



# Predicting daily wind speed using coupled multi-layer perceptron model with water strider optimization algorithm based on fuzzy reasoning and Gamma test

Mohammad Ehteram<sup>1</sup> · Fatemeh Panahi<sup>2</sup> · Nouar AlDahoul<sup>3</sup> · Ali Najah Ahmed<sup>4</sup>  · Yuk Feng Huang<sup>5</sup> · Ahmed Elshafie<sup>6,7</sup>

Accepted: 6 March 2024

© The Author(s), under exclusive licence to Springer-Verlag GmbH Germany, part of Springer Nature 2024

## Abstract

Wind energy is a valuable renewable resource that plays a significant role in electricity generation process. Predicting wind speed (W.S.) is critical for effectively managing wind energy and producing power. This study proposes an improved multi-layer perceptron (MLP) model for W.S. prediction that incorporates a novel optimization algorithm namely water strider algorithm (WSA). The WSA optimizes the MLP parameters to increase the model's accuracy. The MLP-WSA model's predictive capabilities were compared with various algorithms, such as MLP-sine cosine (SCA), MLP-salp swarm (MLP-SSA), Multi-Layer Perceptron-particle swarm optimization (MLP-PSO), and MLP models. Furthermore, we proposed an inclusive multiple model (IMM) that utilizes the outputs of the MLP-WSA to predict W.S. The study utilized fuzzy reasoning to modify MLP models to remove redundant weights and reduce computation time. Finally, to predict W.S, we considered five stations in Malaysia. By utilizing the WSA and Gamma tests, we identified that the IMM model provided the most optimal input for our model. To test the I.P. station, the RMSE of the IMM model was lower than other models. Additionally, the NSE of the IMM model was found to be higher at the B.L. station, indicating superior performance. Furthermore, the IMM model's mean absolute error was notably lower than other models in C.H. station. Overall, the results demonstrate that the combination of the WSA and Gamma tests allowed us to achieve more accurate and efficient predictions with less computation time using fuzzy reasoning.

**Keywords** Multilayer perceptron · Wind speed · Water strider optimization · Gamma test

## 1 Introduction

Energy shortage is a real challenge for different countries. The development and future of each country rely on the control of energy. For fossil fuels, when the consumption increases, the energy reserves decreases and the environmental pollution increases (Zhang et al. 2020). Today, renewable and clean energies such as wind and solar energies are excellent alternatives to fossil fuels for energy production (Samadianfard et al. 2020). The renewable energy has different advantages. It decreases some types of air pollution. Also, renewable energies create economic development and jobs in manufacturing (Liu et al. 2020).

A unique method of desalination using solar energy has been proposed, and its potential effectiveness has been

assessed through a proven model of humidification–dehumidification (Abedi et al. 2023a). It has been demonstrated that if a turbine is used to generate electricity for the desalination system, the plant can supply freshwater to approximately 800 homes. Numerous machine learning regression approaches were utilized to build a surrogate model based on data from the dehumidifier component (Abedi et al. 2023b).

Wind energy, a principle renewable energy, is cost-effective and creates jobs. Also, wind energy is used for developing industries and economies. Wind energy can be used for producing power generation without environmental pollution (Kumar 2020). Wind power generation is an important technology for producing power and developing different countries' industries and economies (Xu et al. 2021). Wind power generation is used to converts wind energy to electric energy. Wind power generation can

Extended author information available on the last page of the article

enhance supplies and facilitate the reclamation of degraded land. To better manage wind power generation, the research works have focused on predicting wind power and wind speed. The W.S. prediction is complex because of the chaotic fluctuations of W.S. There are different methods for predicting W.S., such as physical models, soft computing models (SCMs), and spatial correlation models (Liu et al. 2012). The geographic and geomorphic data are required to establish physical models and different data from the different measurement stations to establish the spatial correlation models. Modelers may encounter challenges in accessing various types of data, including climate, geographic, and geomorphic data. To address this issue, some researchers have turned to using soft computing models (SCMs) to predict various variables. SCMs offer several advantages, such as their high accuracy, ability to handle complex systems, flexibility in coupling with different models and algorithms, and ease of implementation (Ehteram et al. 2021). Notably, researchers have also explored the potential of using SCMs for wind speed prediction. For example, one study demonstrated the effectiveness of combining a hybrid MLP model with self-organizing feature maps to enhance the model's accuracy (Gnana Sheela and Deepa 2013). Research has shown that a coupled model utilizing the multi-layer perceptron (MLP) technique outperforms the standalone MLP model for wind speed (W.S.) prediction. Similarly, the use of coupled PSO with the support vector machine (SVM) produced greater accuracy than the standalone SVM model for predicting W.S. (Kong et al. 2015). Additionally, an artificial neural network (ANN) has demonstrated the ability to accurately predict W.S. with a mean absolute percentage error of 6.48% using altitude, solar radiation air pressure, and air temperature (Ramasamy et al. 2015).

In Kumar and Malik (2016), researchers explored the effectiveness of both the MLP and Generalized Regression Neural Network (GRNN) for predicting wind speed (W.S.). The findings showed that the GRNN outperformed the MLP. Meanwhile, Zhang et al. (2016) conducted a study to evaluate the potential of Gaussian Process Regression (GPR) for W.S. prediction. The results demonstrated that the GPR is more accurate than ANN and SVM techniques.

In Ahmed et al. (2016), the researchers combined the adaptive neuro-fuzzy interface system (ANFIS) with the krill optimization algorithm to predict wind speed (W.S.). They utilized the krill algorithm to optimize the parameters of the ANFIS model. They demonstrated that the combined ANFIS-krill approach enhanced the accuracy of the standalone ANFIS model. In another study, Liu et al. (2018) applied a Convolutional Long Short-Term Memory (CLSTM) to predict W.S. They showed that the CLSTM had better performance compared to the Convolutional neural network.

Researchers in Yu et al. (2018) examined the capacity of a hybrid SCM for predicting wind speed (W.S.). The method involved decomposing the original wind speed history using wavelet transform and using a Recurrent Neural Network (RNN) to extract the more profound features of the data, which were then fed into the SVM model. The results showed that the hybrid model had significant accuracy in W.S. prediction. Similarly, Samadianfard et al. (2020) applied MLP with genetic and whale optimization algorithms to boost the performance of predicting W.S. The optimized MLP models were tested across different climatic regions of Iran, where it was observed that the MLP-Whale Optimization Algorithm (WOA) hybrid model outperformed the standalone MLP model.

Research in Navas et al. (2020) focused on comparing the predictive accuracy of different models including MLP, Radial Basis Function Neural Network (RBFNN), and Categorical Regression for predicting wind speed (W.S.). The study revealed that the MLP had a better accuracy than other models. On the other hand, Sun et al. (2020) looked for the performance of a coupled Multi-Kernel Least Square SVM (MLKSSVM) and Gravitational Search Algorithm (GSA) for W.S. prediction. The GSA was utilized for MLKSSVM's parameters optimization. The results demonstrated that the optimized MLKSSVM model increased the accuracy of W.S. prediction.

While various SVMs have been demonstrated to have a high capacity for predicting wind speed (W.S.), challenges still exist. One major challenge is that the SVMs' structure contains parameters that need to be accurately identified to ensure model accuracy. To address this issue, robust training algorithms are necessary to obtain precise parameter values. Another issue is that most previous studies have focused on comparing the models' performance without exploring how different models could be integrated to achieve improved accuracy. Finally, it is crucial for the ideal model to predict the target variable within a short computational time.

This study aims to tackle the previously mentioned issues through various efforts, including:

- (1) the use of a water strider (WSA) optimization algorithm to train Multiple Linear Regression (MLR) models capable of predicting daily weather station (W.S.) values in five different locations throughout Malaysia. In their introduction of the WSA algorithm, Kaveh and Dadras (2020) explained that it was inspired by the behavior of water striders, a type of insect known for its remarkable ability to walk on the surface of water. According to their findings, this algorithm is highly accurate in solving complex problems and also possesses an ideal balance between exploration and exploitation while

rapidly converging toward optimal solutions. Due to these benefits, the current study employs the WSA.

