

# Wind Power Forecasting Using A Heterogeneous Ensemble of Decomposition-based NNRW Techniques

Pakarat Musikawan<sup>1</sup>, Khamron Sunat<sup>2</sup>, and Yanika Kongsorot<sup>3</sup>,

## ABSTRACT

Accurate and reliable wind power forecasting plays a vital role in the operation and management of power systems. Hence, it has become necessary to research and develop a high-accuracy wind power forecasting model. However, owing to highly nonlinear and non-stationary patterns of wind power time-series, creating a wind forecasting model capable of predicting such series accurately is both complicated and challenging. Aiming at this challenge, this paper introduces a new decomposition-based hybrid model based on multiple decomposition techniques, neural network with random weights (NNRW), and a linear combiner. In our approach, the original time-series is decomposed into a collection of sub-series by different decomposition techniques. Each sub-series is modeled and predicted separately using NNRW. The predicted signals of each decomposition model are then reconstructed independently. Finally, all of the reconstructed results are integrated by the combiner using a linear combination method. The predictive performance of the proposed method was compared with other state-of-the-art techniques in over 12 wind power time-series. The experimental results show that the predictive performance of the proposed method remarkably outperforms the other competitors, proving the developed model to be effective, efficient, and practical.

**Keywords:** Wind power forecasting, Time-series, Neural network with random weights, Decomposition technique, Hybrid model, Ensemble system

## 1. INTRODUCTION

Wind energy is becoming more and more important as a worldwide energy supply. According to the report of the Global Wind Energy Council (GWEC), the global cumulative installed electricity generation capacity from wind power in 2018 was 51.3 gigawatts (GW) and the GWEC forecasts that the capacity of

wind power generation will reach higher than 840 GW by the end of 2022 [1]. Moreover, the report released by the GWEC shows that the cumulative installed wind power capacity could reach 2000 GW by 2030. This illustrates that wind energy has gained greater distinction and has attracted global attention. Wind power is the conversion of energy from the wind into electricity, which is generated by the passing of air-flow through wind turbines. The power generated is therefore dependent upon the wind speed. Accurate wind power forecasting is necessary for power system operations, such as planning, dispatching, and maintenance schedules. However, as wind speed is intermittent and fluctuating, it is not easy to model and predict accurately [2, 3].

In applications of wind power forecasting, decomposition-based hybrid approaches have been proposed based on the combination of decomposition techniques and forecasting models. Different decomposition techniques have been widely applied in the hybrid methods for preprocessing because they can effectively reduce the non-stationary characteristics of the wind power time series [4], including empirical mode decomposition (EMD) [5], variational mode decomposition (VMD) [6], discrete wavelet transform (DWT) [7], wavelet packet decomposition (WPD) [8], and singular spectrum analysis (SSA) [9]. These decomposition methods are used in the data preprocessing stage to decompose the time-series of wind power into several components. Then, a forecasting model is built for each decomposed component. Generally, conventional machine learning algorithms are usually utilized to perform as forecasting models [10]. They have become the most dominant techniques in decomposition-based hybrid approaches, due to their forecasting ability. However, these algorithms require a lot of training time to iteratively find their optimal parameters. Hence, it is necessary to balance forecasting accuracy with the required computational time.

Non-iterative learning approaches have also been proposed to avoid some of the difficulties faced by iterative learning algorithms [11]. A neural network with random weights (NNRW) is a class of non-iterative learning algorithm for training NN with a fixed hidden layer size [12, 13, 14]. The weights and biases within the hidden layer of NNRW are randomly assigned, while the output layer parameters are de-

Manuscript received on February 20, 2020 ; revised on April 10, 2020.

Final manuscript received on April 14, 2020.

<sup>1,2,3</sup>The authors are with the Department of Computer Science, Faculty of Science, Khon Kaen University, Khon Kaen 40002, Thailand, E-mail: <sup>1</sup>pakarat.mus@kkumail.com, <sup>2</sup>skhamron@kku.ac.th, and <sup>3</sup>yanika.k@kkumail.com

DOI: 10.37936/ecti-cit.2020142.239860

terminated by finding the least square solution. Due to their ability to generate a forecasting model with extremely fast learning speed and satisfactory performance, they have attracted the attention of numerous research studies [15, 16, 17], especially in the area of wind energy [18, 19, 20]. According to previous literature [21], it is clear that NNRW can dramatically accelerate the computational speed of the decomposition-based hybrid approach.

Based on the aforementioned research, there have been a lot of successful applications of the individual decomposition technique integrated with NNRW in time-series forecasting. However, the single decomposition-based hybrid approaches often cannot accurately capture the complex relationships existing in the highly nonlinear and non-stationary time-series. By borrowing the idea of ensemble learning that incorporates the advantages of different individual algorithms, this paper proposes a new decomposition-based hybrid approach, named *Heterogeneous Ensemble of Decomposition-based NNRW* (EDNNRW). In our approach, the original time-series is decomposed into a finite number of components through different decomposition techniques. Five prominent decomposition techniques (EMD, VMD, SSA, DWT, and WPD) have been applied in the decomposition process in the preprocessing of our system because they are extensively applied signal decomposition techniques that have been proven to be effective, rapid, and practicable data preprocessing tools in time-series forecasting [22, 23, 24, 25, 26]. To inherit the merits of fast learning, computational simplicity, and good generalization capabilities, four types of NNRW models are utilized to perform as predictors for each decomposed component in the forecasting process. The final forecasting results of each decomposition technique can be reconstructed by adding up all the predicted results. Finally, all of the reconstructed signals are integrated as the ultimate result via a linear combiner, due to its architectural simplicity, fast modeling, and functional approximation capabilities. Simulations on wind power forecasting have demonstrated that the developed model significantly outperforms all comparative algorithms for single and multiple step forecasting.

The three main scientific contributions and novelties of this research are given in the following list:

1. We propose a new decomposition-based hybrid framework integrating multiple decomposition techniques, NNRW, and a linear combiner. This method has not been found in previous studies to the best knowledge of the authors.
2. EMD, VMD, SSA, DWT, and WPD were integrated into the developed framework to decompose the original signals to reduce the non-stationary characteristics as much as possible. This technique has also not been previously published.
3. Four types of NNRW methods were utilized as

predictors of the developed decomposition-based hybrid approach. The impacts of various NNRW methods in the developed model were investigated and documented.

The remainder of this paper is organized as follows: the literature review is presented in Section 2; our proposed method is described in Section 3; our experimental results and performance evaluations are presented in Section 4; and lastly, the conclusions are illuminated in Section 5.

## 2. LITERATURE REVIEW

Many previously published studies have proposed different methods for wind power forecasting. These can be divided into four broad categories [27]: (a) physical methods, (b) statistical methods, (c) intelligent methods, and (d) hybrid methods. Each method, however, is not without its limitations. Physical methods build forecasting methods through physical or meteorological information, such as temperature, pressure, altitude, and so on. Their drawback is that they are very time-consuming [28]. Statistical methods model the predictors through the use of historical data including autoregression (AR), moving average (MA), the combination of AR and MA (ARMA), and AR integrated MA (ARIMA) [29]. Since these models are linear approaches, they are incapable of accurately predicting highly nonlinear or non-stationary time series. Intelligent methods primarily employ machine learning techniques to find the relationship between the input variables and the corresponding output data. Some of these approaches are support vector machine (SVM) [30], artificial neural network (ANN) [31], and ensemble systems [32]. Hybrid methods aggregate various methodologies together. Generally, hybrid approaches combine decomposition-based methods and predictors. They generally have better prediction performance than the previously mentioned approaches. The hybrid approaches provide effective forecasting performance as they combine the advantages of different methodologies, and have thus received increasing attention [33].

The decomposition technique is a powerful tool for reducing the forecast difficulty by converting the original non-stationary time series into several relatively more stationary sub-series. EMD is a self-adaptive analysis technique for the time-domain processing of a nonlinear and non-stationary signal [22]. The EMD decomposes a signal  $\mathbf{x} = [x(1), \dots, x(T)]$  into a finite collection of  $K - 1$  intrinsic mode functions (IMFs) and one residue [22]. The group of IMFs  $\{\mathbf{u}_1, \dots, \mathbf{u}_{K-1}\}$  and the residue  $\mathbf{r}$  can be mathematically expressed as  $\mathbf{x} = \mathbf{r} + \sum_{k=1}^{K-1} \mathbf{u}_k$ . VMD [23] is an adaptive and non-recursive signal decomposition algorithm which is appropriate for analyzing non-stationary signals. The VMD decomposes a signal into  $K$  components with limited bandwidth in the spectral domain. Both the bandwidth and

center frequency of each component are determined by iteratively searching for the optimal solution of a variational problem. DWT [25] is a mathematical technique and powerful tool for analyzing the time-frequency domain. It is well suited for non-stationary signals. The DWT decomposes the signal into a set of *approximation* and *detail* coefficients. The approximation and detail coefficients represent the low and high frequency components, respectively. DWT decomposes only the approximation coefficient at each level. The WPD [26] is a generalized version of DWT which decomposes both the approximation and detail coefficients at each level. SSA [24] is a non-parametric technique which is widely employed in time series analysis. The core purpose of this approach is decomposing an original time-series of data into a sum of sub-series in which each sub-series can be identified as either a trend, quasi-periodic component, or noise.

Jiang et al., 2012 [5] proposed a combination of the EMD, the largest Lyapunov exponent (LLE) prediction method, and the grey forecasting model. The EMD was employed as the data preprocessing approach to decompose the time-series of wind power into various IMF components and one residual component. Then, the LLE method was performed to predict each IMF. Finally, the grey forecasting model was employed to predict the residual component. Zhang et al., 2018 [6] proposed a hybrid prediction model with the VMD and a long short-term memory network (LSTM), called VMD-LSTM. In the first step, the wind power time-series is decomposed into various sub-series using the VMD. In the VMD-LSTM, the LSTM network is exploited to find each sub-series of wind power. Wang et al., 2020 [9] presented a hybrid of SSA and the Laguerre neural network (LNN) optimized by the opposition transition state transition algorithm (OTSTA). The time-series of wind power was decomposed into various sub-series using SSA in the first step. An optimized LNN was built for each sub-series. Catalão et al., 2011 [7] proposed a combination of the DWT and multilayer perceptron (MLP) trained by the Levenberge-Marquardt (LM) algorithm for wind power forecasting in Portugal. In the first stage, the DWT was used to decompose the wind power series into a set of sub-series. Then, the future values of these sub-series were predicted using the LM network. Laouafi et al., 2017 [8] presented a hybrid of the WPD and adaptive neuro-fuzzy inference system (ANFIS) for the prediction of wind power generation in France. As mentioned before, the literature shows that EMD, VMD, SSA, DWT, and WPD can improve the predictive performance of the hybrid approaches in applications of wind power forecasting. The iterative intelligent methods were adopted to perform as predictors in the hybrid methods. These algorithms are very time-consuming because they employ iterative learning methods for tuning their parameters.

NNRW was originally described by Schmidt et al. [12], and they called it a Schmidt neural network (SNN). It is a fast learning approach for training a single hidden layer feedforward network (SLFN) with a fixed hidden layer size. The hidden layer parameters of an SLFN trained by NNRW are randomly generated, whereas the output weights and output biases are analytically determined by finding a least-square solution. Pao et al. [13] proposed a variant version of NNRW named random vector function-link network (RVFL). The RVFL was developed for training a functional-link network (FLN) [34]. In the RVFL, direct connections from the input nodes to the output nodes were allowed. Another version of the RVFL [35] considered the output bias term, which herein we name RVFL\*. Huang et al. [14] proposed a learning algorithm, referred to as extreme learning machine (ELM), for training SLFN. Unlike the original SNN, the bias term within the output layer of the ELM is not considered [11].

