

Received April 2, 2020, accepted April 20, 2020, date of publication April 27, 2020, date of current version May 15, 2020.

Digital Object Identifier 10.1109/ACCESS.2020.2990439

# A Newly Developed Integrative Bio-Inspired Artificial Intelligence Model for Wind Speed Prediction

HAI TAO<sup>1</sup>, SINAN Q. SALIH<sup>2,3</sup>, MANDEEP KAUR SAGGI<sup>4</sup>, ESMAEEL DODANGEH<sup>5</sup>, CYRIL VOYANT<sup>6</sup>, NADHIR AL-ANSARI<sup>7</sup>, ZAHER MUNDHER YASEEN<sup>8</sup>, AND SHAMSUDDIN SHAHID<sup>9</sup>

<sup>1</sup>Computer Science Department, Baoji University of Arts and Sciences, Baoji, China

<sup>2</sup>Institute of Research and Development, Duy Tan University, Da Nang 550000, Vietnam

<sup>3</sup>Computer Science Department, College of Computer Science and Information Technology, University of Anbar, Ramadi, Iraq

<sup>4</sup>Department of Computer Science, Thapar Institute of Engineering and Technology, Patiala 147004, India

<sup>5</sup>Department of Watershed Management, Sari Agricultural Sciences and Natural Resources University, Sari 48181-68984, Iran

<sup>6</sup>SPE Laboratory, University of Corsica UMR 6134, 20000 Ajaccio, France

<sup>7</sup>Civil, Environmental and Natural Resources Engineering, Lulea University of Technology, 97187 Lulea, Sweden

<sup>8</sup>Sustainable Developments in Civil Engineering Research Group, Faculty of Civil Engineering, Ton Duc Thang University, Ho Chi Minh City, Vietnam

<sup>9</sup>School of Civil Engineering, Faculty of Engineering, Universiti Teknologi Malaysia (UTM), Johor Bahru 81310, Malaysia

Corresponding author: Zaher Mundher Yaseen (yaseen@tdtu.edu.vn)

This work was supported by the Key Research and Development Program in Shaanxi Province under Grant 2020GY-078.

**ABSTRACT** Accurate wind speed (WS) modelling is crucial for optimal utilization of wind energy. Numerical Weather Prediction (NWP) techniques, generally used for WS modelling are not only less cost-effective but also poor in predicting in shorter time horizon. Novel WS prediction models based on the multivariate empirical mode decomposition (MEMD), random forest (RF) and Kernel Ridge Regression (KRR) were constructed in this paper better accuracy in WS prediction. Particle swarm optimization algorithm (PSO) was employed to optimize the parameters of the hybridized MEMD model with RF (MEMD-PSO-RF) and KRR (MEMD-PSO-KRR) models. Obtained results were compared to those of the standalone RF and KRR models. The proposed methodology is applied for monthly WS prediction at meteorological stations of Iraq, Baghdad (Station<sub>1</sub>) and Mosul (Station<sub>2</sub>) for the period 1977-2013. Results showed higher accuracy of MEMD-PSO-RF model in predicting WS at both stations with a correlation coefficient ( $r$ ) of 0.972 and  $r = 0.971$  during testing phase at Station<sub>1</sub> and Station<sub>2</sub>, respectively. The MEMD-PSO-KRR was found as the second most accurate model followed by Standalone RF and KRR, but all showed a competitive performance to the MEMD-PSO-RF model. The outcomes of this work indicated that the MEMD-PSO-RF model has a remarkable performance in predicting WS and can be considered for practical applications.

**INDEX TERMS** Wind speed prediction, multivariate empirical mode decomposition, random forest, Kernel Ridge Regression, Iraq region.

## I. INTRODUCTION

The importance of wind speed prediction in wind energy farm operation and maintenance has increased over the years [1], [2]. The sustained increase in the rate of wind turbines erections demands the deployment of optimal dispatching strategy that will guarantee stable power generation by the wind turbines without having much influence on the power

The associate editor coordinating the review of this manuscript and approving it for publication was Xiaowei Zhao.

grid. However, the cost of wind farm operation may be affected by imperfect predictions due to the underlying uncertainties in WS [3], [4]. Similarly, the availability of wind resources must be considered for a maintenance schedule to ensure optimal maintenance for reducing the turbines' production loss [5], [6]. Therefore, accurate WS prediction has grasped research attention in recent years due to its great practical and academic values.

Different prediction models are currently used for prediction of WS in different time horizons. Most of the recent

studies on WS prediction reportedly used intelligent models [7], [8]. Such models are generally helpful in short-range (30 min–6.0 h) WS estimation [7]. There are several versions of these intelligent models, including adaptive neural fuzzy inference system, support vector machine, artificial neural network, etc. In practice, the WS data measured at the turbine locations are used to train these models while the tuning of the model parameters is done using the immediate past observed WS.

Despite the effectiveness of these intelligent models for short-term predictions, they still experience a rapid declination in prediction accuracy with the increase of prediction horizon. Numerical Weather Prediction is the most widely used for medium-term (daily, weekly and monthly) WS prediction to guarantee satisfactory prediction accuracy. The NWP needs intensive computational capacity. Supercomputers are used to NWP which provides predictions only once or twice in a day. Moreover, the accuracy of NWP prediction is lower than the statistical models. Thus, medium- to long-term WS predictions are usually done using the NWP models. Owing to the need for better WS prediction model with good accuracy at multiple horizons, a novel WS prediction technique with enhanced short and medium accuracy should be explored.

Industrialization has predisposed the world to several energy-related problems. Renewable resources are becoming important due to the declining fossil fuel reserves and the adverse impacts of fossil fuel on environment [9]. A major practical renewable energy source is wind energy, hence, its related technologies ought to be thoroughly investigated and developed. Wind has an intermittent characteristic; hence, its energy generation is unstable. Energy conservation and management can be disturbed due to instability in wind energy generation [10], [11]; however, this problem can be addressed by an efficient WS prediction. Several factors influence wind speed, making it difficult to measure the complicated wind-speed features accurately using the simple prediction models. Therefore, much attention has recently been paid to high precision WS prediction techniques.

The past decades have witnessed the development of several numerical WS prediction models. Such models are classified into physical, deterministic and probabilistic models [12], [13]. The physical models use physical features like atmospheric pressure, ambient temperature, and local terrain to approximate WS [14]. Owing to the strong theoretic foundation and tremendous performance of the physical models in WS prediction, they are deployed in the field [15]. Meanwhile, several equations are required in the physical models, making them unsuitable for mid to long-term WS prediction when the computation cost is considered. They are also less capable of short-range WS prediction at the local or station level. Several physical models have recently been developed, for instance, Allen *et al.*, (2017) presented a boundary layer scaling model to predict long-term average near-surface WS [16], a physical-based model for WS prediction in complex terrain was developed by [17].

WS prediction using the statistical models is based on historical data. Statistical models are becoming more popular due to the recent advancements in data science [18], [19]. Statistical models can be categorized into multiple and single data models based on the number of different types of data used. WS prediction with the multiple data models requires a combination of several physical information with some statistical frameworks [20]. Multiple data models often present excellent prediction performances, and therefore, they have attracted the attention of researchers. The observed multiple meteorological data are modelled using Gaussian process regression [21] while a probabilistic approach based on some statistical algorithms (non-parametric) and numerical weather prediction data is used to design the WS prediction model [22]. A new approach based on copula is used by [23] to develop a WS prediction model [23]. Despite the chances of achieving better results using multiple data models, the complication and vagueness of the model can sometimes be increased by the multiple data. The multiple data models are therefore less stable compared to single WS data models. Furthermore, the single WS data models are often associated with low computational complexity and are therefore often suggested for short-term WS prediction [24]. Several algorithms have been deployed in designing single WS data models to ensure maximum utilization of antecedent WS data for prediction of WS; such frameworks include signal processing and time series algorithms [25]–[28]. The time series algorithms are classified as the classical algorithms for WS prediction which contain persistence algorithms. In contrary, the signal processing algorithms are primarily deployed for feature extraction during WS prediction; such algorithms include wavelet decomposition, empirical mode decomposition (EMD), wavelet packet decomposition, and complete ensemble empirical mode decomposition [29].

The WS data has a high degree of non-linearity and non-stationary which make an accurate prediction using a single WS data model challenging. Accuracy of predictive models can be improved using efficient learning and predicted parameters. Various data decompositions methods have been used to overcome this challenge which includes EMD, wavelet decomposition, variational mode decomposition, seasonal adjustment, intrinsic time-scale decomposition and empirical mode decomposition [30]. The EMD model belongs to data-adaptive decomposition techniques has advantages over wavelet transformation methods. EMD methods decompose a time series into a collection of stationary Intrinsic Mode Functions (IMFs) with different frequency bands and a residue-based on local properties of the time series adaptively and efficiently. An EMD model starts by decomposition of a given input predictor which is projected using the adequate lagged sub-series as inputs. Then the predicted series are summed at various time scales to gain the target variable at expected data scales (monthly scale in this study). Therefore, it can efficiently capture the non-linearity and non-stationary of WS time series by decomposing it into several series with independent time resolutions [24].