- (2) If modelers have access to a variety of climate data, predicting W.S. can be relatively straightforward. However, in some cases, particularly in developing countries, data on climate patterns may be limited to just time series data on W.S. In this scenario, modelers must rely on lagged wind speed data to predict wind speed at the current time. This presents a unique challenge as the goal is to establish a highly effective model for predicting wind speed using only limited input data. To tackle this issue, the present study employs an MLP-WSA algorithm that utilizes lagged W.S. data to accurately predict daily wind speed values.
- (3) A new hybrid model has been introduced and referred to as the Gamma Test, which serves as a novel approach to select the optimal input combination when utilizing lagged W.S. data. The WSA algorithm is combined with the Gamma test to find the most proper set of inputs for the MLP model, thereby improving the overall accuracy of predictions.
- (4) To comprehensively evaluate the predictive capabilities of the MLP-WSA used, it is compared with several other variants of MLP such as MLP-Sine Cosine Algorithm (MLP-SCA), MLP-Salp Swarm Algorithm (MLP-SSA), MLP-PSO, and traditional MLP. An integrated multi-model approach is employed to leverage the strengths of each individual model and enhance the accuracy of predictions.
- (5) To increase the efficiency of the MLP variants developed in this study, an approach is adopted to identify and remove redundant weights that do not significantly impact predictions. This helps to reduce computation time and improve overall model performance. A fuzzy reasoning concept is applied to successfully identify and eliminate the unnecessary weights from the MLP models.

This study proposes four key innovations:

- (1) Developing a new hybrid MLP and MLP-WSA model for predicting daily wind speed (W.S.).
- (2) To select the optimal input variables, a novel hybrid Gamma test was created.
- (3) Presenting a comprehensive multi-model approach for enhancing the accuracy of predictions by integrating various models.
- (4) Using the fuzzy reasoning concept to reduce the computational time of the Multi-Layer Perceptron (MLP) models.

Section 2 discusses the material and methods. In Sect. 3, a case study is presented along with its relevant details. Section 4, we present the results of this study. Finally, in Sect. 5, we draw a conclusion based on the findings.

## 2 Materials and methods

### 2.1 Multilayer perceptron (MLP)

MLP is a significant type of ANN. The basic unit of computation in MLP are neurons, which connect to the next layer via weight connections (Muslim et al. 2020b). Incoming data are received by the first layer and processed using activation functions in hidden layers, and finally, the final layer produces the overall result, according to the equation outlined below

$$\text{Out}_k = f_{\text{out}} \left[ \sum_{j=1}^N w_{kj} \times f_h \left[ \sum_{i=1}^n w_{ji} \text{in}_i + B_j \right] + B_k \right], \quad (1)$$

where  $i$ : index of inputs,  $j$ : index of hidden nodes,  $k$  index of outputs,  $w_{ji}$ : the weight connection linked the input to the hidden layer,  $B_k$ : the bias of the output layer,  $B_j$ : the bias of the hidden layer,  $\text{in}_i$ : the inputs,  $N$ : hidden layer's nodes, and  $n$ : number of inputs.  $f_h$ : activation function of the hidden layer, and  $f_{\text{out}}$ : activation function of final layer. Given the success of sigmoid function (SIG) in prior research studies (Banadkooki et al. 2020; Ehteram et al. 2020; Najah Ahmed et al. 2019), it was chosen as the activation

$$f(\text{SIG}) = \frac{1}{1 + e^{-\text{SIG}}}. \quad (2)$$

The process of training the MLP models in this study involves transforming the received signals (SIG) into the activation function through backpropagation. Initially, weights and biases are randomly assigned, and the SIG is then fed into the first layer to generate output values. Subsequently, the error function is calculated to assess the difference between actual and predicted values. In the backward pass, updates are made to weights and biases to decrease the error function. While the backpropagation algorithm is effective for this process, it may converge too slowly or become stuck in local optima. Therefore, optimization algorithms are used to improve the performance of the MLP, as shown in Fig. 1.

The hyperparameters of the MLP are as follows:

Batch size is 16.

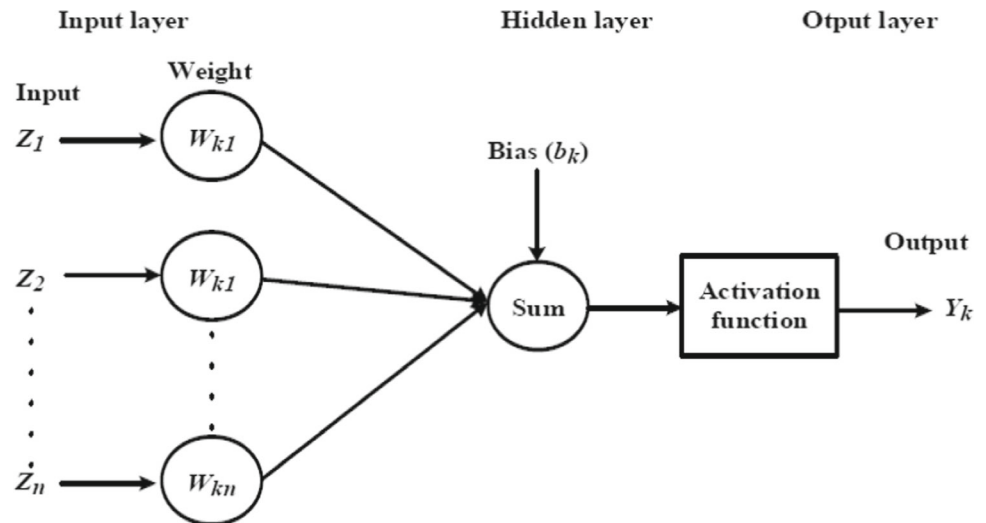
Hidden layers' number is 1.

Hidden layer's nodes are 32.

Activation function in hidden layers is Sigmoid.

Activation function in output layer is Linear.

**Fig. 1** The structure of the MLP model



Optimizer is Stochastic Gradient Decent (SGD).  
Loss function is Mean Squared Error (MSE).

## 2.2 Water strider optimization algorithm (WSOA)

As one of the kinds of insects, the water striders live on water surface top. Water striders claim ownership of specific areas known as territories, which they protect to ensure access to their food and potential mates (Kaveh and Dadras 2020). The social communication of the water striders is performed through provided ripples. The water striders can produce ripples with different amplitudes. The generated ripples are used for different aims, such as sex discrimination and prey locating. The female W.S.s are eager to find the food, while the males are eager to create the mating (Kaveh and Dadras 2020). When the females receive the signals from the male WSs, the response of female W.S.s will be based on the attraction and repulsive signals. Males skate in the females' areas, because the females are eager to find the best location for finding food. In the first stage, the following identifies the initial location of W.S.:

$$X_i^0 = UB + \text{rand}(UB - LB), \quad (3)$$

where  $X_i^0$ : the initial location of W.S.s, UB: the upper bound, and LB: the lower bound. In the next stage, the territories are created by the W.S.s. In this level, the objective function is computed for W.S. Then, the W.S.s are sorted based on the obtained values for their objective function. The W.S.s are divided into  $\frac{\text{number of WSs in each group}}{\text{number of territories}}$  groups. In the next level, the mating behavior is modeled. If  $p$  refers to probability of positive feedback of females to males for mating,  $(1 - p)$  indicates probability of females ignoring the males for

mating. If the females are not eager to the mating, the females get males away. The W.S.s update their location after mating as follows:

$$\begin{cases} X_i^{t+1} = X_i^t + R \cdot \text{rand} \leftarrow (\text{if})(\text{mating})(\text{happens}) \\ X_i^{t+1} = X_i^t + R \cdot (1 + \text{rand}) \end{cases}, \quad (4)$$

where  $X_i^{t+1}$ : the new location of W.S.s,  $R$ : radius ripple wave, and  $\text{rand}$ : random values between 0 and 1

$$R = X_F^{t-1} - X_i^{t-1}, \quad (5)$$

where  $X_F^{t-1}$ : the female W.S., and location  $X_i^{t-1}$ : the male W.S. location. When the W.S. update its position, the objective function is calculated for the W.S. new location. If the new location has not better objective function than the previous location, the W.S. moves to the best location for finding food as follows:

$$X_i^{t+1} = X_i^t + 2 \cdot \text{rand} (X_{\text{best}} - X_i^t), \quad (6)$$

where  $X_{\text{best}}$ : the best location for the W.S. If the W.S. in the new location has not better objective function than the W.S. in the previous location, the W.S. will die and a larva is generated. The location of Larva is as follows:

