Hybrid attention-based temporal convolutional bidirectional LSTM approach for wind speed interval prediction

Bala Saibabu Bommidi[1] · Vishalteja Kosana[1] · Kiran Teeparthi[1] · Santhosh Madasthu[2]

Abstract

Precise wind speed prediction is crucial for the management of the wind power generation systems. However, the stochastic nature of the wind speed makes optimal interval prediction very complicated. In this paper, a hybrid approach consisting of improved complete ensemble empirical mode decomposition with adaptive noise (ICEEMDAN), temporal convolutional network with attention mechanism (ATCN), and bidirectional long short-term memory network (Bi-LSTM) is proposed for wind speed interval prediction (WSIP). First, ICEEMDAN is used to pre-process the raw data by decomposing the wind signal to several intrinsic mode functions. ATCN is used to reduce the uncertainty from the denoised data and extract the important temporal and spatial characteristics. Then, Bi-LSTM is used to forecast the high-quality intervals for the wind speed. Existing approaches observe a decline in the forecasting performance when the time ahead increases. As a result, the hybrid approach is evaluated using 5-min, 10-min, and 30-min ahead WSIP. To evaluate the novelty of the proposed approach, an experiment is conducted utilising wind speed data from the Garden City, Manhattan wind farm. The experimental results demonstrate that the proposed framework outperformed the comparison models with percentage improvements of 36%, 47%, and 17% for 5-min, 10-min, and 30-min ahead WSIP.

Introduction

Background

Fossil fuel exhaustion is causing severe climate change due to the fast expansion of many sectors throughout the world.

✉ Kiran Teeparthi
  kiran.t39@nitandhra.ac.in

  Bala Saibabu Bommidi
  saibabu.dec31@gmail.com

  Vishalteja Kosana
  kosanavishal@gmail.com

  Santhosh Madasthu
  madasthusanthosh@ieee.org

[1] Department of Electrical Engineering, National Institute of Technology Andhra Pradesh, Tadepalligudem, 534101, India

[2] Energy Production, Infrastructure Center (EPIC), University of North Carolina, Charlotte, NC, USA

Climate change is affecting many people around the world. In Awosusi et al. (2022), the authors looked at how carbon emissions are affected by globalisation of trade, rents on natural resources, economic expansion, and financial sector development. The impact of the climatology parameters on the COVID-19 is explained in Ahmadi et al. (2020). The authors of Habeşoğlu et al. (2022) discussed the oil price's impact on the amount of carbon emission levels in Turkey through financial regulation, energy use, and economic expansion. This gains the attention of the world to clean and endless resources such as solar energy, hydro energy, and wind energy, tidal energy (Council 2020). Especially, wind energy is the most popular renewable energy source with rapid development all over the globe. With 93.6 GW of new global wind installations in 2021, brings the total installed wind capacity to 837 GW. Wind turbine regulations control and wind power system dispatch are based on the dynamic wind speed. According to the cubic relation among wind power and speed, even a little change in wind speed causes a noticeable rise in wind power. Therefore, wind speed is essential for producing wind energy. Wind speed forecasting is challenging, nevertheless, due to the intrinsically nonlinear characteristics of wind

speed fluctuations, such as intermittency. In order to increase the usage of wind energy sources, it is crucial and essential to improve the precision of wind speed forecasts. The point forecasting in which the difference between predicted and actual are calculate is focused more in the current literature. Most of the researchers emphases on point forecasting methods. But these point forecasting models possess demerits such as inadequate accuracy problems due to uncertainty in the forecasts and low reliability. However, as the data becomes more complicated, the performance declines. Also, point forecasting approaches, on the other hand, fail to account for uncertainties and do not produce the needed accuracy. To overcome the above demerits of point forecasting, interval prediction is employed, which gives intervals instead of point values.

Literature review

The models utilised for interval prediction in the literature are broadly classified as statistical models and machine learning (ML) models. The statistical models can predict the parameters of error distribution to calculate the upper and lower bounds of a certain confidence interval. The mean and variance of the response variable are predicted using the interval forecasting model in Nix and Weigend (1994) but the coverage of prediction intervals (PI) is quite low for this implemented model. In Khosravi et al. (2011a), the traditional Bayesian approach is developed for interval prediction, but this approach has the demerit of huge computational complexity. In Pullanagari et al. (2018), the interval forecasting approach is implemented using quantile regression. For linear sequences, conventional statistical models perform excellently, but they come up short when applied to non-linear data as they cannot estimate distribution function with a hypothesis, and also have computational complexity problems. Both linear and non-linear data can be handled by artificial intelligence-based machine learning and deep learning algorithms (Ahmadi et al. 2022). The support vector machine is one of the machine learning models that are most frequently employed in this sector. These models are more widely used as a result of the development of AI into time series applications. The performance of AI-based techniques to prediction is nevertheless constrained by the quasi nature of the wind speed. The lower upper bound evaluation (LUBE)-based ML model was developed for the interval prediction. By using optimisation techniques, different variants such as single-objective optimisation (SOO)-based LUBE model (Hu et al. 2017) and multiple objective optimisation (MOO)-based LUBE models (Shrivastava et al. 2016) are implemented. Shallow ML networks in combination with the LUBE framework are also developed for interval prediction. The computational time is unacceptable with the tuning of more hyperparameters

for shallow ML network-based LUBE models. They also fail to provide the necessary information for the proper and efficient decision-making in the power system applications such as load management, spot pricing, and trading. Prediction interval forecasting techniques are used to address these issues since they reduce uncertainty and offer an indicator of accuracy. In comparison with shallow ML models, deep learning methods were given better performance for interval prediction (Khodayar et al. 2018). In Naik et al. (2019), multi-kernel robust ridge regression is used for interval forecasting of wind speed and wind power. The authors, in (Khosravi et al. 2011b; Quan et al. 2014), proposed a technique for constructing prediction intervals in neural network (NN) predictions that is both fast and trustworthy. The authors proposed a lower upper bound estimation (LUBE) method in which a NN with two outputs is built to estimate the prediction interval bounds. For nonparametric prediction intervals of wind power generation, the authors built a novel adaptive bilevel programming (ABP) model using extreme learning machine-based quantile regression in Zhao et al. (2020). The proposed ABP approach tries to reduce the mean interval width when good calibration is used. In He and Zhang (2020), authors used parallel quantile regression neural network wind power probability density forecasting model. This algorithm can improve the efficiency of quantile regression neural network. Results are evaluated by metrics of speed up and parallel efficiency. The authors in Pinson and Kariniotakis (2010) used a fuzzy inference model which allows integrating expertise on the properties of prediction errors for providing conditional interval forecasts. To improve the probabilistic forecast of wind farm levels and regional wind farms, a novel method based on Gaussian processes is developed (Xue et al. 2020). For wind speed interval prediction, a novel hybrid model depending on gated recurrent unit (GRU) with variational mode decomposition (VMD) was developed in Tang et al. (2019). For the forecast interval of wind power, a beta distribution-dependent long short-term memory (LSTM) neural network model has been proposed in Yuan et al. (2019). In the LSTM neural network model, a variation activation function is used, and the Beta distribution parameters are optimised using the PSO method. For wind power forecasting, Niu et al. (2022) use a data-driven strategy based on numerous factors and interval forecasting is done using kernel density estimation using a Gaussian functions. A BiLSTM model that is optimised via an attention mechanism is employed in this paper to increase point predicting accuracy. The authors of Zhang et al. (2020) offer a new interval model which depends on the fast correlation-based filter (FCBF) method, the optimised radial basis function (RBF) model, and the Fourier distribution for wind speed, which blends artificial intelligence techniques with statistical information. An interval prediction model was constructed in Zhang et al.

(2022) using an improved whale optimisation algorithm (IWOA) and a fast learning network (FLN). Adjusting the nonlinear convergence factor, as well as incorporating adaptive inertia weights and a chaos search technique, improved the IWOA's convergence speed and accuracy. The authors of Zhang et al. (2019) employed a multi-objective interval methods that rely on the conditional copula function, in which they completely exploited the correlations between variables to increase prediction accuracy without relying on an assumed probability distribution function. In Heydari et al. (2021), a wind power producer (WPP) in a competitive power market is given an interval prediction algorithm based on a new bidding technique based on optimal scenario making. Based on 39 years of data, the authors of Narayanan Natarajan (2021) check the effectiveness of nine prominent probability distribution methods for an evaluation of wind speed distribution (WSD) at ten sites in Tamil Nadu, India.