In the area of wind energy forecasting, several hybrids of decomposition techniques and NNRW have been proposed by researchers [18, 19, 20]. Abdoos [36] proposed a combination of the Gram-Schmidt orthogonalization (GSO), VMD, and ELM, called VMD-GSO-ELM. The VMD was utilized to decompose the wind power signal into several sub-series, and each decomposed sub-series was utilized to create the training patterns. Then, the GSO was employed as a feature selection method to eliminate irrelevant input features from each training dataset for ELM. Finally, the ELM was used as a forecasting model for each dataset with selected features. The experimental results showed that the VMD-GSO-ELM had faster learning speed than other iterative learning algorithms. Naik et al., 2018 [37] proposed a hybrid EMD and non-iterative learning approach for both wind speed and wind power predictions. Several non-iterative learning methods were selected and compared in this work, including kernel ridge regression (KRR), RVFL, and ELM. The experimental results showed that both a hybrid EMD and KRR (EMD-KRR), as well as a combination of EMD and RVFL (EMD-RVFL), could achieve promising results in applications of both wind speed and wind power forecasting. Their experimental results have also shown that although the predictive performance of EMD-RVFL was slightly lower than that of EMD-KRR, the training speed of EMD-RVFL was much faster than EMD-KRR. Moreover, there are a large number of parameters which must be set in the EMD-KRR algorithm and the EMD-KRR was very sensitive to the values of those parameters. This illustrates that these aforementioned hybrid techniques inherit the advantages of fast learning speed, and good generalization performance from NNRW. Moreover, they have shown excellent results working with time-series for wind power forecasting.

### 3. PROPOSED METHOD

Let  $\{\mathbf{f}_1, \dots, \mathbf{f}_M\}$  denote a set of  $M$  base models, where the  $m$ th base model  $\mathbf{f}_m$  is separately trained on  $\{\mathbf{X}, \mathbf{Y}\} = \{(\mathbf{x}_i, \mathbf{y}_i)\}_{i=1}^N$ . Here,  $\mathbf{x}_i = [x(1), \dots, x(n)] \in \mathbb{R}^n$  is the  $i$ th input sample and  $\mathbf{y}_i \in \mathbb{R}$  is the corresponding desired output. Suppose that the  $(\mathbf{x}_i, \mathbf{y}_i)$  for the base model  $\mathbf{f}_m$  can be decomposed into  $K$  components using the  $m$ th decomposition technique  $\mathcal{D}_m$ , that is  $\{(\hat{\mathbf{x}}_{i,1}^m, \hat{\mathbf{y}}_{i,1}^m), \dots, (\hat{\mathbf{x}}_{i,K}^m, \hat{\mathbf{y}}_{i,K}^m)\}$ .  $\hat{\mathbf{x}}_{i,k}^m$  and  $\hat{\mathbf{y}}_{i,k}^m$  denote the  $k$ th decomposed components of  $\mathbf{x}_i$  and  $\mathbf{y}_i$  using the  $m$  decomposition technique, respectively, where  $\hat{\mathbf{x}}_{i,k}^m = [\hat{x}_{i,k}^m(1), \dots, \hat{x}_{i,k}^m(n)] \in \mathbb{R}^n$ ,  $\hat{\mathbf{y}}_{i,k}^m \in \mathbb{R}$ ,  $\mathbf{x}_i = \sum_{k=1}^K \hat{\mathbf{x}}_{i,k}^m$ , and  $\mathbf{y}_i = \sum_{k=1}^K \hat{\mathbf{y}}_{i,k}^m$ . Here,  $\mathcal{D}_m \in \{\text{EMD, VMD, SSA, WPD, DWT}\}$ , and  $m = 1, \dots, M$ . Therefore, the base model  $\mathbf{f}_m$  for a sample  $\mathbf{x}_j$  with  $K$  components can be expressed as shown in Eq. (1).

$$\mathbf{f}_m(\mathbf{x}_j) = \sum_{k=1}^K f_k^m(\hat{\mathbf{x}}_{j,k}^m) \quad (1)$$

$f_k^m$  is the  $k$ th predictor within  $\mathbf{f}_m$ . In considering the influence of the type of network structure,  $f_k^m$  is given by Eq. (2).

$$f_k^m(\hat{\mathbf{x}}_{j,k}^m) = \sum_{i=1}^L \beta_{i,k}^m \sigma(\hat{\mathbf{x}}_{j,k}^m; \mathbf{w}_{i,k}^m, b_{i,k}^m) + \varphi(\hat{\mathbf{x}}_{j,k}^m, \mu) \quad (2)$$

$\hat{\mathbf{x}}_{j,k}^m$  represents the  $k$ th component decomposed from  $\mathbf{x}_j$  using the  $m$ th decomposition technique. Here,  $\mathbf{w}_{i,k}^m \in \mathbb{R}^n$  denotes the input weights that connect the input layer and the  $i$ th hidden node, and  $b_{i,k}^m$  is the bias of the  $i$ th hidden node. Both the weights and biases within the hidden layer are randomly generated based on a uniform distribution.  $\beta_{i,k}^m$  is the weight connecting the  $i$ th hidden node and the output layer of  $f_k^m$  within  $\mathbf{f}_m$ .  $\varphi$  denotes the structural function, which is adopted to define the type of network structure. The  $\varphi$  for an input  $\hat{\mathbf{x}}_{j,k}^m$  is formulated as shown in Eq. (3).

$$\varphi(\hat{\mathbf{x}}_{j,k}^m, \mu) = \begin{cases} 0, & \text{if } \mu = 0 \text{ (ELM)} \\ \beta_{0,k}^m, & \text{if } \mu = 1 \text{ (SNN)} \\ \sum_{l=L+1}^{L+n} \hat{x}_{j,k}^m(l-L)\beta_{l,k}^m, & \text{if } \mu = 2 \text{ (RVFL)} \\ \beta_{0,k}^m + \sum_{l=L+1}^{L+n} \hat{x}_{j,k}^m(l-L)\beta_{l,k}^m, & \text{if } \mu = 3 \text{ (RVFL*)} \end{cases} \quad (3)$$

$\beta_{0,k}^m$  and  $\beta_{l,k}^m$  represent the output bias and the weight, respectively, that connect the  $l$ th input node and the output layer of the  $\mathbf{f}_m$  for the  $k$ th component. Here, the SLFN structure is adopted if  $\mu \in \{0, 1\}$ , while the FLN structure is employed if  $\mu \in \{2, 3\}$ . The output weights within the  $m$ th base model for estimating the  $k$ th component are determined by the minimization in Eq. (4).

$$\operatorname{argmin}_{\{\beta_{i,k}^m\}} \left\{ \sum_{j=1}^N [\hat{\mathbf{y}}_{j,k}^m - f_k^m(\hat{\mathbf{x}}_{j,k}^m)]^2 \right\} \quad (4)$$

To combine the predicted results of all the base models, all the predicted results of the  $M$  base model are integrated through a linear combination method. Therefore, the ensemble output function can be written as shown in Eq. (5).

$$F(\mathbf{x}_j) = \omega_0 + \sum_{m=1}^M \omega_m \mathbf{f}_m(\mathbf{x}_j) \quad (5)$$

$\omega_m$  denotes the coefficient connecting the  $m$ th base model and the combination layer.

To obtain the optimal  $\{\omega_0, \dots, \omega_M\}$ , the objective function for minimizing the training error can be formulated as shown in Eq. (6).

$$\operatorname{argmin}_{\{\omega_m\}} \left\{ \sum_{j=1}^N \left[ \mathbf{y}_j - \left( \omega_0 + \sum_{m=1}^M \omega_m \mathbf{f}_m(\mathbf{x}_j) \right) \right]^2 \right\} \quad (6)$$

By using Eq. (6), the objective function can be rewritten in the matrix form as shown in Eqs. (7) to (9).

$$\begin{aligned} \mathcal{L} &= (\Phi \omega - \mathbf{Y})^\top (\Phi \omega - \mathbf{Y}) \\ &= \mathbf{Y}^\top \mathbf{Y} + \omega^\top \Phi^\top \Phi \omega - \mathbf{Y}^\top \Phi \omega - \omega^\top \Phi^\top \mathbf{Y} \\ &= \mathbf{Y}^\top \mathbf{Y} + \omega^\top \Phi^\top \Phi \omega - 2\omega^\top \Phi^\top \mathbf{Y} \end{aligned} \quad (7)$$

$$\Phi = \begin{bmatrix} 1 & \mathbf{f}_1(\mathbf{x}_1) & \dots & \mathbf{f}_M(\mathbf{x}_1) \\ \vdots & \vdots & \ddots & \vdots \\ 1 & \mathbf{f}_1(\mathbf{x}_N) & \dots & \mathbf{f}_M(\mathbf{x}_N) \end{bmatrix} \quad (8)$$

$$\omega = \begin{bmatrix} \omega_0 \\ \vdots \\ \omega_M \end{bmatrix} \quad \mathbf{Y} = \begin{bmatrix} \mathbf{y}_1 \\ \vdots \\ \mathbf{y}_N \end{bmatrix} \quad (9)$$

$\Phi \in \mathbb{R}^{N \times (M+1)}$  is the output matrix of the base models.  $\omega \in \mathbb{R}^{(M+1) \times 1}$  and  $\mathbf{Y} \in \mathbb{R}^{N \times 1}$  denote the ensemble weight and the desired output vectors, respectively.

Taking the derivative of  $\mathcal{L}$  with respect to  $\omega$  and setting the derivative equal to zero, we obtain Eq. (10).

$$\begin{aligned} \frac{\partial \mathcal{L}}{\partial \omega} = 0 &\rightarrow 2\Phi^\top \Phi \omega - 2\Phi^\top \mathbf{Y} = 0 \\ &\rightarrow \Phi^\top \Phi \omega = \Phi^\top \mathbf{Y} \end{aligned} \quad (10)$$

Assuming that  $\Phi^\top \Phi$  is an invertible matrix, the optimal least square solution of Eq. (7) is given by Eq. (11).

$$\omega = \Phi^\dagger \mathbf{Y} \quad (11)$$

$\Phi^\dagger = (\Phi^\top \Phi)^{-1} \Phi^\top$  is the generalized pseudoinverse of  $\Phi$ . However,  $\Phi^\top \Phi$  can be a non-invertible matrix. To avoid the non-invertible problem, the SVD is commonly employed to compute the generalized pseudoinverse in all cases [14]. Therefore, the SVD was adopted to compute  $\Phi^\dagger$  in this study.

*Theorem 1* ([38, 39]) Given  $\mathbf{P} \in \mathbb{R}^{n \times m}$  such that  $\mathbf{P}\mathbf{b}$  is the minimum norm least-square solution of  $\mathbf{A}\mathbf{x} = \mathbf{b}$ , where  $\mathbf{A} \in \mathbb{R}^{m \times n}$ , and  $\mathbf{b} \in \mathbb{R}^m$ . It is necessary and sufficient that  $\mathbf{P} = \mathbf{A}^\dagger$ , which is the generalized inverse of  $\mathbf{A}$ .

*Remarks 1:* According to Theorem 1, the proposed ensemble model has following important properties:

- $\mathbf{x}^* = \mathbf{A}^\dagger \mathbf{b}$  is the least-square solution of  $\mathbf{A}\mathbf{x} = \mathbf{b}$

$$\|\mathbf{A}\mathbf{x}^* - \mathbf{b}\| = \|\mathbf{A}\mathbf{A}^\dagger \mathbf{b} - \mathbf{b}\| = \underset{\mathbf{x}}{\operatorname{argmin}} \|\mathbf{A}\mathbf{x} - \mathbf{b}\| \quad (12)$$

- $\mathbf{x}^* = \mathbf{A}^\dagger \mathbf{b}$  has the minimum norm among all the other solutions of  $\mathbf{A}\mathbf{x} = \mathbf{b}$

$$\|\mathbf{x}^*\| = \|\mathbf{A}^\dagger \mathbf{b}\| \leq \|\mathbf{x}\|, \quad \forall \mathbf{x} \in \{\mathbf{x} : \|\mathbf{A}\mathbf{x} - \mathbf{y}\| \leq \|\mathbf{A}\mathbf{z} - \mathbf{y}\|, \forall \mathbf{z} \in \mathbb{R}^n\} \quad (13)$$

- $\mathbf{x}^* = \mathbf{A}^\dagger \mathbf{b}$  is the minimum norm least-squares solution of  $\mathbf{A}\mathbf{x} = \mathbf{b}$ , which is always unique.

The learning process of the proposed decomposition-based hybrid approach is summarized as follows:

1. Each decomposition technique is used to decompose the wind power series data. The time-series data of each decomposition method is decomposed into  $K$  decomposed components. In this step, five single decomposition techniques are adopted separately: EMD, VMD, SSA, DWT, and WPD.
2. The NNRW predictor is built to complete the forecasting computation for each decomposed component of each decomposition technique. In this step, four types of NNRW models are presented: ELM, SNN, RVFL, and RVFL\*.
3. The predicted signals of each decomposition technique are directly summed to built the reconstructed time-series of wind power through Eq. (1).
4. All of the reconstructed results are integrated by a linear combination method using Eq. (5). The weighted coefficients of this combiner can be obtained via Eq. (11).