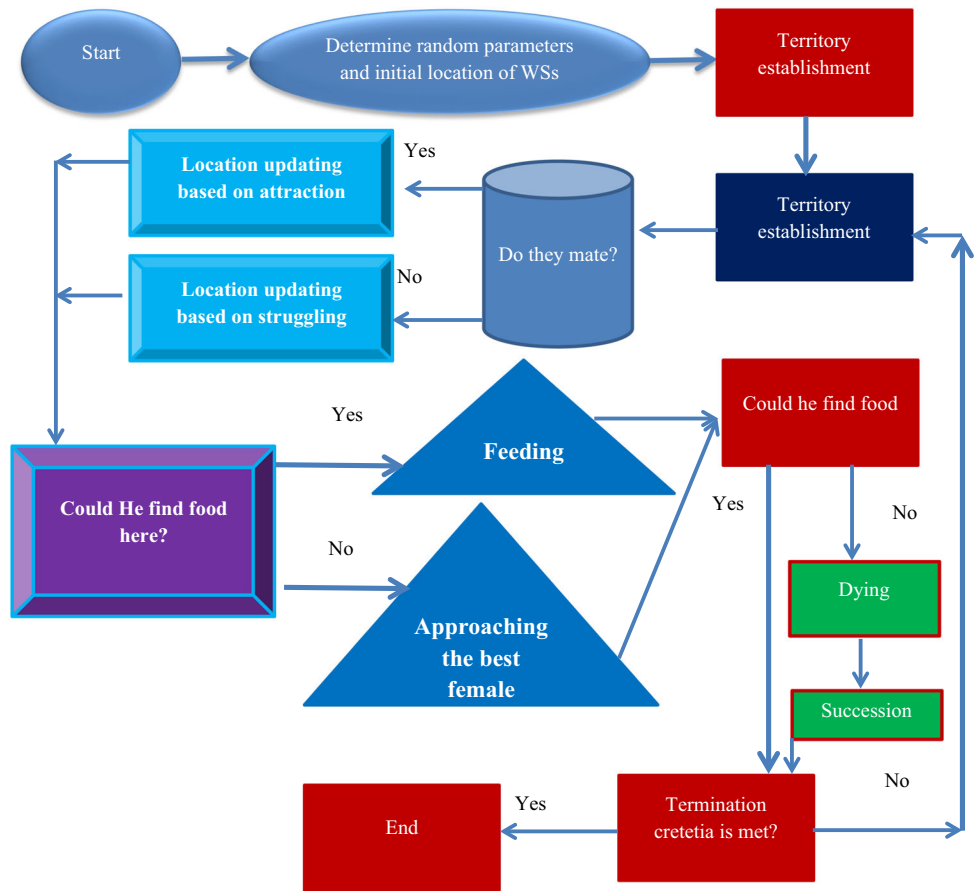
$$X_i^{t+1} = LB_j^t + 2 \cdot \text{rand} (UB_j^t - LB_j^t), \quad (7)$$

where  $LB_j^t$ : lower values of W.S.'s position inside  $j$ th territory and  $UB_j^t$ : upper values of the  $UB_j^t$ . Figure 2 shows the WSA flowchart.

## 2.3 Salp swarm algorithm (SSA)

This algorithm is utilized for various tasks such as feature selection (Tubishat et al. 2021), engineering optimization problems (Salgotra et al. 2021), training SVM (Li et al.

**Fig. 2** The WSA flowchart for optimization problem



2020), and training ANFIS (Mohamadi et al. 2020). Group life is observed for the salps. In each group, there are a leader and followers. A leader guides follower. Each leader updates its location as follows:

$$sa_j^1 = \begin{cases} food_j + \sigma_1((up_j - lo_j)\sigma_2 + lo_j) & \leftarrow \sigma_3 \geq 0 \\ food_j - \sigma_1((up_j - lo_j)\sigma_2 + lo_j) & \leftarrow \sigma_3 < 0 \end{cases}, \quad (8)$$

where  $sa_j^1$ : the location of leader,  $\sigma_1$ ,  $\sigma_2$ , and  $\sigma_3$ : random numbers,  $food_j$ : the location of food source,  $up_j$ : the upper bound, and  $lo_j$ : the lower bound. A balance is provided between the exploration and exploitation as follows:

$$\sigma_1 = 2e^{-\left(\frac{l}{L}\right)^2}, \quad (9)$$

where  $l$ : number of iterations and  $L$ : maximum number of iterations. The follower in each iteration changes its location as follows:

$$sa_j^i = \frac{1}{2} (sa_j^i + sa_j^{i-1}), \quad (10)$$

where  $sa_j^i$ : the location of each follower in  $j$ th dimension. Figure 3 shows the SSA flowchart for optimization problems.

## 2.4 Sine cosine algorithm (SCA)

SCA algorithm was inspired by the sine and cosine functions utilized for different optimization problems such as biomedical signal reconstruction (Daoui et al. 2021), engineering applications (Dhiman 2021), optimal multi-robot path planning (Paikray et al. 2021), feature selection (Neggaz et al. 2020), and image segmentation (Ewees et al. 2020). First, random solutions are created. Then, the final position of solutions is found based on the current location of solutions and destination point

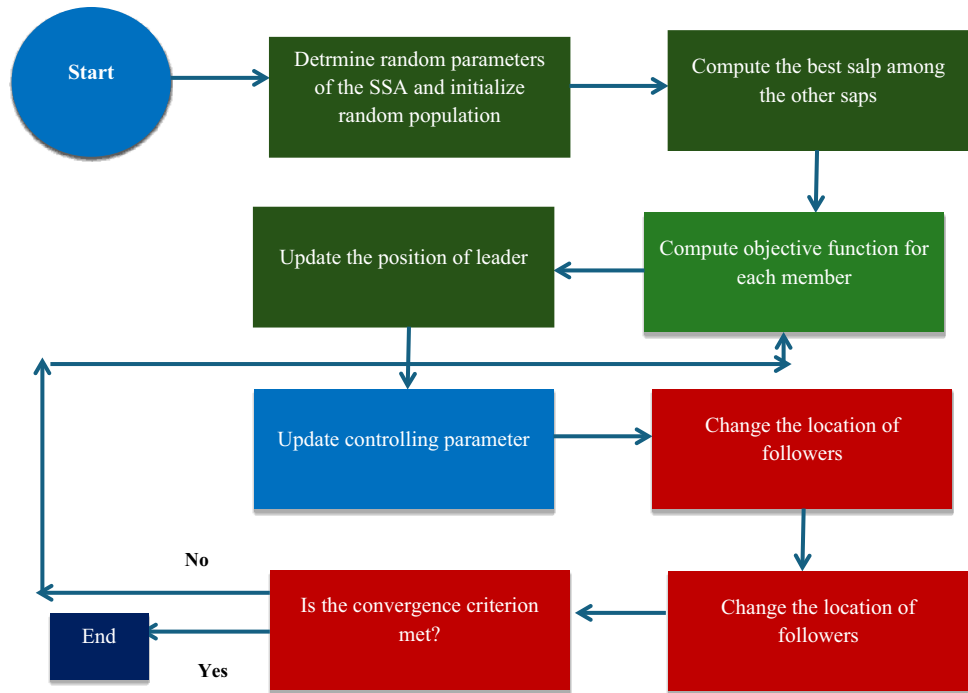
$$so_i^{t+1} = \begin{cases} so_i^t + r_1 \times \sin(r_2) \times |r_3 de_i^t - so_i^t|, & r_4 \leq 0.50 \\ so_i^t + r_1 \times \cos(r_2) \times |r_3 de_i^t - so_i^t|, & r_4 > 0.50 \end{cases}, \quad (11)$$

where  $so_i^{t+1}$ : the new location of  $i$ th solution at iteration  $t + 1$ ,  $de_i^t$ : the location of destination of point,  $r_1$ ,  $r_2$ ,  $r_3$ , and  $r_4$ : random number. The  $r_1$  parameter is responsible for transitioning from exploration to exploitation

$$r_1 = 2 - 2 \times \left(\frac{t}{T}\right), \quad (12)$$

where  $t$ : current iteration and  $T$ : total iterations. Figure 4 shows the SCA flowchart.