The EMD algorithm also overcome the difficulties of wavelet transformation method by fixing the most suitable decomposition levels and specifying the base function [31]. The EMD has been found highly effective in improvement of model accuracy in a broad range of applications for analysing non-linear and nonstationary processes. It has been successfully applied in forecasting different engineering problems such as prediction of evapotranspiration [32], soil water [33], crude oil price [34] and iceberg drift [35].

As WS behaves differently for different time scales, WS model performance can be improved by engaging the appropriate predictors. The WS prediction models can be classified into various categories: data-driven, model-driven and hybrid approaches. Model-driven and data-driven approaches employ metrological information and statistically techniques, respectively to handle the physical properties that influence WS. Therefore, they have their inherent advantages and disadvantages. To overcome the limitations of model-driven and data-driven approaches in handling the challenges of stochastic and intermitted nature of WS in its accurate prediction, a hybrid approach can be used [36].

Besides, it is required to optimize the parameters of hybrid models for improvement of their prediction accuracy. In recent decades, many optimization methods are being adopted for wind forecasting models such as particle swarm optimization (PSO) and Genetic algorithm (GA). Among them, the PSO has been widely used in recent years to optimize the parameters of different models in many fields.

Novel models through hybridization of multivariate empirical mode decomposition (MEMD) with random forest (RF) and Kernel Ridge Regression (KRR), and a parameter optimization algorithm known as Particle swarm optimization algorithm (PSO) are proposed in this study for accurate prediction of WS. The proposed model directed in a way to benefit the advantages of other soft computing techniques for improved tuning of the MEMD model. Random forest (RF) [37] and Kernel Ridge Regression (KRR) [38] algorithms were used to avoid the overfitting the MEMD model. The regression method adopted is the nonlinear KRR method, which has shown particularly attractive for its simple implementation, fast processing and accuracy [39]. Nonetheless, the accurate prediction of the model parameters is also a requirement for optimal model performance which has been done using PSO. The key objectives of this study include evaluation of the performance of the MEMD-PSO-RF and MEMD-PSO-KRR in WS prediction and benchmark the results by comparing the performance of newly developed models with Standalone RF and KRR models.

## II. CASE STUDY AND DATASET DESCRIPTION

Despite the arid to semi-arid climate in most part of Iraq, the Tigris basin still ranges from semi-humid to semi-arid in the headwaters to the north and south, respectively [40]. Hence, any event of drought in the future is expected to adversely affect the already limited water resources of the nation and this will affect the socio-ecological system of the Tigris

Basin's that is home to over 18 million people [41], [42]. The persistent increase in temperature is continuously boosting surface water scarcity and reducing aquifers' water tables, indicating an ongoing drought condition which can worsen with time. Using climate forecasting models, it has been predicted that drought and temperature are on the increase in the region and the condition may soon become unsustainable [43]. Drought-related issues are closely related to the balance between temperature and precipitation [44]. In the Tigris, the annual rainfall ranges from 400 to 600 mm; however, the range from the downstream to the upstream reaches is about it ranges from 150 to 800 mm, respectively. According to the Iraqi meteorological monitoring stations, there have been high rates of evaporation in Mosul and Baghdad. From 1960-2009, Baghdad has recorded a mean July temperature range of 23.5 to 44°C while the rate of annual rainfall has been 244 mm; the annual evaporation rate has been 3200 mm. Regarding Mosul, the observed July temperature range has been from 24.8 to 43°C, with annual precipitation rate of 729 mm and evaporation rate of 3900 mm.

In this research, multiple hydrometeorological variables including sunshine radiation (SS), rainfall ( $R_{nf}$ ), minimum air temperature ( $T_{min}$ ), maximum air temperature ( $T_{max}$ ), evaporation (ET) and relative humidity (RH) to predict the monthly wind speed (WS), were used to build the proposed predictive model. Historical data over 1977-2013 with monthly scale data without any missing data at Baghdad and Mosul meteorological station were used in this investigation (Figure 1). The statistical characteristics of the used historical data are reported in Table 1. The selection of this region was due to the lake of advanced technologies that support the hydrologist and climatologist and thus the proposed advanced soft computing model can provide a remarkable assistance for simulating the WS. It is highly essential for energy sources management, monitoring and assessment. Also, the studied Iraq region is considered as developing country with a high potential sources of wind energy for exploitation. Meanwhile the region is one of the primary sources of dust generation and emissions with the resulting dust affecting the neighboring countries [45]. Hence such an inspection can produce high aspect of benefit for environmental engineering projects, e.g., the risk analysis of the dust storms as well as prediction of the future wind energy sources is available for exploitation.

## III. METHODOLOGY

### A. THE RANDOM FOREST MODEL

Bootstrapping is principally an ensemble modelling approach which provides prediction using decision trees [32], [33]. The RF algorithm is a decision tree-based machine learning approach which developed ensembles using a random bagging technique [47]. Every node is connected randomly by selecting well-known predictors to increase predictive performance and avoid overfitting [48]. The RF model can be constructed using the following steps:

TABLE 1. The statistical characteristics of meteorological data used in this study.

Sites	Baghdad						Mosul					
Long. (°E)	44.36						43.15					
Lat. (°N)	33.31						36.34					
Elevation (m)	39						32000					
Meteorological predictors	Mean	Max	Min	Std.	Skew	Kurt	Mean	Max	Min	Std.	Skew	Kurt
Sunshine radiation (SS)	8.4	13.3	2.7	2.8	-0.0	-1.2	9.0	13.5	3.5	2.3	0.1	-1.1
Rainfall (R <sub>nr</sub> )	30.8	205.6	0.0	39.9	1.6	2.3	10.1	148.7	0.0	16.4	3.0	14.9
Minimum air temperature (T <sub>min</sub> )	12.7	45.0	-1.0	8.1	0.3	-0.9	14.8	28.6	0.1	7.8	-0.1	-1.4
Maximum air temperature (T <sub>max</sub> )	27.8	46.4	8.3	11.1	0.0	-1.5	30.6	47.7	12.3	10.4	-0.1	-1.4
Evaporation (ET)	182.3	695.5	17.0	138.5	0.7	-0.1	279.4	727.6	45.6	175.6	0.4	-1.0
Relative humidity (RH)	52.2	88.0	19.0	21.3	0.0	-1.5	44.4	96.0	20.0	17.8	0.5	-1.0
Target variable												
Wind speed (WS)	1.3	3.2	0.2	0.6	0.1	-0.6	3.2	6.1	1.2	0.8	0.4	0.3



FIGURE 1. The location of the investigated meteorological station.

- i. Embedding the inputs to generate n-number of trees (i.e.  $n_{trees}$ ) using bootstrapping.
- ii. Select the maximum number of split predictors using a random sample of inputs ( $m_{try}$ ) based on unpruned regression tree.
- iii. Combine the estimates of  $n_{trees}$  in terms of aggregations to predict WS.

**B. THE MEMD APPROACH**

The multivariate empirical mode decomposition (MEMD) is a self-adaptive method capable of handling the hurdles of mode alignment. The mathematical structure of MEMD is an improved variant of EMD which is defined as:

$$\Phi(\alpha) = \Re_l(\alpha) + \sum_{k=1}^l C_k(\alpha) \tag{1}$$

In Eq. (1),  $\Phi(\alpha)$ ,  $C_k(\alpha)$  and  $R_l(\alpha)$  are the input variable, the  $k^{th}$  Intrinsic Mode Functions (IMF) and residue factor respectively. The MEMD designed by [37] demarcates the multiple inputs into IMFs using White Gaussian noise [50]. The mean  $\aleph(\alpha)$  can be computed as:

$$\aleph(\alpha) = \frac{1}{t} \sum_{s=1}^t e^{\theta_s}(\alpha) \tag{2}$$

The term  $e^{\theta_s}(\alpha)$  is called the envelope curves with  $t$ .

$$\Re(\alpha) = \Phi(\alpha) - \aleph(\alpha) \tag{3}$$

In Eq. (3),  $R(\alpha)$  is a multivariate IMF. The application of MEMD can be found in signal processing [46], [47] and solar radiation [53].