## 4. CASE STUDY AND RESULTS DISCUSSION

### 4.1 Datasets specification and preparation

Twelve actual wind power datasets were retrieved from the *50Hertz Transmission GmbH* website. They are available at <https://www.50hertz.com/>. These datasets were collected over 12 months from January 1, 2018, to December 31, 2018, in Germany. These

data series were recorded at an interval of 15 minutes. The series of the wind power datasets were continuously recorded with the exception of March 25, 2018, from 2:00 to 2:45, and November 26, 2018, at 12:00 and 12:15, when data was not collected. The specification and statistical information including mean, maximum (Max.), minimum (Min.), standard deviation (SD), skewness (Skew.), and kurtosis (Kurt.) values of each dataset are detailed in Table 1.

**Table 1:** Statistical information for the wind power datasets.

Dataset	#Sample	Mean	Max.	Min.	SD	Skew.	Kurt.
Jan	2976	6092.09	14354.70	90.48	4266.99	0.30	1.78
Feb	2688	2771.60	10329.50	11.67	2283.86	0.91	3.02
Mar	2972	4458.47	13775.59	85.37	3522.37	0.81	2.75
Apr	2880	3886.14	12935.60	25.05	2865.42	0.49	2.28
May	2976	2988.28	11406.52	89.84	2097.93	1.05	4.22
Jun	2880	2545.68	12204.81	59.69	2071.62	1.79	6.50
Jul	2976	1911.56	6085.28	15.07	1440.31	0.90	2.94
Aug	2976	2567.85	9687.07	90.10	1963.07	1.11	3.90
Sept	2880	3307.30	13037.33	127.15	2844.73	1.24	3.97
Oct	2976	4884.26	15382.38	11.03	4060.25	0.78	2.61
Nov	2878	3837.43	12527.33	90.10	2943.53	0.83	2.79
Dec	2976	6388.35	15672.40	275.67	4022.16	0.47	2.26

All of the experiments were performed using 30 independent runs. In each run, the dataset was divided into training and test sets, in which the first 80% was designated for training, and the last 20% was assigned for testing. In our experiment, the min-max normalization method was adopted to scale the time-series data to values in the range of 0 to 1.

### 4.2 Evaluation metrics

To evaluate the predictive performance of different comparative algorithms, three well-known error measurement indices were considered. These metrics include the root mean square error (RMSE), mean absolute error (MAE), and mean absolute percentage error (MAPE). The details of these error measurements are given in Table 2.

**Table 2:** Evaluation metrics for measuring predictive accuracy.

Criteria	Formula
Root mean square error (RMSE)	$\sqrt{\frac{1}{N} \sum_{i=1}^N (\hat{y}_i - y_i)^2}$
Mean absolute error (MAE)	$\frac{1}{N} \sum_{i=1}^N  \hat{y}_i - y_i $
Mean absolute percentage error (MAPE)	$\frac{1}{N} \sum_{i=1}^N \left  \frac{\hat{y}_i - y_i}{y_i} \right  \times 100$

To further evaluate the enhancement of model  $A$  over model  $B$ , the improvement percentage of each criterion was exploited to illustrate the promotion degree, which can be expressed as shown in Eq. (14).

$$\mathcal{I}_\nu = \left| \frac{E_A - E_B}{E_B} \right| \times 100 \quad (14)$$

$\nu$  is the RMSE, MAE, or MAPE. Here,  $E_A$  and  $E_B$  are the evaluated values of model  $A$  and  $B$ , respectively, using measure  $\nu$ .

### 4.3 Comparative algorithms and parameter settings

To verify the effectiveness of the proposed EDNNRW, ten comparative algorithms were selected for comparison with the proposed method. The selected algorithms were ELM [14], SNN [12], RVFL [13], RVFL\* [35], VMD-WRELM [18], EMD-RVFL [37], CVAELM [19], WPD-EMD-ELM [40], CEEMDAN-ANN [41], and VMD-GSO-ELM [36].

Following previous studies [42, 28, 43], three-level decomposition of WPD was applied in this study. Since the three-level WPD provides eight frequency bands, the maximum number of decomposed components for each decomposition method was eight. For all algorithms, an additive sigmoid function was applied as the nonlinear mapping activation function for the hidden layer. The 15-minute historical data values of the wind power series in the past day (24 hours) were considered as the input for prediction of the desired value. The maximum number of lag orders (features) was empirically set to  $24 \times 4 = 96$ . The other parameters of each competing algorithm were set to the same as those used in the corresponding published research.

### 4.4 Comparison of statistical error measures

The predictive performance comparisons of the different algorithms in one, three, and five step ahead forecasting for the wind power predictions are tabulated in Tables 3 to 8. As shown in these tables, we find that the proposed EDNNRW<sub>RVFL</sub> and EDNNRW<sub>RVFL\*</sub> produce a relatively better forecasting accuracy than the other comparative algorithms in most cases. The average RMSE, MAE, and MAPE of the proposed EDNNRW<sub>ELM</sub> and EDNNRW<sub>SNN</sub> are generally lower than those of the ELM, SNN, VMD-WRELM, CVAELM, WPD-EMD-ELM, VMD-GSO-ELM, EMD-RVFL, and CEEMDAN-ANN. From these tables, it can be observed that RVFL, RVFL\* and the decomposition-based RVFL methods (EMD-RVFL, EDNNRW<sub>RVFL</sub>, and EDNNRW<sub>RVFL\*</sub>) have good forecasting abilities. This indicates that the direct connections between the input layer and the output layer can significantly improve the predictive performance of these models. Interestingly, we observed that the proposed EDNNRW<sub>RVFL</sub> and EDNNRW<sub>RVFL\*</sub> approaches do not need a large number of hidden nodes to attain good predictive performance. In Tables 6 to 8, we observed that the predictive performance of all the competitors decreases as the number of  $n$ -step ahead increases. This indicates that it gets harder to accurately capture the complex relationships existing in the multi-step ahead forecasting as  $n$  increases.

For multiple-comparison tests, the Friedman statistical test was employed to perform multiple-comparison tests for multiple-problem analysis, as suggested in [44]. Under the null hypothesis, the performance of all  $k$  competitors are equivalent, so their average ranks  $\mathcal{R}_j$  over all  $\mathcal{N}$  benchmarks should be equal. The Friedman statistic ( $\chi_F^2$ ) can be calculated as shown in Eq. (15).

$$\chi_F^2 = \frac{12\mathcal{N}}{k(k+1)} \left[ \sum_{j=1}^k \mathcal{R}_j^2 - \frac{k(k-1)^2}{4} \right] \quad (15)$$

$\mathcal{R}_j = \frac{1}{\mathcal{N}} \sum_{i=1}^{\mathcal{N}} r_{i,j}$ , and  $r_{i,j}$  represents the rank of the  $j$ th of  $k$  algorithms on the  $i$ th of  $\mathcal{N}$  benchmarks. The  $\chi_F^2$  is distributed according to the chi-square or  $\chi^2$ -distribution with  $k-1$  degree of freedom whenever the values of  $\mathcal{N}$  and  $k$  are sufficiently large. As a rule of a thumb,  $\mathcal{N} > 10$ , and  $k > 5$  [44].

Iman and Davenport [45] showed that the  $\chi_F^2$  is undesirably conservative, and presented an improved version of the  $\chi_F^2$ , called the Iman-Davenport test ( $F_F$ ) which is computed with Eq. (16).

$$F_F = \frac{(\mathcal{N}-1)\chi_F^2}{\mathcal{N}(k-1) - \chi_F^2} \quad (16)$$

The  $F_F$  is distributed according to the  $F$ -distribution with  $k-1$  and  $(k-1)(\mathcal{N}-1)$  degrees of freedom.

If the null-hypothesis is rejected, which means that the differences among the competitors are statistically significant, the Nemenyi post-hoc test can be applied to compare all the competitors with each other as previously suggested in [46, 47]. The performance of two among  $k$  competitors are considered to be significantly different if the difference of their corresponding average ranks is greater than the critical difference ( $C_D$ ). The value of the  $C_D$  for the Nemenyi post-hoc test is computed with Eq. (17).

$$C_D = q_\alpha \sqrt{\frac{k(k+1)}{6\mathcal{N}}} \quad (17)$$

$q_\alpha$  is the critical value that is based on the studentized range statistic divided by  $\sqrt{2}$ .  $\alpha$  is the significance level and was set to be 0.05 in this study.

In this experiment, the number of competitors is 13. There are twelve time-series datasets for this experiment. For each dataset, one, three, and five step ahead forecasting were considered. Three different sizes of hidden layers with 25, 50, and 100 nodes were tested. Three evaluation metrics were utilized. Therefore,  $\mathcal{N} = 12 \times 3^3 = 324$ , and  $k = 14$ .

In our case, the  $\chi_F^2$  value for this experiment is equal to 3213.41, and thus the  $F_F$  value is 1039.40. The critical value for the  $F$ -distribution with  $14-1 = 13$  and  $(14-1)(324-1) = 4199$  degrees of freedom at a 0.05 significance level is 1.72. Because the value of  $F_F$  is greater than the critical value of the  $F$ -distribution,