**Fig. 3** The SSA flowchart for optimization problem



**2.5 Particle swarm optimization (PSO)**

The PSO operates on the principle of sharing information among particles, which makes it a straightforward approach with numerous benefits, such as easy implementation, computational efficiency, and simplicity of concept. Due to its effectiveness, PSO has been utilized in various problem-solving contexts, including but not limited to environmental economic dispatch (Xin-gang et al. 2020), ANN training (Darwish et al. 2020), sports image detection (Lei et al. 2021), and particle filter noise reduction (Chen et al. 2020). Initially, we defined the starting position of particles and random parameters of PSO. The

objective function is then computed for each particle, followed by the updating of the location and velocity of particles in accordance with the equations provided below

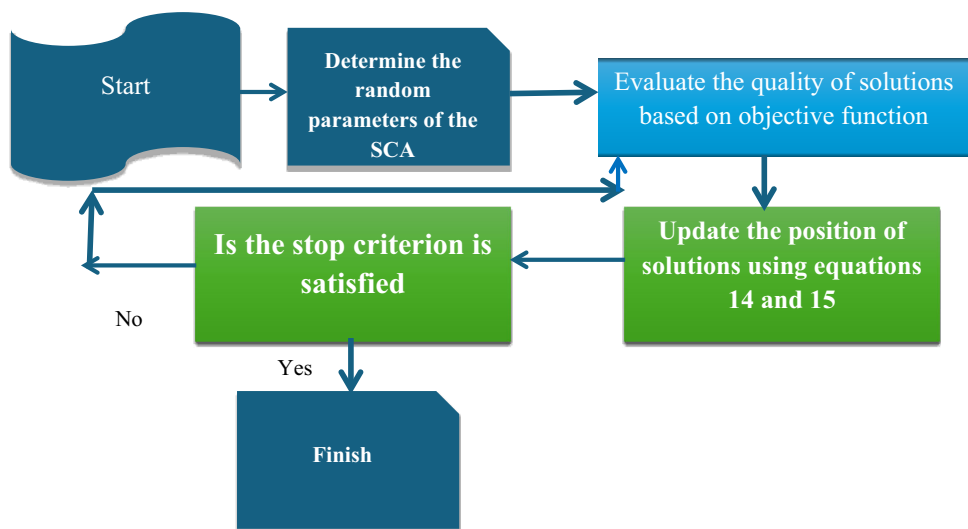
$$po_{ij}^{t+1} = po_{ij}^t + v_{ij}^{t+1} \tag{13}$$

$$ve_{ij}^{t+1} = wve_{ij}^t + \theta_1 r_1 (po_{ij}^{p(t)} - p_{ij}^t) + \theta_2 r_2 (po_{ij}^{g(t)} - p_{ij}^t). \tag{14}$$

**2.6 Inclusive multiple model**

The hybrid models of the current study are considered as competitive models. The previous research works utilized

**Fig. 4** The SCA flowchart for optimization problem





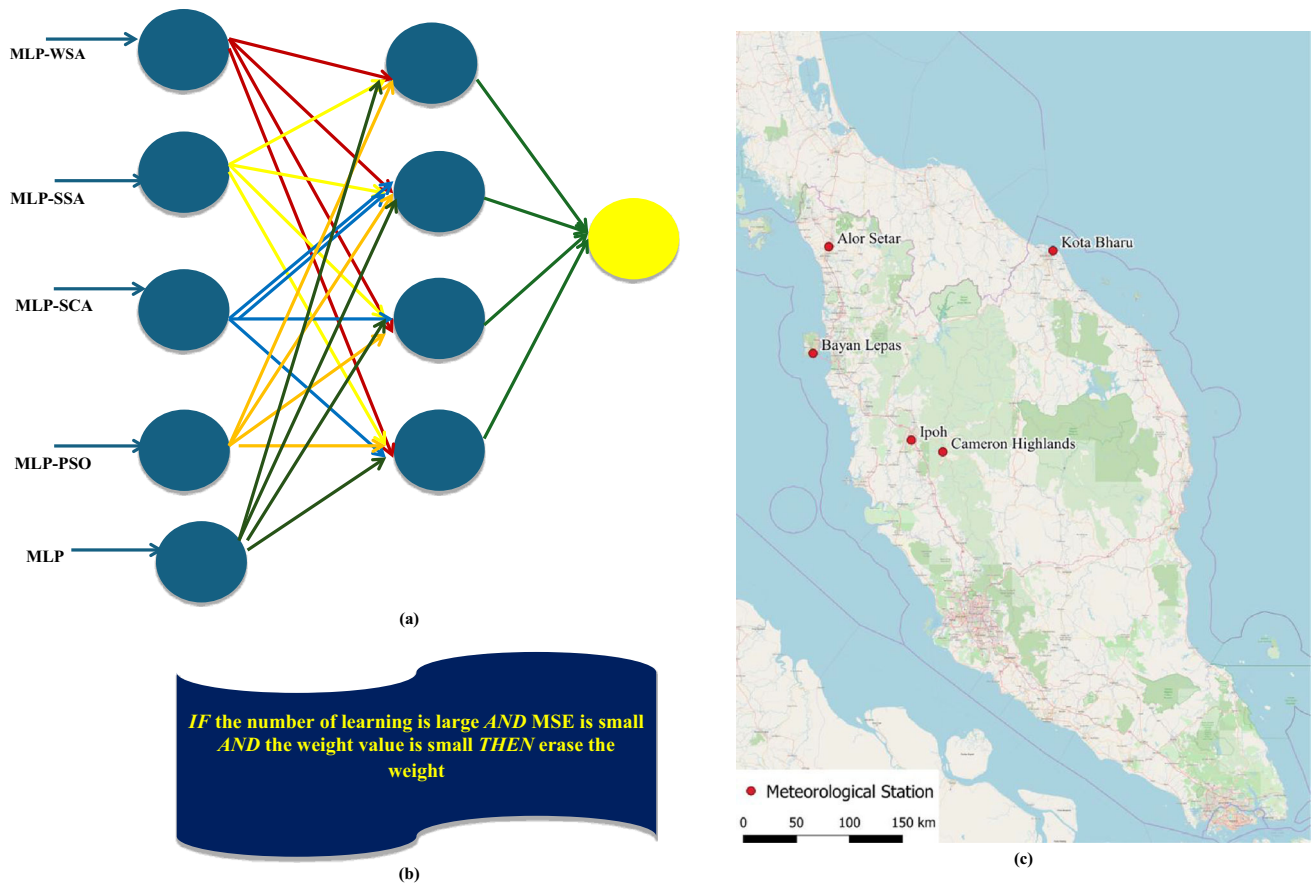


Fig. 5 a The structure of the IMM, b Fuzzy reasoning conception, c The location of case study

different models for predicting W.S. and determined the worst and best model. If the modelers generate synergy among multiple different models, the final outputs will be based on different models' advantages. Also, the modelers can ensure to extract all of the required information for predicting W.S. based on contributing all models. In this study, first, W.S. is predicted based on different hybrid and Standalone MLP. Then, each MLP model's outputs as the lower order modeling results are used as the input to an ANN as inclusive multiple model (IMM). The application of IMM model for predicting groundwater level and CO<sub>2</sub> emission was successful for previous studies (Shabani et al. 2021; Khatibi et al. 2017). Figure 5a shows IMM structure.

### 2.7 Fuzzy reasoning

The weak weights in the standalone and hybrid MLP structure are considered the redundant weights. To identify these weights, three rules are used. The rules are observed in Fig. 5b. In the starting simulation process, the values of weights are small. Thus, the learning cycle rule is proposed to avoid removing these weights. If this rule is satisfied, the second rule is RMSE. If the RMSE does not decrease, the

weights are considered redundant weights. The third rule, which involves the weight rules, is employed, because the first and second rules are less effective in dealing with complex data and high levels of noise. As a result, increasing the number of weak weights can lead to redundancy, which necessitates their removal. For the first, second, and third rules, a monotonically increasing, decreasing, and decreasing sigmoid functions are used, respectively to serve as membership functions. Minimum values for each of these membership functions are selected, and this minimum value is multiplied by the weight being removed.

### 3 Case study

In this study, five stations, namely, Alor Setar (AS), Bayan Lepas (B.L.), Cameron Highlands (C.H.), Ipoh (I.P.), and Kota Bharu (K.B.), were chosen for predicting WS. Figure 5c shows the location of stations. The Peninsular Malaysia has a typical tropical climate whereby it is warm and humid throughout the year with relatively lower wind speed in its upper part (Hwang et al. 2019). The five

meteorological stations located at the upper part of Peninsular Malaysia were chosen as the sites of interest for this investigation. The selected stations were Cameron Highlands (CH) (4° 28' N, 101° 22' E), Alor Setar (AS) (6° 12' N, 100° 24' E), Kota Bharu (KB) (6° 10' N, 102° 18' E), Bayan Lepas (BL) (5° 18' N, 100° 16' E), and Ipoh (IP) (4° 34' N, 101° 06' E). These stations represent the wind speed condition in the low land of the upper part of Peninsular Malaysia, since they are located nearby or inside the airports of the respective areas except for Cameron highlands station. Figure 6 shows the W.S. time series. The AS station has a tropical monsoon climate based on the Koppen climate. The average low and high temperatures of the AS are 32 °C and 23 °C, respectively. The climate of B.L. is tropical, and the average temperature of the B.L. is 26 °C. The annual rainfall of B.L. is 2552 mm. A tropical rainforest climate is observed in the C.H. The mean annual temperature of the C.H. station is 18 °C. A tropical rainforest climate is observed for the I.P. The average temperature of the I.P. is 28 °C. The wettest and driest months of the I.P. are October and January. The tropical monsoon climate is observed in the K.B. station. The station experiences heavier rainfall from August through January.