In order to improve accuracy and better account for the non-stationary characteristics of wind variables, hybrid models (Ahmadi et al. 2022; Ghoushchi et al. 2017) for such WSF have been developed. Intelligent models and signal decomposition techniques are some of the ways utilised to generate hybrid models (Cui et al. 2020; Sun and Xiaoxuan Wang 2022; Zhang and Pan 2020; Gupta et al. 2021). The ensemble empirical mode decomposition was used to pre-process the data (Artin et al. 2021; Cui et al. 2020). The bat algorithm was then used to optimise the back propagation neural network's connection weights and thresholds for forecasting. In Sun and Xiaoxuan Wang (2022), the wind speed sequence is decomposed using the wavelet transform, and the resulting detailed coefficients are further decomposed by using symplectic geometry mode decomposition. BPNN is optimised by using the marine predator's algorithm is then applied for WSF. To produce a decomposition result, variational mode decomposition (VMD) is applied to the original wind speed data in Zhang and Pan (2020), and the combined prediction approach using elman radial basis function is employed for prediction. In Zhang and Wang (2022), instead of initialising the parameters, an optimisation algorithm is used. In this paper, an improved PSO algorithm is used. And a rolling training prediction method is employed for wind speed prediction (WSP). In Barenya et al. (2022), authors utilised an hourly wind speed data for WSP and WSP is carried out by using wavelet kernel-based least square twin support vector regression. In Chen (2022), CEEMDAN is used in combination of singular value decomposition (SVD) to deconstruct and denoise the actual data, after which optimised Elman and ARIMA models are employed to forecast the wind speed components.

A hybrid framework is proposed in this paper for wind speed interval prediction to address the following challenges. The present literature focuses on point forecasting, in which the difference between anticipated and actual values is computed. However, as the data becomes more complicated, the performance declines. Also, point forecasting approaches, on the other hand, fail to account for uncertainties and do not produce the needed accuracy. They also fail to provide the necessary information for the proper and efficient decision-making in the power system applications such as load management, spot pricing, and trading. Prediction interval forecasting techniques are used to address these issues since they reduce uncertainty and offer an indicator of accuracy. A hybrid framework is proposed in this paper for wind speed interval prediction (WSIP) to address these challenges. The primary contributions of this study are as follows: (1) A novel hybrid ICEEMDAN-ATCN-BiLSTM approach is proposed by integrating the neural network architecture in the LUBE framework. (2) An efficient data preprocessing ICEEMDAN algorithm reduces noise in the input data and enhances the signal-to-noise ratio. ATCN extracts the important and dominating spatial and temporal features from the denoised wind speed. Bi-LSTM model interprets the important features bidirectionally to forecast the high-quality prediction intervals (PIs). (3) The integration of attention mechanism to the TCN layers enhanced feature extraction.

The rest of the manuscript is organised as follows: The proposed approach methodology and working are presented in Section "Proposed hybrid approach for wind speed interval prediction". Section "Experimental results" discusses the experimental results as well as a comparison to other methodologies. Section "Conclusions" outlines the conclusions.

Proposed hybrid approach for wind speed interval prediction

A novel hybrid framework using ICEEMDAN, TCN with attention (ATCN), and BiLSTM approach is proposed to enhance the quality of the WSIP. This section demonstrates the architecture of the proposed hybrid approach. The proposed hybrid approach is mainly divided into three sections: presupposition and noise elimination using ICEEMDAN, feature extraction using ATCN, and BiLSTM for WSIP. The proposed approach is illustrated in Fig. 1.

Presupposition and data decomposition

Prediction interval forecasting differs from deterministic forecasting. As a result, there is no way to train forecasting networks directly. To address this problem, the boundaries of the input wind speed data must be presupposed for the training to develop a framework for the WSIP using a construction interval strategy. U_i and L_i are the upper and
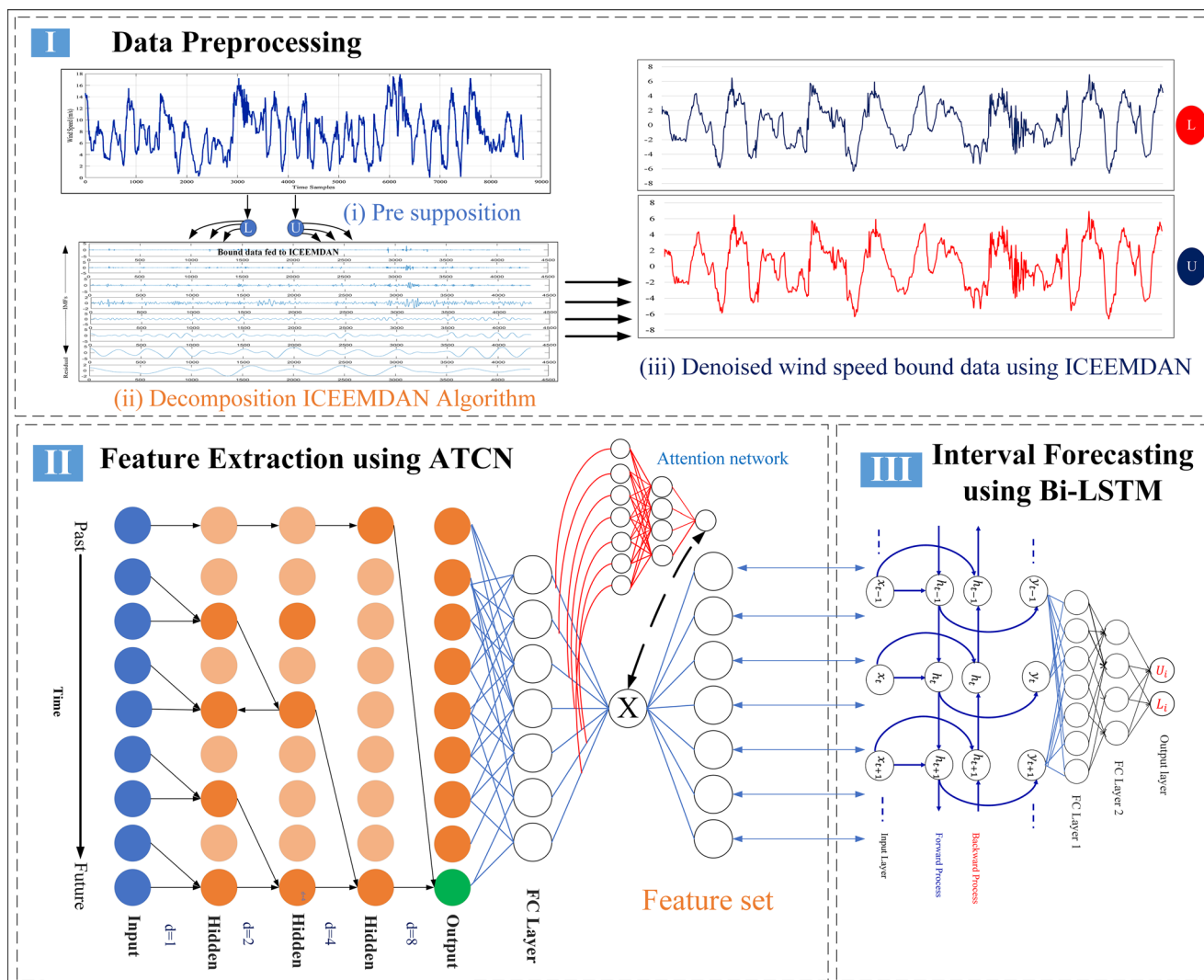
Fig. 1 Complete framework of the proposed approach for WSIP

lower bound of the wind speed data, formulated using Eqs. 1 and 2.

$$U_i = x + R_c \qquad (1)$$

$$L_i = x - R_c \qquad (2)$$

where x is the wind speed point, and R_c is the bound coefficient which is calculated by Eq. 3. α is the width coefficient pertaining to [0, 1].

$$R_c = \alpha * (\max(x) - \min(x)) \qquad (3)$$

However, the upper and lower bounds formed are highly complex and non-linear making the WSIP more difficult. Thus, the ICEEMDAN decomposition method is used to decompose the signals to produce the denoised upper and lower bounds of the wind speed data for WSIP.