**Table 3:** Comparison of the RMSE for single-step ahead forecasting on the wind power datasets.

Algorithm	#Node	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sept	Oct	Nov	Dec
ELM	25	1077.8899	508.1707	527.9063	799.7369	350.5450	399.1899	472.0114	571.3072	638.9294	865.6619	669.4292	634.2632
	50	640.8990	354.2666	339.6481	533.5646	238.8330	241.0928	336.8541	381.3924	438.9887	595.6127	419.7418	451.4413
	100	454.1925	246.6544	217.2271	399.1912	160.9137	159.7372	252.7164	270.6281	317.4356	435.8968	323.5809	297.5217
SNN	25	1065.3440	492.8111	524.0015	776.6720	345.3606	382.4041	469.6051	548.7256	624.3776	858.9303	659.9267	625.1785
	50	633.9775	352.7503	337.9543	530.1204	236.5219	236.4948	333.0792	375.4767	430.1154	587.4965	414.6198	451.7378
	100	446.3632	246.3696	216.0351	399.2304	160.4021	157.8549	250.3847	270.2166	314.8736	431.3529	322.0117	297.0513
RVFL	25	125.6721	107.3927	83.8238	128.3062	87.1840	77.8762	82.0211	91.0204	95.0954	130.8247	90.8343	85.9773
	50	127.5425	108.9456	84.8563	133.5205	87.2764	80.1918	83.5758	93.0134	98.8170	135.5078	98.9546	88.0795
	100	134.3287	117.7148	88.2966	150.4055	88.5737	85.7771	89.5043	101.3499	113.0885	152.1445	120.6946	94.2397
RVFL*	25	125.8285	107.4328	83.8803	128.5537	87.1982	77.8577	82.0705	91.0582	95.1461	131.0043	91.0161	86.0730
	50	127.5367	108.9495	84.8750	133.9194	87.3500	80.2569	83.6643	93.0815	99.0298	135.7172	99.2281	88.2107
	100	134.3888	117.6728	88.2968	150.4742	88.6466	85.8525	89.6685	101.4722	114.0421	152.4270	120.8202	94.3402
VMD-WRELM	25	813.6324	625.3517	660.2039	1239.6683	404.7806	480.5825	561.3909	709.6684	673.3550	821.5793	563.0511	700.0600
	50	685.8568	653.7556	538.0673	1233.5268	336.3726	432.7172	655.0061	669.0840	646.9663	830.3886	536.6075	609.1592
	100	691.3892	717.4448	611.0715	1396.3089	318.2711	423.6853	711.3816	776.9645	660.1612	886.2147	514.1605	632.5330
EMD-RVFL	25	344.5322	233.1965	89.1176	328.8172	154.3878	143.9844	566.5544	196.6954	62.6973	163.336	217.9763	1529.3758
	50	376.6307	230.4858	89.1513	336.6596	157.7874	142.4879	569.9852	200.121	63.1513	205.1379	227.4321	1926.329
	100	374.5026	385.8259	87.5015	349.8421	160.3974	283.7791	723.8319	203.7918	64.8406	415.6875	290.8698	2001.6393
CVAELM	25	889.6870	560.7026	693.9573	1787.8637	366.4071	425.0032	462.9788	681.3635	1350.4911	1178.0192	501.7303	708.9714
	50	1108.0442	573.0517	813.2144	2965.5171	382.5068	453.0623	653.9089	838.4289	2391.7360	2785.2966	1223.6144	764.7904
	100	1525.5777	1172.7979	1349.8339	7406.1711	436.6313	479.0533	1951.2137	3094.4727	13309.0908	7244.2236	3756.7232	1268.4420
WPD-EMD-ELM	25	907.3720	307.7888	1272.7869	800.6908	359.6207	244.3873	304.3849	707.8707	1114.4254	1075.6277	155.6857	517.0279
	50	913.7809	165.8250	1027.8840	572.7348	211.2338	139.6895	214.8595	280.2406	973.1849	617.3614	109.8208	356.9260
	100	607.2151	110.5022	941.8276	501.1405	157.1439	84.8838	150.7874	176.0198	1015.8936	426.0218	70.8776	432.6715
CEEMDAN-ANN	25	1905.8862	1075.3154	1272.1952	1513.8517	659.0357	746.7405	823.1969	986.3195	1295.0372	1694.6829	1086.0782	1333.3237
	50	1764.5858	1110.7394	1154.1826	1504.8091	651.8736	742.9498	798.8099	955.1060	1262.1357	1984.9719	1029.1094	1312.6860
	100	1896.8751	1257.1964	1127.0232	1638.7462	689.8876	742.8222	891.4711	1106.2863	1341.9202	1863.1557	1174.8174	1314.3089
VMD-GSO-ELM	25	881.2108	345.5984	278.9920	510.9522	226.6726	260.9036	245.2514	353.3819	397.4226	566.3508	533.7089	432.2714
	50	666.3918	232.9827	187.5592	399.7784	139.2565	172.6345	162.2051	244.7778	310.0748	400.7444	409.9243	315.5440
	100	539.3314	171.8538	138.8488	330.9926	105.3167	113.3443	125.9376	182.4463	234.0031	304.4986	326.5897	275.8758
EDNNRW <sub>ELM</sub>	25	505.0901	232.2337	275.9698	508.7423	165.0745	169.9321	222.9359	274.8745	321.3706	413.5750	323.9303	993.4474
	50	373.3667	145.0110	174.2149	381.1074	89.8081	96.8357	184.2217	174.3016	165.9073	243.5601	177.2189	1461.4089
	100	241.0919	118.7535	102.7784	280.3973	49.4380	61.4337	136.0190	130.0768	119.5172	160.1496	150.1401	956.7388
EDNNRW <sub>SNN</sub>	25	500.0111	233.0730	267.8660	485.8635	158.7690	166.3369	220.5385	274.1589	316.3518	404.8543	234.2711	1042.1312
	50	372.5830	147.1491	168.7073	380.1015	88.1280	96.7649	183.0695	174.5345	162.5399	242.3527	171.8997	2646.3118
	100	238.8608	121.5439	103.0565	284.0253	49.0589	61.2850	134.4357	133.1456	119.3373	160.0442	149.9189	1158.5729
EDNNRW <sub>RVFL</sub>	25	7.8874	4.6234	4.1081	5.4890	3.2465	4.9058	5.0988	<b>3.6484</b>	<b>4.1831</b>	4.8844	<b>3.0902</b>	5.0826
	50	7.8479	4.9563	4.1367	5.5655	3.2560	4.9385	5.2417	3.7373	4.3464	4.9898	3.2849	6.3320
	100	7.9060	5.7693	4.2774	6.2961	3.2945	5.1106	5.6658	4.1276	4.7403	5.4126	3.9753	26.2878
EDNNRW <sub>RVFL</sub> *	25	<b>7.6434</b>	<b>4.6755</b>	<b>4.0416</b>	<b>5.4695</b>	<b>3.2286</b>	<b>4.8631</b>	<b>4.9818</b>	3.7185	4.2110	<b>4.8694</b>	3.1984	<b>5.0437</b>
	50	7.6572	4.8798	4.0999	5.5545	3.2530	4.8929	5.1580	3.8004	4.3726	4.9759	3.4146	8.3492
	100	7.7463	5.6926	4.2486	6.2588	3.3120	5.0712	5.5942	4.2098	4.7692	5.3775	4.0908	34.2599

**Table 4:** Comparison of the MAE for single-step ahead forecasting on the wind power datasets.

Algorithm	#Node	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sept	Oct	Nov	Dec
ELM	25	815.8571	386.9769	379.2840	597.7381	275.4511	299.3291	365.1250	430.5173	503.4171	685.3628	437.8410	475.5373
	50	476.1355	265.3042	238.5636	389.2144	186.9524	178.9438	258.8831	282.6584	338.7535	471.2694	273.5344	336.7140
	100	318.6395	182.0499	150.4487	281.1206	124.6839	116.9624	190.2233	194.2716	239.7398	338.6025	206.3221	223.7423
SNN	25	805.9921	376.1213	374.4639	575.4442	270.8105	287.2023	362.9652	413.8325	490.1372	678.1006	432.5123	468.1485
	50	471.6339	263.3988	235.8666	385.3006	185.2271	175.6564	256.0916	278.4988	330.9019	464.3356	269.4299	336.2978
	100	314.9176	181.1263	149.9874	279.8402	124.2716	115.6614	188.7571	193.6917	237.4048	335.4531	204.6273	222.9956
RVFL	25	90.4460	68.5560	58.3082	90.3931	63.4333	57.4272	56.9510	66.4128	71.7640	99.5213	64.4714	64.8439
	50	91.8332	70.8246	59.0609	94.6415	63.6101	59.3000	58.7233	68.2711	74.4287	103.4544	70.0130	66.8110
	100	96.9845	80.3043	61.7505	107.8839	64.8475	63.2891	64.9440	74.1059	84.8875	116.6331	83.8304	71.5084
RVFL*	25	90.5517	68.5904	58.3620	90.5826	63.4583	57.4146	57.0295	66.4555	71.8004	99.6885	64.5660	64.9213
	50	91.8301	70.8405	59.1242	94.8865	63.7009	59.3602	58.8425	68.3238	74.5785	103.5803	70.1905	66.9383
	100	97.0713	80.3327	61.7846	108.0634	64.9111	63.3118	65.0848	74.1730	85.7452	116.8851	83.9385	71.5928
VMD-WRELM	25	645.3862	488.3664	419.9084	898.3230	321.3213	368.1130	423.4135	551.6838	540.1781	665.3102	442.6742	505.8623
	50	542.0940	490.4546	357.1310	831.2453	263.3951	330.7619	468.1280	515.8237	513.7614	644.2939	419.1064	437.4046
	100	540.3123	515.3404	385.3128	933.6569	247.9951	330.2487	513.4009	585.0001	523.6549	675.0130	397.5713	441.9259
EMD-RVFL	25	75.3936	50.3246	34.1455	68.4339	46.9466	48.0777	112.1531	52.8543	36.7794	70.8037	52.5839	261.5778
	50	77.7332	57.9072	34.3903	70.767	47.4587	50.2838	117.17685	53.9788	38.2965	84.4715	55.2208	306.808
	100	78.5829	161.5351	34.9845	83.9616	48.4928	157.541	139.2714	57.4025	41.94	205.9874	77.6723	318.7244
CVAELM	25	695.6507	450.1853	515.8671	1138.3536	286.5855	335.8334	350.4484	515.9827	937.6208	806.1816	373.9924	536.6474
	50	827.2152	449.3084	569.3605	1658.5690	302.5415	358.1443	459.7391	574.4164	1482.9795	1542.7116	788.1099	545.8501
	100	1056.4935	848.6936	836.1761	3536.0303	335.6564	354.2607	1045.3076	1658.8732	5639.0706	3610.6457	2076.0189	770.3259
WPD-EMD-ELM	25	551.4060	238.8270	726.3307	541.2890	262.7741	191.2006	227.5170	477.9326	607.0901	558.8497	116.7970	307.1884
	50	436.15											

**Table 5:** Comparison of the MAPE for single-step ahead forecasting on the wind power datasets.

Algorithm	#Node	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sept	Oct	Nov	Dec
ELM	25	19.0714	10.8339	22.5434	28.6360	18.9860	14.0783	31.0597	21.0864	17.5301	13.2591	12.3757	15.0837
	50	11.4012	7.7611	13.6877	18.1869	12.8655	8.4980	21.5434	13.5902	11.2088	8.7281	7.6228	12.7084
	100	7.3628	5.1955	8.1340	11.8284	8.5794	5.3886	15.2657	8.8876	7.8598	6.2778	5.3055	8.4465
SNN	25	18.7381	10.5954	21.6196	27.0012	18.5100	13.6347	30.6579	20.1414	17.0254	13.1517	12.2841	15.1862
	50	11.1968	7.7215	13.4755	17.8115	12.7293	8.3300	21.3890	13.3809	10.9761	8.6173	7.4367	12.7792
	100	7.2769	5.1600	8.1464	11.6038	8.5189	5.3385	15.2433	8.8558	7.7955	6.1855	5.1900	8.4155
RVFL	25	2.0527	1.8040	2.6274	3.4194	3.8771	2.6052	3.8113	2.7626	2.3726	1.6197	2.1295	1.3873
	50	2.0836	1.8728	2.6891	3.7222	3.7944	2.6917	3.9910	2.9054	2.4615	1.7270	2.2502	1.6064
	100	2.2106	2.1505	2.9182	4.4795	3.8627	2.8499	4.5932	3.2672	2.7752	2.0574	2.4832	1.9035
RVFL*	25	2.0568	1.8054	2.6350	3.4216	3.8698	2.6055	3.8193	2.7663	2.3736	1.6236	2.1309	1.3934
	50	2.0848	1.8736	2.6959	3.7350	3.8002	2.6955	3.9982	2.9096	2.4656	1.7309	2.2546	1.6244
	100	2.2128	2.1505	2.9260	4.4809	3.8658	2.8510	4.6113	3.2719	2.7926	2.0620	2.4872	1.9142
VMD-WRELM	25	22.8170	13.7216	24.1591	40.2863	25.0575	21.2942	41.7686	39.9625	21.6897	13.3948	17.8272	15.7598
	50	21.1606	14.2085	22.0551	32.2838	19.2460	17.6767	46.7381	33.3791	20.1960	12.5352	16.6224	15.1090
	100	21.6433	14.7236	22.9083	33.1818	16.8616	16.8434	49.8259	33.1635	20.4968	12.8901	15.4296	14.0311
EMD-RVFL	25	1.3975	1.0782	1.4728	2.0294	2.2979	1.8153	5.7781	2.9774	1.3509	1.6984	1.7889	4.5273
	50	1.4366	1.2361	1.5394	2.1337	2.3216	1.9205	6.3674	3.0474	1.4394	2.0549	1.8622	5.6718
	100	1.4518	3.5867	1.6182	2.5089	2.3753	8.191	7.7054	3.2238	1.6746	4.3943	2.4587	7.9021
CVAELM	25	25.7810	13.1272	35.3606	45.6340	18.0660	17.0745	31.7731	27.3987	32.0918	14.3997	14.6504	17.7185
	50	25.2309	13.0346	35.9238	57.6773	20.1139	17.8089	43.9052	30.1831	42.4472	26.9993	23.4337	21.1156
	100	26.4516	24.7665	47.2477	106.1666	22.7581	16.5511	95.9618	93.0120	114.1415	61.9417	44.1566	32.8827
WPD-EMD-ELM	25	30.2686	6.7387	42.5932	47.3757	15.9584	9.6716	19.2425	29.6637	20.4512	10.0715	4.8729	11.7394
	50	21.7418	3.5348	27.8078	32.4750	8.9213	5.3562	13.2426	12.1165	22.7088	6.3604	3.4004	8.3056
	100	14.1987	2.2665	18.5993	25.9765	5.7635	3.0920	9.2606	8.0996	13.7385	4.1275	2.1544	8.3349
CEEMDAN-ANN	25	49.0757	24.5001	91.8048	55.3903	37.5744	28.8084	58.3581	44.4531	39.8705	27.5509	33.0076	32.8728
	50	48.1933	26.1839	83.5642	57.8535	36.9776	29.7961	57.6330	42.3377	37.9342	31.9632	32.4556	
	100	53.5551	30.0813	75.3938	62.0737	40.0481	29.6424	60.8706	50.3714	39.5476	28.9557	35.1428	32.5400
VMD-GSO-ELM	25	11.1375	7.4982	13.0015	15.3906	11.2672	9.5049	14.9598	14.8694	10.7605	8.6001	7.4976	10.0999
	50	7.9357	4.8664	8.2622	11.8099	6.6775	6.1988	9.6686	10.8448	8.2195	5.8902	5.3546	8.7225
	100	6.1516	3.2854	5.8359	9.2806	5.0796	4.2643	7.3995	7.8700	6.3872	4.6048	4.1634	8.8351
EDNNRW <sub>ELM</sub>	25	13.4722	4.9183	11.7116	12.8007	8.4419	5.8104	14.0766	12.4318	8.2873	6.3834	4.3132	23.6246
	50	9.9242	3.0448	6.1558	7.5497	4.9608	3.1859	9.7110	7.0780	4.2891	3.6854	2.6477	26.6362
	100	5.3058	2.2510	3.2762	5.1028	2.6780	1.8377	6.6828	4.6075	2.9293	2.4248	1.6971	16.9545
EDNNRW <sub>SNN</sub>	25	13.3533	4.8803	10.9287	12.1103	8.1190	5.6027	13.8343	12.1533	8.0956	6.2799	4.1623	25.9442
	50	9.4506	3.0710	6.0146	7.3917	4.8923	3.1554	9.7471	7.0901	4.2053	3.6608	2.5630	40.8008
	100	5.1107	2.2849	3.2209	5.0950	2.6435	1.8230	6.6751	4.6705	2.9221	2.4248	1.6925	16.9076
EDNNRW <sub>RVFL</sub>	25	0.0980	0.0889	0.1330	<b>0.1501</b>	0.1552	0.1068	0.2011	<b>0.1290</b>	0.1090	0.0661	0.0740	<b>0.0615</b>
	50	0.0985	0.0954	0.1357	0.1559	0.1573	0.1095	0.2168	0.1335	0.1135	0.0677	0.0774	0.1195
	100	0.1067	0.1089	0.1378	0.1814	0.1616	0.1201	0.2601	0.1508	0.1255	0.0756	0.0850	0.5204
EDNNRW <sub>RVFL*</sub>	25	<b>0.0965</b>	<b>0.0888</b>	<b>0.1292</b>	0.1506	<b>0.1539</b>	<b>0.1061</b>	<b>0.1931</b>	0.1312	<b>0.1071</b>	<b>0.0643</b>	<b>0.0720</b>	0.0688
	50	0.0977	0.0942	0.1335	0.1568	0.1567	0.1098	0.2131	0.1347	0.1121	0.0667	0.0762	0.1452
	100	0.1065	0.1083	0.1373	0.1808	0.1613	0.1222	0.2549	0.1523	0.1272	0.0744	0.0846	0.6853

