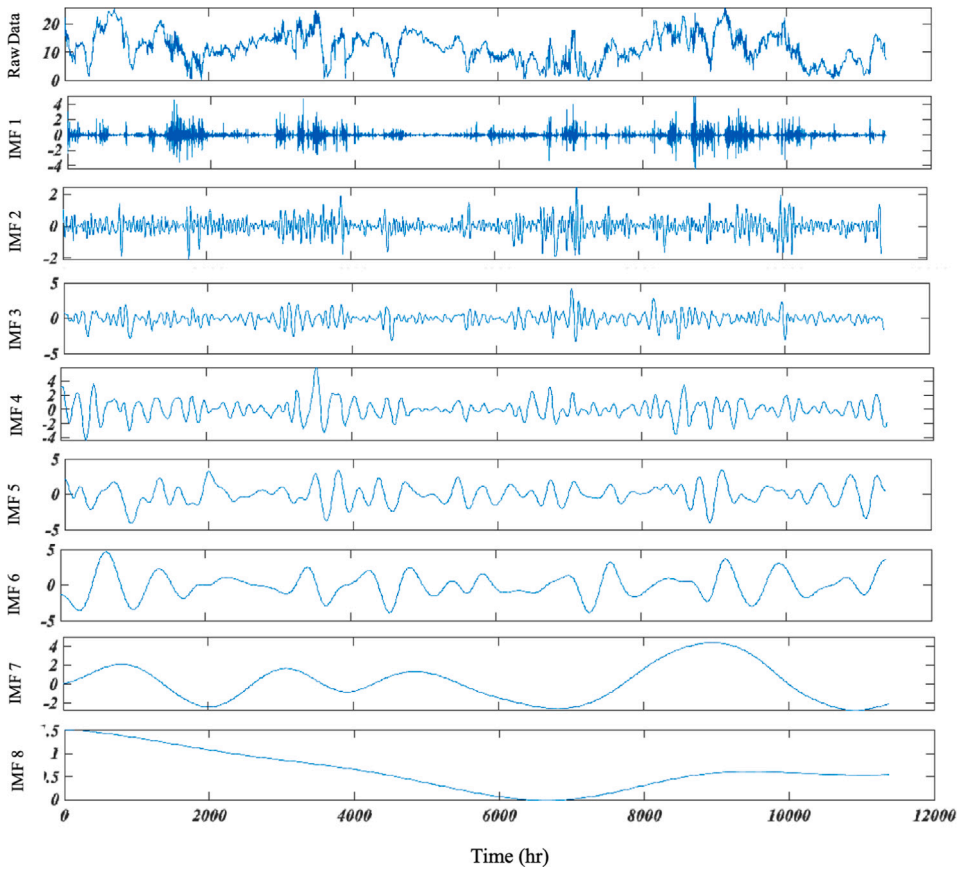

Fig. 8. Decomposed Intrinsic Modal Functions (IMF) sub-series from original signal using CEEMDAN.

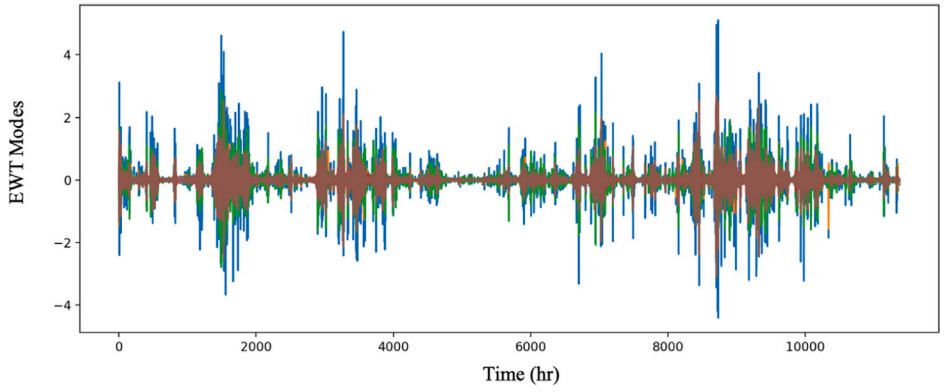

Fig. 9. Secondary signal decomposition result of IMF 1 using EWT (Different colours represent the different modes extracted from the IMF 1 signal).