The ensemble empirical mode decomposition (EEMD) has been developed to solve the mode mixing problem occurred in the empirical mode decomposition (EMD). However, the presence of the residual noise in the EEMD affects the performance. Hence, complete ensemble empirical mode decomposition with adaptive noise (CEEMDAN) is developed. But, intrinsic mode functions (IMFs) produced consist of the residual noise and spurious modes. Thus, improved CEEMDAN is implemented to address the disadvantages of the CEEMDAN approach. Figure 2 shows the flowchart for ICEEMDAN approach (Bouhalais and Nouioua 2021). By lowering the number of trials, this ICEEMDAN overcomes the following issues:

a) mode mixing problem,
b) frequency aliasing problem,
c) residual noise problem.

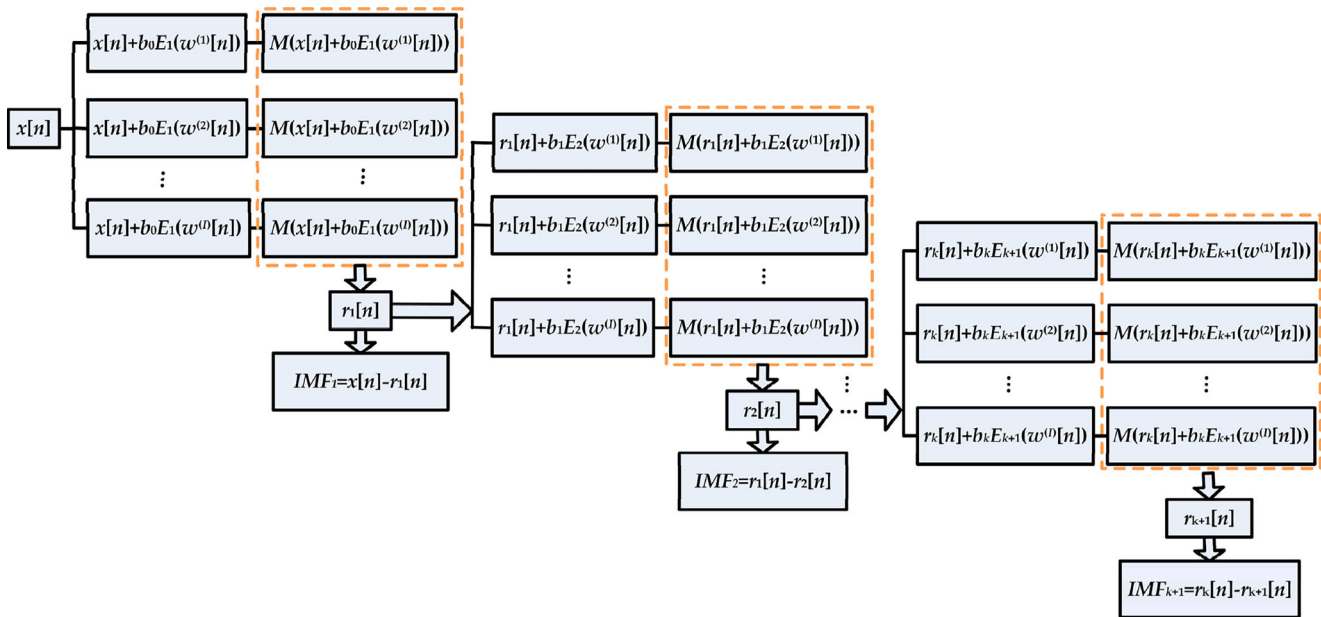In this approach, the wind speed data bound is decomposed into eight IMFs. As IMF1 is a very nonlinear

Fig. 2 Flowchart of improved CEEMDAN (Bouhalais and Nouioua 2021)

signal, it is discarded, and the other IMFs are combined to generate the wind speed data without any external noise.

Attention TCN-based Bi-LSTM approach for enhanced prediction interval forecasting

The proposed forecasting network consists of two steps: feature extraction using ATCN and PI forecasting using BiLSTM. The two steps are discussed in the following subsections.

Feature extraction using attention-based TCN method

For forecasting the prediction intervals, optimal feature extraction is essential to decrease the uncertainty and enhance the PI's quality. For this reason, ATCN layers are used for the feature extraction. The temporal convolutional network (TCN) is derived from the base convolutional neural network (CNN). The difference is the convolution operation. They are causal in the TCN whereas non-causal in the CNN model. TCNs use causal convolutions on the sequential data for the feature extraction. In this sequential method, the sequential inputs (X) are mapped to the sequential outputs (Y) through a non-linear mapping function (f). The main principle of the causal condition is that the output prediction (y_t) is depended only on the past data ($x_0, x_1, ..., x_t$) but not on the future data ($x_{t+1}, x_{t+2}, ..., x_T$) at the instant t. In this study, the input of the ATCN is the denoised data and the output represents the features extracted from the denoised wind speed.

$$\bar{Y}_0, \bar{Y}_1, \bar{Y}_2, ..., \bar{Y}_t = f(X_0, X_1, X_2, ..., X_t) \tag{4}$$

The TCN is trained in the supervised manner for decreasing the loss function L. There is no data leakage from the future since the output length is equivalent to the input length. Dilated convolutions are present in the TCN to exponentially extend the receptive field, allowing for more past data to be included when forecasting. There is no data leakage from the future data. Equation 5 represents the dilated convolution function. The representation of the dilated convolution with $k=2$ and $d=1,2,4,8$ is shown in Fig. 3.

$$(X_d^* f)(s) = \sum_{j-1}^{i=1} f(i).X_{s-d.i} \tag{5}$$

where x is the 1-dimensional sequence; k is the filter size; $*$ is the convolution operator; and d is the dilation factor.

Normally high dilation factor and filter size are used to produce bigger receptive field to interpret more information from the past useful in long predictions. The fully connected 1D layer, and various residual blocks are present sequentially in the TCN for training the network. The dilated convolutions are performed using Eq. 5. For feature extraction, a generic residual block made up of layers of causal convolution is used for each layer. The output from the last residual block are used as the input to the model's next layer. And the output from the last residual block are fed to the fully connected (FC) layers present in the TCN model. These FC layers convert the high-dimensional features to lower dimensional features representing the features. The ReLU activation function is used in the residual blocks. Batch normalisation and spatial dropouts
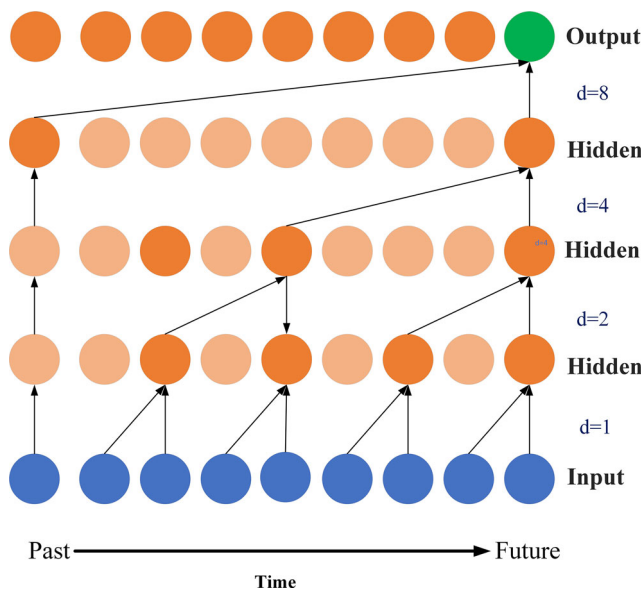
Fig. 3 Representation of dilated convolution with d=1,2,4,8 and filter k= 2

are used for the regularisation. The output vector of the FC layer is the OVec, in which each value representing the feature from the denoised wind speed data. Enhancing the spatial and temporal feature extraction property of the TCN, this study used attention mechanism (AM) to amplify the important features and suppress the irrelevant features from the vector OVec. The primary principle of the AM is to imitate the human brain's ability to comprehend things consciously. The implementation of the AM is represented in Eqs. 6 to 8.