**C. PARTICLE SWARM OPTIMIZATION (PSO)**

PSO meta-heuristic algorithm introduced by [54] is a soft computing optimization method that is theorized to optimize a problem through the searching for the best candidate solution. The PSO algorithm is invented by replicating the social behavior of animals in a bunch e.g., birds and fishes [55]. In this algorithm, a bunch of creatures, that are called particles, spread in the search area [56]. Every single particle approximates its situation relative to the target position. They fine-tune their location and velocity using the current situation and the best position they were already in and the situation of the best particles in the bunch [57]:

$$V_{id}^t = wV_{id}^{t-1} + c_1r_1(P_{id}^t - X_{id}^t) + c_2r_2(P_{gd}^t - X_{id}^t) \tag{4}$$

$d = 1, 2, 3, \dots, D$

where  $X_{id}^t$  indicates the location of the particle  $i$  in iteration  $t$ ,  $V_{id}^t$  is the velocity of particle  $i$  in iteration  $t$ ,  $P_{id}^t$  is the best location of the particle  $i$ ,  $P_{gd}^t$  is the global best position of particle  $i$ ,  $w$  expresses the inertia weight,  $c_1$  expresses the cognitive learning factor,  $c_2$  expresses the social learning factor, and  $r_1$  and  $r_2$  denote the random values in [0,1].

The basic steps for implementing the algorithm are as follow: **Step 1:** Generate the initial swarm and assessing it.

**Step 2:** Evaluation of fitness of every single particle within the bunch.

**Step 3:** Update velocity of every single particle according to Eq. 4.

**Step 4:** Update the position for each particle by the following equation:

$$X_{id}^{t+1} = X_{id}^t + V_{id}^t \quad (5)$$

**Step 5:** The algorithm is stopped when the termination criterion is satisfied or returned to Step 2.

#### D. THE KRR MODEL

The Kernel Ridge Regression (KRR) is a kernel-based regression technique [58] that handle the over-fitting problems by adopting regularization and the kernel procedure in non-linear input variables. Mathematically, The KRR is described using the following mathematical structure:

$$\arg \min \frac{1}{r} \sum_{s=1}^r \|g_s - h_s\|^2 + \beta \|g\|_H^2 \quad (6)$$

$$g_s = \sum_{l=1}^r \beta_l \Psi(z_l, z_s) \quad (7)$$

In Eq. (6),  $\|\cdot\|_H$  represents the Hilbert normed space [58]. Eq. (7) can be rewritten in the following form:

$$(\Lambda + \lambda r I) = w \quad (8)$$

$$\tilde{w} = \sum_{s=1}^r \beta_l \Psi(z_l, \tilde{z}) \quad (9)$$

$\Lambda$  can be designed using  $\Lambda_{l,s} = \Psi(z_l, z_s)$  which is max kernel matrix and  $y$  is the input  $r \times 1$  regress and vector, and  $\beta$  is the  $r \times 1$  unknown solution vector. In training phase, KRR is estimated  $\beta$  using Eq. (9) which utilized later in validation to calculate the regression of unidentified sample  $\tilde{z}$  in Eq. (9). Different kinds of kernels (for example linear, polynomial and Gaussian) can be used to achieve better performance [59], [60]. The mathematical formulation of these kernels can be expressed a following:

$$\Psi(z_l, z_s) = z_l^T \cdot z_s \quad (10)$$

$$\Psi(z_l, z_s) = \left( z_l^T \cdot z_s + R \right)^D \quad (11)$$

$$\Psi(z_l, z_s) = \exp \left( -\frac{\|z_l - z_s\|^2}{(2\rho^2)} \right) \quad (12)$$

#### E. THE PERFORMANCE METRICS

The performance of MEMD-PSO-RF and other models were measured using statistical metrics. The mathematical representation of these metrics [61]–[63] and [64]–[69] is given below:

I. Correlation coefficient ( $r$ ):

$$r = \left( \frac{\sum_{i=1}^N (WS_o - \overline{WS_o}) (WS_p - \overline{WS_p})}{\sqrt{\sum_{i=1}^N (WS_o - \overline{WS_o})^2} \sqrt{\sum_{i=1}^N (WS_p - \overline{WS_p})^2}} \right) \quad (13)$$

II. Willmott's Index ( $E_{WI}$ ):

$$E_{WI} = 1 - \left[ \frac{\sum_{i=1}^N (WS_p - WS_o)^2}{\sum_{i=1}^N (|WS_p - \overline{WS_o}| + |WS_o - \overline{WS_p}|)^2} \right], \quad 0 \leq E_{WI} \leq 1 \quad (14)$$

III. Nash-Sutcliffe efficiency (ENS) agreement:

$$E_{NS} = 1 - \left[ \frac{\sum_{i=1}^N (WS_o - WS_p)^2}{\sum_{i=1}^N (\overline{WS_o} - \overline{WS_p})^2} \right], \quad 0 \leq E_{NS} \leq 1 \quad (15)$$

IV. Root mean square error (RMSE):

$$RMSE = \sqrt{\frac{1}{N} \sum_{i=1}^N (WS_p - WS_o)^2} \quad (16)$$

V. Mean absolute error (MAE):

$$MAE = \frac{1}{N} \sum_{i=1}^N |(WS_p - WS_o)| \quad (17)$$

VI. Legates and McCabe's (ELM):

$$E_{LM} = 1 - \left[ \frac{\sum_{i=1}^N |WS_p - WS_o|}{\sum_{i=1}^N |WS_o - \overline{WS_o}|} \right], \quad 0 \leq E_{LM} \leq 1 \quad (18)$$

VII. Relative root mean squared percentage error (RRMSPE; %):

$$RRMSPE = \frac{1}{N} \sum_{i=1}^N \left| \frac{(WS_p - WS_o)}{WS_o} \right| \times 100 \quad (19)$$

VIII. Relative mean absolute percentage error (RMAPE; %):

$$RMAPE = \frac{1}{N} \sum_{i=1}^N \left| \frac{(WS_p - WS_o)}{WS_o} \right| \times 100 \quad (20)$$

In Eqs. (13)–(20),  $WS_o$  and  $WS_p$  indicate the observed and predicted values of WS,  $\overline{WS_o}$  and  $\overline{WS_p}$  are the observed and predicted average WS and  $N$  is the total value of data records.

#### F. MEMD-PSO-RF MODEL DEVELOPMENT

The MEMD-PSO-RF model was implemented using MATLAB R2016b (The Math Works Inc. USA) programming language. The simulations were performed in operating Pentium 4, 2.93 GHz dual-core Central Processing Unit. The datasets are divided straight into training 70% and testing 30% periods. Climate information including SS),  $R_{nf}$ ,  $T_{min}$ ,  $T_{max}$ , ET and RH were incorporated to design MEMD-PSO-RF for WS prediction as described in the following phases.

**TABLE 2.** Parameters used in selecting IMFs for training period using PSO method.

Station	Max. Iteration	Population Size	Training period	
			No. of total IMFs	No. of selected IMFs
Station <sub>1</sub>	10	20	54	16
Station <sub>2</sub>	10	20	54	16

1) THE MEMD PHASE

The MEMD model is applied to demarcate the predictor data into respective IMFs and residuals. Additionally, the predefined parameters include the ensemble number ( $N = 500$ ) and the amplitude of the added white noise ( $\epsilon$ ) between -0.2 and 0.2. Total fifty-four IMFs (Table 1) for WS in station<sub>1&2</sub> were demarcated where every single predictor has IMFs = 9. It is worth to highlight that the MEMD is subjected with several tuning parameters including tolerance and threshold values, stop vector, stopping criteria and total projection, as reported in Table 2 to attain the equal IMFs for both training and testing subsets.

2) THE PSO PHASE

The PSO algorithm is used for the selection of the utmost suitable IMFs for the only training set. The predefined set of parameters includes maximum iterations (=10) and population size (=20). The number of fix selected IMFs is retained to 16 that were pre-defined before executing the PSO algorithm. The same number of IMFs are adopted for the testing set following the training IMFs.

3) THE NORMALIZATION PHASE

The training and testing sets are normalized between 0 and 1 using Eq. (20) [70] to overcome large differences in the data [70]:

$$\Theta_{norm} = \frac{(\Theta - \Theta_{min})}{(\Theta_{max} - \Theta_{min})} \tag{21}$$

In Eq. (20),  $\Theta$  denotes the input/output,  $\Theta_{min}$  is the smallest and  $\Theta_{max}$  is the largest magnitude of the data, and  $\Theta_{norm}$  is the desired normalized point.

4) THE RF PHASE

The last modelling phase is the employment of RF algorithm to predict WS and investigate its ability in WS prediction in the monthly timescale. The designated IMFs (for training period) were embedded in RF model. Some pre-defined parameter set (i.e., 1000 trees and 5 predictors) needs to establish using hit and trial approach in both training and testing stages. The equivalent number of IMFs are used for the testing set to validate the MEMD-PSO-RF model. Further,

the MEMD-PSO-RF, standalone RF and standalone KRR models were also benchmarked. Figure 2 demonstrates the diagrammatic presentation of MEMD-PSO-RF model.

IV. APPLICATION RESULTS AND ANALYSIS

In this study, a new hybrid intelligence framework for wind forecasting is proposed. The relative performance of different models was evaluated to validate the performance of the proposed model.

