Hindawi
Journal of Electrical and Computer Engineering
Volume 2022, Article ID 8498021, 12 pages
https://doi.org/10.1155/2022/8498021

Research Article

Pearson Autocovariance Distinct Patterns and Attention-Based Deep Learning for Wind Power Prediction

W. G. Jency and J. E. Judith

Department of Computer Science and Engineering, Noorul Islam Centre for Higher Education, Thuckalay, Kumaracoil, Tamil Nadu 629180, India

Correspondence should be addressed to W. G. Jency; jencywg@gmail.com

Received 20 February 2022; Revised 4 April 2022; Accepted 6 April 2022; Published 22 April 2022

Academic Editor: Jit S. Mandeep

Swift development in wind power and extension of wind generation necessitates significant research in numerous fields. Due to this, wind power is weather dependent; it is fluctuating and is sporadic over numerous time periods. Hence, timely wind power prediction is perceived as an extensive contribution to well-grounded wind power prediction with complex patterns. In addition, a number of wind power prediction methods have been developed. For proper planning and operation of power systems with complicated patterns, wind power prediction in an accurate and timely manner is essential. This paper presents a wind power prediction method with feature selection and prediction called, Pearson Autocovariance Distinct Patterns and Attention-based Deep Learning (PACDP-ADL). In the deep learning environment, feature selection plays a crucial aspect and a prediction task. A Pearson Autocovariance Feature Selection model is used for identifying necessary features for wind power prediction and reduces the complexity of the model. Next, an Attention-based Long Short-Term Memory Wind Prediction algorithm is employed to retain required patterns and forget irrelevant patterns to acquire more satisfactory prediction precision. The proposed PACDP-ADL method is validated by utilizing the wind power data with various performance metrics such as wind power accuracy, wind power time, and true positive rate compared with the state-of-the-art method.

1. Introduction

In recent years, wind being a renewable, sustainable, and environmentally friendly energy resource has received awareness from a heightening number of countries owing to its low-cost operation and maintenance. As a result, it results in swiftly minimizing the generation cost concerning electricity.

With exceptional recognition of patterns in an automatic fashion and mapping of big data in a nonlinear manner, the deep learning technique is said to be utilized in wind power prediction. However, noteworthy truths are that in situ estimated wind data are moderately high priced and beyond reach. Therefore, an association between steps is discarded in most multistep wind power predictions. Encoder-Decoder Long Short-Term Memory (EDLSTM) was proposed in [1] to wind power prediction by comprehensively encapsulating and presenting an integrated time series model based on physics and data wind data augmentation method.

By integrating physics with data-oriented wind data, sequential input and sequential output for wind power prediction were attained with minimum error. However, with the complex pattern discovery of the sequential input, accurate and timely wind power prediction is said to be a major issue that has to be concentrated on. To address this aspect, first, relevant features addressing overfitting are performed by introducing the Pearson Autocovariance Feature Selection model. With this feature selection model, only highly relevant features present in the complex patterns are selected, therefore corroborating the objective.

A hybrid model for wind power prediction in a short-term fashion employing variational mode decomposition (VMD), the K-means clustering, and long short-term memory (LSTM) network called VMD-K-means-LSTM was proposed in [2]. Initially, the VMD algorithm split the wind power data into finite stationary intrinsic mode functions (IMFs). Each stationary IMF was then unfolded by means of a trajectory matrix with which K-means were applied to split the given trajectory matrix into a clustering sample set.

The distances between clustering centers were estimated, and finally, optimal LSTM selected all sublayers to acquire the final prediction result, therefore minimizing the error rate. Despite the minimum error rate, the precision or the accurate wind power prediction was not made. To ensure this, an attention mechanism has been introduced in the deep learning model that in turn retains required patterns in the cell or memory state while removing the irrelevant patterns from the cell state and moving them into the forget gate, therefore ensuring maximum sensitivity at minimum time.

Despite improvement observed in minimizing error, the accuracy and time with which the prediction was made involving complex patterns were less focused. To address this issue, a Pearson Autocovariance Distinct Patterns and Attention-based Deep Learning (PACDP-ADL) method is designed that predicts the wind power with maximum accuracy and minimum time even in the presence of complex patterns.

The main contributions of this paper are as follows.

(1) We propose a novel Pearson Autocovariance Feature Selection model that uses the autocovariance function to address overfitting and then with the autocorrelation function identifies the minimum subset of the significant informative features, therefore, paving way for normalization, which guarantees accurate and relevant feature selection in a computationally efficient manner.

(2) We propose a model of wind power prediction involving complex patterns, using Attention-oriented Long Short-Term Memory-based Wind Power Prediction that with the consideration of complex patterns in the input organizes an efficient mapping by means of attention mechanism, therefore increasing true positive rate significantly.

(3) We propose a method Pearson Autocovariance Distinct Patterns and Attention-based Deep Learning (PACDP-ADL) for wind power prediction to ensure the quality of service in terms of accuracy and time and accelerate the solving speed involving complex patterns by improving the true positive rate.

(4) Conduct extensive simulation and performance analysis of the proposed method. In addition, we investigate how the three wind power prediction methods have an influence on the performance in terms of the wind power prediction accuracy, wind power prediction time, and true positive rate.

The rest of this paper is organized as follows. Section 2 discusses related works. Section 3 introduces the proposed Pearson Autocovariance Distinct Patterns and Attention-based Deep Learning (PACDP-ADL) for wind power prediction. Section 4 presents an experimental evaluation of the proposed method. Section 5 discusses the comparative analysis with two state-of-the-art methods. Finally, the paper is concluded in section 6.

1.1. Motivation. Wind power prediction has an essential function in maintaining the stability of the complete power grid. The forecast of wind power corresponds to the estimation of the predictable production of one or more wind turbines. Wind power prediction of accuracy is very higher and the time of wind power prediction is lesser. In wind power, rapid development enhances wind generation that needs serious research within different fields. The various parameters are optimized by applying the cuckoo search optimization algorithm (CSO) for ensuring forecasting accuracy. The renewable energy source is increased and employs the various fields. But, the uncertainty in the power system is increased considerably to the occurrence of wind power variability. Wind power transmission is an important element of smart grid energy management systems. Motivated by this fact, the novel Pearson Autocovariance Distinct Patterns is developed using wind power prediction.

2. Related Works

In the past decades in developed countries, wind energy has swiftly progressed in several countries as a predominantly renewable energy source. However, the sporadic characteristic of wind energy brings out uncertainty in wind power generation that may result in several issues connected to an imbalance in power, frequency continuity of power networks, issues related to power quality, and difficulty in the power system planning and scheduling. Owing to these facts, high-precision wind power prediction has gained awareness globally.

A hybrid forecasting model for wind power forecasting with the purpose of enhancing the prediction performance called long short-term memory network-enhanced forget-gate network (LSTM-EFG) was proposed in [3]. In this method, appropriate parameters are optimized by utilizing a cuckoo search optimization algorithm (CSO), therefore ensuring forecasting accuracy.