An in-depth analysis of the prediction error was conducted, focusing on the error distribution and its impact on model predictions. The mean absolute error (MAE) and root mean square error (RMSE) metrics were used to quantify the prediction accuracy. The results showed that the hybrid model significantly reduced errors compared to traditional methods. The error analysis highlighted the model's robustness in handling various wind speed patterns, indicating its reliability for practical applications (see Fig. 8).

4.1. Pre-processing using CEEMDAN and EWT

To build a forecasting model with high performance, it is critical to fully analyse and consider the features of the original time series. In this work, the CEEMDAN approach is first used to decompose the original wind speed time series to reduce the non-stationary and non-linear characteristics. As presented in Fig. 9, the raw data in addition to the IMFs extracted from the original wind speed data are shown from highest to lowest frequency. Results show that each decomposition contains its characteristics, reflecting the different oscillatory nature in the time series provided as shown in the literature for other examples. For this experiment, 8 total IMF components are generated, named IMF1 to IMF8, with IMF1 representing the decomposition with the highest frequency, additionally, IMF1 has the most detailed information of the original time series. Contrarily, IMF 8 represents the decomposition with the lowest frequency, which presents the variational trend of the wind speed series. Forecasting results of the IMFs show good performance, except for IMF1 due to its high-frequency oscillatory nature. Therefore, IMF1 requires additional pre-processing due to its poor prediction performance in addition to its importance.

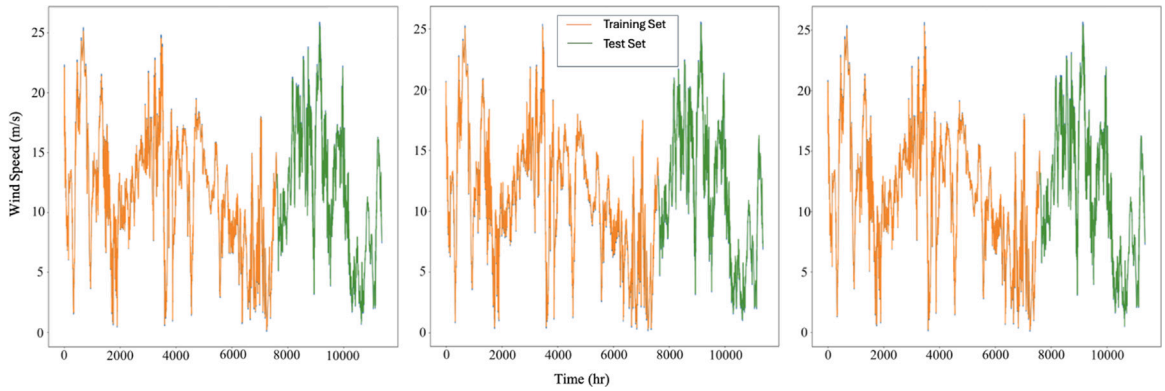

Fig. 10. Prediction result for all three location points, where the orange line represents the training points and the green represents the test points.