$$Va_i = \text{attten}(OVec_i) \tag{6}$$

$$\mathcal{E}_i = \text{softmax}(Va_i) = \frac{\exp(Va_i)}{\sum_{i=q}^{Q} \exp(Va_i)} \tag{7}$$

$$\mathcal{X} = \mathcal{E}^T * OVec \tag{8}$$

The function atten() is used for obtaining the importance of the individual element in the vector OVec. The importance of the each element is denoted by the new vector Va_i, where i denotes the index value of the element in the vector OVec. Then, the Va_i is normalised using the softmax function obtaining the attention weight vector $\mathcal{E}$. Q indicates the vector Va length. The weight vector consists of the weights corresponding to the importance of the element in the vector OVec. More important feature will have more weight, and vice versa. Finally, the attention vector $\mathcal{X}$ is produced by multiplication of $\mathcal{E}$ and OVec. Thus, this final feature vector is fed to the Bi-LSTM that interprets the features for WSIP.

## Wind speed interval prediction by Bi-LSTM

The features extracted with the ATCN model from the denoised wind speed data are given as the input to the bidirectional LSTM (Bi-LSTM) model for interpreting the features and forecasting. The LSTM is a recurrent neural network which is developed to work with the long-term sequences by integrating gate mechanism and memory unit. However, the LSTM transmits the information in only one way, i.e. it interprets only the past information. Hence, a derivative of LSTM, i.e. BiLSTM is developed to consider both the future sequence and past sequence information, i.e. it transmits the information bidirectionally. It is divided into forward LSTM and backward LSTM to extract the features and it concatenates the hidden features to achieve the extraction bidirectionally. In one LSTM, the input sequence is fed and to the other LSTM, the reverse sequence is fed. The implementation process of the Bi-LSTM is shown in the equations below, where Eq. 9 represents the implementation for the forward LSTM, and Eq. 10 for the backward LSTM. The structure of BiLSTM is represented in Fig. 4.

$$\begin{aligned}
\overrightarrow{i_t} &= \sigma(\overrightarrow{W_{ih}}\overrightarrow{h_{t-1}} + \overrightarrow{W_{ix}}\overrightarrow{x_t} + \overrightarrow{W_{ic}}\overrightarrow{c_{t-1}} + \overrightarrow{b_i}) \\
\overrightarrow{o_t} &= \sigma(\overrightarrow{W_{oh}}\overrightarrow{h_{t-1}} + \overrightarrow{W_{ox}}\overrightarrow{x_t} + \overrightarrow{W_{oc}}\overrightarrow{c_{t-1}} + \overrightarrow{b_o}) \\
\overrightarrow{f_t} &= \sigma(\overrightarrow{W_{fh}}\overrightarrow{h_{t-1}} + \overrightarrow{W_{fx}}\overrightarrow{x_t} + \overrightarrow{W_{fc}}\overrightarrow{c_{t-1}} + \overrightarrow{b_f}) \\
\overrightarrow{\tilde{c}_t} &= \tanh(\overrightarrow{W_{ch}}\overrightarrow{h_{t-1}} + \overrightarrow{W_{cx}}\overrightarrow{x_t} + \overrightarrow{b_c}) \\
\overrightarrow{c_t} &= \overrightarrow{f_t} * \overrightarrow{c_{t-1}} + \overrightarrow{i_t} * \overrightarrow{\tilde{c}_t} \\
\overrightarrow{h_t} &= \overrightarrow{o_t} * \tanh(\overrightarrow{c_t})
\end{aligned} \tag{9}$$

$$\begin{aligned}
\overleftarrow{i_t} &= \sigma(\overleftarrow{W_{ih}}\overleftarrow{h_{t+1}} + \overleftarrow{W_{ix}}\overleftarrow{x_t} + \overleftarrow{W_{ic}}\overleftarrow{c_{t+1}} + \overleftarrow{b_i}) \\
\overleftarrow{o_t} &= \sigma(\overleftarrow{W_{oh}}\overleftarrow{h_{t+1}} + \overleftarrow{W_{ox}}\overleftarrow{x_t} + \overleftarrow{W_{oc}}\overleftarrow{c_{t+1}} + \overleftarrow{b_o}) \\
\overleftarrow{f_t} &= \sigma(\overleftarrow{W_{fh}}\overleftarrow{h_{t+1}} + \overleftarrow{W_{fx}}\overleftarrow{x_t} + \overleftarrow{W_{fc}}\overleftarrow{c_{t+1}} + \overleftarrow{b_f}) \\
\overleftarrow{\tilde{c}_t} &= \tanh(\overleftarrow{W_{ch}}\overleftarrow{h_{t+1}} + \overleftarrow{W_{cx}}\overleftarrow{x_t} + \overleftarrow{b_c}) \\
\overleftarrow{c_t} &= \overleftarrow{f_t} * \overleftarrow{c_{t+1}} + \overleftarrow{i_t} * \overleftarrow{\tilde{c}_t} \\
\overleftarrow{h_t} &= \overleftarrow{o_t} * \tanh(\overleftarrow{c_t})
\end{aligned} \tag{10}$$

where i_t and h_t indicate the input and hidden layer vector at the time t. Similarly, h_{t-1}, c_{t-1} represents the hidden layer and memory cell value at the time $t-1$. b_i, b_o, b_f, and b_c represent the bias of the input, output, forget gate, and memory cell respectively. W indicates the weight matrices for the different gates such as input, cell state, output, and forget gates. $\tanh$ and σ represent the activation functions. The $\rightarrow$ and $\leftarrow$ represent the forward and backward propagations in the LSTM network. The output of the ATCN is a feature set of optimal characteristics from the input denoised data. Thus, to interpret the features for WSIP, Bi-LSTM is adopted, because of its high performance in

Fig. 4 Representation of Bi-LSTM
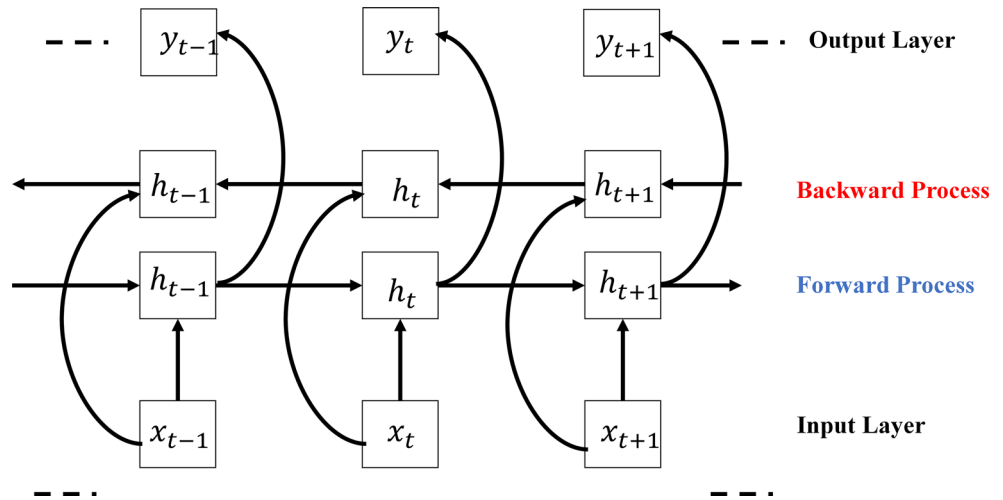
time series forecasting applications. In the Bi-LSTM, the hidden layer is formed by concatenating the forward and backward propagation of LSTM as shown in Eq. 11.

$$\beta_t = \overleftarrow{h_t} + \overrightarrow{h_t} \tag{11}$$

The features extracted by the ATCN are the inputs and and the outputs are corresponding lower (L_i) and upper (U_i) bound of the data point, i.e. $y_i = [L_i, U_i]$. However, in point or deterministic forecasting, the loss function of the network used is mean squared error (MSE). However, the MSE loss function is ineffective for predicting intervals. This paper proposes a novel loss function technique to solve this challenge, keeping the optimisation of PI as the criteria. Two evaluation criteria, prediction interval coverage probability (PICP) and coverage width (CW), are used to construct the custom loss function. The loss function is formulated in Eq. 12, where θ represents the parameters such as weights and bias of the network.

$$\text{Loss function} = \text{argmin} \begin{cases} \text{CW}(\theta) \\ \frac{1}{\text{PICP}(\theta)} \end{cases} \tag{12}$$

The loss function is developed in such a way that the proposed network can predict the optimal intervals while taking into account both PICP and CW. When the PICP is high and the CW is low, the PIs are regarded to be ideal. As a result, the loss function is developed using these two criteria. The weights are modified during the network's training to achieve the best result. A case study is performed for evaluating the proposed approach.