A survey of wind energy prediction using machine language was investigated in [4]. However, the presence of extreme weather conditions and disturbances in waves results in changes in pitch angle and compromises accurate wind power prediction. To address this issue, a hybrid system integrating fuzzy logic and deep learning was proposed in [5] that with the aid of an intelligent controller, therefore, minimizes training time.

In recent years, wind, a renewable energy source, has been increased and found to be of use in various fields. However, the uncertainty in the power system has also increased significantly owing to the occurrence of wind power variability, which is why wind power transmission has become an essential element of smart grid energy management systems.

In [6], a novel hybrid wind power prediction method was designed. In this method first, cascaded deep learning was applied with the purpose of extracting implicit meteorological and temporal characteristics by means of two-layer mode decomposition. Next, empirical mode decomposition was used with the purpose of splitting the original time series into intrinsic mode functions (IMFs). Finally, variational mode decomposition (VMD) was applied to further split IMF1 sublayers into numerous subseries, therefore ensuring accuracy in both short-term and long-term prediction.

With the uncertain nature, accuracy gets to be compromised. To address this aspect, tree-based learning algorithms were applied in [7], therefore, enhancing forecasting accuracy. However, with the high penetrative nature, the training time also provides unsatisfactory performance. To address this issue, the Convolution-based Spatial-Temporal Wind Power Predictor (CSTWPP) was developed in [8]. The high nonlinearity training process was said to be accelerated.

In the recent few years, a swift development has been seen in wind power. Also, large-scale integration of wind power is found to be demanding in terms of both power grid functioning and governance. Unanticipated deviations of wind generation may result in the increased costs involved in operating, increasing requirements concerning the reserves, and finally constituting possible risks to system reliability. Hence, wind speed prediction is an elementary chore done in all wind farms as a proportion of their operation management.

An elaborate review of the prevailing forecasting methods and moreover their performance evaluation was analyzed in [9]. Also, the materials and methods utilized for enhancing prediction accuracy and mechanisms to address forecasting issues were also explored.

In [10], an integrated method was designed on the basis of an extreme learning machine (ELM) and an error correction with the purpose of predicting wind power on a short-term time scale. With the integrated mechanism, an error was reduced significantly. However, fast training was not focused. To address this issue, a hybrid coral reefs optimization method employing an extreme learning machine was proposed in [11].

To ensure both reliability and security, wind speed forecasting has to be done in an accurate manner. However, it is still a demanding chore owing to the laboriously undetermined and volatile characteristics of wind speed. In consequence, a novel deep learning-based model combining Discrete Wavelet Packet Transform (DWPT) and Bidirectional Long Short-Term Memory (BLSTM) was deployed in [12] to accurately acquire deep temporal features and learn the time-varying correlation of wind speed time series. Therefore, both forecasting performances were increased in addition to reducing learning complexity.

Despite improvement in learning complexity, the error factor was not focused on the above-discussed methods. To concentrate on this issue, a wind power uncertainty forecasting model was proposed in [13]. With the unstable and stochastic nature, still, wind power prediction is facing significant problems. In [14], optimization of wind power features using particle swarm optimization model and classification employing support vector machine was designed that in turn resulted in improvement of prediction accuracy.

Owing to the reason that the errors involved during the forecasting cannot be accurately modeled utilizing distribution probability functions, a powerful method called Lower Upper Bound Estimation (LUBE) was proposed in [15] to formulate prediction intervals. Moreover, the LUBE method utilized a novel framework on the basis of integration of PIs to mitigate the execution uncertainty of neural networks (NNs) employed in the LUBE method.

In addition, a novel fuzzy-based cost function was also designed with the objective of possessing more flexibility in fine-tuning NN metrics for the construction of PIs. However, precision was not ensured. To address this factor, Empirical Model Decomposition (EMD) was integrated with General Regression Neural Network (GRNN) [16] to achieve precision.

A hybrid Cuckoo Search Ensemble Empirical Model Decomposition and Feedforward Neural Network (CSEEMD-FNN) was proposed in [17] with the purpose of ensuring multistep ahead prediction of wind speed. Here, EEMD was employed as a measure for data-cleaning that in turn eliminated the high-frequency noise embedded in wind speed sample data. Moreover, CS optimization was utilized to acquire the best parameters in the FNN, therefore, ensuring accuracy.

An approach was designed in [18] on the basis of Sparse Bayesian Learning (SBL) and numerical weather prediction (NWP) for probabilistic wind power forecasting with minimum error and maximum forecasting accuracy in the horizon of 1–24 hours. Artificial neural networks were applied in [19] to predict wind speed for Tehran, Iran, with reasonable accuracy. Multifrequency components were included in [20] based on variational mode decomposition (VMD) for accurate wind power prediction.

Reliable and accurate planning and scheduling of wind farms and power grids were introduced in [21] to ensure that sustainable use of wind energy can be better achieved with the use of precise and accurate prediction models. However, due to the highly chaotic, intermittent, and stochastic behavior of wind, which means a high level of difficulty when predicting wind speed and, consequently, wind power, the evolution of models capable of narrating data of such complexity is an emerging area of research.

A new data-driven model that combined the variational mode decomposition (VMD) and the prediction models is proposed in [22] for daily streamflow forecasting. The prediction models include the autoregressive moving average (ARMA), the gradient boosting regression tree (GBRT), the support vector regression (SVR), and the backpropagation neural network (BPNN).

A comprehensive review of decomposition-based wind forecasting methods in order was designed in [23] to explore their effectiveness. Decomposition-based hybrid forecasting models are classified into different groups based on the decomposition methods, such as wavelet, empirical mode decomposition, seasonally adjust methods, variational mode decomposition, intrinsic time-scale decomposition, and Bernaola Galvan algorithm.

3. Methodology

In this section, the proposed method Pearson Autocovariance Distinct Patterns and Attention-based Deep Learning

(PACDP-ADL) method for wind power prediction is elaborated on in detail. The modeling process of PACDP-ADL involves feature selection and prediction as depicted in Figure 1. Figure 1 shows the block diagram of the Pearson Autocovariance Distinct Patterns and Attention-based Deep Learning (PACDP-ADL) method for wind power prediction.

As shown in the figure, first, autocovariance theory along with the Pearson Correlation Coefficient is utilized to analyze the relationship between the input-output mapping feature variables in the sample database and realize the selection of optimal features with the presence of complex patterns. Second, the selected relevant features for distinct patterns are utilized as input for wind power prediction.

An attention mechanism based on long short-term memory is then developed, and the predictive patterns for wind power are determined. Finally, the dataset, consisting of LV active power (LVAP), wind speed (WS), theoretical power curve (TPC), and wind direction (WD) for 10-minute time intervals from a wind turbine's SCADA system that is working and generating power in Turkey, was used to verify the method proposed in this work. Multistep predictions with different numbers of wind data were carried out, and the results were analyzed and discussed.