Tables 4 and 5 provide a summary of the overall performance of the models in predicting WS at two locations in term of eight statistical measures during training and testing stages. A general inspection reveals all of the investigated models have reasonably good performance in reproducing the WS of Station<sub>1</sub> compared to Station<sub>2</sub>. For instance, the RMSE values at Station<sub>1</sub> were in the range of 0.175–270 and 0.186–408 during the training and testing phases respectively, while those were 0.282-0.609 and 0.216-0.451 at Station<sub>2</sub>. This reflects the complexity of the relationships among the climate variables at Station<sub>1</sub> compared to that at Station<sub>2</sub> which increases the uncertainty of WS predictions at Station<sub>2</sub>.

The results presented in Tables 4 and 5 showed that the highest determination coefficient ( $R^2$ ) in WS prediction during both training ( $R^2 = 0.94$ ) and testing ( $R^2 = 0.944$ ) periods by MEMD-PSO-RF model at Station<sub>1</sub>. The MEMD-PSO-RF model also attained the best results in terms of the rest of the performance evaluation measures during training (RMSE = 0.1860, MAE = 0.1480, WI = 0.92, NSE = 0.88) and testing (RMSE = 0.17533, MAE = 0.13943, WI = 0.94, NSE = 0.90). The MEMD-PSO-RF model was also found to perform best in WS prediction at Station<sub>2</sub>. The performance metrics of MEMD-PSO-RF at Station<sub>2</sub> during both training (RMSE = 0.281, MAE = 0.2227, R = 0.971, WI = 0.9197, NSE = 0.888) and testing (RMSE = 0.216, MAE = 0.169, R = 0.971, WI = 0.920, NSE = 0.893) revealed best performance of the model in WS prediction.

The Standalone RF and MEMD-PSO-KRR hybrid model also showed competitive results to the MEMD. PSO-RF model in WS prediction at both the study locations. However, the Standalone KRR model was not found to perform well. The Standalone RF model showed better performance in predicting WS at Station<sub>2</sub> compared to MEMD-PSO-KRR model during both of the training (R = 0.954, WI = 0.896, NSE = 0.862, RMSE = 0.312, MAE = 0.243,  $E_{LM} = 0.633$ , RRMSPE = 9.807, RMAPE = 8.776) and testing periods (R = 0.947, WI = 0.882, NSE = 0.849, RMSE = 0.257, MAE = 0.192,  $E_{LM} = 0.633$ , RRMSPE = 8.30, RMAPE = 6.59). The standalone RF model also showed a slightly better performance compared to MEMD-PSO-KRR model during the training phase at Station<sub>1</sub>. The Standalone KRR model was found as the worst model in predicting the WS at both Station<sub>1</sub>(R = 0.778, WI = 0.562, NSE = 0.461, RMSE = 0.408, MAE = 0.321, LM = 0.297, RRMSPE = 26.19, RMAPE = 26.25) and Station<sub>2</sub> (R = 0.770, WI = 0.607, NSE = 0.535, RMSE = 0.451, MAE = 0.353, LM = 0.326, RRMSPE = 14.56, RMAPE = 11.89). The best prediction

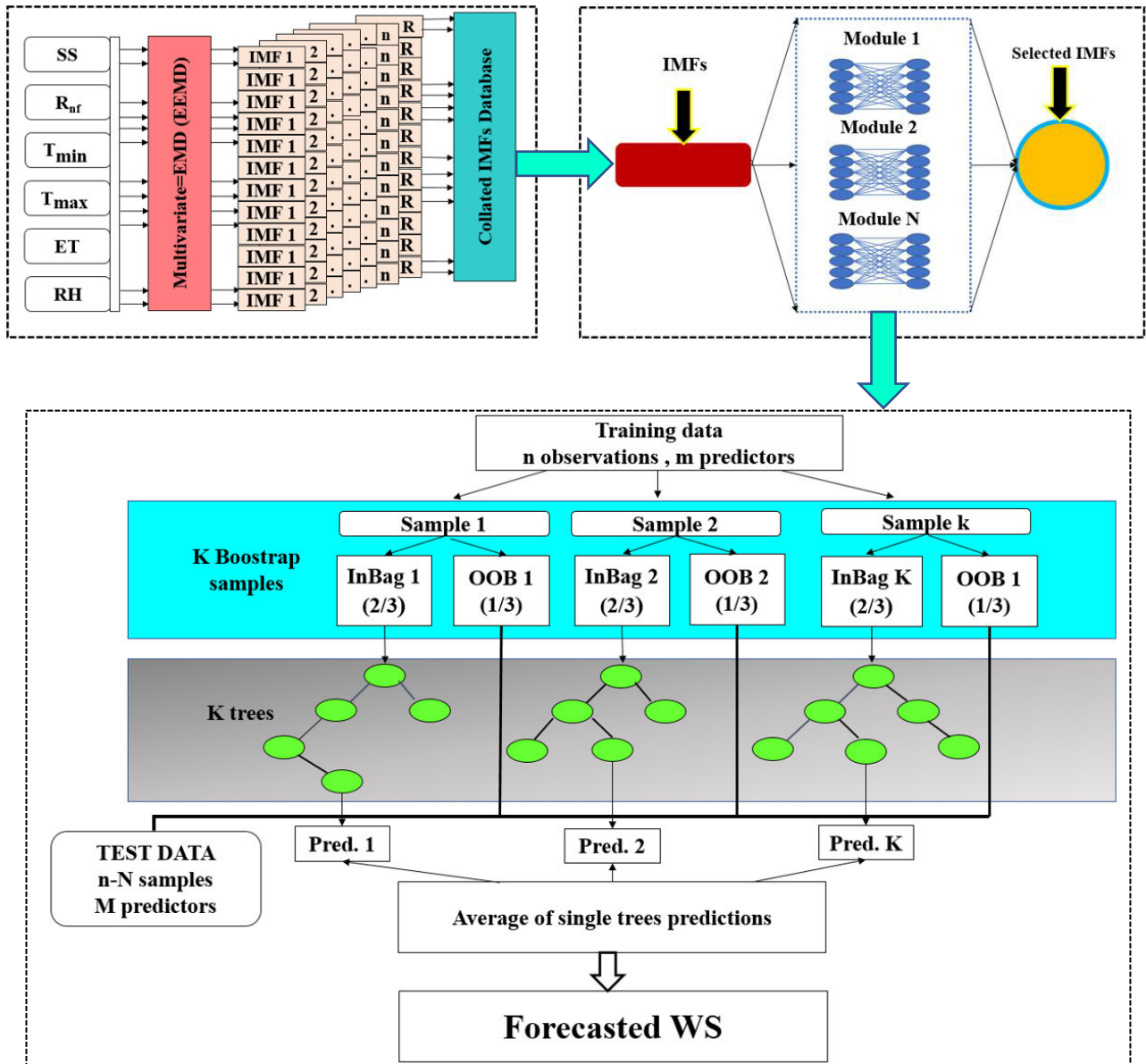


FIGURE 2. Schematic view of the newly developed MEMD-PSO-RF model.

TABLE 3. Design parameters involved to decomposed IMFs and residuals for training and testing period using MEMD method.

Sites	Training period				Testing period			
	No. of total projections	Stop vector		No. of total (IMFs & Res.)	No. of total projections	Stop vector		No. of total (IMFs & Res.)
		tolerance values	threshold			tolerance values	threshold	
Station 1	14	[0.05 0.05]	0.5	54	14	[0.03 0.03]	0.3	54
Station 2	90	[0.03 0.03]	0.3	54	90	[0.04 0.04]	0.3	54

of WS after MEMD-PSO-RF model was achieved using the standalone RF model.

The MEMD-PSO-KRR model was found to perform better compared to standalone RF during testing in term of MSE,

RMSE and NSE, while standalone RF model was found to perform better compared to MEMD-PSO-KRR during training. Standalone KRR model performed worst in term of all metrics at both the station. However, it was found to perform

TABLE 4. The training modeling performance of the both investigated stations.

Station <sub>1</sub>	r	WI	NSE	RMSE	MAE	E <sub>LM</sub>	RRMSPE	RMAPE
Standalone RF	0.952	0.877	0.826	0.240	0.191	0.609	19.647	26.044
Standalone KRR	0.695	0.498	0.400	0.446	0.366	0.249	36.500	49.611
MEMD-PSO-RF	0.970	0.941	0.907	0.175	0.139	0.714	14.335	17.771
MEMD-PSO-KRR	0.887	0.864	0.780	0.270	0.215	0.558	22.083	25.927
Station <sub>2</sub>	r	WI	NSE	RMSE	MAE	E <sub>LM</sub>	RRMSPE	RMAPE
Standalone RF	0.954	0.896	0.862	0.312	0.243	0.633	9.807	8.776
Standalone KRR	0.726	0.517	0.476	0.609	0.476	0.281	19.114	17.136
MEMD-PSO-RF	0.972	0.920	0.888	0.282	0.223	0.664	8.838	7.791
MEMD-PSO-KRR	0.843	0.805	0.705	0.457	0.358	0.459	14.338	12.151

TABLE 5. The testing performance of models in both the investigated stations.