Experimental results

A wind speed dataset of 2013 year from a wind farm located in Garden City, Manhattan, is used in this study (NREL 2022). A map of the investigated area is shown in Fig. 5. The wind speed is obtained in 5-min ahead samples. However, to evaluate the novelty of the proposed approach for WSIP, wind speed is re-sampled into 10-min and 30-min ahead samples. The characteristics of the wind farm are shown in Table 1. Literature demonstrated that the performance of the forecasting models decreases with the increase in the ahead. Hence, the proposed approach is tested with three different time interval data. The original input data are divided into testing, training data, where training data consists of the first 80% data, and the rest data is 20% data is used as test data.

Evaluation criteria for optimal PI

Evaluation criteria such as mean squared error (MSE), root mean squared error (RMSE), mean absolute error (MAE), and mean absolute percentage error (MAPE) are used in the point forecasting. However, in the prediction interval forecasting applications, the performance cannot be evaluated using conventional metrics. To deal with this, evaluation indices such as prediction interval coverage probability (PICP), PI normalised root-mean-square width (PINRW), and mean prediction interval width (MPIW) are used to evaluate the PI quality (Tang et al. 2019), (Khosravi et al. 2011a). PICP and PINRW represent the reliability and precision of PI. MPIW indicates the mean width of the PIs. The evaluation indices are formulated in Eqs. 13 to 15.

$$\text{PICP} = \frac{1}{n} \sum_{j=1}^{n} C_j, \ C_j = \begin{cases} 1, y_j \in [L_j, U_j] \\ 0, y_j \notin [L_j, U_j] \end{cases} \tag{13}$$

$$\text{PINRW} = \frac{1}{R} \sqrt{\frac{1}{n} \sum_{j=1}^{n} (U_j - L_j)^2} \tag{14}$$
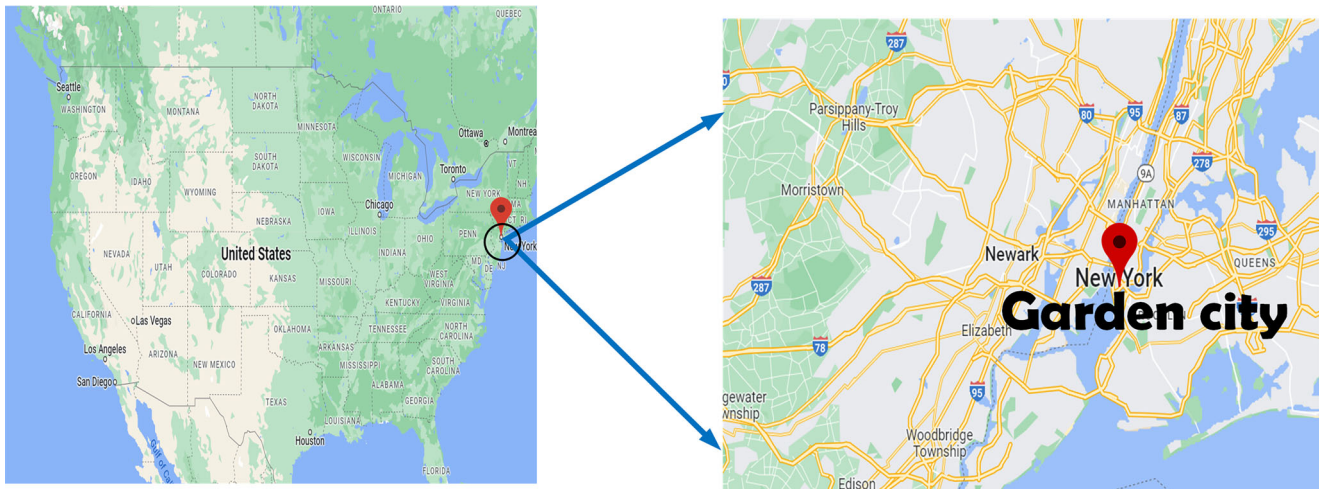
Fig. 5 Map of the Garden city, Manhattan, USA

$$\text{MPIW} = \frac{1}{n} \sum_{j=1}^{n} (L_j - U_j) \tag{15}$$

where n is the number of samples and upper and lower bounds of the jth prediction interval are denoted by U_i and L_i, respectively. R represents the range of the wind speed data. C_i indicates whether the actual wind speed data at jth point lies in the predicted interval $[L_j, U_j]$. An optimal PI should have higher PICP and lower PINRW (narrow PI). Because of their inverse relationship, PICP and PINRW cannot provide a perfect evaluation of PI quality. Thus, a new criteria, coverage width criterion (CWC), is developed by combining the PICP and PINRW as shown in Eq. 16.

$$\text{CWC} = (1 + \eta_1 \text{PINRW})(1 + \gamma(\text{PICP})e^{-\eta_2(\text{PICP}-\mu)}) \tag{16}$$

$$\gamma(\text{PICP}) = \begin{cases} 0, \text{PICP} \geq \mu \\ 1, \text{PICP} < \mu \end{cases} \tag{17}$$

where η and μ are the crucial hyperparameters controls the CWC index. η is used to magnify the variation of the PICP with respect to μ, which is the confidence level of the interval. Therefore, this hybrid index is used to evaluate the quality of the PIs. Smaller the CWC, better the quality.

Table 1 Characteristics of Garden city, Manhattan wind farm

Time	Median	SD	Mean	Skewness	Max	Min	Range
5 min	7.63	3.62	7.66	0.152	17.20	0.25	16.99
10 min	7.62	3.70	7.72	0.252	17.81	0.03	17.77
30 min	7.69	3.79	7.85	0.261	19.44	0.03	19.41

Results and discussions

To evaluate the performance of the proposed approach, in this paper, two categories of approaches are compared. The benchmark methods, such as ATCN-BiLSTM, MLP, LSTM, and CNN, fall into the first group of models. All these models from the first category are hybridised using the ICEEM-DAN algorithm for denoising. As a result, ICEEMDAN-CNN, ICEEMDAN-LSTM, and ICEEMDAN-MLP fall into the second hybrid category of models. The proposed approach as well as the reference models are evaluated for WSIP, and the results are given in Tables 2, 3, and 4. Figure 6 illustrates the WSIP result by the proposed approach. Figure 7 illustrates the comparison of WSIP by all the hybrid approaches.