we reject the null hypothesis that the predictive performance of all competitors are statistically equivalent.

Based on this null hypothesis rejection, the Nemenyi post-hoc test was conducted to determine whether the predictive performances of two among  $k$  competitors are significantly different. The value of  $q_\alpha$  for a 0.05 significance level is  $q_{0.05} = 3.354$ , which can be confirmed in standard statistical textbooks. Thus, the value of the  $C_D$  is equal to  $3.354\sqrt{\frac{14(14+1)}{1944}} \approx 1.102$ . The statistical results of the post-hoc analyses for the wind forecasting are presented using a critical difference diagram, as shown in Figure 1. In this figure, the algorithms with higher ranks (lower numbers) are preferable to those with lower ranks (higher numbers). Statistically equivalent algorithms are grouped into a *clique*, represented by a red horizontal bar.

In Figure 1, the overall performance of RVFL, RVFL\*, EDNNRW<sub>ELM</sub>, and EDNNRW<sub>SNN</sub> were comparable. This figure shows that the overall predictive performance of the proposed EDNNRW<sub>RVFL</sub>, and EDNNRW<sub>RVFL\*</sub> were significantly superior to the other competitors. The overall predictive performance of VMD-GSO-ELM and EMD-RVFL were significantly better than those of VMD-WRELM, CVAELM, WPD-EMD-ELM, and CEEMDAN-ANN. Interestingly, we observed that the FLN family

approaches (RVFL and RVFL\*) have good average ranks when compared with ELM, SNN, VMD-WRELM, CVAELM, WPD-EMD-ELM, VMD-GSO-ELM, and CEEMDAN-ANN. It is noteworthy that the proposed EDNNRW<sub>RVFL\*</sub> and EDNNRW<sub>RVFL</sub> had statistically significantly better average ranks than the other comparative algorithms, and their clique was located far from the other cliques with large gaps.

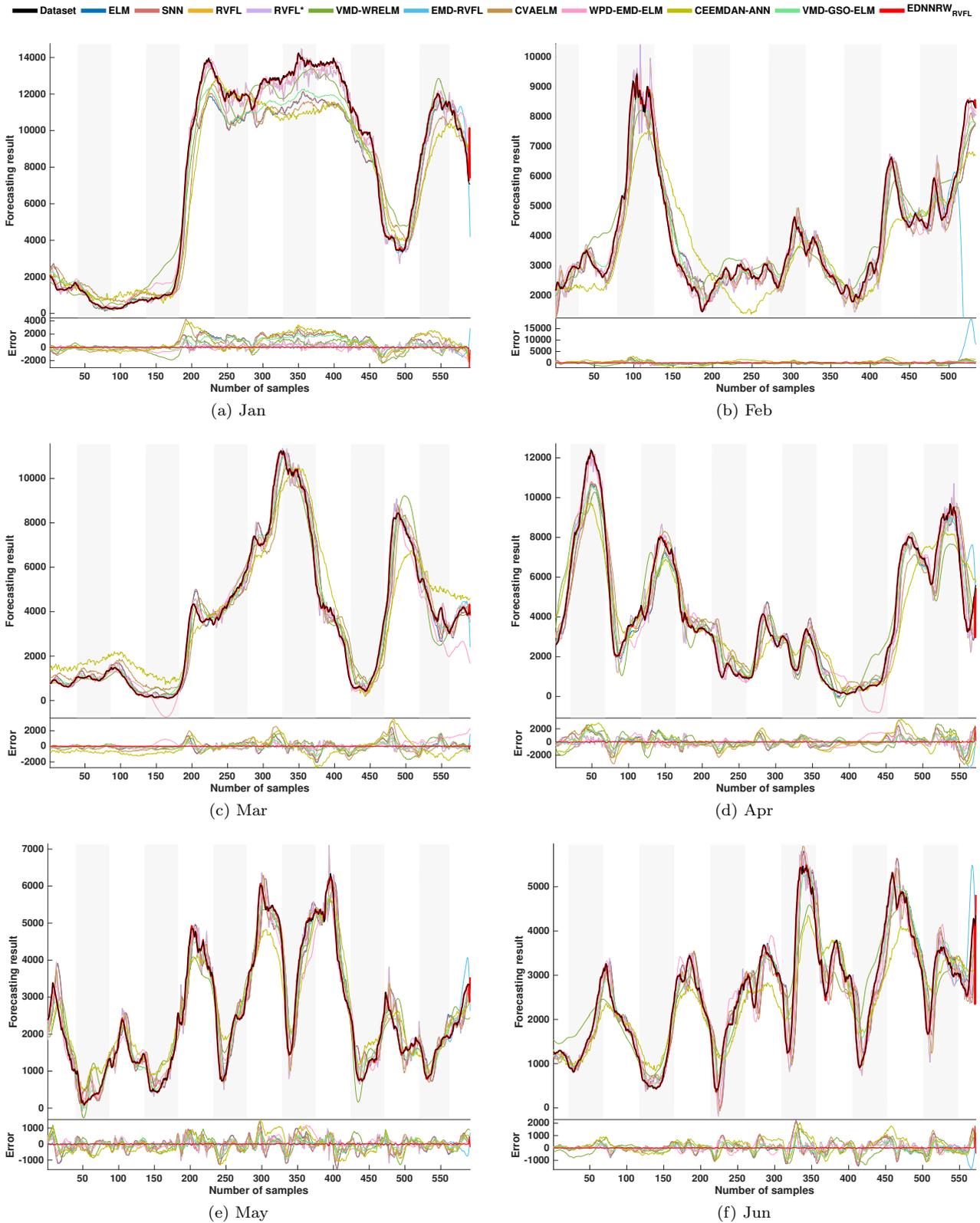
The forecasting results and the corresponding residual errors of the different competitors in five step ahead forecasting for the wind power forecasts are depicted in Figure 2. From this figure, it can be observed that the proposed EDNNRW<sub>RVFL</sub> has good forecasting abilities and its residual errors are closer to zero than those of the other comparative algorithms in most cases.

#### 4.5 Comparison of improvement percentages

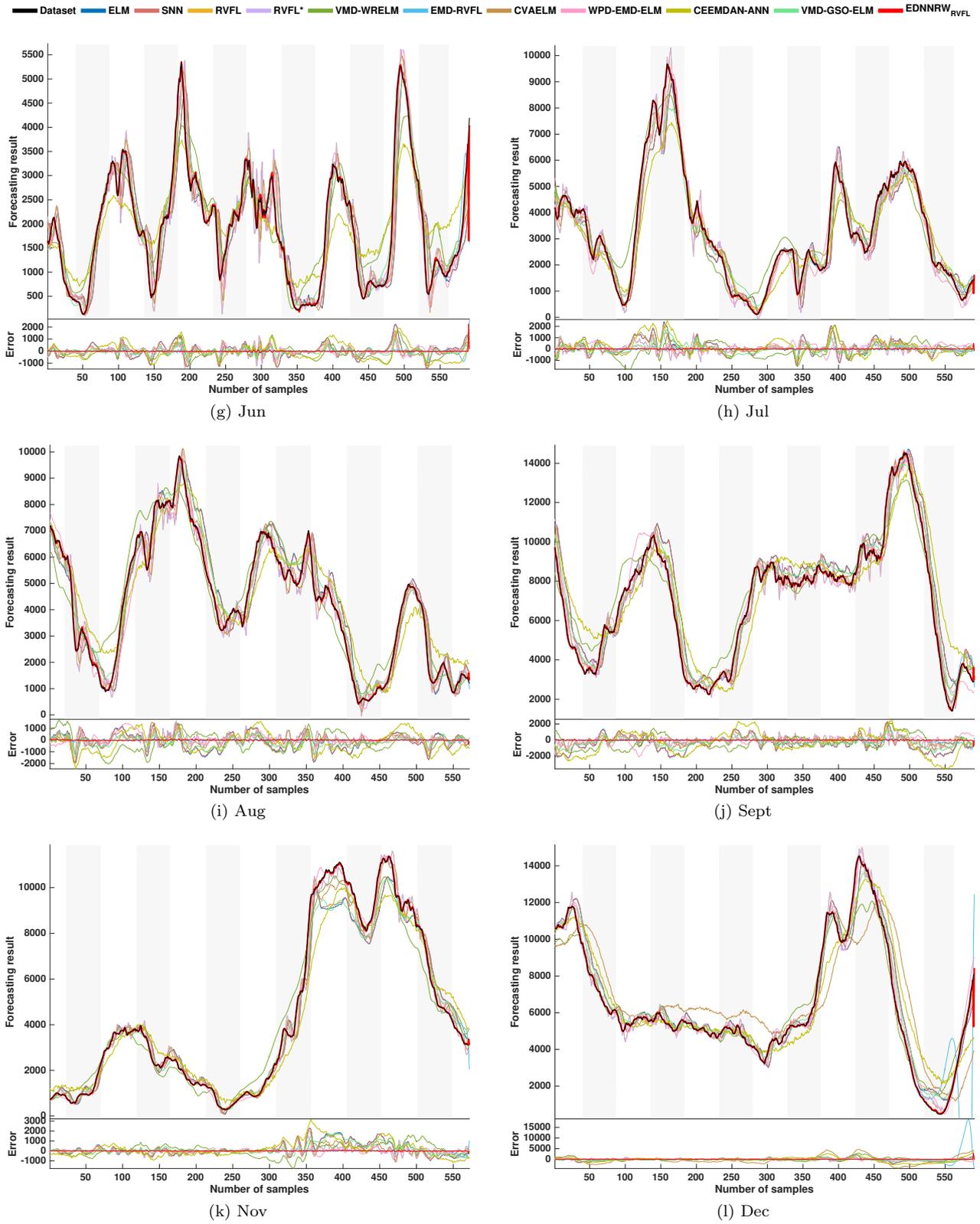
To further exhibit the effectiveness of the proposed EDNNRW<sub>RVFL</sub>, the improvement percentages in terms of RMSE, MAE, and MAPE indices were used for analysis. The improvement percentages of the proposed EDNNRW<sub>RVFL</sub> over ELM, SNN, RVFL, RVFL\*, VMD-WRELM, EMD-RVFL, CVAELM, WPD-EMD-ELM, VMD-GSO-ELM, and CEEMDAN-ANN in terms of RMSE, MAE, and MAPE for the wind power forecast in one, three, and







**Fig.2:** Results of five step ahead forecasting and the corresponding residual errors for the wind power datasets. Grey shaded regions represent the intervals at night time.