3.1. Pearson Autocovariance Feature Selection. Wind power prediction refers to an estimate of the expected production of one or more wind turbines. Here, the wind speed prediction is specifically denoted in terms of available wind power, representing the power production potential over a time interval. However, the involvement of complex patterns elaborates the extraction of the input-output mapping function of the prediction process and then minimizes its overall performance. Therefore, feature selection is a prerequisite to wind power prediction.

The feature selection model selects a variable subset from the input data and hence reduces the influence of the irrelevant data and provides good results. Wind speed is predicted on the basis of the input data acquired from the wind turbine dataset. To do so, each sample of wind speed is contemplated as a target, and its previous most relevant input data is obtained via feature selection, therefore, reducing the data training or process time and hence improving the predictor overall performance.

In this work, the Pearson Autocovariance Feature Selection model is employed to select the most relevant input set for training even in the presence of complex patterns. The autocovariance function is utilized here to select the most relevant samples and therefore aids in determining the most significant candidate input. Next, high correlated samples are obtained by means of the Pearson Correlation function for further prediction process. Figure 2 shows the structure of the Pearson Autocovariance Feature Selection model.

As shown in the figure, to be more specific, 5000 hourly lagged values of LV active power, wind speed, wind direction, and theoretical power curve from different dates and times are served as the candidate input data. Given that the recurrent neural network is very sensitive to the diversity of input involving complex patterns and that the unreliability of the wind will create outliers in the wind power prediction, feature selection is performed with the wind power data to obtain the most relevant input data.

$$S(F_t) = \left[\frac{F_t - \mu}{\sigma}\right]. \tag{1}$$

From (1), 'F_t' represents the wind turbine power data at time 't', and the mean wind turbine power data acquired at different time intervals is estimated as follows.

$$\mu = \frac{1}{N}\sum_{t=1}^{N}[F_t]. \tag{2}$$

From (2), the mean value 'μ' is estimated based on the wind turbine power data obtained at different time intervals, therefore acquiring candidate input data. Next, with the acquired mean wind turbine power data at different time intervals, candidate input data are then processed by the autocorrelation function with the objective of identifying a minimum subset of the significant informative features to be integrated into the proposed method. The autocorrelation minimum subsets between times interval 't_1' and 't_2' are mathematically stated as follows.

$$R_{XY}(t_1, t_2) = \text{Exp}\left[F_{t_1}, P_{\text{WTO}}\right]. \tag{3}$$

From (3), autocovariance results of wind speed exhibit the relationship between input data and the target value, where 'Exp' represents the expected value and 'P_{WTO}' denotes the wind turbine output for the corresponding inputs 'X' and 'Y', respectively. To estimate the wind energy, the wind turbine output is utilized and mathematically stated as follows.

$$P_{\text{WTO}} = \frac{1}{2} * \rho * A * (WS)^3 - PC. \tag{4}$$

From (4), 'P_{WTO}' specifies the output generated by the wind turbine obtain using the air density 'ρ', area 'A', wind speed 'WS', and the power coefficient 'PC', respectively. Finally, the autocovariance functions between times interval 't_1' and 't_2' are mathematically formulated as follows.

$$K_{XY}(t_1, t_2) = \left[\left(F_{t_1} - \mu_{t_1}\right) * \left(F_{t_2} - \mu_{t_2}\right) * P_{\text{WTO}}\right]. \tag{5}$$

From (5), resultant data with a high autocovariance function is selected as an input to normalize the autocovariance function. This is performed by applying the Pearson Correlation Coefficient.

$$\rho_{XY}(t_1, t_2) = \frac{K_{XY}(t_1, t_2)}{\sigma^2}. \tag{6}$$

From the resultant value of (6), highly correlated features from the prediction engine are selected as input and considered for further processing. The pseudocode representation of the Pearson Autocovariance Feature Selection is given in Algorithm 1.

As given in the above Pearson Autocovariance Feature Selection with the objective of selecting the relevant feature with minimum time and maximum accuracy even in the

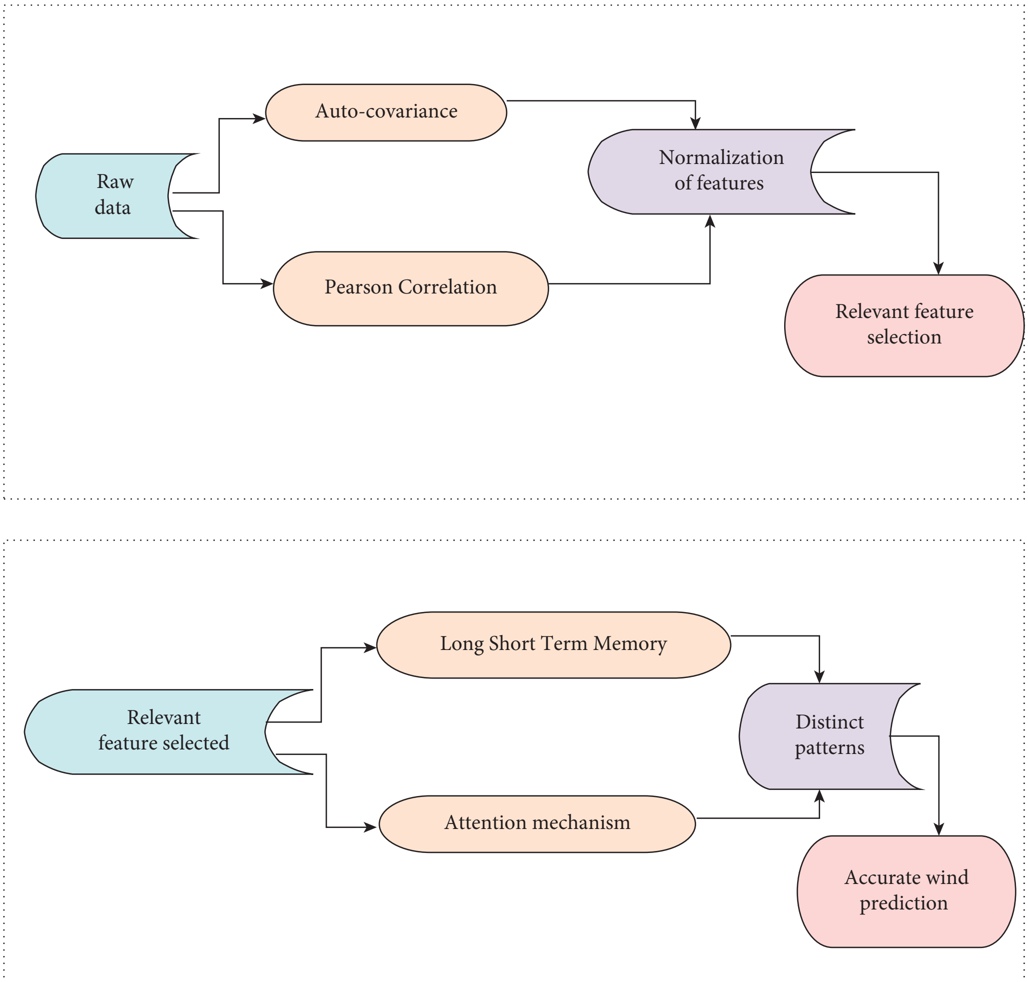

Figure 1: Block diagram of the Pearson Autocovariance Distinct Patterns and Attention-based Deep Learning (PACDP-ADL) for wind power prediction.

presence of complicated patterns, two functions are invoked. First, owing to the fact that the wind faced by each wind turbine interacts in time and space, the wind turbine power is said to be closely correlated to the wind at its location. Therefore, autocorrelation and autocovariance for each sample of wind data at two different time intervals are obtained to find how finely they are correlated with each other. In time, winds in different spatial positions at the same time can also affect each other due to the occurrence of complicated patterns. Hence, Pearson Correlation is applied to eliminate the irrelevant features, therefore ensuring accurate and timely relevant features.