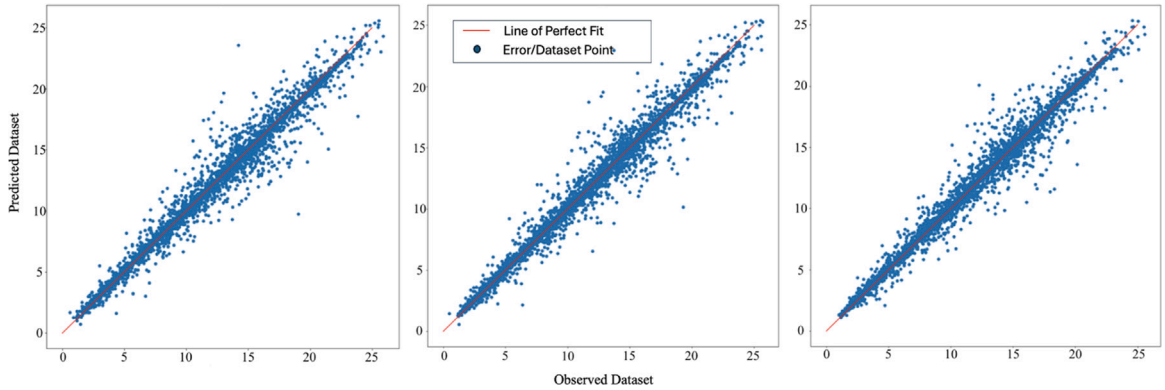

Fig. 11. Plots of observed data against predicted data from the proposed model for all three data locations.

Table 3
Configurations and accuracy of the Gaussian Process Regression models.

WRF Point	Metric	Value
Point 1	MAE (Training)	0.4
	MAE (Test)	0.48
	MSE (Training)	0.56
	MSE (Test)	0.79
	RMSE (Training)	0.75
	RMSE (Test)	0.88
Point 3	MAE (Training)	0.41
	MAE (Test)	0.49
	MSE (Training)	0.55
	MSE (Test)	0.76
	RMSE (Training)	0.74
	RMSE (Test)	0.87
Point 3	MAE (Training)	0.38
	MAE (Test)	0.51
	MSE (Training)	0.59
	MSE (Test)	0.83
	RMSE (Training)	0.77
	RMSE (Test)	0.91

To improve the status of IMF1, EWT is employed as a secondary signal decomposition approach to further decompose that signal. The result is 6 decomposed modes of IMF1 (shown in Fig. 9) and all IMFs and modes, except IMF1, are then fed into the BiLSTM model for predictions (see Fig. 10 and Table 3).

4.2. Predictions using BiLSTM optimized using GWO

Following the pre-processing of the three furthest offshore points from the WRF simulation using CEEMDAN and EWT, data is fed into Bidirectional Long Short Term Memory (BiLSTM) neural networks optimized using Grey Wolf Optimizer (GWO). The performance of the model is optimized, by adjustment of the weights, biases, and number of layers for each of the neural networks. The performance of the model is explored by hiding 40% of the data, and the remaining 60% is used in training the neural networks. Fig. 11 show the prediction results of training and testing the three signals and a close-up on the generated results. The performance of predictions is evaluated using 3 evaluation metrics, MAE, MSE, and RMSE. The variations of speed values of the evaluation metrics for the proposed model are 0.4, 0.56, and 0.75 m/s for the training section of the data and 0.48, 0.79, and 0.88 m/s, respectively. Fig. 11 shows the observed vs. predicted curve for three data points.

4.3. Proposed model compared to lidar observations and the data fusion model

The entirely numerical experiment that is performed above revealed that building hybrid models that employ more than one pre-processing algorithm, combined with advanced and optimized deep neural networks, is sufficient for forecasting signals of numerically simulated wind speed time series. However, in wind resource assessment, the objective is not only to predict data accurately but to predict data as accurately as possible relative to the physically measured data that is later observed through measurements. For the majority of experiments, models are trained and tested on simulation data, since they are easier and cheaper to obtain than physical data. Contrarily, simulation data has low-fidelity to wind speed measurement using for example lidars, but it delivers a continuous stream of data across large spaces for long periods, with low cost compared to lidar measurements. A possible solution to the problem of scarcity of lidar data on the one hand and the low-fidelity of numerical simulations, on the other hand, is proposed