**Fig.2:** Results of five step ahead forecasting and the corresponding residual errors for the wind power datasets. Grey shaded regions represent the intervals at night time (Cont.).

predictive performance of the decomposition-based hybrid model.

- The performance of the proposed  $EDNNRW_{RVFL}$  is relatively superior to the other comparative algorithms in terms of forecasting capability, thereby indicating a significant improvement exists in the predictive performance of the  $EDNNRW_{RVFL}$ .

#### 4.6 Comparison of computational times

All experiments were conducted in the MATLAB environment and run on a personal computer with an Intel Core i7-3370 3.40 GHz processor, 8 GB of RAM, and Windows 7 x64 operating system. The computational times of all competitors were obtained using the *tic* and *toc* commands in the MATLAB program. The average computational time of each algorithm across all the problems is shown in Figure 3. This figure shows that the computational times of the NNRW algorithms (ELM, SSN, RVFL, and RVFL\*) were faster among all competing algorithms, due to the benefit of random weights generation and the closed-form least-squares solution. The computational speeds of the proposed  $EDNNRW$  algorithms were much faster than that of the CEEMDANN-ANN. In the CEEMDANN-ANN, the estimators must be iteratively fine-tuned by the back-propagation algorithm to obtain the optimal weight parameters. Consequently, the CEEMDANN-ANN was the most time-consuming technique. This supports our hypothesis that the algorithm for training the predictors in the decomposition-based method should be a non-iterative learning approach. As seen in Figure 4, the proposed  $EDNNRW_{ELM}$ ,  $EDNNRW_{SNN}$ ,  $EDNNRW_{RVFL}$ , and  $EDNNRW_{RVFL*}$  achieved good trade-offs between predictive performance and computational speed compared to other decomposition-based hybrid methods. Although the computational speeds of the proposed  $EDNNRW_{RVFL}$  and  $EDNNRW_{RVFL*}$  were slower than some decomposition-based approaches, the forecasting accuracy obtained by the proposed methods is dramatically increased. In practical applications, the additional accuracy is worth the extra computational time.

#### 5. CONCLUSION

We developed an improved decomposition-based hybrid approach for wind power forecasting using EMD, VMD, SSA, DWT, WPD, NNRW, and a linear combiner. In our approach, each decomposition technique is applied to decompose the original time-series data into a collection of components. The NNRW is then exploited as an estimator for each decomposed component. After the reconstruction of the predicted values, the reconstructed results of all of the decomposition techniques are combined with a linear combiner. The main advantage of our approach is that

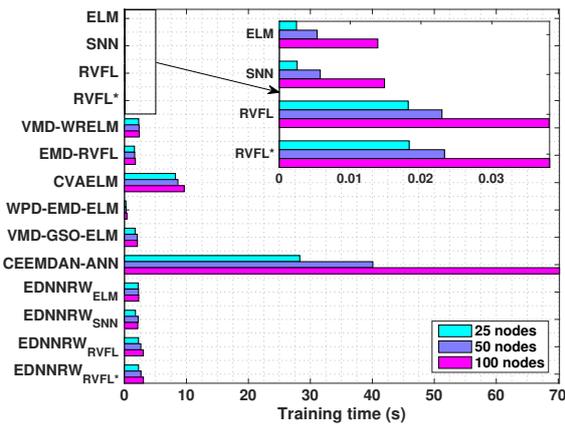


Fig.3: Computation time of different comparative methods.

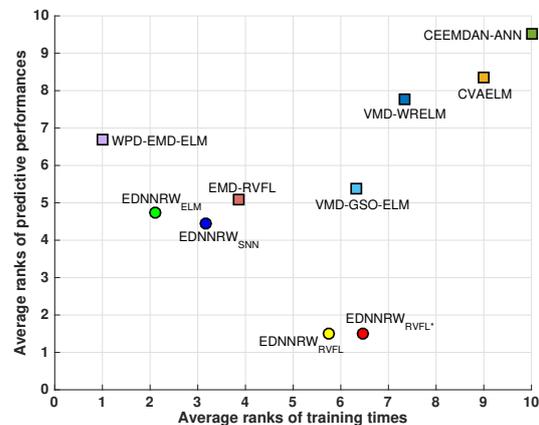


Fig.4: Trade-offs between predictive performance and computational time of different decomposition-based hybrid methods.

the valuable characteristics of several decomposition techniques are combined.

The experimental results lead to the following conclusions:

- The proposed  $EDNNRW_{ELM}$  and  $EDNNRW_{SNN}$  have good average ranks and were significantly superior to the other decomposition-based ELM methods and single models with a confidence of 95%. This indicates that the heterogeneous combination of different decomposition-based models can improve the forecasting capability of the proposed model.
- When comparing both the  $EDNNRW_{RVFL}$  and  $EDNNRW_{RVFL*}$  with the  $EDNNRW_{ELM}$  and  $EDNNRW_{SNN}$ , the forecasting accuracies of the former methods were higher than the latter approaches. The FLN family approaches (RVFL and RVFL\*) generated greater forecasting accuracy for the developed decomposition-based hybrid framework.

**Table 9:** Improvement percentages of the RMSE results of EDNNRW<sub>RVFL</sub> over the other competitors.

EDNNRW <sub>RVFL</sub> vs	n-step	Dataset											
		Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sept	Oct	Nov	Dec
ELM	1-step	98.80 %	98.50 %	98.67 %	98.92 %	98.56 %	97.86 %	98.42 %	98.97 %	98.94 %	99.09 %	99.12 %	98.36 %
	3-step	93.81 %	96.09 %	96.35 %	94.94 %	96.24 %	92.10 %	92.82 %	96.40 %	96.55 %	97.02 %	96.36 %	92.34 %
	5-step	86.30 %	92.63 %	93.50 %	87.28 %	92.75 %	84.99 %	83.21 %	92.46 %	93.16 %	93.73 %	93.13 %	75.77 %
SNN	1-step	98.78 %	98.49 %	98.65 %	98.91 %	98.55 %	97.81 %	98.40 %	98.95 %	98.93 %	99.08 %	99.11 %	98.35 %
	3-step	93.74 %	96.06 %	96.33 %	94.90 %	96.22 %	92.02 %	92.74 %	96.36 %	96.49 %	96.98 %	96.34 %	92.33 %
	5-step	86.21 %	92.61 %	93.47 %	87.18 %	92.69 %	84.91 %	83.11 %	92.43 %	93.11 %	93.69 %	93.07 %	75.71 %
RVFL	1-step	94.30 %	95.59 %	95.19 %	95.96 %	96.34 %	93.99 %	94.01 %	96.08 %	95.71 %	96.36 %	96.55 %	93.15 %
	3-step	84.79 %	93.72 %	93.42 %	90.61 %	93.96 %	87.28 %	86.61 %	93.82 %	93.63 %	93.98 %	93.61 %	84.21 %
	5-step	74.78 %	90.38 %	90.83 %	82.57 %	90.28 %	79.36 %	75.66 %	90.13 %	90.48 %	90.55 %	90.44 %	62.95 %
RVFL*	1-step	94.30 %	95.59 %	95.19 %	95.96 %	96.35 %	94.00 %	94.02 %	96.09 %	95.72 %	96.36 %	96.55 %	93.16 %
	3-step	84.82 %	93.74 %	93.42 %	90.65 %	93.96 %	87.29 %	86.63 %	93.83 %	93.65 %	94.01 %	93.64 %	84.26 %
	5-step	74.87 %	90.39 %	90.84 %	82.61 %	90.30 %	79.36 %	75.73 %	90.15 %	90.51 %	90.57 %	90.45 %	63.10 %
VMD-WRELM	1-step	98.98 %	99.25 %	99.28 %	99.56 %	99.08 %	98.90 %	99.19 %	99.48 %	99.33 %	99.39 %	99.33 %	99.03 %
	3-step	93.80 %	97.22 %	97.64 %	97.34 %	96.65 %	93.90 %	95.09 %	97.64 %	97.42 %	97.78 %	97.00 %	94.07 %
	5-step	84.38 %	93.36 %	94.93 %	91.82 %	91.82 %	83.49 %	85.61 %	93.79 %	93.81 %	94.32 %	92.92 %	77.10 %
EMD-RVFL	1-step	97.94 %	98.05 %	95.34 %	98.35 %	97.96 %	97.02 %	99.02 %	98.13 %	93.10 %	97.41 %	98.51 %	99.47 %
	3-step	84.95 %	96.59 %	86.69 %	91.80 %	92.40 %	94.00 %	94.18 %	91.80 %	90.09 %	98.06 %	95.32 %	97.46 %
	5-step	61.30 %	96.27 %	80.01 %	87.09 %	84.86 %	89.80 %	87.63 %	80.65 %	88.35 %	97.57 %	91.75 %	92.26 %
CVAELM	1-step	99.01 %	99.13 %	99.30 %	99.68 %	98.99 %	98.99 %	98.99 %	99.42 %	99.62 %	99.56 %	99.34 %	99.03 %
	3-step	93.80 %	97.25 %	97.59 %	98.36 %	96.60 %	92.92 %	94.78 %	97.37 %	99.02 %	98.49 %	97.15 %	94.78 %
	5-step	89.88 %	91.49 %	95.62 %	90.63 %	90.12 %	85.43 %	76.04 %	92.04 %	91.53 %	91.55 %	91.03 %	90.26 %
WPD-EMD-ELM	1-step	98.43 %	96.77 %	99.42 %	99.01 %	98.43 %	96.16 %	97.41 %	98.46 %	99.35 %	98.98 %	96.13 %	98.04 %
	3-step	92.72 %	91.53 %	98.24 %	95.63 %	95.54 %	85.32 %	89.40 %	94.87 %	98.02 %	96.79 %	86.70 %	91.09 %
	5-step	84.67 %	84.64 %	96.82 %	89.93 %	91.47 %	71.16 %	77.11 %	89.74 %	96.45 %	93.84 %	75.19 %	74.19 %
CEEMDAN-ANN	1-step	99.58 %	99.55 %	99.64 %	99.64 %	99.51 %	99.33 %	99.37 %	99.62 %	99.64 %	99.71 %	99.66 %	99.51 %
	3-step	97.29 %	98.29 %	98.63 %	97.74 %	98.13 %	96.35 %	96.08 %	98.25 %	98.57 %	98.75 %	98.46 %	97.09 %
	5-step	92.57 %	95.89 %	96.86 %	92.38 %	95.42 %	90.26 %	88.30 %	95.32 %	96.24 %	96.85 %	96.10 %	88.35 %
VMD-GSO-ELM	1-step	98.85 %	97.78 %	97.73 %	98.59 %	97.71 %	96.93 %	96.84 %	98.43 %	98.49 %	98.67 %	99.08 %	98.03 %
	3-step	93.23 %	92.37 %	92.85 %	91.92 %	92.53 %	86.27 %	83.59 %	93.71 %	94.19 %	94.91 %	95.70 %	88.73 %
	5-step	83.30 %	82.66 %	85.30 %	76.75 %	84.70 %	69.48 %	59.19 %	84.97 %	86.20 %	87.65 %	90.12 %	62.14 %