Station <sub>1</sub>	r	WI	NSE	RMSE	MAE	E <sub>LM</sub>	RRMSPE	RMAPE
Standalone RF	0.954	0.882	0.828	0.231	0.183	0.599	14.80	14.67
Standalone KRR	0.778	0.562	0.461	0.408	0.321	0.297	26.19	26.25
MEMD-PSO-RF	0.972	0.926	0.888	0.186	0.148	0.676	11.95	11.95
MEMD-PSO-KRR	0.933	0.922	0.866	0.204	0.165	0.639	13.07	11.81
Station <sub>2</sub>	r	WI	NSE	RMSE	MAE	LM	RRMSPE	RMAPE
Standalone RF	0.947	0.882	0.849	0.257	0.192	0.633	8.30	6.59
Standalone KRR	0.770	0.607	0.535	0.451	0.353	0.326	14.56	11.89
MEMD-PSO-RF	0.971	0.920	0.893	0.216	0.169	0.678	6.98	5.72
MEMD-PSO-KRR	0.921	0.897	0.845	0.260	0.203	0.613	8.40	6.86

slightly better at Station<sub>1</sub> with respect to (MSE = 0.167, RMSE = 0.408, MAE = 0.321, R = 0.778, WI = 0.562, and NSE = 0.461).

Figures 3a and b present the results using violin plots [71] obtained for all the models during testing phase at Station<sub>1</sub> and Station<sub>2</sub> respectively. A violin plot is a kind of box plot combined with kernel density plots to illustrate the probability distribution of a given data time series. It can be seen that MEMD-PSO-RF, MEMD-PSO-KRR and Standalone RF performed the best while standalone KRR performed worst. The figures also show that MEMD-PSO-RF model reproduced the shape of the density mass function of observed WS better in testing phase compared to training phase.

Figures 4a and b demonstrate the results obtained by the four models at Station<sub>1</sub> and Station<sub>2</sub> through the scatter plots of actual and predicted WS during testing phase. The figures show a very good agreement between the actual and predicted WS values for MEMD-PSO-RF model. Nonetheless, the overall performance of the evaluated models at both the stations revealed MEMD-PSO-RF model provides better accuracy during both training and testing, while the standalone KRR model performed worst during both the phases. It is important to note that the practicality of MEMD method in prediction of WS to increase the forecasting capacity of the RF model is a key development of this paper.

The predictive accuracy confirmed that MEMD-PSO-RF model can deliver healthier predictions of WS compared to other models in the study regions. The results also revealed that MEMD-PSO-RF is effective in extracting features from climatological variables in a tangible way. The performance of MEMD-PSO-RF also revealed that the PSO algorithm has

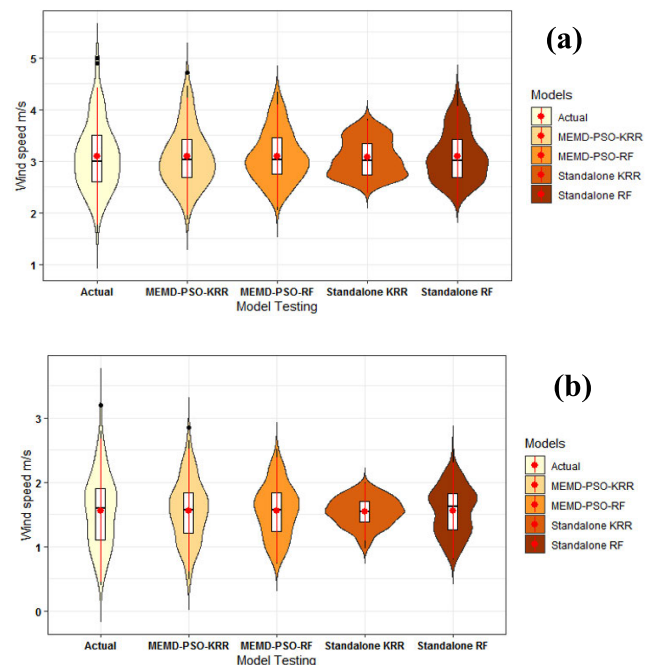


FIGURE 3. Estimated wind speed for observed, MEMD-PSO-KRR, MEMD-PSO-RF, Standalone KRR and Standalone RF models over the testing phase at a) Station<sub>1</sub> and b) Station<sub>2</sub>.

advantages in indicating the pertinent features to assist the RF in predicting WS time-series. In addition to the overall performance of MEMD-PSO-RF also confirmed the appropriateness of PSO in sorting out relevant IMFs with the assessment criteria of MEMD-PSO-RF method (i.e., Tables 1-2). Therefore, it remarkably improved the performance of



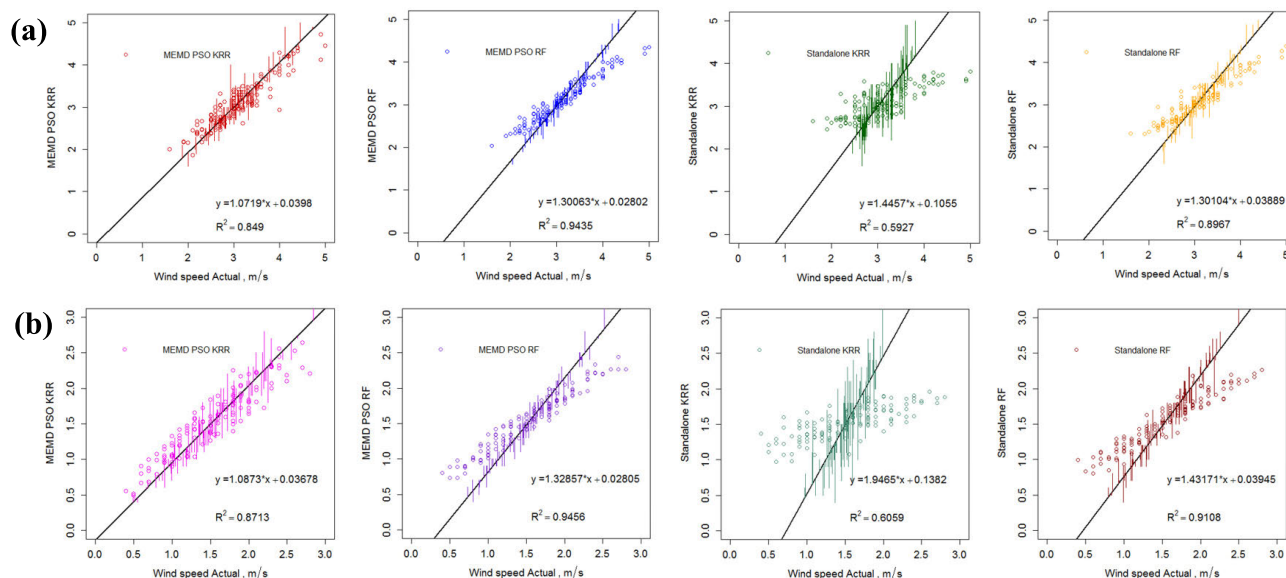


FIGURE 4. The scatter plots between the observed and estimated wind speed for all applied models over the testing period a) Station<sub>1</sub> and b) Station<sub>2</sub>.

MEMD-PSO-RF compared to MEMD-PSO-KRR and standalone counterpart models.

Since the artificial intelligence models exclusively depend on past data that may significantly affect the ‘learning’ and forecasting process, the outcomes of the present study established that an appropriate feature selection should be performed carefully before implementation of data-driven models. The MEMD is successfully classified and segregate the relevant features inside the climatological inputs to establish a more consistent physical foundation for a particular artificial intelligence method. The usefulness of MEMD in the present study is the concurrent data pre-processing of numerous climatological predictors. The MEMD can identify concurrently the signal’s main frequency to capture the respective features. This finding collaborates with the findings in [72]–[76]. Another key perception is that a smaller number of predictors with competitive accuracy is a parsimonious and computationally good model which is possible to achieve using MEMD-PSO-RF.

The feasibility of the MEMD method to predict WS is a major progression in this study. It improved the predicting ability of the RF and KRR models. It is apparent that better understandings of the physical procedure were given to the hybrid model, mainly by the MEMD method to the artificial intelligence model effectively capture the information in the meteorological variables in modelling WS.