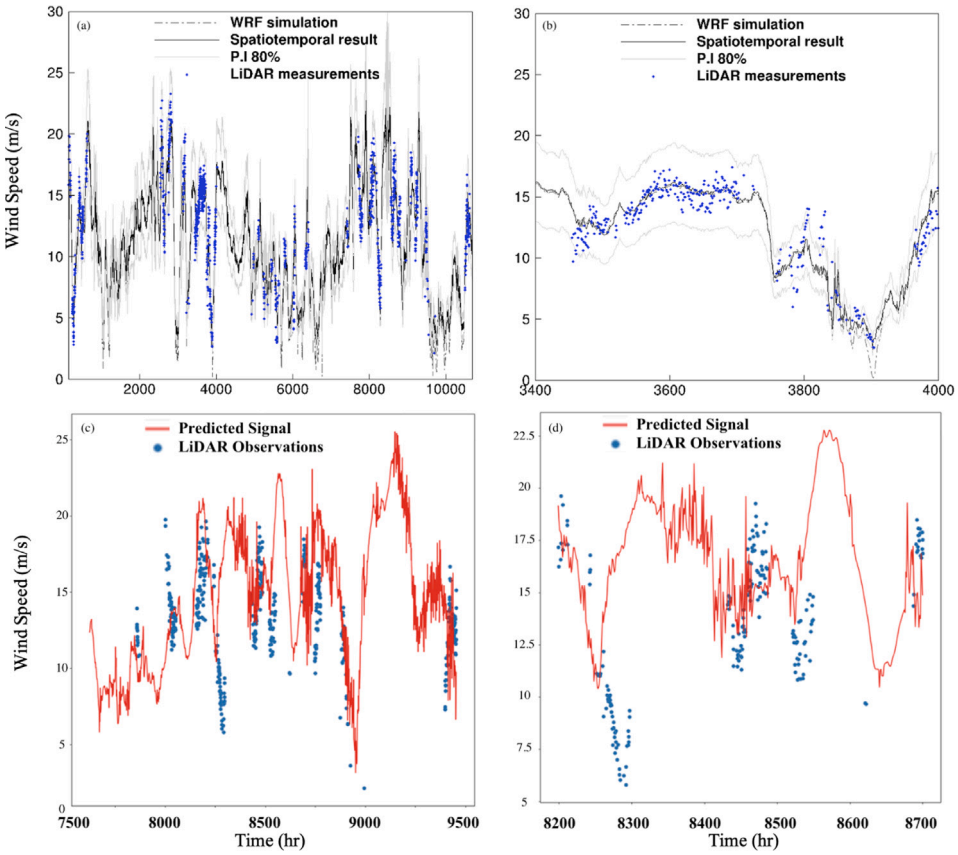

Fig. 12. Panels (a) and (b) show predicted signal and close-up of the test data using the proposed model (CEEMDAN+EWT)+ (BiLSTM + GWO) compared to lidar measurements. Panels (c) and (d) show the predicted signal and close-up of the test data using the proposed data fusion model of multi-fidelity Gaussian process regression against lidar measurements.

in the second approach in which lidar measurements are merged with WRF simulations in a hybrid solution for offshore wind resource assessment using a data fusion approach. For this work, the significance of data fusion with high fidelity is demonstrated. We compare results from a model trained with only simulation data to a model trained with both lidar measurements and simulation data of the measured locations.

Fig. 12 panels (a) and (b) show results from the (CEEMDAN + EWT) + (BiLSTM + GWO) model compared to lidar measurements. Results show that despite the model performing effectively in predicting the WRF simulation time series, performance is poor compared to the high-fidelity lidar measurements, where the average RMSE is 1.36 m/s. Alternately, the model trained with lidar measurements showed significant out-performance, where the RMSE is 0.63 m/s, which is a 53% improvement in performance. In panels (c) and (d) a presentation of the results from the model that follows the data fusion method using lidar measurements as additional input information. Results show that the model using data fusion outperforms the model that does not consider lidar data but relies solely on weather forecasting numerical simulation.

5. Conclusions

Current wind energy resource assessment methods often struggle with non-stationary and non-linear characteristics of wind speed signals. Traditional single decomposition techniques are inadequate for accurately forecasting high-frequency signals. Existing models also fail to fully leverage high-fidelity lidar measurements, limiting their predictive capabilities. This study addressed these limitations by proposing a hybrid model that combines secondary decomposition techniques with data fusion, providing a more robust and accurate assessment of wind resources.