From Tables 2, 3, and 4 and Figs. 6 and 7, the hybrid model ATCN-BiLSTM achieved the best results compared to the first category models like CNN, MLP, and LSTM. In the 5-min ahead WSIP, ATCN-BiLSTM has an MPIW

Table 2 Comparison of 5-min ahead WSIP statistical indices between benchmark approaches and proposed hybrid framework

Model	MPIW	PINRW	PICP	CWC
ATCN-BiLSTM	1.9298	0.1355	0.9072	0.1398
CNN	1.9423	0.1427	0.8642	0.1535
LSTM	2.3403	0.1459	0.8235	0.4054
MLP	2.5208	0.1466	0.8176	0.4296
ICEEMDAN-CNN	1.8532	0.1335	0.9232	0.1366
ICEEMDAN-LSTM	1.9309	0.1365	0.8769	0.1423
ICEEMDAN-MLP	2.0204	0.1447	0.8444	0.3518
Proposed approach	1.3885	0.1021	0.9793	0.0880

Table 3 Comparison of 10-min ahead WSIP statistical indices between benchmark approaches and proposed hybrid framework

Model	MPIW	PINRW	PICP	CWC
ATCN-BiLSTM	2.3415	0.1437	0.8715	0.4388
CNN	2.7256	0.1533	0.8340	0.8954
LSTM	2.7074	0.1521	0.8478	0.8853
MLP	2.7769	0.1556	0.8280	1.0144
ICEEMDAN-CNN	2.6565	0.1496	0.8528	0.7005
ICEEMDAN-LSTM	2.5327	0.1482	0.8669	0.6706
ICEEMDAN-MLP	2.6311	0.1488	0.8635	0.6833
Proposed approach	1.8904	0.1107	0.9555	0.2306

value of 1.9298, which is the least in the indices of first category models. Similarly, in PINRW, PICP, and CWC, this hybrid model achieved the best results compared to benchmark approaches. The CNN model occupies the second spot with a PICP of 0.8642. The MLP occupies the last spot with relatively higher CWC and PINRW indices. ATCN-BiLSTM approach, on the other hand, continues to dominate throughout the 10-min and 30-min ahead WSIP. LSTM and CNN approaches occupied the second spot in 10-min and 30-min WSIP, respectively. Therefore, the hybrid ATCN-BiLSTM's dominant performance is confined to the first category models in the 5-min, 10-min, and 30-min WSIP. However, in the 10-min and 30-min WSIP, the hybrid ATCN-BiLSTM is having better performance than the comparative hybrid models using ICEEMDAN. Category 2 models with data denoising technique achieved better results than the category 1 models without denoising technique. In the second category of models, the proposed approach achieved the best results. For example, considering the 5-min ahead WSIP, the proposed approach observed a very high PICP, i.e. 0.9793, which is around 97%. The proposed approach achieved a lower CWC value of 0.0880, which is 36% lesser than the second-best

Table 4 Comparison of 30-min ahead WSIP statistical indices between benchmark approaches and proposed hybrid framework

Model	MPIW	PINRW	PICP	CWC
ATCN-BiLSTM	2.7503	0.1397	0.8229	0.7152
CNN	2.7748	0.1419	0.7898	0.9722
LSTM	2.7709	0.1416	0.7778	1.1797
MLP	2.7984	0.1441	0.6874	1.2411
ICEEMDAN-CNN	2.7511	0.1406	0.7994	0.9386
ICEEMDAN-LSTM	2.7416	0.1346	0.8568	0.5416
ICEEMDAN-MLP	2.7630	0.1425	0.7615	1.0331
Proposed approach	2.2097	0.1008	0.9141	0.4502

value. Similarly, in MPIW and PINRW, the proposed approach observed best indices of 1.3885 and 0.1021, respectively. 25% and 24% improvement is achieved by the proposed approach in the MPIW and PINRW over the second best model. ICEEMDAN-based CNN observed the second-best results among the two categories of models. ATCN-BiLSTM model at the third position followed them. In 5-min interval, the improvement percentage of proposed approach over ATCN-BiLSTM model is 28%, 25%, and 37% in terms of MPIW, PINRW, and CWC. Comparing the 10-min ahead WSIP indices, the proposed approach achieved the best values of indices. For example, the CWC is 0.2306, which is 47% lesser compared to the second-best approach available. The hybrid ATCN-BiLSTM approach occupied the second spot, and MLP occupied the last rank in the 10-min ahead WSIP. The improvement of 19% and 23% in the MPIW, and PINRW is achieved by the proposed approach over the second best model in the 10-min ahead WSIP. However, as the time ahead increased from 5-min to 10-min ahead, the performance of all the models declined. On the other hand, the proposed approach observed betterment in the performance from 5-min to 10-min ahead WSIP. Coming to the 30-min ahead WSIP, the proposed approach maintained stable performance with a CWC of 0.4502. The proposed approach achieved the highest improvement percentage in terms of the evaluation indices in the 30-min ahead WSIP. For instance, the improvement percentages are 19%, 25%, and 17% in the MPIW, PINRW, and CWC respectively. MLP model ranked last in the indices of 30-min ahead WSIP among all the comparative approaches. In 30-min interval, the improvement percentage of proposed approach over ATCN-BiLSTM model is 20%, 28%, and 37% in terms of MPIW, PINRW, and CWC. The percentage improvement of proposed model over all other individual and hybrid models is shown in Table 5.

Based on the evaluation indices, it is evident that the hybridisation of category 1 models with ICEEMDAN resulted in the enhanced performance in WSIP. This is due to the denoising of highly non-linear signals. From Fig. 6, the actual wind speed lies precisely in the prediction interval obtained by the proposed approach, respectively. However, there is a slight decline in the quality of the prediction interval from 5-min to 30-min ahead. But, from Fig. 7, it is very apparent that the proposed model's prediction interval quality is quite optimal compared to the other reference models, which are having an unstable performance for 5-, 10-, and 30-min ahead WSIP.

The proposed approach's performance is also evaluated in terms of testing and training time efficiency with the comparative CNN and LSTM models. The proposed framework's training time is 820.52 s, while the testing
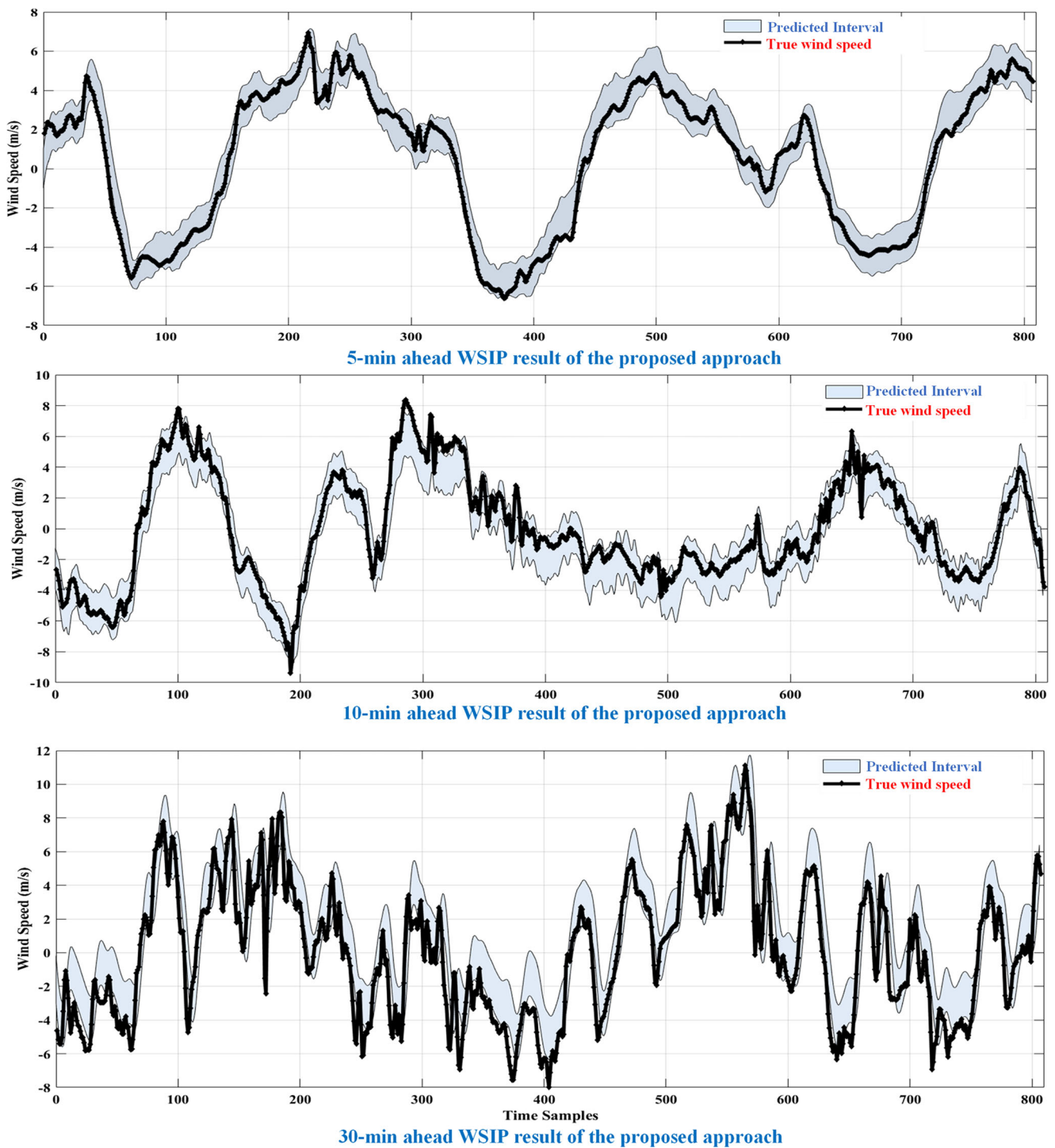
Fig. 6 WSIP result of proposed approach using Garden city wind farm data