The primary purpose of implementing MEMD in this study is its self-adaptive nature which involves minor human effort in factorizing IMFs. The MEMD conducts the data-driven based time-frequency investigation of multiple inputs by considering nonlinear behaviours via dynamical process [49]. Other major advantages of MEMD is its ability to handle the mode alignment concerns very efficiently [77]. Therefore,

the MEMD-PSO-RF has the potential for WS prediction and management systems. The proposed MEMD-PSO-RF model can be used as a WS modelling system for improving the efficiency of wind energy farm.

### V. CONCLUSION

This study provided new insights into environmental modelling by introducing the innovative integrative intelligent-based data-driven models, MEMD-PSO-RF and MEMD-PSO-KRR for WS modelling. The proposed models with enhanced short- and medium-term prediction accuracy were applied for modelling monthly WS at Baghdad and Mosul of Iraq. Several meteorological variables were used to build four models, MEMD-PSO-RF, MEMD-PSO-KRR, and standalone RF and standalone KRR. Predictive accuracy of the proposed models was evaluated using several performance measures.

Results indicated the superiority of MEMD-PSO-RF model in reproducing the WS time series in both the Baghdad and Mosul stations. The MEMD-PSO-KRR and standalone RF models also showed good performance. However, the standalone KRR was found to perform unsatisfactorily at both the stations. Since most of the meteorological variables used in this study as input are readily available in most regions, WS prediction using MEMD-PSO-RF model is feasible for practical applications. The results of these models can aid the practitioners to determine the windy areas for deployment of wind energy systems. Iraq often suffers from dust storms; the use of such models can also help in dust storm risk management. The models can help the authorities to determine the dust storm-prone areas and to adopt the appropriate strategies for dust storm mitigation. In future, the models developed in this study can be employed with a geographical information system for spatial prediction of WS over the whole country.

## ACKNOWLEDGMENT

Authors appreciate the climate data provider “Ministry of Agriculture and Water resources, Iraq”. The appreciation is extended to the respected editors and reviewers for their constructive comments.

## REFERENCES

- [1] X. Jia, C. Jin, M. Buzza, W. Wang, and J. Lee, “Wind turbine performance degradation assessment based on a novel similarity metric for machine performance curves,” *Renew. Energy*, vol. 99, pp. 1191–1201, Dec. 2016.
- [2] H. B. Azad, S. Mekhilef, and V. G. Ganapathy, “Long-term wind speed forecasting and general pattern recognition using neural networks,” *IEEE Trans. Sustain. Energy*, vol. 5, no. 2, pp. 546–553, Apr. 2014.
- [3] S. E. Haupt, G. Wiener, Y. Liu, B. Myers, J. Sun, D. Johnson, and W. Mahoney, “A wind power forecasting system to optimize power integration,” in *Proc. ASME 5th Int. Conf. Energy Sustainability A, B, C*, Jan. 2011, pp. 2215–2222.
- [4] X. Jia, M. Zhao, Y. Di, P. Li, and J. Lee, “Sparse filtering with the generalized  $l_p/l_q$  norm and its applications to the condition monitoring of rotating machinery,” *Mech. Syst. Signal Process.*, vol. 102, pp. 198–213, Mar. 2018.
- [5] A. Kovács, G. Erdős, Z. J. Viharos, and L. Monostori, “A system for the detailed scheduling of wind farm maintenance,” *CIRP Ann.*, vol. 60, no. 1, pp. 497–501, 2011.
- [6] H. Leng, X. Li, J. Zhu, H. Tang, Z. Zhang, and N. Ghadimi, “A new wind power prediction method based on ridgelet transforms, hybrid feature selection and closed-loop forecasting,” *Adv. Eng. Informat.*, vol. 36, pp. 20–30, Apr. 2018.
- [7] S. S. Soman, H. Zareipour, O. Malik, and P. Mandal, “A review of wind power and wind speed forecasting methods with different time horizons,” in *Proc. North Amer. Power Symp.*, Sep. 2010, pp. 1–8.
- [8] A. Tascikaraoglu and M. Uzunoglu, “A review of combined approaches for prediction of short-term wind speed and power,” *Renew. Sustain. Energy Rev.*, vol. 34, pp. 243–254, Jun. 2014.
- [9] M. A. Abdelkareem, M. El Haj Assad, E. T. Sayed, and B. Soudan, “Recent progress in the use of renewable energy sources to power water desalination plants,” *Desalination*, vol. 435, pp. 97–113, Jun. 2018.
- [10] J. Wan, G. Ren, J. Liu, Q. Hu, and D. Yu, “Ultra-short-term wind speed prediction based on multi-scale predictability analysis,” *Cluster Comput.*, vol. 19, no. 2, pp. 741–755, Jun. 2016.
- [11] J. Naik, P. Satapathy, and P. K. Dash, “Short-term wind speed and wind power prediction using hybrid empirical mode decomposition and kernel ridge regression,” *Appl. Soft Comput.*, vol. 70, pp. 1167–1188, Sep. 2018.
- [12] H. Liu, Z. Duan, Y. Li, and H. Lu, “A novel ensemble model of different mother wavelets for wind speed multi-step forecasting,” *Appl. Energy*, vol. 228, pp. 1783–1800, Oct. 2018.
- [13] Y. Jiang, N. Zhao, L. Peng, and S. Liu, “A new hybrid framework for probabilistic wind speed prediction using deep feature selection and multi-error modification,” *Energy Convers. Manage.*, vol. 199, Nov. 2019, Art. no. 111981.
- [14] E. Cadenas, W. Rivera, R. Campos-Amezcuca, and C. Heard, “Wind speed prediction using a univariate ARIMA model and a multivariate NARX model,” *Energies*, vol. 9, no. 2, p. 109, 2016.
- [15] Z. Zhang, L. Ye, H. Qin, Y. Liu, C. Wang, X. Yu, X. Yin, and J. Li, “Wind speed prediction method using shared weight long short-term memory network and Gaussian process regression,” *Appl. Energy*, vol. 247, pp. 270–284, Aug. 2019.
- [16] D. J. Allen, A. S. Tomlin, C. S. E. Bale, A. Skea, S. Vosper, and M. L. Gallani, “A boundary layer scaling technique for estimating near-surface wind energy using numerical weather prediction and wind map data,” *Appl. Energy*, vol. 208, pp. 1246–1257, Dec. 2017.
- [17] M. O. Mughal, M. Lynch, F. Yu, B. McGann, F. Jeanneret, and J. Sutton, “Wind modelling, validation and sensitivity study using weather research and forecasting model in complex terrain,” *Environ. Model. Softw.*, vol. 90, pp. 107–125, Apr. 2017.
- [18] A. Azad, M. Manoochehri, H. Kashi, S. Farzin, H. Karami, V. Nourani, and J. Shiri, “Comparative evaluation of intelligent algorithms to improve adaptive neuro-fuzzy inference system performance in precipitation modelling,” *J. Hydrol.*, vol. 571, pp. 214–224, Apr. 2019.
- [19] A. Azad, H. Kashi, S. Farzin, V. P. Singh, O. Kisi, H. Karami, and H. Sanikhani, “Novel approaches for air temperature prediction: A comparison of four hybrid evolutionary fuzzy models,” *Meteorol. Appl.*, 2019.
- [20] D. Ulkat and M. E. Günay, “Prediction of mean monthly wind speed and optimization of wind power by artificial neural networks using geographical and atmospheric variables: Case of Aegean region of Turkey,” *Neural Comput. Appl.*, vol. 30, no. 10, pp. 3037–3048, Nov. 2018.
- [21] V. Hoolohan, A. S. Tomlin, and T. Cockerill, “Improved near surface wind speed predictions using Gaussian process regression combined with numerical weather predictions and observed meteorological data,” *Renew. Energy*, vol. 126, pp. 1043–1054, Oct. 2018.
- [22] X. Zhao, J. Liu, D. Yu, and J. Chang, “One-day-ahead probabilistic wind speed forecast based on optimized numerical weather prediction data,” *Energy Convers. Manage.*, vol. 164, pp. 560–569, May 2018.
- [23] Y. Wang, H. Ma, D. Wang, G. Wang, J. Wu, J. Bian, and J. Liu, “A new method for wind speed forecasting based on copula theory,” *Environ. Res.*, vol. 160, pp. 365–371, Jan. 2018.
- [24] Y. Dong, J. Wang, and Z. Guo, “Research and application of local perceptron neural network in highway rectifier for time series forecasting,” *Appl. Soft Comput.*, vol. 64, pp. 656–673, Mar. 2018.
- [25] D. Zhang, X. Peng, K. Pan, and Y. Liu, “A novel wind speed forecasting based on hybrid decomposition and online sequential outlier robust extreme learning machine,” *Energy Convers. Manage.*, vol. 180, pp. 338–357, Jan. 2019.
- [26] M. Jamil and M. Zeeshan, “A comparative analysis of ANN and chaotic approach-based wind speed prediction in India,” *Neural Comput. Appl.*, vol. 31, no. 10, pp. 6807–6819, Oct. 2019.
- [27] J. Wang and Y. Li, “An innovative hybrid approach for multi-step ahead wind speed prediction,” *Appl. Soft Comput.*, vol. 78, pp. 296–309, May 2019.
- [28] M.-R. Chen, G.-Q. Zeng, K.-D. Lu, and J. Weng, “A two-layer nonlinear combination method for short-term wind speed prediction based on ELM, ENN, and LSTM,” *IEEE Internet Things J.*, vol. 6, no. 4, pp. 6997–7010, Aug. 2019.
- [29] H. Liu, Z. Duan, F.-Z. Han, and Y.-F. Li, “Big multi-step wind speed forecasting model based on secondary decomposition, ensemble method and error correction algorithm,” *Energy Convers. Manage.*, vol. 156, pp. 525–541, Jan. 2018.
- [30] Z. Qian, Y. Pei, H. Zareipour, and N. Chen, “A review and discussion of decomposition-based hybrid models for wind energy forecasting applications,” *Appl. Energy*, vol. 235, pp. 939–953, Feb. 2019.
- [31] S. Adarsh, A. P. John, R. N. Anagha, A. Abraham, M. P. Afia, K. K. Arathi, and A. Azeem, “Developing stage-discharge relationships using multivariate empirical mode decomposition-based hybrid modeling,” *Appl. Water Sci.*, vol. 8, no. 8, Dec. 2018.
- [32] S. Adarsh, S. Sanah, K. K. Murshida, and P. Nooramol, “Scale dependent prediction of reference evapotranspiration based on multi-variate empirical mode decomposition,” *Ain Shams Eng. J.*, vol. 9, no. 4, pp. 1839–1848, Dec. 2018.
- [33] W. Hu and B. C. Si, “Soil water prediction based on its scale-specific control using multivariate empirical mode decomposition,” *Geoderma*, vols. 193–194, pp. 180–188, Feb. 2013.
- [34] K. He, R. Zha, J. Wu, and K. Lai, “Multivariate EMD-based modeling and forecasting of crude oil price,” *Sustainability*, vol. 8, no. 4, p. 387, 2016.
- [35] L. E. Andersson, M. F. Aftab, F. Scibilia, and L. Imsland, “Forecasting using multivariate empirical mode decomposition—Applied to iceberg drift forecast,” in *Proc. 1st Annu. IEEE Conf. Control Technol. Appl. (CCTA)*, Aug. 2017, pp. 1097–1103.
- [36] N. D. Bokde, A. Feijoo, N. Al-Ansari, and Z. M. Yaseen, “A comparison between reconstruction methods for generation of synthetic time series applied to wind speed simulation,” *IEEE Access*, vol. 7, pp. 135386–135398, 2019.
- [37] L. Breiman, “Random Forests,” *Mach. Learn.*, vol. 45, no. 1, pp. 5–32, 2001.
- [38] C. Robert, “Machine learning, a probabilistic perspective,” *Chance*, vol. 27, no. 2, pp. 62–63, 2014.
- [39] F. Douak, F. Melgani, and N. Benoudjit, “Kernel ridge regression with active learning for wind speed prediction,” *Appl. Energy*, vol. 103, pp. 328–340, Mar. 2013.
- [40] J. Chenoweth, P. Hadjinicolaou, A. Bruggeman, J. Lelieveld, Z. Levin, M. A. Lange, E. Xoplaki, and M. Hadjikakou, “Impact of climate change on the water resources of the eastern mediterranean and middle east region: Modeled 21st century changes and implications,” *Water Resour. Res.*, vol. 47, no. 6, Jun. 2011.