This novel hybrid approach decomposed the wind speed time series using two pre-processing approaches to reduce high-frequency noise and decrease non-stationarity and non-linearity in the signal. The IMFs and modes from the simulation data preprocessed with CEEMDAN and EWT, respectively, are fed into BiLSTM neural networks optimized by a GWO algorithm to predict the wind speed time series. Results from the model show high performance and accuracy in mimicking the simulated time series (WRF) with errors ranging from 0.48 m/s to 0.91 m/s, however, the simulation data are not a good representation of the actual wind speed observed, as they are considered of poor quality compared to high-fidelity data (lidar), the error increases by at least 30% when compared to the observed lidar measurements. Contrarily, the data fusion model using few lidar measurements generates predictions that are 40% more accurate than the hybrid model in that sense. In this study, we predicted wind speeds at a time scale of 10-minute intervals in a time series format.

In the field of wind resource assessment, and specifically, when predicting wind speed data, the task is to get better forecasts relative to the high-fidelity data, as the main objective of the process is to estimate the power generated from the wind turbine or farm. Minor differences in the wind speed would have a significant impact on the estimation of wind power since the formula for calculating the output power depends on the cube of the wind speed. Thus, a complete dependence on simulations may risk calculating power values that are far off real values. Contrarily, data fusion of lidar measurements and simulations could provide a solution that combines the merits of both techniques for the assessment of wind resources and hence provides results that are more accurate compared to the observed wind speeds.

The experiments applied a one-step-ahead forecasting horizon, with plans to extend to multi-step forecasting in future work. The current setup ensures high accuracy for immediate predictions, critical for operational decision-making in wind farms.

CRediT authorship contribution statement

Basem Elshafei: Writing – original draft, Methodology, Conceptualization. **Atanas Popov:** Writing – review & editing, Validation, Supervision, Funding acquisition, Conceptualization. **Donald Giddings:** Writing – review & editing, Validation, Supervision, Project administration, Investigation, Funding acquisition, Conceptualization.

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.

Acknowledgements

This work has received funding from the European Union Horizon 2020 research and innovation programme, UK under the Marie Skłodowska-Curie grant agreement No 777717 and the future and emerging technologies programme, UK with agreement No. 828799.