### 3.1 Input sensitivity with Gamma test

As observed in Table 1,  $2^{11}-1$  input combinations can be combined for predicting W.S. based on the lagged input values. Thus, it is necessary to choose the best input combination based on the lagged W.S.s. The Gamma test is one of the powerful preprocessing methods for choosing the best input combination. They utilized Gamma test in different domains such as predicting evaporation (Allawi et al. 2019), predicting groundwater level (Sharafati et al. 2020), estimating evapotranspiration (El-Shafie et al. 2013), estimating solar radiation (Jumin et al. 2021), and predicting streamflow (AlDahoul et al. 2023). In the Gamma test, the relationship between the inputs and outputs is as follows:

$$y = f(x) + r, \quad (15)$$

where  $x$ : input,  $y$ : output,  $f(x)$ : smooth function, and  $r$ : the error term. The  $\Gamma$  in the Gamma test describes the variance of observations. The Gamma test acts based on the  $i$ th input's closet neighbor ( $N[i, k]$ ,  $1 \leq k \leq p$ ,  $p$ : the maximum number of neighbors). To compute the  $\Gamma$ , the value of  $\xi_M(k)$  should be computed by

$$\xi_M(k) = \frac{1}{M} \sum_{i=1}^M |x_{N[i,k]} - X_i|^2, \quad (16)$$

where  $M$ : number of observations. In the next level, the value of the

$$\gamma_M(k) = \frac{1}{M} \sum_{i=1}^M |y_{N[i,k]} - y_i|, \quad (17)$$

where  $y_{N[i,k]}$ : the output value corresponding to the  $k$ th neighborhood of  $x_i$ . Finally, the  $\Gamma$  is calculated as follows:

$$\gamma = A\xi + \Gamma. \quad (18)$$

Another index in the Gamma test is  $V_{\text{ratio}}$

$$V_{\text{ratio}} = \frac{\Gamma}{\sigma(y)}, \quad (19)$$

where  $\sigma(y)$ : the output variance. The lowest values of the  $V_{\text{ratio}}$  and  $\Gamma$  show the best input combination. However, it is difficult to compute  $\Gamma$  for  $2^{11}-1$  input combinations. To satisfy the process of selection of the best input combination, the WSA is coupled with the Gamma test. First, the name of input variables is inserted as the initial population of WSA. Then, the random combinations of the inputs are generated based on the initial population of WSA. In fact, each WSA shows a random combination of inputs. Then, the  $\Gamma$  is computed for each member as the objective function. The operators of the WSA based on Sect. 2.1 are used to update the value of agents. The optimization process is continued until the  $\Gamma$  is converged to the least value.

### 3.2 Hybrid MLP and optimization algorithms

In this research, the optimization algorithms are used to set the MLP parameters as follows:

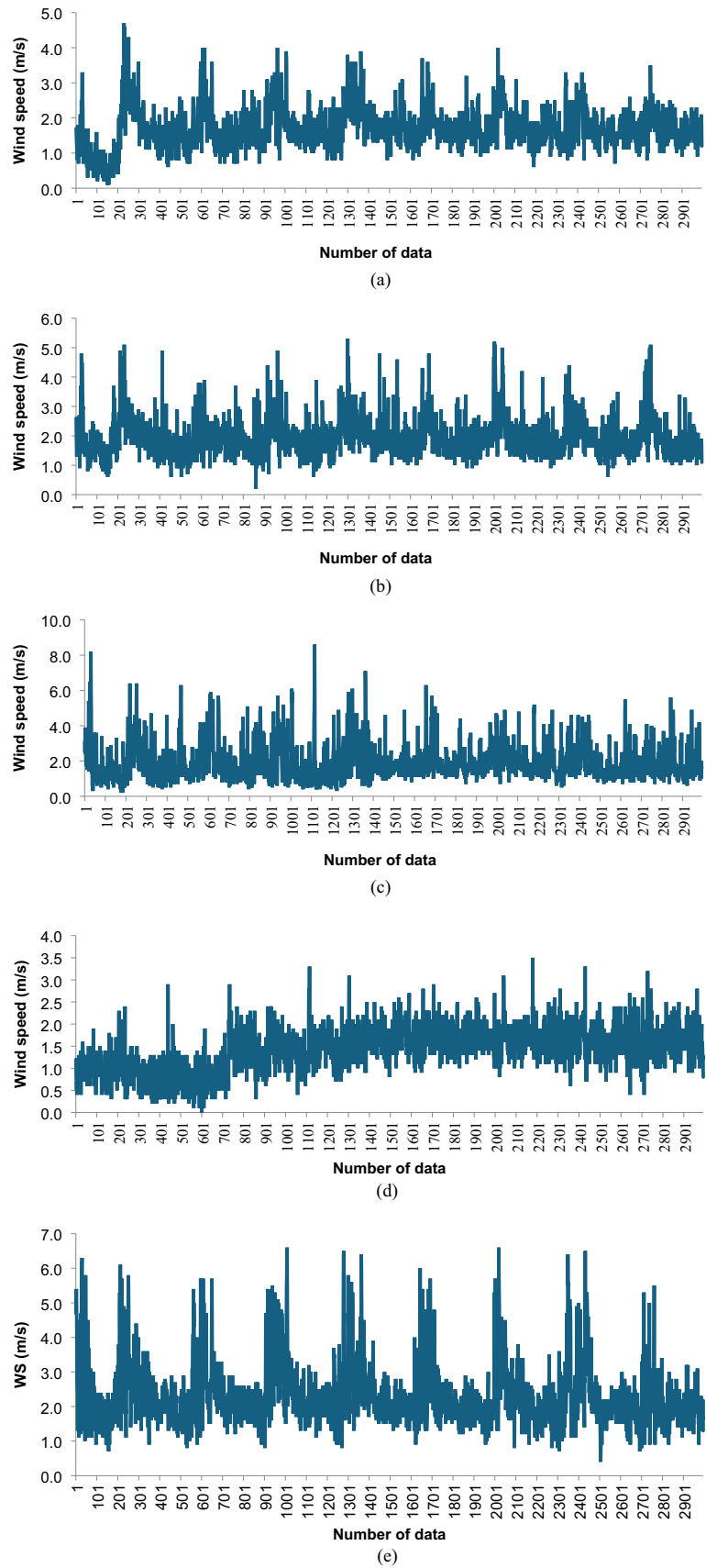
1. First, the data are split into 30% for testing and 70% for training levels. This fraction is utilized, because it makes the least value of RMSE error. In this study, RMSE error is considered as the objective function. The data were collected from January 2000 to September 2009.
2. The initial values of weights and biases are initialized.
3. The MLP runs for the training data.
4. If the stop criterion is satisfied, the MLP is used for the testing stage; otherwise, it is hybridized with the optimization algorithm.
5. The initial population of algorithms is initialized based on the random value of weights and biases.
6. The RMSE for each agent of optimization algorithms is calculated as the objective function.
7. The operators of the algorithms are utilized to change the values of weights and biases.
8. Move to step 3, after checking the convergence criterion and found it met; otherwise, move to step 6.

In this work, the indexes used to evaluate the models are as follows:

1. Nash Sutcliffe efficiency (Yafouz et al. 2021)



**Fig. 6** The WS time series for **a** Alor Seta, **b** Bayan Lepas, **c** Cameron Highlands, **d** Ipoh and **e** Kota Bharu



**Table 1** The input and output data to the MLP models

Wind speed	$WS_{t-1}, WS_{t-2}, WS_{t-3}, WS_{t-4}, WS_{t-5}, WS_{t-6}, WS_{t-7}, WS_{t-8}, WS_{t-9}, WS_{t-10}, WS_{t-11}$
	$WS_{t-1}$ : the 1-day lagged WS
	$WS_{t-2}$ : the 2-day lagged WS
	$WS_{t-3}$ : the 3-day lagged WS
	$WS_{t-4}$ : the 4-day lagged WS
	$WS_{t-5}$ : the 5-day lagged WS
	$WS_{t-6}$ : the 6-day lagged WS
	$WS_{t-7}$ : the 7-day lagged WS
	$WS_{t-8}$ : the 8-day lagged WS
	$WS_{t-9}$ : the 9-day lagged WS
	$WS_{t-10}$ : the 10-day lagged WS
	$WS_{t-11}$ : the 11-day lagged WS
	Output: WS ( $t$ ): a 1-day ahead forecast of WS

$$NSE = 1 - \frac{\sum_{i=1}^N (WS_{ob} - WS_{es})}{\sum_{i=1}^N (WS_{ob} - \overline{WS})}. \quad (20)$$

2. Root-mean-square error (RMSE) (Osman et al. 2021)

$$RMSE = \sqrt{\frac{1}{n} \sum_{n=1}^N (WS_{ob} - WS_{es})^2}. \quad (21)$$

3. Mean absolute error (MAE) (Abba et al. 2020)

$$MAE = \frac{1}{N} \sum_{i=1}^N |WS_{es} - WS_{ob}|. \quad (22)$$

4. Scatter index (S.I.) (Muslim et al. 2020a)

$$SI = \frac{RMSE}{\overline{WS_{ob}}} \quad (23)$$

( $SI < 0.10$ : excellent performance,  $0.10 < SI < 0.20$ : good performance,  $0.20 < SI < 0.30$ : fair performance,  $SI > 0.30$ : poor performance).