**Table 10:** Improvement percentages of the MAE results of EDNNRW<sub>RVFL</sub> over the other competitors.

EDNNRW <sub>RVFL</sub> vs	n-step	Dataset											
		Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sept	Oct	Nov	Dec
ELM	1-step	99.04 %	98.60 %	98.80 %	98.96 %	98.60 %	98.49 %	98.85 %	98.94 %	99.00 %	99.11 %	99.09 %	98.83 %
	3-step	96.90 %	96.42 %	96.68 %	96.58 %	96.51 %	95.85 %	96.10 %	96.54 %	96.65 %	97.20 %	96.58 %	96.15 %
	5-step	93.71 %	93.21 %	94.02 %	92.43 %	93.79 %	92.95 %	92.65 %	93.06 %	93.46 %	94.49 %	93.63 %	89.48 %
SNN	1-step	99.03 %	98.59 %	98.79 %	98.95 %	98.59 %	98.46 %	98.84 %	98.92 %	98.98 %	99.10 %	99.08 %	98.83 %
	3-step	96.87 %	96.39 %	96.66 %	96.56 %	96.49 %	95.81 %	96.06 %	96.51 %	96.59 %	97.17 %	96.55 %	96.14 %
	5-step	93.67 %	93.18 %	94.01 %	92.36 %	93.74 %	92.90 %	92.61 %	93.03 %	93.39 %	94.45 %	93.58 %	89.46 %
RVFL	1-step	95.44 %	95.33 %	95.65 %	96.06 %	96.21 %	95.79 %	95.36 %	95.97 %	95.86 %	96.32 %	96.70 %	95.12 %
	3-step	92.46 %	94.00 %	93.94 %	93.88 %	94.31 %	93.54 %	92.43 %	93.94 %	93.67 %	94.26 %	94.12 %	92.47 %
	5-step	88.93 %	90.93 %	91.45 %	90.07 %	91.72 %	90.68 %	89.42 %	90.75 %	90.60 %	91.73 %	91.25 %	84.62 %
RVFL*	1-step	95.44 %	95.33 %	95.66 %	96.06 %	96.22 %	95.79 %	95.37 %	95.97 %	95.87 %	96.32 %	96.71 %	95.13 %
	3-step	92.47 %	94.02 %	93.95 %	93.90 %	94.31 %	93.54 %	92.45 %	93.94 %	93.68 %	94.29 %	94.14 %	92.49 %
	5-step	88.98 %	90.93 %	91.46 %	90.09 %	91.73 %	90.69 %	89.46 %	90.76 %	90.63 %	91.75 %	91.25 %	84.66 %
VMD-WRELM	1-step	99.25 %	99.31 %	99.31 %	99.56 %	99.11 %	99.26 %	99.40 %	99.49 %	99.39 %	99.40 %	99.42 %	99.27 %
	3-step	97.10 %	97.43 %	97.63 %	98.14 %	96.93 %	97.01 %	97.31 %	97.85 %	97.56 %	97.91 %	97.49 %	96.91 %
	5-step	93.36 %	93.79 %	94.73 %	94.98 %	93.12 %	92.85 %	93.79 %	94.58 %	94.16 %	95.05 %	94.17 %	89.62 %
EMD-RVFL	1-step	94.49 %	94.97 %	92.48 %	94.79 %	94.91 %	95.92 %	97.62 %	94.88 %	91.82 %	95.76 %	96.02 %	98.60 %
	3-step	83.86 %	94.63 %	86.00 %	91.86 %	90.11 %	95.72 %	94.07 %	88.72 %	89.17 %	96.63 %	91.97 %	96.64 %
	5-step	72.51 %	93.93 %	79.72 %	90.88 %	84.43 %	94.36 %	92.37 %	80.12 %	86.92 %	96.49 %	88.51 %	93.71 %
CVAELM	1-step	99.25 %	99.21 %	99.40 %	99.67 %	99.03 %	99.14 %	99.26 %	99.42 %	99.61 %	99.52 %	99.40 %	99.29 %
	3-step	97.05 %	97.53 %	97.94 %	98.84 %	96.87 %	96.55 %	97.13 %	97.53 %	98.94 %	98.41 %	97.43 %	97.39 %
	5-step	95.65 %	91.97 %	96.06 %	94.53 %	91.54 %	93.28 %	89.26 %	92.65 %	91.63 %	92.70 %	92.01 %	95.86 %
WPD-EMD-ELM	1-step	98.48 %	97.01 %	99.29 %	98.97 %	98.33 %	97.34 %	98.02 %	98.45 %	99.16 %	98.64 %	96.33 %	97.90 %
	3-step	95.49 %	92.19 %	97.97 %	96.82 %	95.51 %	92.31 %	93.85 %	95.19 %	97.42 %	96.07 %	88.17 %	93.67 %
	5-step	91.71 %	85.75 %	96.40 %	93.61 %	92.20 %	86.54 %	89.47 %	90.76 %	95.31 %	92.97 %	98.57 %	84.55 %
CEEMDAN-ANN	1-step	99.69 %	99.61 %	99.71 %	99.67 %	99.53 %	99.56 %	99.56 %	99.63 %	99.68 %	99.72 %	99.70 %	99.66 %
	3-step	98.73 %	98.50 %	98.89 %	98.57 %	98.30 %	98.24 %	97.95 %	98.40 %	98.66 %	98.85 %	98.67 %	98.60 %
	5-step	96.80 %	96.37 %	97.39 %	95.69 %	96.16 %	95.83 %	95.09 %	95.85 %	96.50 %	97.32 %	96.70 %	95.15 %
VMD-GSO-ELM	1-step	99.10 %	97.97 %	97.95 %	98.64 %	97.75 %	97.84 %	97.67 %	98.41 %	98.63 %	98.75 %	98.97 %	98.47 %
	3-step	96.57 %	92.92 %	93.42 %	94.46 %	92.96 %	92.70 %	90.71 %	94.00 %	94.53 %	95.40 %	95.43 %	93.89 %
	5-step	92.24 %	83.68 %	86.22 %	85.61 %	86.87 %	85.39 %	81.34 %	86.17 %	87.20 %	89.42 %	89.89 %	81.44 %

**Table 11:** Improvement percentages of the MAPE results of EDNNRW<sub>RVFL</sub> over the other competitors.

EDNNRW <sub>RVFL</sub> vs	n-step	Dataset											
		Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sept	Oct	Nov	Dec
ELM	1-step	99.03 %	98.63 %	98.94 %	99.00 %	98.69 %	98.56 %	98.86 %	98.91 %	98.92 %	99.14 %	98.92 %	98.66 %
	3-step	96.76 %	96.45 %	96.81 %	96.80 %	96.61 %	96.22 %	96.17 %	96.35 %	96.34 %	97.20 %	96.66 %	95.49 %
	5-step	92.38 %	93.09 %	93.79 %	93.21 %	93.88 %	93.69 %	93.12 %	92.24 %	92.85 %	94.09 %	94.00 %	84.85 %
SNN	1-step	99.01 %	98.62 %	98.93 %	98.97 %	98.67 %	98.53 %	98.86 %	98.89 %	98.90 %	99.12 %	98.89 %	98.66 %
	3-step	96.73 %	96.42 %	96.76 %	96.77 %	96.59 %	96.19 %	96.14 %	96.33 %	96.29 %	97.15 %	96.63 %	95.49 %
	5-step	92.32 %	93.07 %	93.75 %	93.13 %	93.82 %	93.64 %	93.10 %	92.21 %	92.78 %	94.04 %	93.94 %	84.84 %
RVFL	1-step	95.19 %	95.06 %	95.34 %	95.78 %	95.96 %	95.87 %	94.56 %	95.47 %	95.48 %	96.10 %	96.60 %	92.98 %
	3-step	92.13 %	93.87 %	93.27 %	93.99 %	94.09 %	94.15 %	91.78 %	93.33 %	92.97 %	93.81 %	94.54 %	89.90 %
	5-step	86.41 %	90.69 %	90.08 %	90.52 %	91.54 %	91.78 %	89.55 %	89.07 %	89.46 %	90.73 %	91.87 %	76.76 %
RVFL*	1-step	95.20 %	95.07 %	95.35 %	95.78 %	95.96 %	95.87 %	94.57 %	95.48 %	95.49 %	96.10 %	96.60 %	93.04 %
	3-step	92.14 %	93.89 %	93.28 %	94.02 %	94.08 %	94.16 %	91.80 %	93.34 %	92.98 %	93.86 %	94.54 %	90.00 %
	5-step	86.43 %	90.71 %	90.14 %	90.54 %	91.55 %	91.78 %	89.63 %	89.10 %	89.49 %	90.74 %	91.85 %	76.97 %
VMD-WRELM	1-step	99.53 %	99.32 %	99.44 %	99.53 %	99.21 %	99.39 %	99.51 %	99.61 %	99.45 %	99.45 %	99.53 %	99.19 %
	3-step	98.25 %	97.46 %	97.79 %	98.03 %	97.18 %	97.57 %	97.86 %	98.31 %	97.69 %	97.95 %	98.17 %	96.25 %
	5-step	95.40 %	93.88 %	94.46 %	94.77 %	93.54 %	94.28 %	95.31 %	95.42 %	94.19 %	94.75 %	95.82 %	84.38 %
EMD-RVFL	1-step	92.88 %	93.49 %	91.71 %	92.68 %	93.35 %	95.39 %	96.55 %	95.63 %	92.24 %	96.75 %	96.09 %	97.67 %
	3-step	81.13 %	93.08 %	88.99 %	91.04 %	87.21 %	95.42 %	92.80 %	88.54 %	90.47 %	97.75 %	92.61 %	96.32 %
	5-step	64.19 %	91.90 %	87.25 %	88.99 %	79.69 %	94.13 %	91.87 %	76.78 %	88.73 %	97.44 %	89.47 %	94.32 %
CVAELM	1-step	99.48 %	99.23 %	99.57 %	99.67 %	99.05 %	99.23 %	99.33 %	99.46 %	99.58 %	99.53 %	99.47 %	99.23 %
	3-step	97.85 %	97.59 %	98.35 %	98.76 %	96.89 %	96.98 %	97.41 %	97.67 %	98.75 %	98.38 %	97.90 %	97.13 %
	5-step	94.81 %	91.52 %	96.01 %	94.70 %	91.63 %	94.07 %	89.47 %	91.26 %	90.64 %	91.86 %	93.39 %	93.60 %
WPD-EMD-ELM	1-step	99.23 %	97.02 %	99.43 %	99.43 %	98.11 %	97.58 %	98.11 %	98.73 %	99.12 %	98.71 %	97.36 %	98.11 %
	3-step	97.71 %	92.16 %	98.45 %	98.22 %	94.87 %	93.09 %	94.13 %	95.73 %	97.10 %	96.09 %	92.81 %	93.82 %
	5-step	95.01 %	85.60 %	97.25 %	96.26 %	91.05 %	87.91 %	90.34 %	91.17 %	94.39 %	92.48 %	87.32 %	81.78 %
CEEMDAN-ANN	1-step	99.79 %	99.62 %	99.83 %	99.72 %	99.58 %	99.61 %	99.60 %	99.69 %	99.69 %	99.74 %	99.76 %	99.60 %
	3-step	99.14 %	98.56 %	99.31 %	98.82 %	98.45 %	98.46 %	98.18 %	98.54 %	98.68 %	98.90 %	99.04 %	98.20 %
	5-step	97.50 %	96.49 %	98.15 %	96.45 %	96.41 %	96.40 %	95.90 %	95.94 %	96.54 %	97.32 %	97.63 %	92.44 %
VMD-GSO-ELM	1-step	98.68 %	97.84 %	98.34 %	98.56 %	97.70 %	98.06 %	97.61 %	98.66 %	98.53 %	98.76 %	98.50 %	98.60 %
	3-step	94.92 %	92.49 %	94.21 %	94.27 %	92.74 %	93.63 %	90.84 %	94.86 %	94.29 %	95.27 %	94.34 %	93.77 %
	5-step	86.78 %	82.59 %	86.92 %	85.96 %	86.16 %	87.55 %	82.79 %	87.67 %	87.15 %	88.76 %	87.92 %	78.96 %

3. The proposed EDNNRW<sub>RVFL</sub> and EDNNRW<sub>RVFL</sub>\* ranked higher and significantly outperformed the comparative algorithms with a 0.05 significance level.

Future research directions and further possible improvements to this work include:

1. The type of prediction model selected has a significant influence on the predictive performance. Thus, the reservoir computing model and other variants of recurrent NN with random weights [48] should be further investigated.
2. The number of lag orders and structure size of NNRW within the proposed method are user-specified parameters. Therefore, further work on how to automatically determine the optimal lag orders and node sizes is worth further investigation.