References

[1] Global wind report 2021 - Global Wind Energy Council - GWEC (no date), Available at: https://gwec.net/global-wind-report-2021/.
[2] Chen Jian, Zhang Yu, Xu Zhongyun, Li Chun. Flow characteristics analysis and power comparison for two novel types of vertically staggered wind farms. Energy 2023;263:126141. http://dx.doi.org/10.1016/j.energy.2022.126141, Part E.
[3] Ma Yixiang, Yu Lean, Zhang Guoxing. Short-term wind power forecasting with an intermittency-trait-driven methodology. Renew Energy 2022;198:872–83. http://dx.doi.org/10.1016/j.renene.2022.08.079.
[4] Bentsen Lars Ødegaard, Warakagoda NaradaDilp, Stenbro Roy, Engelstad Paal. Spatio-temporal wind speed forecasting using graph networks and novel transformer architectures. Appl Energy 2023;333:120565. http://dx.doi.org/10.1016/j.apenergy.2022.120565.
[5] Wang Yun, Zou Runmin, Liu Fang, Zhang Lingjun, Liu Qianyi. A review of wind speed and wind power forecasting with deep neural networks. Appl Energy 2021;304:117766. http://dx.doi.org/10.1016/j.apenergy.2021.117766.
[6] Ahmad Tanveer, Chen Huanxin. A review on machine learning forecasting growth trends and their real-time applications in different energy systems. Sustainable Cities Soc 2020;54:102010. http://dx.doi.org/10.1016/j.scs.2019.102010.
[7] Song J, Wang J, Lu H. A novel combined model based on advanced optimization algorithm for short-term wind speed forecasting. Appl Energy 2018;215:643–58.
[8] Liu H, Mi X, Li Y. Smart multi-step deep learning model for wind speed forecasting based on variational mode decomposition, singular spectrum analysis, lstm network and elm. Energy Convers Manage 2018;159:54–64.
[9] Neshat M, Nezhad MM, Abbasnejad E, Mirjalili S, Tjernberg LB, Garcia DA, Alexander B, Wagner M. A deep learning-based evolutionary model for shortterm wind speed forecasting: A case study of the lillgrund offshore wind farm. Energy Convers Manage 2021;236:114002.
[10] Zounemat-Kermani M. Hourly predictive Levenberg–Marquardt ANN and multi linear regression models for predicting of dew point temperature. Meteorol Atmos Phys 2012;117:181–92. http://dx.doi.org/10.1007/s00703-012-0192-x.
[11] Chen Y, Zhang S, Zhang W, Peng J, Cai Y. Multifactor spatio-temporal correlation model based on a combination of convolutional neural network and long short-term memory neural network for wind speed forecasting. Energy Convers Manage 2019;185:783–99.
[12] Wu Z, Xiao L. A secondary decomposition based hybrid structure with meteorological analysis for deterministic and probabilistic wind speed forecasting. Appl Soft Comput 2019;85:105799.
[13] Liu H, X-w Mi, Y-f Li. Wind speed forecasting method based on deep learning strategy using empirical wavelet transform, long short term memory neural network and elman neural network. Energy Convers Manage 2018;156:498–514.
[14] Li L-L, Chang Y-B, Tseng M-L, Liu J-Q, Lim MK. Wind power prediction using a novel model on wavelet decomposition-support vector machines-improved atomic search algorithm. J Clean Prod 2020;270:121817.
[15] Luo L, Li H, Wang J, Hu J. Design of a combined wind speed forecasting system based on decomposition-ensemble and multi-objective optimization approach. Appl Math Model 2021;89:49–72.
[16] Wu C, Wang J, Chen X, Du P, Yang W. A novel hybrid system based on multi-objective optimization for wind speed forecasting. Renew Energy 2020;146:149–65.
[17] Qu Zongxi, Mao Wenqian, Zhang Kequan, Zhang Wenyu, Li Zhipeng. Multi-step wind speed forecasting based on a hybrid decomposition technique and an improved back-propagation neural network. Renew Energy 2019;133:919–29. http://dx.doi.org/10.1016/j.renene.2018.10.043.
[18] Liu Hui, Tian Hong-qi, Liang Xi-feng, Li Yan-fei. Wind speed forecasting approach using secondary decomposition algorithm and elman neural networks. Appl Energy 2015;157:183–94. http://dx.doi.org/10.1016/j.apenergy.2015.08.014.
[19] Huang NE, Shen Z, Long SR, Wu MC, Shih HH, Zheng Q, Yen N-C, Tung CC, Liu HH. The empirical mode decomposition and the Hilbert spectrum for nonlinear and non-stationary time series analysis. Proc: Math Phys Eng Sci 1998;454(1971):903–95, http://www.jstor.org/stable/53161.