3.2. Attention-Oriented Long Short-Term Memory-Based Wind Power Prediction. Wind power prediction depends on the estimation of wind speed. Due to the distinct changes found in the patterns, i.e., cyclical pattern and daily pattern, accurate prediction of wind power is a too laborious and cumbersome process. Hence, it is coherent that significant modification and wind energy resources application necessitate exact and absolute information on the wind features. With the accurate and relevant feature selection made using the Pearson Autocovariance Feature Selection algorithm, the actual wind power prediction is made in this section by applying the Attention-oriented Long Short-Term Memory-based Wind Power Prediction model. Figure 3 shows the structure of the Attention-oriented Long Short-Term Memory-based Wind Power Prediction model.

The long short-term memory wind power forecasting on multiple scales does not concentrate on the pattern changes. Though the assistance of cellular states that determined the states that should be preserved and those which should be forgotten was utilized [1], with the absence of pattern change observations, the true positive involved in wind power prediction was not focused. To address this issue, an Improved Long Short-Term Memory-based Wind Power Prediction model is proposed in this work, with the consideration of patterns in the input vector to organize a mapping between the input and output vector, thereby enabling the prediction potentiality of the wind power network model.

For the wind power prediction model, the input vector contains a random pattern with the features present in '$FS = (FS_1, FS_2, \ldots, FS_t)$', therefore, corresponding to the input gate as given below. With the wind power signal being

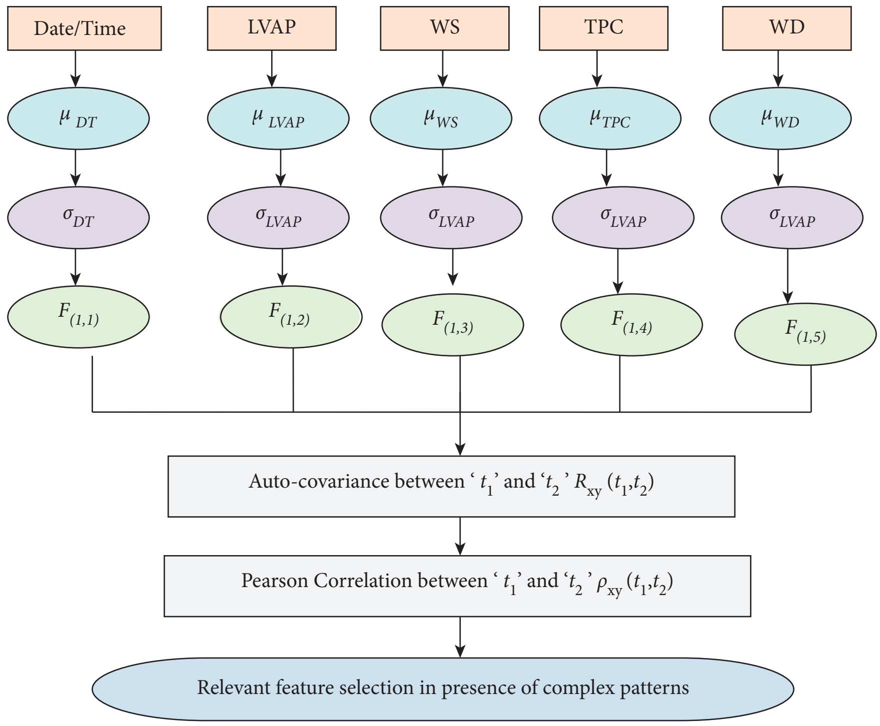

Figure 2: Structure of Pearson Autocovariance Feature Selection model.

Input: dataset ‘DS‘, sample data ‘$D = D_1, D_2, \ldots, D_n$‘, feature set ‘$FS(t) = FS_1, FS_2, .., FS_n$’
Output: Accurate and timely relevant feature selection
Step 1: **Initialize** samples ‘N‘, time interval ‘t_1‘, ‘t_2’
Step 2: **Initialize** air density ‘ρ‘, area ‘A‘, power coefficient ‘PC’
Step 3: **Begin**
Step 4: **For** each dataset ‘DS’ with sample data ‘D’
Step 5: Preprocess wind power data as in equation (1)
Step 6: Obtain the mean wind power data as in equation (2)
Step 7: Estimate autocorrelation between time interval ‘t_1‘, ‘t_2’ as in equations (3) and (4)
Step 8: Estimate autocovariance between time interval ‘t_1’, ‘t_2’ as in equation (5)
Step 9: Obtain highly correlated data using Pearson Correlation Coefficient as in equation (6)
Step 10: **Return** relevant feature selected ‘FS’
Step 11: **End for**
Step 12: **End**

Algorithm 1: Pearson Autocovariance Feature Selection.

a single-dimensional time series data, in our work, a sliding window is utilized to model multidimensional time-series data as the input vector. Let us assume that the present time is ‘t‘; the wind power is ‘P_t‘; then, the input vector is represented as ‘$X_t = (P_{t-n+1}, P_{t-n+2}, \ldots, P_t)$’ and learns the patterns ‘Pat’ at time ‘t’ with which the input gate is formulated as given below.

$$I = \sigma\left(W_I X_t \mathrm{Pat}_t + U_I h_{t-1} + B_I\right). \tag{7}$$

From th (7), ‘W’ denotes the weight matrix, ‘W_I’ is the correlation matrix between previously hidden layer input ‘X_t’ and memory, ‘B_I‘ is the input bias vector, and the patterns ‘Pat_t‘ here represent three types, namely, short-term fluctuation, long term fluctuation, or random signal, respectively. Then, with the input gate ‘I‘, the hidden layer sequentially estimates the activation values of the three gates, forget gate ‘F’ and output gate ‘O’ and memory gate ‘C_t’ as given below.

$$F = \sigma\left(W_F X_t \mathrm{Pat}_t + U_F h_{t-1} + B_F\right). \tag{8}$$

From (8), ‘W_F’ is the correlation between previously hidden layers, forget gate, and memory, and ‘B_F ’ represents the bias vector for the respective forget gate. Next, forget gate ‘F’ and input gate ‘I’ are utilized to control how many