## ACKNOWLEDGMENT

This work was supported by the Graduate Education of Computer and Information Science Interdisciplinary Research Grant from Department of Computer Science, Faculty of Science, Khon Kaen University #002/2556.

## References

- [1] Global Wind Energy Council, "Global wind report 2018," 2019.
- [2] B. Bhandari, K.-T. Lee, G.-Y. Lee, Y.-M. Cho, and S.-H. Ahn, "Optimization of hybrid renewable energy power systems: A review," *International journal of precision engineering and manufacturing-green technology*, vol. 2, no. 1, pp. 99–112, 2015.
- [3] E. B. Ssekulima, M. B. Anwar, A. Al Hinai, and M. S. El Moursi, "Wind speed and solar irradiance forecasting techniques for enhanced renewable energy integration with the grid: a review," *IET Renewable Power Generation*, vol. 10, no. 7, pp. 885–989, 2016.
- [4] H. Liu, X. Mi, and Y. Li, "Comparison of two new intelligent wind speed forecasting approaches based on wavelet packet decomposition, complete ensemble empirical mode decomposition with adaptive noise and artificial neural networks," *Energy Conversion and Management*, vol. 155, pp. 188–200, 2018.
- [5] X. An, D. Jiang, M. Zhao, and C. Liu, "Short-term prediction of wind power using emd and chaotic theory," *Communications in Nonlinear Science and Numerical Simulation*, vol. 17, no. 2, pp. 1036–1042, 2012.
- [6] L. Han, R. Zhang, X. Wang, A. Bao, and H. Jing, "Multi-step wind power forecast based on vmd-lstm," *IET Renewable Power Generation*, vol. 13, no. 10, pp. 1690–1700, 2019.
- [7] J. P. d. S. Catalão, H. M. I. Pousinho, and V. M. F. Mendes, "Short-term wind power forecasting in portugal by neural networks and wavelet transform," *Renewable energy*, vol. 36, no. 4, pp. 1245–1251, 2011.
- [8] A. Laouafi, M. Mordjaoui, A. Medoued, T. E. Boukelia, and A. Ganouche, "Wind power forecasting approach using neuro-fuzzy system combined with wavelet packet decomposition, data

- preprocessing, and forecast combination framework,” *Wind Engineering*, vol. 41, no. 4, pp. 235–244, 2017.
- [9] C. Wang, H. Zhang, and P. Ma, “Wind power forecasting based on singular spectrum analysis and a new hybrid laguerre neural network,” *Applied Energy*, vol. 259, p. 114139, 2020.
- [10] Z. Qian, 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.
- [11] W. Cao, X. Wang, Z. Ming, and J. Gao, “A review on neural networks with random weights,” *Neurocomputing*, vol. 275, pp. 278–287, 2018.
- [12] W. F. Schmidt, M. A. Kraaijveld, and R. P. Duin, “Feedforward neural networks with random weights,” in *Proceedings., 11th IAPR International Conference on Pattern Recognition. Vol.II. Conference B: Pattern Recognition Methodology and Systems*, pp. 1–4, IEEE, 1992.
- [13] Y.-H. Pao, G.-H. Park, and D. J. Sobajic, “Learning and generalization characteristics of the random vector functional-link net,” *Neurocomputing*, vol. 6, no. 2, pp. 163–180, 1994.
- [14] G.-B. Huang, Q.-Y. Zhu, and C.-K. Siew, “Extreme learning machine: Theory and applications,” *Neurocomputing*, vol. 70, no. 1, pp. 489–501, 2006.
- [15] L. Tang, Y. Wu, and L. Yu, “A non-iterative decomposition-ensemble learning paradigm using rvfl network for crude oil price forecasting,” *Applied Soft Computing*, vol. 70, pp. 1097–1108, 2018.
- [16] E. Fijani, R. Barzegar, R. Deo, E. Tziritis, and S. Konstantinos, “Design and implementation of a hybrid model based on two-layer decomposition method coupled with extreme learning machines to support real-time environmental monitoring of water quality parameters,” *Science of the total environment*, vol. 648, pp. 839–853, 2019.
- [17] R. Prasad, R. C. Deo, Y. Li, and T. Maraseni, “Soil moisture forecasting by a hybrid machine learning technique: Elm integrated with ensemble empirical mode decomposition,” *Geoderma*, vol. 330, pp. 136–161, 2018.
- [18] N. Huang, C. Yuan, G. Cai, and E. Xing, “Hybrid short term wind speed forecasting using variational mode decomposition and a weighted regularized extreme learning machine,” *Energies*, vol. 9, no. 12, pp. 1–19, 2016.
- [19] T. Peng, J. Zhou, C. Zhang, and Y. Zheng, “Multi-step ahead wind speed forecasting using a hybrid model based on two-stage decomposition technique and adaboost-extreme learning machine,” *Energy Conversion and Management*, vol. 153, pp. 589–602, 2017.
- [20] H. Liu, H.-Q. Tian, and Y.-F. Li, “Four wind speed multi-step forecasting models using extreme learning machines and signal decomposing algorithms,” *Energy conversion and management*, vol. 100, pp. 16–22, 2015.
- [21] L. Tang, Y. Wu, and L. Yu, “A randomized-algorithm-based decomposition-ensemble learning methodology for energy price forecasting,” *Energy*, vol. 157, pp. 526–538, 2018.
- [22] N. E. Huang, Z. Shen, S. R. Long, M. C. Wu, H. H. Shih, Q. Zheng, N.-C. Yen, C. C. Tung, and H. H. Liu, “The empirical mode decomposition and the hilbert spectrum for nonlinear and non-stationary time series analysis,” *Proceedings of the Royal Society of London. Series A: Mathematical, Physical and Engineering Sciences*, vol. 454, no. 1971, pp. 903–995, 1998.
- [23] K. Dragomiretskiy and D. Zosso, “Variational mode decomposition,” *IEEE transactions on signal processing*, vol. 62, no. 3, pp. 531–544, 2014.
- [24] D. S. Broomhead and G. P. King, “Extracting qualitative dynamics from experimental data,” *Physica D: Nonlinear Phenomena*, vol. 20, no. 2–3, pp. 217–236, 1986.
- [25] S. G. Mallat, “A theory for multiresolution signal decomposition: the wavelet representation,” *IEEE Transactions on Pattern Analysis and Machine Intelligence*, no. 7, pp. 674–693, 1989.
- [26] R. R. Coifman and M. V. Wickerhauser, “Entropy-based algorithms for best basis selection,” *IEEE Transactions on information theory*, vol. 38, no. 2, pp. 713–718, 1992.
- [27] P. Du, J. Wang, W. Yang, and T. Niu, “A novel hybrid model for short-term wind power forecasting,” *Applied Soft Computing*, vol. 80, pp. 93–106, 2019.
- [28] Y. Li, H. Shi, F. Han, Z. Duan, and H. Liu, “Smart wind speed forecasting approach using various boosting algorithms, big multi-step forecasting strategy,” *Renewable Energy*, vol. 135, pp. 540–553, 2019.
- [29] G. E. P. Box and G. M. Jenkins, *Time Series Analysis, Forecasting and Control*. San Francisco, CA: Holden-Day, 1970.
- [30] L.-L. Li, X. Zhao, M.-L. Tseng, and R. R. Tan, “Short-term wind power forecasting based on support vector machine with improved dragonfly algorithm,” *Journal of Cleaner Production*, vol. 242, p. 118447, 2020.
- [31] A. P. Marugán, F. P. G. Márquez, J. M. P. Perez, and D. Ruiz-Hernández, “A survey of artificial neural network in wind energy systems,” *Applied energy*, vol. 228, pp. 1822–1836, 2018.
- [32] Y. Ren, P. Suganthan, and N. Srikanth, “Ensemble methods for wind and solar power forecasting—a state-of-the-art review,” *Renewable and Sustainable Energy Reviews*, vol. 50, pp. 82–91, 2015.

- [33] A. Tascikaraoglu and M. Uzunoglu, "A review of combined approaches for prediction of short-term wind speed and power," *Renewable and Sustainable Energy Reviews*, vol. 34, pp. 243–254, 2014.
- [34] Y.-H. Pao and Y. Takefuji, "Functional-link net computing: Theory, system architecture, and functionalities," *Computer*, vol. 25, no. 5, pp. 76–79, 1992.
- [35] G. Li, P. Niu, X. Duan, and X. Zhang, "Fast learning network: A novel artificial neural network with a fast learning speed," *Neural Computing and Applications*, vol. 24, no. 7-8, pp. 1683–1695, 2014.
- [36] A. A. Abdoos, "A new intelligent method based on combination of vmd and elm for short term wind power forecasting," *Neurocomputing*, vol. 203, pp. 111–120, 2016.
- [37] J. Naik, P. Satapathy, and P. Dash, "Short-term wind speed and wind power prediction using hybrid empirical mode decomposition and kernel ridge regression," *Applied Soft Computing*, vol. 70, pp. 1167–1188, 2018.
- [38] C. R. Rao, S. K. Mitra, *et al.*, *Generalized inverse of a matrix and its applications*. New York: Springer, 1972.
- [39] D. Serre, *Matrices: Theory and Applications*. New York: Springer, 2002.
- [40] H. Liu, X. Mi, and Y. Li, "An experimental investigation of three new hybrid wind speed forecasting models using multi-decomposing strategy and elm algorithm," *Renewable energy*, vol. 123, pp. 694–705, 2018.
- [41] Y. Ren, P. Suganthan, and N. Srikanth, "A comparative study of empirical mode decomposition-based short-term wind speed forecasting methods," *IEEE Transactions on Sustainable Energy*, vol. 6, no. 1, pp. 236–244, 2014.
- [42] H. Yin, Z. Dong, Y. Chen, J. Ge, L. L. Lai, A. Vaccaro, and A. Meng, "An effective secondary decomposition approach for wind power forecasting using extreme learning machine trained by crisscross optimization," *Energy conversion and management*, vol. 150, pp. 108–121, 2017.
- [43] Y. Zhang, J. Le, X. Liao, F. Zheng, and Y. Li, "A novel combination forecasting model for wind power integrating least square support vector machine, deep belief network, singular spectrum analysis and locality-sensitive hashing," *Energy*, vol. 168, pp. 558–572, 2019.
- [44] J. Demšar, "Statistical comparisons of classifiers over multiple data sets," *Journal of Machine learning research*, vol. 7, pp. 1–30, 2006.
- [45] R. L. Iman and J. M. Davenport, "Approximations of the critical region of the fbietkan statistic," *Communications in Statistics-Theory and Methods*, vol. 9, no. 6, pp. 571–595, 1980.
- [46] M. Lindauer, J. N. van Rijn, and L. Kotthoff, "The algorithm selection competitions 2015 and 2017," *Artificial Intelligence*, vol. 272, pp. 86–100, 2019.
- [47] I. A. Carvalho, "On the statistical evaluation of algorithmic's computational experimentation with infeasible solutions," *Information Processing Letters*, vol. 143, pp. 24–27, 2019.
- [48] Y. Rizk and M. Awad, "On extreme learning machines in sequential and time series prediction: A non-iterative and approximate training algorithm for recurrent neural networks," *Neurocomputing*, vol. 325, pp. 1–19, 2019.



**Pakarat Musikawan** received his B.S. degree in information technology from Ubon Ratchathani University, Thailand. He received his M.S. and Ph.D. degrees in computer science from Khon Kaen University, Thailand. His research interests include machine learning, artificial intelligence, computational intelligence, soft computing, data mining, and knowledge-based systems.



**Khamron Sunat** received his B.S. degree in chemical engineering from Chulalongkorn University, Thailand, in 1989. He received his M.S. degree in computational science in 1998 and his Ph.D. degree in computer science from Chulalongkorn University in 2004. He now works as a lecturer in the Department of Computer Science at Khon Kaen University, Thailand. His research interests are in neural networks, pattern recognition,

computer vision, soft computing, fuzzy systems, evolutionary computing and, optical computing.



**Yanika Kongsorot** received her B.S., M.S. and Ph.D. degrees in computer science from Khon Kaen University, Thailand. Her research interests include machine learning, pattern recognition, data mining, intelligent information processing, and decision support systems.