[20] Sun Na, Zhou Jianzhong, Chen Lu, Jia Benjun, Tayyab Muhammad, Peng Tian. N adaptive dynamic short-term wind speed forecasting model using secondary decomposition and an improved regularized extreme learning machine. Energy 2018;165:939–57. http://dx.doi.org/10.1016/j.energy.2018.09.180, Part B.
[21] Torres ME, Colominas MA, Schlotthauer G. A complete ensemble empirical mode decomposition with adaptive noise. In: Proceedings of the 36th IEEE international conference on acoustics, speech and signal processing. ICASSP, vol. 22–27, Prague: Czech Republic; 2011, p. 4144–7.
[22] Jianming Hu, Jianzhou Wang. Short-term wind speed prediction using empirical wavelet transform and Gaussian process regression. Energy 2015;93.
[23] Duan Jikai, Zuo Hongchao, Bai Yulong, Duan Jizheng, Chang Mingheng, Chen Bolong. Short-term wind speed forecasting using recurrent neural networks with error correction. Energy 2021;217:119397. http://dx.doi.org/10.1016/j.energy.2020.119397.
[24] Hochreiter Sepp. Jürgen Schmidhuber; long short-term memory. Neural Comput 1997;9(8):1735–80. http://dx.doi.org/10.1162/neco.1997.9.8.1735.
[25] Yuan Xiaohui, Chen Chen, Jiang Min, Yuan Yanbin. Prediction interval of wind power using parameter optimized beta distribution based LSTM model. Appl Soft Comput 2019;82:105550. http://dx.doi.org/10.1016/j.asoc.2019.105550.
[26] Banik A, Behera C, Sarathkumar TV, Goswami AK. Uncertain wind power forecasting using LSTM-based prediction interval. IET Renew Power Gen 2020;14:2657–67. http://dx.doi.org/10.1049/iet-rpg.2019.1238.
[27] Elshafei Basem, Peña Alfredo, Xu Dong, Ren Jie, Badger Jake, Pimenta Felipe M, Giddings Donald, Mao Xuerui. A hybrid solution for offshore wind resource assessment from limited onshore measurements. Appl Energy 2021;298:117245. http://dx.doi.org/10.1016/j.apenergy.2021.117245, ISSN 0306 2619.
[28] Huang L, Li L, Wei X, et al. Short-term prediction of wind power based on BiLSTM–CNN–WGAN-GP. Soft Comput 2022;26(1060):7–10621. http://dx.doi.org/10.1007/s00500-021-06725-x.
[29] Singla P, Duhan M, Saroha S. An ensemble method to forecast 24-h ahead solar irradiance using wavelet decomposition and BiLSTM deep learning network. Earth Sci Inform 2022;15:291–306. http://dx.doi.org/10.1007/s12145-021-00723-1.
[30] Siami-Namini S, Tavakoli N, Namin AS. The performance of LSTM and BiLSTM in forecasting time series. In: 2019 IEEE international conference on big data (big data). Los Angeles, CA, USA; 2019, p. 3285–92. http://dx.doi.org/10.1109/BigData47090.2019.9005997.
[31] Liang Tao, Zhao Qing, Lv Qingzhao, Sun Hexu. A novel wind speed prediction strategy based on Bi-LSTM, MOOFADA and transfer learning for centralized control centers. Energy 2021;230:120904. http://dx.doi.org/10.1016/j.energy.2021.120904.
[32] Mirjalili Seyedali, Mirjalili SeyedMohammad, Lewis Andrew. Grey wolf optimizer. Adv Eng Softw 2014;69:46–61. http://dx.doi.org/10.1016/j.advengsoft.2013.12.007.
[33] Perdikaris P, Raissi M, Damianou A, Lawrence ND, Karniadakis GE. Nonlinear information fusion algorithms for data-efficient multi-fidelity modelling. Proc R Soc A 2017;473.
[34] Miltiadis Alamaniotis, Georgios Karagiannis. Integration of Gaussian processes and particle swarm optimization for very-short term wind speed forecasting in smart power. Int J Monit Surv Technol Res 2018;5.
[35] Kennedy MC, O'Hagan A. Predicting the output from a complex computer code when fast approximations are available. Biometrika 2000;87.
[36] Edward Rasmussen Carl, KI Williams Christopher. Gaussian processes for machine learning. 2018, http://dx.doi.org/10.7551/mitpress/3206.001.0001.
[37] Yunhua Li, Lina Ling, Jiantao Chen. Combined grey prediction fuzzy control law with application to road tunnel ventilation system. J Appl Res Technol 2015;13.
[38] Rogier Floors, Alfredo Peña, Guillaume Lea, Nikola Vasiljević, Elliot Simon, Michael Courtney. The RUNE experiment-a database of remote-sensing observations of near-shore winds. Remote Sens 2016;8.
[39] Floors R, Hahmann AN, Peña A. Evaluating mesoscale simulations of the coastal flow using lidar measurements. J Geophys Res: Atmos 2018;123.