- [41] N. Al-Ansari, A. A. Ali, and S. Knutsson, "Present conditions and future challenges of water resources problems in Iraq," *J. Water Resource Protection*, vol. 6, no. 12, pp. 1066–1098, 2014.
- [42] J. Lelieveld, P. Hadjinicolaou, E. Kostopoulou, J. Chenoweth, M. El Maayar, C. Giannakopoulos, C. Hannides, M. A. Lange, M. Tanarhte, E. Tyrlis, and E. Xoplaki, "Climate change and impacts in the eastern mediterranean and the middle east," *Climatic Change*, vol. 114, nos. 3–4, pp. 667–687, Oct. 2012.
- [43] N. Abbas, S. A. Wasimi, N. Al-Ansari, and N. Sultana, "Water resources problems of Iraq: Climate change adaptation and mitigation," *J. Environ. Hydrol.*, vol. 26, no. 6, pp. 1–6, 2018.
- [44] R. Moazenzadeh, B. Mohammadi, S. Shamsirband, and K.-W. Chau, "Coupling a firefly algorithm with support vector regression to predict evaporation in northern Iran," *Eng. Appl. Comput. Fluid Mech.*, vol. 12, no. 1, pp. 584–597, Jan. 2018.
- [45] F. A. Vishkaee, C. Flamant, J. Cuesta, L. Oolman, P. Flamant, and H. R. Khalesifard, "Dust transport over Iraq and northwest Iran associated with winter Shamal: A case study," *J. Geophys. Res., Atmos.*, vol. 117, no. D3, Feb. 2012.
- [46] R. E. Schapire, Y. Freund, P. Bartlett, and W. S. Lee, "Boosting the margin: A new explanation for the effectiveness of voting methods," *Ann. Statist.*, vol. 26, no. 5, pp. 1651–1686, Oct. 1998.
- [47] L. Breiman, "Bagging predictors," *Mach. Learn.*, vol. 24, no. 2, pp. 123–140, Aug. 1996.
- [48] L. E. O. Breiman, *Random Forest.pdf*. Norwell, MA, USA: Kluwer, 2001.
- [49] N. Rehman and D. P. Mandic, "Multivariate empirical mode decomposition," *Proc. Roy. Soc. A, Math. Phys. Eng. Sci.*, vol. 466, no. 2117, pp. 1291–1302, 2010.
- [50] N. Ur Rehman and D. P. Mandic, "Filter bank property of multivariate empirical mode decomposition," *IEEE Trans. Signal Process.*, vol. 59, no. 5, pp. 2421–2426, May 2011.
- [51] J.-R. Huang, S.-Z. Fan, M. Abbod, K.-K. Jen, J.-F. Wu, and J.-S. Shieh, "Application of multivariate empirical mode decomposition and sample entropy in EEG signals via artificial neural networks for interpreting depth of anesthesia," *Entropy*, vol. 15, no. 12, pp. 3325–3339, 2013.
- [52] D. P. Mandic, N. U. Rehman, Z. Wu, and N. E. Huang, "Empirical mode decomposition-based time-frequency analysis of multivariate signals: The power of adaptive data analysis," *IEEE Signal Process. Mag.*, vol. 30, no. 6, pp. 74–86, Nov. 2013.
- [53] R. Prasad, M. Ali, P. Kwan, and H. Khan, "Designing a multi-stage multivariate empirical mode decomposition coupled with ant colony optimization and random forest model to forecast monthly solar radiation," *Appl. Energy*, vol. 236, pp. 778–792, Feb. 2019.
- [54] J. Kennedy and R. Eberhart, "Particle swarm optimization," in *Proc. Int. Conf. IEEE Neural Netw.*, vol. 4, Nov./Dec. 1995, pp. 1942–1948.
- [55] E. Assareh, M. A. Behrang, M. R. Assari, and A. Ghanbarzadeh, "Application of PSO (particle swarm optimization) and GA (genetic algorithm) techniques on demand estimation of oil in Iran," *Energy*, vol. 35, no. 12, pp. 5223–5229, Dec. 2010.
- [56] A. Azad, H. Karami, S. Farzin, S.-F. Mousavi, and O. Kisi, "Modeling river water quality parameters using modified adaptive neuro fuzzy inference system," *Water Sci. Eng.*, vol. 12, no. 1, pp. 45–54, Mar. 2019.
- [57] O. Kisi, A. Azad, H. Kashi, A. Saedian, S. A. A. Hashemi, and S. Ghorbani, "Modeling groundwater quality parameters using hybrid neuro-fuzzy methods," *Water Resour. Manage.*, vol. 33, no. 2, pp. 847–861, Jan. 2019.
- [58] Y. Zhang, J. Duchi, and M. Wainwright, "Divide and conquer kernel ridge regression," in *Proc. Conf. Learn. Theory*, 2013.
- [59] V. Vovk, "Kernel ridge regression," in *Empirical Inference*. Berlin, Germany: Springer, 2013, pp. 105–116.
- [60] A. El Alaoui and M. W. Mahoney, "Fast randomized kernel ridge regression with statistical guarantees?" in *Proc. Adv. Neural Inf. Process. Syst. (NIPS)*, 2015, pp. 775–783.
- [61] B. C. Yen, "Discussion and closure: Criteria for evaluation of watershed models," *J. Irrigation Drainage Eng.*, vol. 121, no. 1, pp. 130–132, Jan. 1995.
- [62] American Society of Civil Engineers Task Committee, "Artificial neural networks in hydrology. II: Hydrologic applications," *J. Hydrol. Eng.*, vol. 5, no. 2, pp. 124–137, 2000.
- [63] ASCE Task Committee, "Criteria for evaluation of watershed models," *J. Irrigation Drainage Eng.*, vol. 119, no. 3, pp. 429–442, May 1993.
- [64] C. W. Dawson, R. J. Abrahart, and L. M. See, "HydroTest: A Web-based toolbox of evaluation metrics for the standardised assessment of hydrological forecasts," *Environ. Model. Softw.*, vol. 22, no. 7, pp. 1034–1052, Jul. 2007.
- [65] R. C. Deo, X. Wen, and F. Qi, "A wavelet-coupled support vector machine model for forecasting global incident solar radiation using limited meteorological dataset," *Appl. Energy*, vol. 168, pp. 568–593, Apr. 2016.
- [66] D. R. Legates and G. J. McCabe, "Evaluating the use of 'goodness-of-fit' measures in hydrologic and hydroclimatic model validation," *Water Resour. Res.*, vol. 35, no. 1, pp. 233–241, Jan. 1999.
- [67] C. J. Willmott, "Some comments on the evaluation of model performance," *Bull. Amer. Meteorol. Soc.*, vol. 63, no. 11, pp. 1309–1313, Nov. 1982.
- [68] C. J. Willmott, "On the validation of models," *Phys. Geogr.*, vol. 2, no. 2, pp. 184–194, 1981.
- [69] C. J. Willmott, "On the evaluation of model performance in physical geography," in *Spatial Statistics and Models*. Dordrecht, The Netherlands: Springer, 1984, pp. 443–446.
- [70] C.-W. Hsu, C.-C. Chang, and C.-J. Lin, "A practical guide to support vector classification," *BJU Int.*, vol. 101, no. 1, pp. 1396–1400, 2008.
- [71] J. L. Hintze and R. D. Nelson, "Violin plots: A box plot-density trace synergism," *Amer. Statistician*, vol. 52, no. 2, pp. 181–184, May 1998.
- [72] A. Badr and A. Fahmy, "A proof of convergence for ant algorithms," *Inf. Sci.*, vol. 160, nos. 1–4, pp. 267–279, Mar. 2004.
- [73] C. Garcia, O. Herrera, F. Triguero, and T. Stützle, "A review on the ant colony optimization metaheuristic: Basis, models and new trends," *Mathware Soft Comput.*, vol. 9, nos. 2–3, 2002.
- [74] R. J. Mullen, D. Monekosso, S. Barman, and P. Remagnino, "A review of ant algorithms," *Expert Syst. Appl.*, vol. 36, no. 6, pp. 9608–9617, 2009.
- [75] G. Singh, N. Kumar, and A. K. Verma, "Ant colony algorithms in MANETS: A review," *J. Netw. Comput. Appl.*, vol. 35, no. 6, pp. 1964–1972, Nov. 2012.
- [76] J. D. Sweetlin, H. K. Nehemiah, and A. Kannan, "Feature selection using ant colony optimization with tandem-run recruitment to diagnose bronchitis from CT scan images," *Comput. Methods Programs Biomed.*, vol. 145, pp. 115–125, Jul. 2017.
- [77] D. Looney and D. P. Mandic, "Multiscale image fusion using complex extensions of EMD," *IEEE Trans. Signal Process.*, vol. 57, no. 4, pp. 1626–1630, Apr. 2009.