time is 0.041 s. When compared to the proposed approach, the CNN model requires 778.21 s for training and 0.033 s for testing. Similarly, the LSTM model took 801.23 s to train and 0.039 s to test. However, since the training is done offline and just once, the quantity of training time achieved is acceptable. The testing time is also significantly fast, with predictions taking 0.3 ms. Furthermore, when

the evaluation indices are taken into account, the proposed approach achieves the best WSIP performance indices as already presented through Tables 2, 3, and 4. The proposed approach predicts high-quality intervals because it uses effective feature extraction, denoising, and feature interpretation to forecast intervals. As a result, the proposed approach meets the criterion for real-time WSIP.
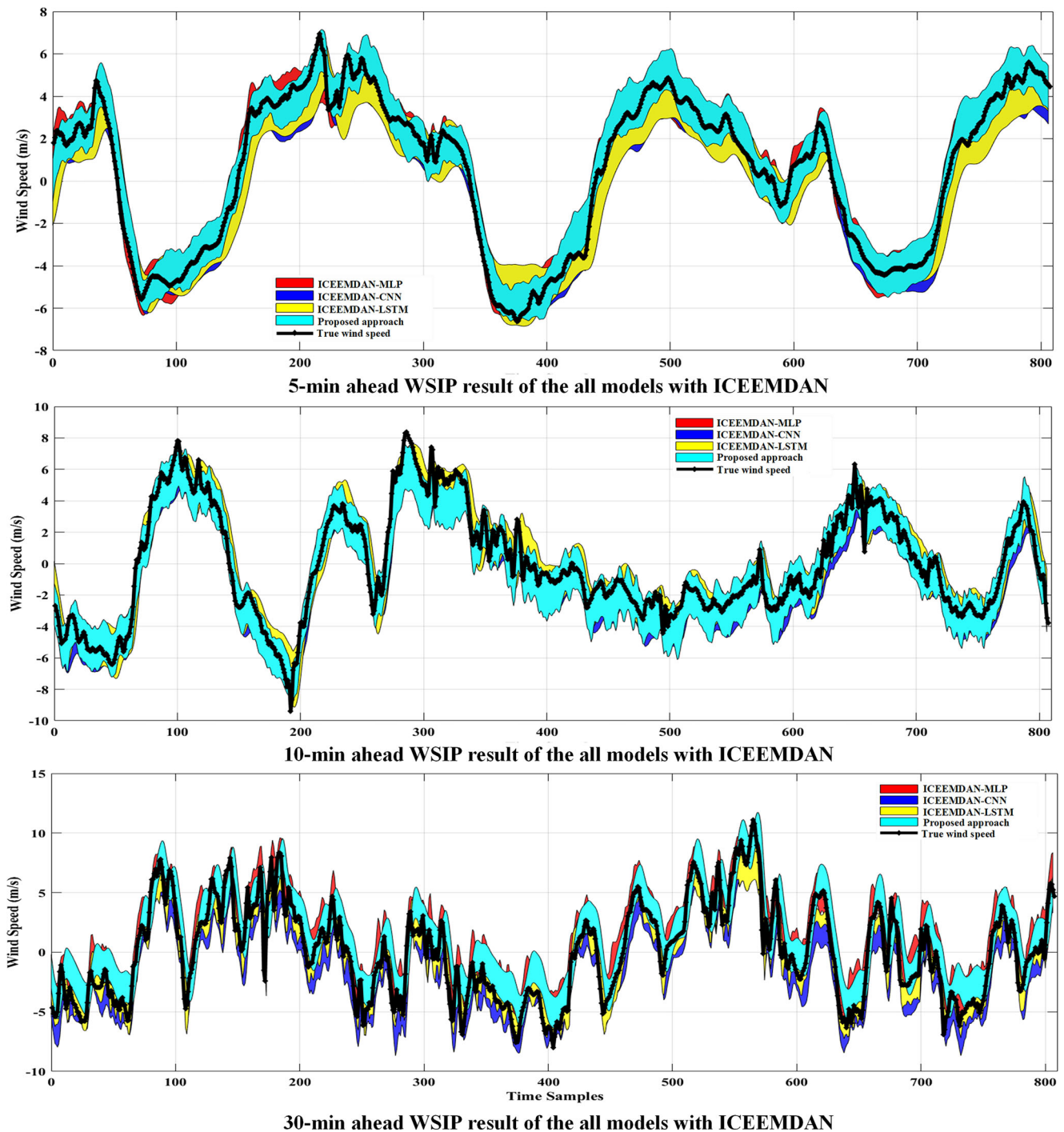
Fig. 7 Illustration of WSIP result of all models with ICEEMDAN algorithm for Garden city wind farm

The experimental findings from the WSIP results of all available methods reveal the following critical points.

(1) The approaches such as CNN, LSTM, and MLP failed to maintain the consistency for the 5-, 10-, and 30-min ahead WSIP.

(2) But, it is also observed that the proposed approach is consistently leading in all the three forecasting results with the best indices.

(3) The evaluation indices clearly show the efficiency of hybridisation using the ICEEMDAN method. Category 2 models are outperforming category 1 models in terms of prediction interval quality.

Table 5 Percentage improvement of proposed approach over all other models

Model	5-min interval			10-min interval			30-min interval		
	MPIW	PINRW	CWC	MPIW	PINRW	CWC	MPIW	PINRW	CWC
Proposed vs ATCN-BiLSTM	28%	25%	37%	19%	23%	47%	20%	28%	37%
Proposed vs CNN	29%	28%	43%	31%	28%	74%	20%	29%	54%
Proposed vs LSTM	41%	30%	78%	30%	27%	74%	20%	29%	62%
Proposed vs MLP	45%	30%	80%	32%	29%	77%	21%	30%	64%
Proposed vs ICEEMDAN-CNN	25%	24%	36%	29%	26%	67%	20%	28%	52%
Proposed vs ICEEMDAN-LSTM	28%	25%	38%	25%	25%	66%	19%	25%	17%
Proposed vS ICEEMDAN-MLP	31%	29%	75%	28%	26%	66%	20%	29%	56%

(4) The quality of the prediction intervals of the benchmark approaches hybridised with ICEEMDAN, on the other hand, is not adequate.

(5) The proposed approach achieved an increase in the improvement percentage from 5-min WSIP to 30-min ahead WSIP.