5. Uncertainty with 95% confidence level ( $U_{95}$ ) (Jumin et al. 2020)

$$U_{95} = \sqrt{(SD^2 + RMSE^2)}, \quad (24)$$

where SD: the standard deviation of the difference,  $WS_{es}$ : estimate values,  $WS_{ob}$ : observed values,  $\overline{WS_{ob}}$ : average observed values, and  $N$ : number of samples. The highest values of NSE are ideal, and the lowest values of MAE, RMSE, and  $U_{95}$  are ideal.

## 4 Results and discussion

### 4.1 Optimization algorithms' parameters

Obtaining precise values of random parameters is critical for achieving optimal performance. This requires

computing the variance of the objective function in relation to variations in the random parameters. An analysis was conducted on the variance of the objective function across different domains of random parameters in the AS station, and the results are presented in Table 2. Based on these results, it was concluded that the best population size for the WSA is 40, and as a result, best population size of RMSE became the lowest. Furthermore, the ideal value for the maximum number of iterations in WSA was found to be 200, producing the least RMSE value. Similarly, for the SSA, SCA, and PSO algorithms, the best population size was found to be 40, 40, and 60, respectively. The optimal random parameters for other stations were also established, as shown in Table 3.

### 4.2 The best input tuning for predictive models

According to Table 4, the first to third-best input combinations for each station are shown. From WS ( $t - 1$ ), ... to WS ( $t - 6$ ) combination was set to be the optimal input combination for predicting water level at the AS and CH stations. Similarly, from WS ( $t - 1$ ), ... to WS ( $t - 5$ ) combination was found to be the best input combination for the BL, LB, and IP stations.

The coupling of the Gamma test with an optimization algorithm provides a convenient way to automatically determine the best input combination for predicting various target variables, without the need for manual computation of different input combinations. Therefore, this hybridized Gamma test proves to be a highly effective tool for selecting optimal inputs in models.

### 4.3 Accuracy comparison for various models

As shown in Fig. 7a–d, when examining the testing results of the models in the AS station, the MLP-WSA model outperformed the MLP-SSA, MLP-SCA, MLP-PSO, and

**Table 2** Determining random parameters in the AS station

a: WSA							
Population size	Objective function	Maximum number of iterations	Objective function				
20	4.25	100	5.89				
40	3.78	200	3.79				
60	3.90	300	4.59				
80	4.44	400	5.25				
b: SSA							
Population size	Objective function	Maximum number of iterations	Objective function	$\sigma_3$	Objective function	$\sigma_2$	Objective function
20	4.89	100	4.95	0.20	4.95	0.30	4.36
40	3.98	200	3.98	0.40	4.21	0.50	4.12
60	4.02	300	4.21	0.60	3.98	0.70	3.98
80	4.56	400	4.45	0.80	4.02	0.90	4.11
c: SCA							
Population size	Objective function	Maximum number of iterations	Objective function	$r_2$	Objective function	$r_4$	Objective function
20	5.78	100	5.79	$\pi/3$	5.82	0.3	5.78
40	5.12	200	5.02	$2\pi/3$	5.02	0.5	5.22
60	5.02	300	5.24	$\pi$	5.46	0.7	5.04
80	5.21	400	5.36	$4\pi/3$	4.65	0.90	5.46
d: PSO							
Population size	Objective function	Maximum number of iterations	Objective function	$\theta_1$	Objective function	$\theta_2$	Objective function
20	6.23	100	6.91	1.4	6.92	1.4	6.91
40	5.89	200	5.89	1.6	6.23	1.6	6.42
60	5.95	300	6.12	2.0	6.12	2.0	6.12
80	6.12	400	6.23	2.2	6.24	2.2	6.32

MLP models in terms of accuracy. The  $U_{95}$  of the MLP-WSA, MLP-SSA, MLP-SCA, MLP-PSO, and MLP models were 17%, 19%, 20%, 22%, and 24%, respectively. The comparison between the accuracy of the models and the IMM model demonstrated that the IMM improved the accuracy and decreased the RMSE by 1.5%, 3.2%, 5.9%, 8.03%, and 23.7%, compared to MLP-WSA, MLP-SCA, MLP-SSA, MLP-PSO, and MLP, respectively. Both the IMM and MLP-WSA models showed the highest NSE values, and the MLP model obtained the highest value of the  $U_{95}$ .

Figure 7a–d presents the outcomes of the evaluation stage at the B.L. site, revealing that the MLP-WSA model achieved an RMSE of 2.78 (m/s), whereas the MLP-SSA, MLP-SCA, MLP-PSO, and MLP models yielded RMSE values of 3.88 m/s, 4.12 m/s, 4.78 m/s, and 4.98 m/s,

respectively. Furthermore, it was determined that the MLP-WSA model exhibited superior performance compared to other models. Moreover, the implementation of the IMM model demonstrated that it could enhance the accuracy of all models by incorporating information from each model. The NSE value of the IMM model was found to be 0.92, whereas the NSE values for the MLP-WSA, MLP-SSA, MLP-SCA, MLP-PSO, and MLP models were 0.90, 0.86, 0.82, 0.80, and 0.78, respectively.

Figure 7a–d reveals the outcomes of the models during the testing phase at the C.H. station. Notably, the IMM model outperformed the other models with a substantially reduced MAE of 6%, 22%, 25%, 39%, and 41%, respectively, compared to the MLP-WA, MLP-SCA, MLP-SSA, MLP-PSO, and MLP models. Furthermore, the IMM model achieved the highest NSE value, while the MLP model

**Table 3** Determining random parameters in the different stations

Station	WSA	SSA	SCA	PSO
BL	Population size: 40, maximum number of iterations: 200	Population size: 40, maximum number of iterations: 200, $\sigma_3 = 0.5$ and $\sigma_2 = 0.4$	Population size: 40, maximum number of iterations: 200, $r_2 = 2\frac{\pi}{3}$ and $\sigma_2 = 0.7$	Population size: 40, maximum number of iterations: 200, $\theta_1 = 2.00$ $\theta_2 = 2.00$
CH	Population size: 40, maximum number of iterations: 300	Population size: 40, maximum number of iterations: 200, $\sigma_3 = 0.5$ and $\sigma_2 = 0.4$	Population size: 60, maximum number of iterations: 200, $r_2 = 2\frac{\pi}{3}$ and $\sigma_2 = 0.7$	Population size: 40, maximum number of iterations: 300, 2.00 $\theta_2 = 2.00$
IP	Population size: 40, maximum number of iterations: 200	Population size: 40, maximum number of iterations: 200, $\sigma_3 = 0.5$ and $\sigma_2 = 0.6$	Population size: 40, maximum number of iterations: 200, $r_2 = 2\frac{\pi}{3}$ and $\sigma_2 = 0.5$	Population size: 60, maximum number of iterations: 300, 2.00 $\theta_2 = 2.00$
KB	Population size: 40, maximum number of iterations: 200	Population size: 40, maximum number of iterations 3, $\sigma_3 = 0.5$ and $\sigma_2 = 0.4$	Population size: 40, maximum number of iterations: 200, $r_2 = 2\frac{\pi}{3}$ and $\sigma_2 = 0.7$	Population size: 40, maximum number of iterations: 300, 2.00 $\theta_2 = 2.00$