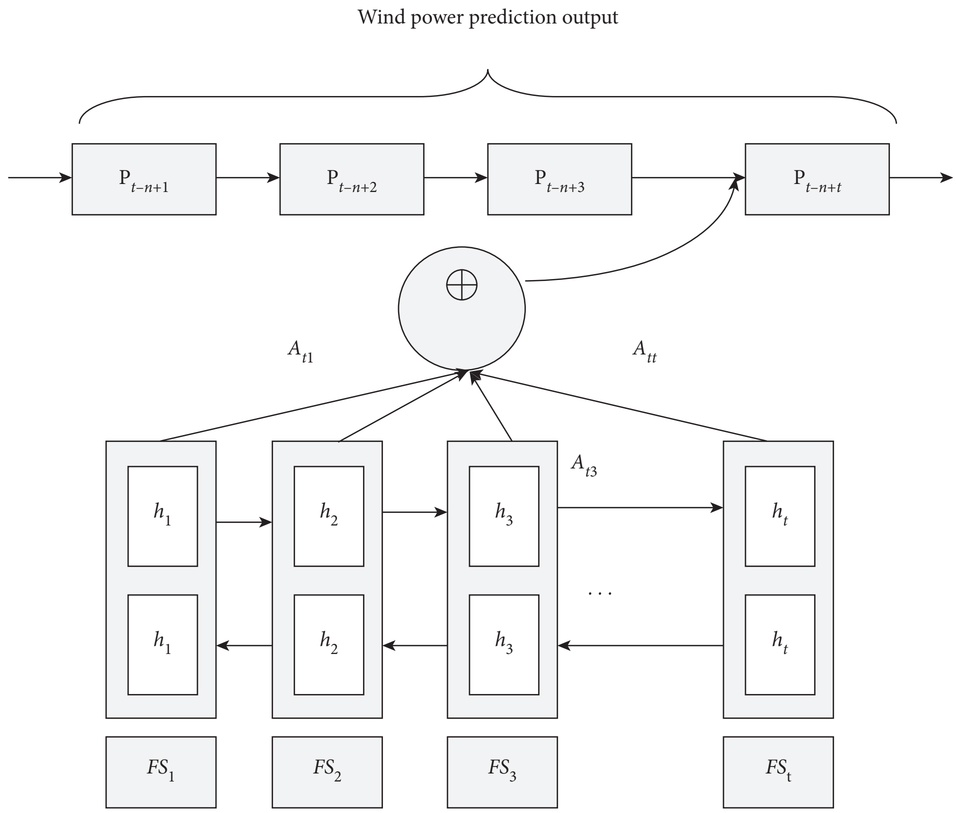

Figure 3: Structure of Attention-oriented Long Short-Term Memory-based Wind Power Prediction model.

historical patterns 'C_{t-1}' are forgotten and how many new patterns 'e' are saved in order to update the cell state 'C_t' which is mathematically stated as given below.

$$C_t = F_t \otimes C_{t-1} \oplus I \otimes e + R. \tag{9}$$

From (9), to circumvent the influence of arbitrary elements (i.e., the features) on the patterns in the memory, a parameter 'R' is introduced while updating the cell state, therefore enhancing wind power prediction even in presence of complex patterns. Owing to the fact that arbitrary pattern is also an element of wind power prediction, it is inappropriate to disregard the arbitrary element on the patterns in the forecast output. Therefore, the Attention-oriented Long Short-Term Memory-based Wind Power Prediction model in our work models the output gate as given below.

$$O_t = W_{OX} X_t \mathrm{Pat}_t + B_O, \tag{10}$$

$$h_t = O_t \otimes \tanh(C_t). \tag{11}$$

From (10) and (11), 'W_{OX}' represents the correlation between the previously hidden layer output gate and input vector corresponding to different patterns 'Pat_t', and 'B_O' represents the bias vector for the respective output gate. Finally, the attention vector is utilized in distributing distinct weights 'W_C' according to the cell state at time 't' for different patterns better predict the wind power according to time.

$$A_t = f(C_t, h_t) = \tanh(W_C[C_t, h_t]). \tag{12}$$

The design of the above attention mechanism in the LSTM 'A_t' highly predictive wind power even in the presence of complex patterns with better sensitivity can be attained. The pseudocode representation of the Attention-oriented Long Short-Term Memory-based Wind Power Prediction is given in Algorithm 2.

As given in the above algorithm with the objective of accurately predicting the wind power data even in the presence of complex patterns, Attention-oriented Long Short-Term Memory-based Wind Power Prediction is designed. First, long short-term memory-based prediction convolving for different patterns are obtained, and accordingly, changes are made in the forget gate and cell state. Next, with the aid of the attention mechanism, according to time and distinct patterns, accurate predictions are made.

4. Experimental Settings

The efficiency of the proposed method is validated by conducting experimentation with Native Java Codes. Experimental evaluation is carried out using factors such as wind power prediction accuracy, wind power prediction time, and true positive rate with respect to the number of data.

To perform a fair comparison, experimentation with the number of data is performed with the proposed PACDP-ADL and two existing methods, Fuzzy-based Multidimensional Resource Scheduling and Queuing Network Encoder-Decoder Long Short-Term Memory (EDLSTM) [1] and

Input: Dataset 'DS', sample data '$D = D_1, D_2, \ldots, D_n$'
Output: Wind power data correct prediction
Step 1: **Initialize** feature selected 'FS', time 't', wind power is 'P_t'
Step 2: **Begin**
Step 3: **For** each feature selected 'FS' at time 't' with wind power 'P_t'
Step 4: **For** each input vector '$X_t = (P_{t-n+1}, P_{t-n+1}, \ldots, P_t)$' with patterns 'Pat'
Step 5: Formulate the input gate as given in equation (7)
Step 6: Formulate forget gate as in equation (8)
Step 7: Update the cell state 'C_t' as in equation (9)
Step 8: Estimate the output gate as in equation (10)
Step 9: Model attention mechanism as in equation (11)
Step 10: According to the attention mechanism result make changes in the cell state
Step 11: **Return** wind power evaluation
Step 12: **End for**
Step 13: **End**

Algorithm 2: Attention-oriented Long Short-Term Memory-based Wind Power Prediction.

variational mode decomposition (VMD), the K-means clustering and long short-term memory (LSTM) VMD-K-means-LSTM [2]. The wind power prediction is performed with the SCADA data of a wind turbine in Turkey taken from the https://www.kaggle.com/berkerisen/wind-turbine-SCADA-dataset. From the dataset, the number of data is collected. The collected data are used for extracting the feature and classification. The Pearson Autocovariance Feature Selection is employed for choosing the most relevant input set for training even in the complex patterns. Next, Attention-based Long Short-Term Memory Wind Prediction algorithms are used for required patterns as well as forget irrelevant patterns. In this dataset, cross-validation has a training and testing dataset. In the training dataset, it occupies 18%, and in the testing dataset, it has 20% which are predicted by wind power.