HAI TAO received the B.Sc. degree from the Department of Computer and Information Science, Northwest University of Nationalities, in 2004, the M.S. degree from the School of Mathematics and Statistics, Lanzhou University, in 2009, and the Ph.D. degree from the Faculty of Computer System and Software Engineering, Universiti Malaysia Pahang, in 2012. He is currently an Associate Professor with the Baoji University of Arts and Sciences. His current research interests include machine learning, the Internet of Things, and optimization computation.



SINAN Q. SALIH received the B.Sc. degree in information systems from the University of Anbar, Anbar, Iraq, in 2010, the M.Sc. degree in computer sciences from Universiti Tenaga National (UNITEN), Malaysia, in 2012, and the Ph.D. degree in soft modeling and intelligent systems from Universiti Malaysia Pahang (UMP). His current research interests include optimization algorithms, nature-inspired metaheuristics, machine learning, and feature selection problem for real world problems.



ing. Her research areas of a topic are crop modeling and irrigation water management.

**MANDEEP KAUR SAGGI** received the bachelor's degree in computer science and engineering from Punjab Technical University and the M.Tech. degree in computer science engineering from D. A. V. University, Jalandhar. She is currently pursuing the Ph.D. degree in computer science engineering with the Thapar Institute of Engineering and Technology, Patiala. Her area of interest include in machine learning, big data analytics, quantum computing, network security, and cloud computing.



**ESMAEEL DODANGEH** received the Ph.D. degree in watershed science and engineering from Sari Agricultural Science and Natural Resources University (SANRU), Sari, Iran. His research interests include hydrological modeling, extreme hydrological (drought and flood) frequency analysis, machine learning, and uncertainty analysis of bivariate hydrological frequency analysis. He is very interested in non-linear and complex hydrological systems.



process, ARMA, and Bayesian neural networks), to classical prediction tools with recursive formalism (Kalman filter), and to methodologies combining time series models and surface solar irradiation. He also works about the Monte Carlo algorithms to model the dose deposition of atmospheric neutrons and the modeling of the radiobiological effects.

**CYRIL VOYANT** received the Qualification Diploma degree in radiological and medical physics in 2003 and the Ph.D. degree in the use of artificial intelligence in solar radiation forecasting in 2011. He is currently a Medical Physicist at the Castelluccio Hospital, Ajaccio, France, and an Invited Professor at the Renewable Energy Laboratory in Corte, University of Corsica, France. His research activities are related to the stochastic modelling tools (ANN, SVM, SVR, Gaussian



His research interests are mainly in geology, water resources, and environment. He served as the dean and the head of department for several academic administrative posts. His publications include more than 424 articles in international/national journals, chapters in books, and 13 books. He executed more than 60 major research projects in Iraq, Jordan, and U.K. He received several scientific and educational awards, among them is the British Council on its 70th Anniversary awarded him top 5 scientists in Cultural Relations. He holds one patent on physical methods for the separation of iron oxides. He has supervised more than 66 postgraduate students at Iraq, Jordan, U.K., and Australia universities. He is a member of several scientific societies, e.g., International Association of Hydrological Sciences, Chartered Institution of Water and Environment Management, and Network of Iraqi Scientists Abroad and the Founder and President of the Iraqi Scientific Society for Water Resources. He is also a member of the editorial board of ten international journals.

**NADHIR AL-ANSARI** received the B.Sc. and M.Sc. degrees from the University of Baghdad, in 1968 and 1972, respectively, and the Ph.D. degree in water resources engineering from Dundee University, in 1976. He is currently a Professor at the Department of Civil, Environmental and Natural Resources Engineering, Lulea University of Technology, Sweden. Previously, he worked at Baghdad University, from 1976 to 1995, and then at Al-Bayt University, Jordan, from 1995 to 2007.



analytics. He has published over 100 research articles in international journals with a Google Scholar H-Index of 25, and a total of 1900 citations.

**ZAHER MUNDHER YASEEN** received the master's and Ph.D. degrees from the National University of Malaysia (UKM), Malaysia, in 2012 and 2017, respectively. He is currently a Senior Lecturer and a Senior Researcher in the field of civil engineering at Ton Duc Thang University. He is major in hydrology, water resources engineering, hydrological processes modeling, environmental engineering, and climate. In addition, he has major interest in machine learning and advanced data



statistical and mathematical tools for innovative solutions of hydrological problems for adaptation to global environmental changes. He published about 100 research articles in internationally reputed indexed journals and five academic books. His major research achievements in recent years include forecasting water demand in a holistic way, downscaling climate using an alternative approach for better projection of climate change, modeling economical impacts of global warming induced changes in water resources, modeling droughts during crop-growing seasons, projection of drought severity-area-frequency curves for adaptation to climate change impacts on droughts and water stress, multi-criteria decision making for selection of adaptation and mitigation strategies, and determination of unidirectional trends to distinguish global climate change from natural variability. He has successfully completed about ten national and international research projects, as a Principal Investigator.

**SHAMSUDDIN SHAHID** is currently an Associate Professor at the Department of Hydraulic and Hydrology, Faculty of Civil Engineering, Universiti Teknologi Malaysia (UTM). He is also the Head of the Integrated Water Resources Management (IWRM) research group, UTM. His major research interests include water resources management, climate change impacts and adaptation to water resources, statistical hydrology, hydrological disasters, and groundwater hydrology. He uses

...