**Table 4** Selection of the best input combination based on improved Gamma test

Combination	$\Gamma$	$V_{\text{ratio}}$
<i>AS station</i>		
WS ( $t - 1$ ), WS ( $t - 2$ ), WS ( $t - 3$ ), WS ( $t - 4$ ), WS ( $t - 5$ ), WS ( $t - 6$ )	$2.12 \times 10^{-3}$	$5.76 \times 10^{-3}$
WS ( $t - 1$ ), WS ( $t - 2$ ), WS ( $t - 3$ ), WS ( $t - 4$ ), WS ( $t - 5$ )	$2.55 \times 10^{-3}$	$6.92 \times 10^{-3}$
WS ( $t - 1$ ), WS ( $t - 2$ ), WS ( $t - 4$ ), WS ( $t - 3$ )	$3.24 \times 10^{-3}$	$8.78 \times 10^{-3}$
<i>BL station</i>		
WS ( $t - 1$ ), WS ( $t - 2$ ), WS ( $t - 3$ ), WS ( $t - 4$ ), WS ( $t - 5$ )	$4.24 \times 10^{-3}$	$8.72 \times 10^{-3}$
WS ( $t - 1$ ), WS ( $t - 2$ ), WS ( $t - 3$ ), WS ( $t - 4$ )	$6.32 \times 10^{-3}$	$12.96 \times 10^{-3}$
WS ( $t - 1$ ), WS ( $t - 2$ ), WS ( $t - 3$ )	$7.21 \times 10^{-3}$	$14.79 \times 10^{-3}$
<i>CH station</i>		
WS ( $t - 1$ ), WS ( $t - 2$ ), WS ( $t - 3$ ), WS ( $t - 4$ ), WS ( $t - 5$ ), WS ( $t - 6$ )	$3.12 \times 10^{-3}$	$5.14 \times 10^{-3}$
WS ( $t - 1$ ), WS ( $t - 2$ ), WS ( $t - 3$ ), WS ( $t - 4$ ), WS ( $t - 5$ )	$4.55 \times 10^{-3}$	$7.49 \times 10^{-3}$
WS ( $t - 1$ ), WS ( $t - 2$ ), WS ( $t - 4$ ), WS ( $t - 3$ )	$6.89 \times 10^{-3}$	$11.29 \times 10^{-3}$
<i>IP station</i>		
WS ( $t - 1$ ), WS ( $t - 2$ ), WS ( $t - 3$ ), WS ( $t - 4$ ), WS ( $t - 5$ )	$2.21 \times 10^{-3}$	$5.14 \times 10^{-3}$
WS ( $t - 1$ ), WS ( $t - 2$ ), WS ( $t - 3$ ), WS ( $t - 4$ )	$1.42 \times 10^{-3}$	$3.30 \times 10^{-3}$
WS ( $t - 1$ ), WS ( $t - 2$ ), WS ( $t - 3$ )	$5.19 \times 10^{-3}$	$12.07 \times 10^{-3}$
<i>KB station</i>		
WS ( $t - 1$ ), WS ( $t - 2$ ), WS ( $t - 3$ ), WS ( $t - 4$ ), WS ( $t - 5$ )	$1.12 \times 10^{-3}$	$3.22 \times 10^{-3}$
WS ( $t - 1$ ), WS ( $t - 2$ ), WS ( $t - 3$ ), WS ( $t - 4$ )	$2.45 \times 10^{-3}$	$7.03 \times 10^{-3}$
WS ( $t - 1$ ), WS ( $t - 2$ ), WS ( $t - 3$ )	$3.46 \times 10^{-3}$	$9.94 \times 10^{-3}$

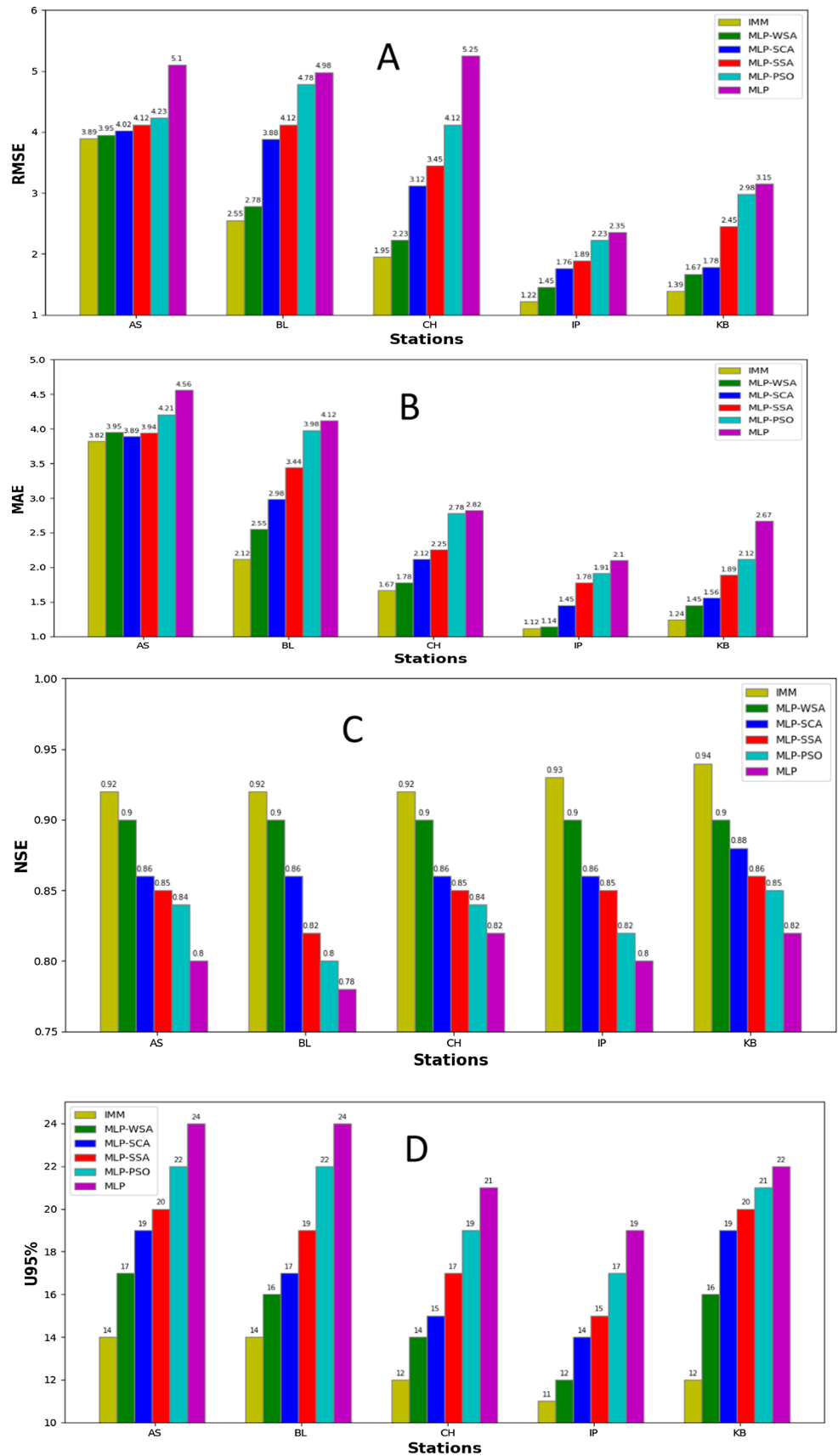
The first-best input combination to the third-best input combination

attained the lowest NSE. Conversely, the MLP model recorded the lowest U95, indicating the lowest accuracy compared to the other models.