In wind turbines, SCADA Systems are measured, and data like wind speed, wind direction, generated power, etc. for 10 minutes intervals are obtained. This file was taken from a wind turbine SCADA system that is working and generating power in Turkey. Wind speed (m/s) at the hub height of the turbine is obtained (the wind speed that the turbine uses for electricity generation). The theoretical power values of the turbine are produced by the wind speed that is given by the turbine manufacturer. The wind direction at the hub height of the turbine is obtained (wind turbines turn to this direction automatically). An overall of 10 simulation runs are performed for three distinct performance metrics wind power prediction accuracy, wind power prediction time, and true positive rate, respectively.

5. Performance Metrics

In order to compare the efficiency of the PACDP-ADL method, we use several metrics to evaluate their performance. The following metrics are used:

(i) Wind power prediction accuracy
(ii) Wind power prediction time
(iii) True positive rate

5.1. Performance Analysis of Wind Power Prediction Accuracy. It refers to the accuracy factor involved during the prediction of wind power involving distinct patterns. It is measured in terms of percentage (%). This is mathematically stated as given below.

$$\mathrm{WPP}_{\mathrm{acc}} = \sum_{i=1}^{n} \frac{D_{AP}}{D_i} * 100. \tag{13}$$

From (13), the wind power prediction accuracy '$\mathrm{WPP}_{\mathrm{acc}}$' is measured on the basis of the wind data involved in the simulation process 'D_i' and the accurately predicted data 'D_{AP}' during the experimentation.

Initially, we start by evaluating the wind power prediction accuracy across different methods, proposed PACDP-ADL, and two existing methods, EDLSTM [1] and VMD-K-means-LSTM [2]. The three methods utilize wind data acquired from the SCADA wind turbine dataset to begin our experiments without affecting operation. However, for fair comparisons, similar wind sample data has been applied for all three methods. There is more improvement in wind power prediction accuracy when employing the PACDP-ADL method upon comparison with EDLSTM [1] and VMD-K-means-LSTM [2] as given in Table 1.

To explore the influence of wind power prediction accuracy efficiency on the PACDP-ADL method with the help of the Pearson Autocovariance Feature Selection algorithm, the simulations were performed by varying the wind sample data in Figure 4. It also shows that the application of the Pearson Autocovariance Feature Selection algorithm extensively provides competitive results compared to the state-of-the-art methods, namely, EDLSTM [1] and VMD-K-means-LSTM [2]. The wind power prediction accuracy efficiency using the PACDP-ADL method decreases with the increase in the number of wind sample data but comparatively performs better than the state-of-the-art methods. This is because the autocovariance function using the PACDP-ADL method selects the relevant feature even in the presence of complex patterns based on the mean of wind turbine power data obtained at different time intervals. Also with the application

Table 1: Wind power prediction accuracy versus the number of data.

Number of data	Wind power prediction accuracy (%)		
	PACDP-ADL	EDLSTM	VMD-kmeans-LSTM
500	98	97	96
1000	97.35	93.85	92.15
1500	97.15	93	90.15
2000	97	92.35	90
2500	96.75	92.15	89.55
3000	96.35	92	89.35
3500	96.15	91	88
4000	95.85	89	86.55
4500	95.25	88.55	84.25
5000	95	88	83

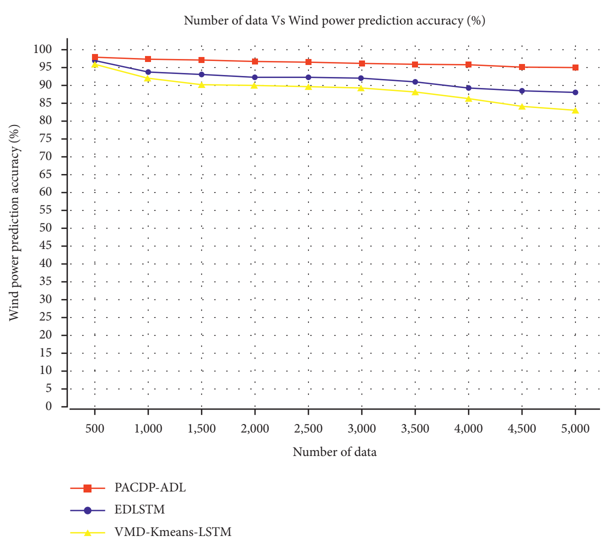

Figure 4: Graphical representation of wind power prediction accuracy.

of the Pearson Correlation Coefficient, significant informative features are selected that are closely correlated to the wind at its location. This in turn improves the wind power prediction accuracy even for complex patterns using the PACDP-ADL method by 5% compared to EDLSTM [1] and 9% compared to VMD-K-means-LSTM [2], respectively.

5.2. Performance Analysis of Wind Power Prediction Time. It refers to the time involved in the prediction process for wind power involving distinct processes. It is measured in terms of milliseconds (ms). This is mathematically stated as given below.

$$\mathrm{WPP}_{\mathrm{time}} = \sum_{i=1}^{n} D_i * \mathrm{Time}\left[O_t + h_t\right]. \tag{14}$$

From (14), the wind power prediction time '$\mathrm{WPP}_{\mathrm{time}}$' is measured based on the wind data involved in the experimentation 'D_i' and the time consumed in involving the prediction based on the output and hidden gate 'Time $[O_t + h_t]$' at time instance 't', respectively.

Second, we have evaluated the wind power prediction time using the proposed method, PACDP-ADL, and two state-of-the-art methods, EDLSTM [1] and VMD-K-means-LSTM [2] with the assistance of wind sample data in the range of 500 and 5000. To ensure consistency between three different methods, similar wind sample data has been applied for all three methods. There is more saving in wind power prediction time when using the PACDP-ADL method compared to the results obtained from EDLSTM [1] and VMD-K-means-LSTM [2] as given in Table 2.

Figure 5 shows the comparison of the proposed PACDP-ADL method, EDLSTM [1], and VMD-K-means-LSTM [2] based on their wind power prediction time. The average wind power prediction time is measured by selecting relevant features of wind turbines. The number of sample wind data is varied in the range from 500 to 5000 for simulation purposes. It is clear that the average wind power prediction time is increased for all methods with the increase in the number of sample wind data. But the proposed PACDP-ADL method achieves better performance on average wind power prediction time when compared to existing methods. In order to improve the average wind power prediction time, the PACDP-ADL method uses the Pearson Autocovariance Feature Selection that selects the relevant feature subset according to covariance and correlation function. The Pearson Correlation function performs efficient feature selection by addressing overfitting that is said to occur due to complex patterns. Additionally using the PACDP-ADL method, due to the reason that wind in different spatial positions at the same time also has a negative influence on wind power prediction due to the occurrence of complicated patterns by applying the Pearson Correlation, this eliminates the irrelevant features, therefore reducing wind power prediction time using PACDP-ADL method by 18% compared to [1] and 27% compared to [2].