Conclusions

The primary prerequisite for wind energy grid management is accurate wind speed predictions. As noted in the literature, point forecasting fails to account for uncertainties and does not produce needed information for power system operations. As a result, in this paper, a novel approach consisting of ICEEMDAN and TCN with attention mechanism and Bi-LSTM is proposed for improved accuracy of WSIP. Effective elimination of auxiliary noise, feature extraction plays a crucial role in the forecasting performance. In the proposed approach, ICEEMDAN is used for decomposing the signal to eliminate the auxiliary noise, ATCN is used to extract important features from the decomposed wind speed, and Bi-LSTM forecasts the accurate prediction intervals. Addressing the decline of performance of the models with an increase in ahead values, the proposed approach is tested using 5-min, 10-min, and 30-min ahead WSIP. A comparative analysis is performed using two categories of models to evaluate the proposed approach performance. The evaluation indices from the experiment indicate that the proposed approach's performance is consistent for 5-min, 10-min, and 30-min WSIP. The feasibility and performance of the proposed approach are investigated and confirmed during the experiments. The experimental results indicate the dominating performance of the proposed approach for 5-min, 10-min, and 30-min ahead WSIP. The proposed approach offers an improvement of 36%, 47%, and 17% for three time intervals WSIP.

The future work would demonstrate the application of the optimisation techniques for enhancing the performance of the proposed approach.

Author contribution Bala Saibabu Bommidi: conceptualisation, methodology, software. Vishalteja Kosana: visualisation, methodology, performed the experiments. Kiran Teeparthi: writing — review and editing, supervision, validation. Santhosh Madasthu: supervision, validation, investigation.

Declarations

Ethics approval and consent to participate Not applicable.

Consent for publication Not applicable.

Competing interests The authors declare no competing interests.

References

Ahmadi M, Sharifi A, Dorosti S, Ghoushchi S, Ghanbari N (2020) Investigation of effective climatology parameters on COVID-19 outbreak in Iran, vol 729

Ahmadi M, Soofiabadi M, Nikpour M, Naderi H, Abdullah L, Arandian B (2022) Developing a deep neural network with fuzzy wavelets and integrating an inline PSO to predict energy consumption patterns in urban buildings. Mathematics 10(8):1270

Artin J, Valizadeh A, Ahmadi M, Kumar SA, Sharifi A (2021) Presentation of a novel method for prediction of traffic with climate condition based on ensemble learning of neural architecture search (NAS) and linear regression. Complexity, 2021

Awosusi AA, Xulu NG, Ahmadi M, Rjoub H, Altuntaş M, Uhunamure SE, Akadiri SS, Kirikkaleli D (2022) The sustainable environment in Uruguay: the roles of financial development, natural resources, and trade globalization. Front Environ Sci 10:875577

Barenya BH, Deepak G, Narayanan N (2022) Wavelet kernel least square twin support vector regression for wind speed prediction.

Environmental Science and Pollution Research. https://doi.org/10.1007/s11356-022-18655-8

Bouhalais ML, Nouioua M (2021) The analysis of tool vibration signals by spectral kurtosis and ICEEMDAN modes energy for insert wear monitoring in turning operation. The International Journal of Advanced Manufacturing Technology, pp 1–13

Chen YZY (2022) Application of hybrid model based on CEEMDAN, SVD, PSO to wind energy prediction. Environ Sci Pollut Res 29:22661–22674

Council GWE (2020) Global wind report 2019 released on april 2020

Cui Y, Chenchen H, Cui Y (2020) A novel compound wind speed forecasting model based on the back propagation neural network optimized by bat algorithm. Environ Sci Poll Res 27:7353–7365

Gupta D, Natarajan N, Berlin M (2021) Short-term wind speed prediction using hybrid machine learning techniques. Environmental Science and Pollution Research

Zhang G, Li Z, Zhang K, Zhang L, Hua X, Wang Y (2019) Multi-objective interval prediction of wind power based on conditional copula function. Journal of Modern Power Systems and Clean Energy 7:802–812

Ghoushchi S, Sharifi A, Ahmadi M, Maghami MR (2017) Statistical study of seasonal storage solar system usage in Iran. J Sol Energy Res 2:39–44

Habeşoğlu O, Samour A, Tursoy T, Ahmadi M, Abdullah L, Othman M (2022) A study of environmental degradation in Turkey and its relationship to oil prices and financial strategies: novel findings in context of energy transition. Frontiers in Environmental Science, 220

He Y, Zhang W (2020) Probability density forecasting of wind power based on multi-core parallel quantile regression neural network. Knowl-Based Syst 209:106431

Heydari A, Memarzadeh G, Astiaso Garcia D, Keynia F, Santoli L (2021) Interval prediction algorithm and optimal scenario making model for wind power producers bidding strategy. Optim Eng, 22

Hu M, Hu Z, Yue J, Zhang M, Hu M (2017) A novel multi-objective optimal approach for wind power interval prediction. Energies 10(4):419

Khodayar M, Wang J, Manthouri M (2018) Interval deep generative neural network for wind speed forecasting. IEEE Trans Smart Grid 10(4):3974–3989

Khosravi A, Mazloumi E, Nahavandi S, Creighton D, Van Lint J (2011a) Prediction intervals to account for uncertainties in travel time prediction. IEEE Trans Intell Transp Syst 12(2):537–547

Khosravi A, Nahavandi S, Creighton D, Atiya AF (2011b) Lower upper bound estimation method for construction of neural network-based prediction intervals. IEEE Trans Neural Netw 22(3):337–346

Naik J, Dash PK, Dhar S (2019) A multi-objective wind speed and wind power prediction interval forecasting using variational modes decomposition based multi-kernel robust ridge regression. Renew Energy 136:701–731

Narayanan Natarajan MVSR (2021) Evaluation of suitability of wind speed probability distribution models: a case study from Tamil Nadu, India. Environmental Science and Pollution Research

Niu D, Sun L, Yu M, Wang K (2022) Point and interval forecasting of ultra-short-term wind power based on a data-driven method and hybrid deep learning model. Energy, 124384

Nix DA, Weigend AS (1994) Estimating the mean and variance of the target probability distribution. In: Proceedings of 1994 ieee international conference on neural networks (ICN 94), vol 1. IEEE, pp 55–60

NREL (2022) N. R. E. L. www.nrel.gov accessed on 6 January

Pinson P, Kariniotakis G (2010) Conditional prediction intervals of wind power generation. IEEE Trans Power Syst 25(4):1845–1856

Pullanagari RR, Kereszturi G, Yule I, Irwin M (2018) Determining uncertainty prediction map of copper concentration in pasture from hyperspectral data using qunatile regression forest. In: IGARSS 2018-2018 IEEE International Geoscience and Remote Sensing Symposium, pp 3809–3811. IEEE

Quan H, Srinivasan D, Khosravi A (2014) Short-term load and wind power forecasting using neural network-based prediction intervals. IEEE Trans Neural Netw Learn Syst 25(2):303–315

Shrivastava NA, Lohia K, Panigrahi BK (2016) A multiobjective framework for wind speed prediction interval forecasts. Renew Energy 87:903–910

Tang G, Xue X, Saeed A, Hu X (2019) Short-term wind speed interval prediction based on ensemble GRU model. IEEE Transactions on Sustainable Energy, 1–1

Sun W, Xiaoxuan Wang BT (2022) Multi-step wind speed forecasting based on a hybrid decomposition technique and an improved back-propagation neural network. Environmental Science and Pollution Research

Xue H, Jia Y, Wen P, Farkoush SG (2020) Using of improved models of Gaussian processes in order to regional wind power forecasting. J Clean Prod 262:121391

Zhang Y, Pan G (2020) A hybrid prediction model for forecasting wind energy resources. Environ Sci Pollut Res 27:19428–19446

Zhang Y, Wang S (2022) An innovative forecasting model to predict wind energy. Environmental Science and Pollution Research

Yuan X, Chen C, Jiang M, Yuan Y (2019) Prediction interval of wind power using parameter optimized beta distribution based LSTM model. Appl Soft Comput 82:105550

Zhang D, Chen Z, Zhou Y (2022) Wind power interval prediction based on improved whale optimization algorithm and fast learning network. Journal of Electrical Engineering & Technology, 17

Zhang Y, Pan G, Zhao Y, Li Q, Wang F (2020) Short-term wind speed interval prediction based on artificial intelligence methods and error probability distribution. Energy Conversion and Manag 224:113346

Zhao C, Wan C, Song Y (2020) An adaptive bilevel programming model for nonparametric prediction intervals of wind power generation. IEEE Trans Power Syst 35(1):424–439

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.