Additionally, the PACDP-ADL method uses data utilization history that assists in obtaining the resource estimation based on the prior evaluation by using multiple regressions. As a result, optimum outputs are obtained close to the target output (i.e., average resource optimization time). Hence, the TWLR-IQ method reduces average resource optimization time by 19% when compared to the existing PA-KJF [1] and 32% when compared to the existing D-JStorm [2], respectively.

5.3. Performance Analysis of True Positive Rate. It is measured as the ratio of the number of wind data objects that are correctly predicted using the deep learning algorithm. It is mathematically formulated as given below.

$$TPR = \sum_{i=}^{n} \frac{TP}{TP + FN}. \tag{15}$$

From (15), the true positive rate 'TPR' is measured based on the true positive rate (i.e., wind power data correctly

Table 2: Wind power prediction time versus the number of data.

Number of data	Wind power prediction time (ms)		
	PACDP-ADL	EDLSTM	VMD-kmeans-LSTM
500	307.5	367.5	392.5
1000	335.5	395.55	445.55
1500	385.15	455.15	515.25
2000	415	495.15	605.25
2500	455.55	585.55	655.15
3000	535.15	635.15	725.35
3500	585.25	680.25	795.15
4000	615.35	735.55	825.15
4500	635.15	815.55	880
5000	650.25	900.15	985

Table 3: True positive rate versus the number of data.

Number of data	True positive rate		
	PACDP-ADL	EDLSTM	VMD-kmeans-LSTM
500	0.95	0.92	0.9
1000	0.93	0.85	0.81
1500	0.91	0.83	0.79
2000	0.9	0.82	0.77
2500	0.88	0.8	0.75
3000	0.86	0.78	0.71
3500	0.84	0.76	0.69
4000	0.83	0.72	0.66
4500	0.82	0.7	0.65
5000	0.81	0.68	0.62

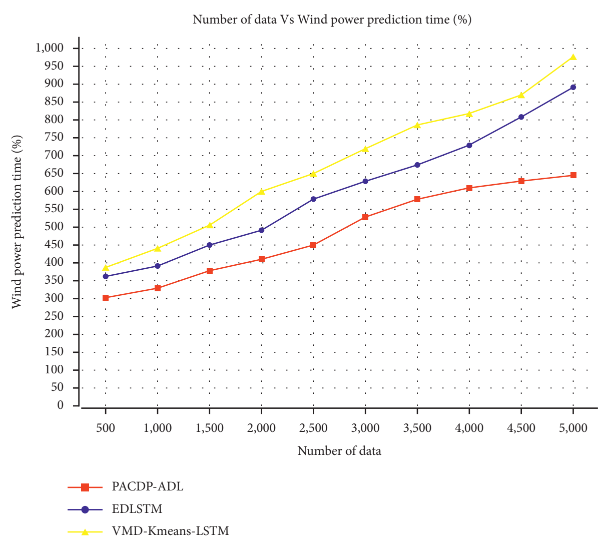

Figure 5: Graphical representation of wind power prediction time.

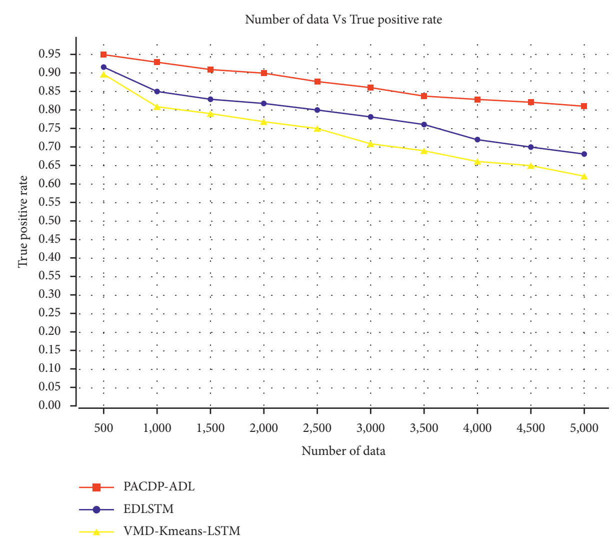

Figure 6: Graphical representation of true positive rate.

predicted) '*TP*' and the false negative rate (i.e., wind power data is wrongly predicted) '*FN*', respectively.

Third, the true positive rate involved in the process of wind power prediction is estimated by employing the proposed method, PACDP-ADL, and two state-of-the-art methods, EDLSTM [1] and VMD-K-means-LSTM [2]. To provide a fair comparison between the three methods, the wind sample data are selected at random but the same number of data ranges between 500 and 5000 for all the three methods. True positive rate applying the PACDP-ADL method is found to be greatly improved when compared to EDLSTM [1] and VMD-K-means-LSTM [2] as provided in Table 3.

Figure 6 shows the performance of the true positive rate calculated using the proposed PACDP-ADL method and compared with existing methods, namely, EDLSTM [1] and VMD-K-means-LSTM [2]. The number of wind sample data is varied in the range from 500 to 5000 for simulation purposes. It is clear that the true positive rate gets decreased for all methods with the increase in the number of wind sample data. But the proposed PACDP-ADL method achieves better performance on a true positive rate when compared to existing methods. From Figure 6, it is clear that the true positive rate is improved gradually for the proposed PACDP-ADL method when compared to other existing methods. This efficient improvement of true positive rate is achieved using the Attention-oriented Long Short-Term Memory-based Wind Power Prediction algorithm. Here, initially, long short-term memory-based wind power prediction for different patterns is acquired and then upon changes in different time instances changes were made in the forget gate and cell state. Followed by which, using the attention mechanism, predictions were made taking into consideration both the time and distinct patterns. Hence, the true positive rate is improved in the proposed PACDP-ADL method by 11% when compared to existing EDLSTM [1] and 203% when compared to VMD-K-means-LSTM [2], respectively.

6. Discussion

The Pearson Autocovariance Distinct Patterns and Attention-based Deep Learning (PACDP-ADL) is developed. Feature selection plays a vital feature and prediction task. The Pearson Autocovariance Feature Selection model is used for identifying essential features for wind power prediction and minimizes the complexity. Next, an Attention-based Long Short-Term Memory Wind Prediction algorithm is employed to retain required patterns and forget irrelevant patterns to acquire more satisfactory prediction precision.

The Encoder-Decoder Long Short-Term Memory (EDLSTM) is designed to wind power prediction by presenting the time series scheme on physics and data wind data augmentation method. By integrating physics with data-oriented wind data, sequential input and sequential output for wind power prediction were maintained by lesser minimum error. But, complex pattern discovery of the sequential input is a major issue that has to be concentrated on correct and timely wind power prediction.

A hybrid scheme for wind power prediction within a short-term fashion employing Variation Mode Decomposition (VMD), the K-means clustering, and long short-term memory (LSTM) network, termed VMD-K-means-LSTM was introduced. Initially, the VMD algorithm divides the wind power data within finite stationary IMFs. Every stationary IMF is unfolded by trajectory matrix with K-means being utilized for dividing trajectory matrix into clustering sample set. Distances among clustering centers were estimated. Lastly, optimal LSTM was selected for all sublayers to acquire the final prediction result for minimizing the error rate.

7. Conclusion

In this section, wind power generation is directly associated with the weather conditions, and the prediction of wind power remains the major factor at the level of the wind farm. With the development of numerical weather prediction, deep learning, and artificial intelligence, wind power prediction demonstrates the characteristics of high generation time and reduces accuracy factors owing to complex patterns. Unquestionably, wind farms frequently experience such emergent complicated patterns which require to be estimated swiftly and optimally. PACDP-ADL method is used for wind power prediction. Initially, relevant features in the complicated patterns occurring in different time instances are designed to obtain accurate relevant features required for further prediction process. Then, an Attention-oriented Long Short-Term Memory-based Wind Power Prediction algorithm for analyzing wind power is presented for different patterns to guarantee the true positive rate involved in the prediction process. Experiments are performed to compare our proposed PACDP-ADL method with various emergent requirements in wind farms with higher true positive rates, better accuracy, and lesser time involved with the state-of-the-art methods.

Data Availability

SCADA data of a wind turbine in Turkey are available.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

References

[1] H. Chen, Y. Birkelund, and Q. Zhang, "Data-augmented sequential deep learning for wind power forecasting," *Energy Conversion and Management*, vol. 248, 2021.

[2] Z. Sun, S. Zhao, and J. Zhang, "Short-term wind power forecasting on multiple scales using VMD decomposition, K-means clustering and LSTM principal computing," *IEEE Access*, vol. 7, 2019.

[3] A. Shobana Devi, G. Maragatham, K. Boopathi, and A. G. Rangaraj, "Hourly day-ahead wind power forecasting with the EEMD-CSO-LSTMEFG deep learning technique," *Soft Computing–A Fusion of Foundations, Methodologies and Applications*, Springer, Berlin, Germany, 2020.

[4] F. Elyasichamazkoti and A. Khajehpoor, "Application of machine learning for wind energy from design to energy-water nexus: a survey," *Energy Nexus*, vol. 2, 2021.

[5] J. Enrique Sierra-Garcia and M. Santos, "Deep learning and fuzzy logic to implement a hybrid wind turbine pitch control," *Neural Computing and Applications*, vol. 33, 2021.

[6] H. Yin, Z. Ou, S. Huang, and A. Meng, "A cascaded deep learning wind power prediction approach based on a two-layer of mode decomposition," *Energy*, vol. 189, 2019.

[7] A. Ahmad, M. Nabipour, B. Mohammadi-Ivatloo, Ali Mordai Amani, S. Rho, and Md Jalil Piran, "Long-term wind power forecasting using tree-based learning algorithms," *IEEE Access*, vol. 8, 2020.

[8] T. Hu, W. Wu, Q. Guo, H. Sun, L. Shi, and X. Shen, "Very short-term spatial and temporal wind power forecasting: a deep learning approach," *CSEE Journal Of Power and Energy Systems*, vol. 6, no. 2, 2020.

[9] Madasthu Santhosh, Chintham Venkaiah, and D. M. Vinod Kumar, "Current Advances and Approaches in Wind Speed and Wind Power Forecasting for Improved Renewable Energy Integration: A Review," *Wiley*, vol. 2, no. 2, 2020.

[10] Z. Li, L. Ye, Y. Zhao, X. Song, J. Teng, and J. Jin, "Short Term Wind Power Prediction Based on Extreme Learning Machine with Error Correction", Protection and Control of Modern Power Systems," *Springer*, vol. 1, 2016.

[11] S. Salcedo-Sanz, A. Pastor-Sánchez, L. Prieto, A. Blanco-Aguilera, and R. García Herrera, "Feature Selection in Wind Speed Prediction Systems Based on a Hybrid Coral Reefs Optimization – Extreme Learning Machine Approach," *Energy Conversion and Management, Elsevier*, vol. 87, 2014.

[12] A. Dolatabadi, H. Abdeltawab, Y. Abdel-Rady, and I. Mohamed, "Hybrid Deep Learning-Based Model for Wind Speed Forecasting Based on DWPT and Bidirectional LSTM Network," *IEEE Access*, vol. 8, 2020.

[13] I. González-Aparicio and A. Zucker, "Impact of Wind Power Uncertainty Forecasting on the Market Integration of Wind Energy in Spain," *Applied Energy, Elsevier*, vol. 159, 2015.

[14] Yi Zhang, H. Sun, and Y. Guo, "Wind Power Prediction Based on PSO-SVR and Grey Combination Model," *IEEE Access*, vol. 7, 2019.

[15] A. Kavousi-Fard, A. Khosravi, and S. Nahavandi, "A new fuzzy-based combined prediction interval for wind power forecasting," *IEEE Transactions on Power Systems*, vol. 31, 2015.

[16] S. Vidya and E. Srie Vidhya Janani, "Tabu Search algorithm based general regression neural network for long term wind speed predictions," *Journal for Control, Measurement, Electronics, Computing and Communications*, vol. 61, Dec 2020.

[17] J. Wang, H. Jiang, B. Han, and Q. Zhou, "An Experimental Investigation of FNN Model for Wind Speed Forecasting Using EEMD and CS", Mathematical Problems in Engineering,"

Hindawi Publishing Corporation, vol. 2015, Article ID 464153, 13 pages, 2015, https://doi.org/10.1155/2015/464153.

[18] K. Pan, Q. Zheng, and N. Chen, "Probabilistic Short-Term Wind Power Forecasting Using Sparse Bayesian Learning and NWP", Mathematical Problems in Engineering," *Hindawi Publishing Corporation*, vol. 2015, Article ID 785215, 11 pages, 2015.

[19] F. Fazelpour, N. Tarashkar, and M. A. Rosen, "Short-term wind speed forecasting using artificial neural networks for Tehran, Iran," *International Journal of Energy and Environment Engineering*, Springer, vol. 7, Sep 2016.

[20] G. Zhang, H. Liu, J. Zhang et al., "Wind power prediction based on variational mode decomposition multi-frequency combinations," *Journal of Modern Power Systems and Clean Energy*, vol. 9, Springer, Dec 2018.

[21] N. Bokde, A. Feijóo, D. Villanueva, and K. Kulat, "A review on hybrid empirical mode decomposition models for wind speed and wind power prediction," *Energies*, vol. 12, no. 2, p. 254, 2019.

[22] H. Hu, J. Zhang, and T. Li, "A comparative study of VMD-based hybrid forecasting model for nonstationary daily streamflow time series," *Complexity*, vol. 2020, Article ID 4064851, 21 pages, 2020.

[23] Q. Zheng, Y. Pei, H. Zareipour, and N. Chen, "A review and discussion of decomposition-based hybrid models for wind energy forecasting applications," *Applied Energy*, vol. 235, pp. 939–953, 2019.