Figure 7a–d presents the accuracy during the testing phase at the I.P. station. It is evident that the IMM model

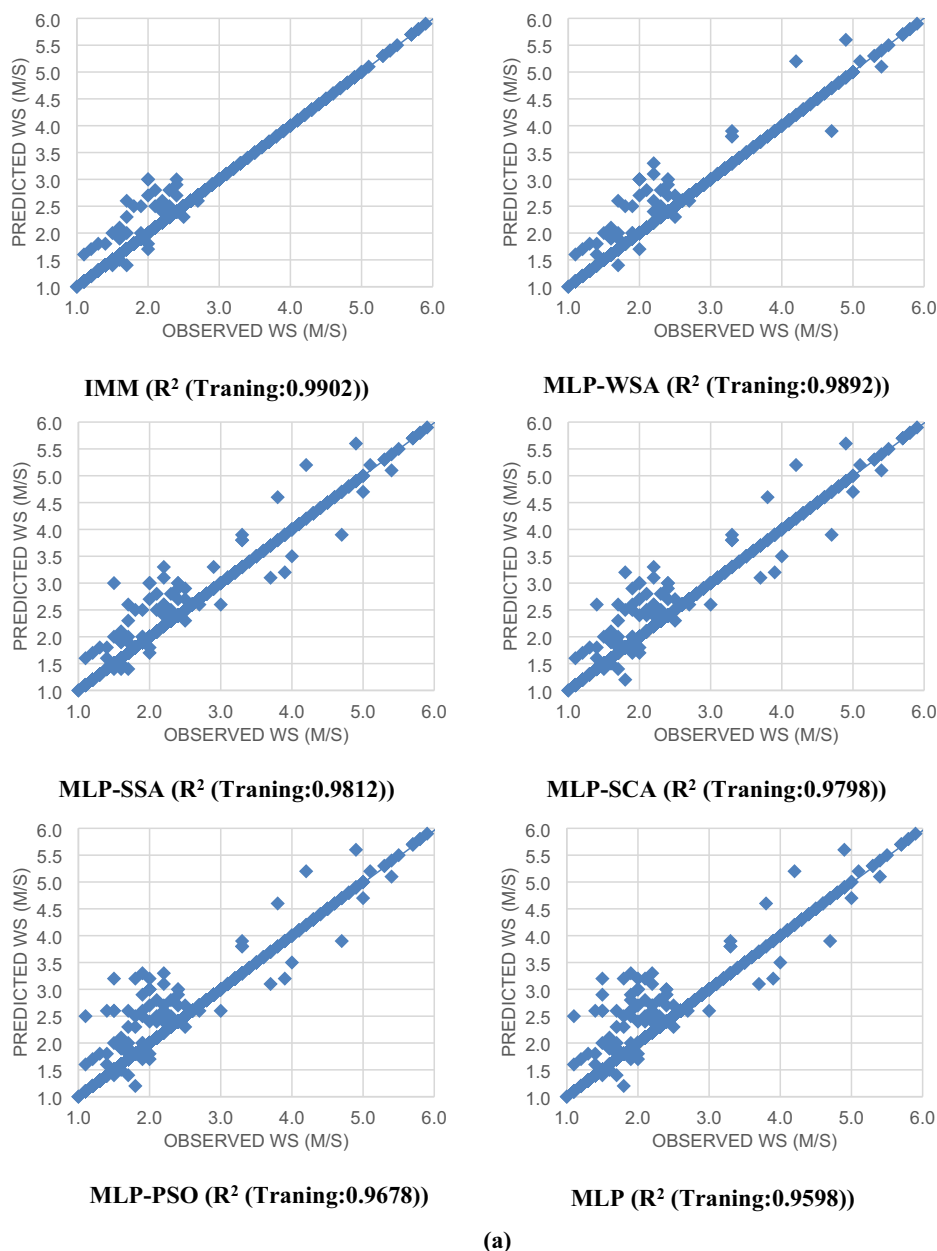
had the lowest RMSE of 1.22 m/s, which indicates its high accuracy compared to other models. On the other hand, the MLP-WSA, MLP-SSA, MLP-SCA, MLP-PSO, and MLP models had relatively higher RMSE values of 1.45 m/s, 1.76 m/s, 1.89 m/s, 2.23 m/s, and 2.35 m/s, respectively.

**Fig. 7** Comparison between various models in terms of RMSE, MAE, NSE, and U95% in five stations





**Fig. 8** Scatterplots for different models in **a** AS station, **b** BL station, **c** CH station, **d** IP station, **e** KB station



The IMM and MLP-WSA models demonstrated superior performance with the highest NSE and lowest  $U_{95}$  values recorded, respectively.

Figure 7a–d presents the performance of the models during the testing phase at the K.B. station. It was observed that the RMSE of the IMM model was significantly lower than the other models. The IMM model achieved a 17%, 22%, 44%, 54%, and 55% reduction in RMSE compared to the MLP-WSA, MLP-SSA, MLP-SCA, MLP-PSO, and MLP models, respectively. Additionally, the NSE value of the MLP-WSA model was higher (0.90) compared to the MLP-SSA, MLP-SCA, MLP-PSO, and MLP models, which had NSE values of 0.88, 0.86, 0.85, and 0.82,

respectively. Therefore, the hybrid MLP models showed superior performance than the standalone MLP models.

Figure 8 illustrates the scatterplots for the testing stages at the A.S., B.L., C.H., I.P., and K.B. stations, respectively. In Fig. 8a, the IMM model demonstrated the best performance with a testing  $R^2$  value of 0.9891, while the MLP-WSA model achieved superior accuracy among the other hybrid and standalone MLP models, with an  $R^2$  value of 0.9816. In Fig. 8b, the IMM and MLP models exhibited the best and worst accuracy with testing  $R^2$  values of 0.989 and 0.9451, respectively. In Fig. 8c, the IMM and MLP-WSA models show the highest testing  $R^2$  values of 0.9894 and 0.9860, respectively, while MLP-SSA, MLP-SCA, MLP-

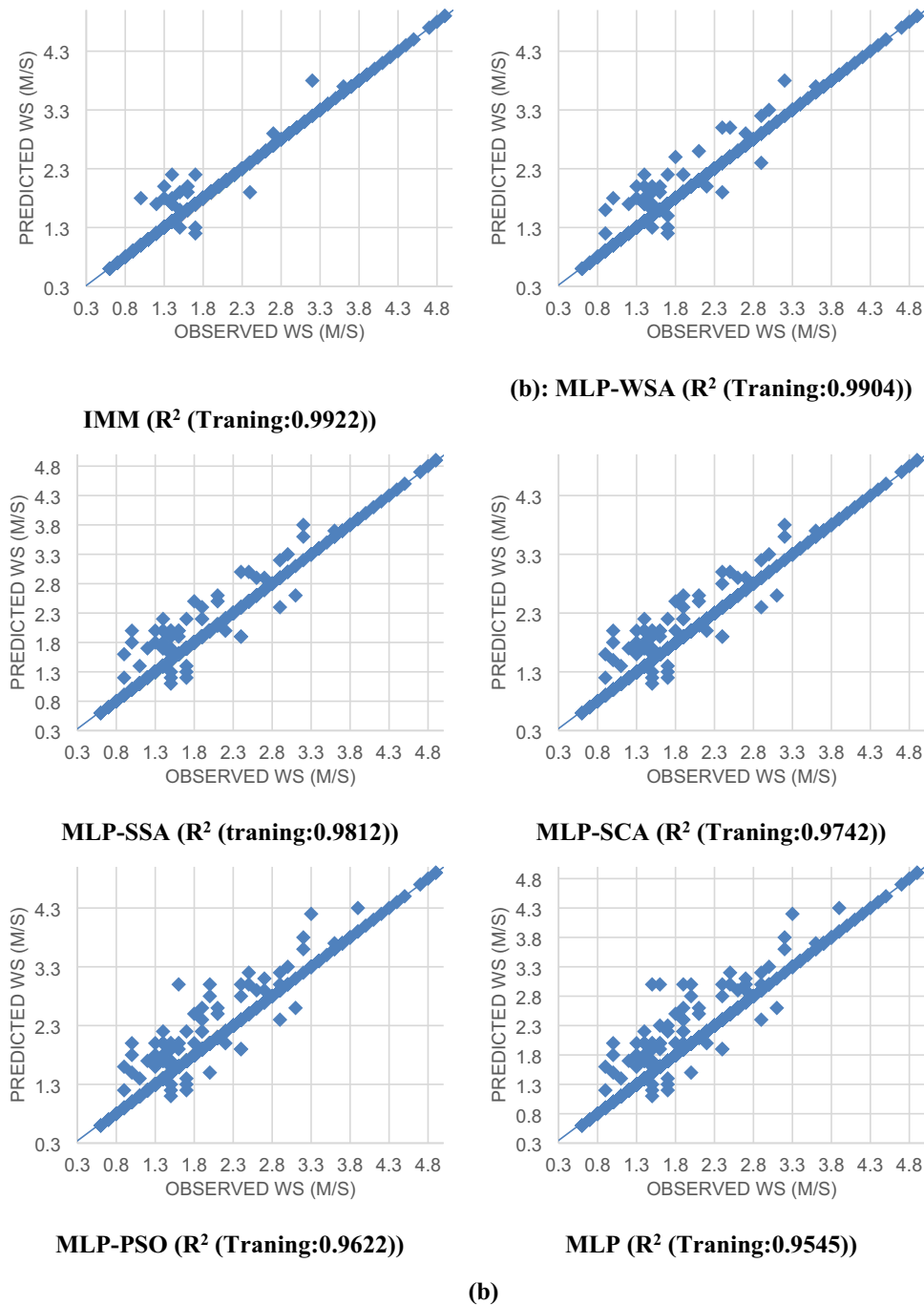


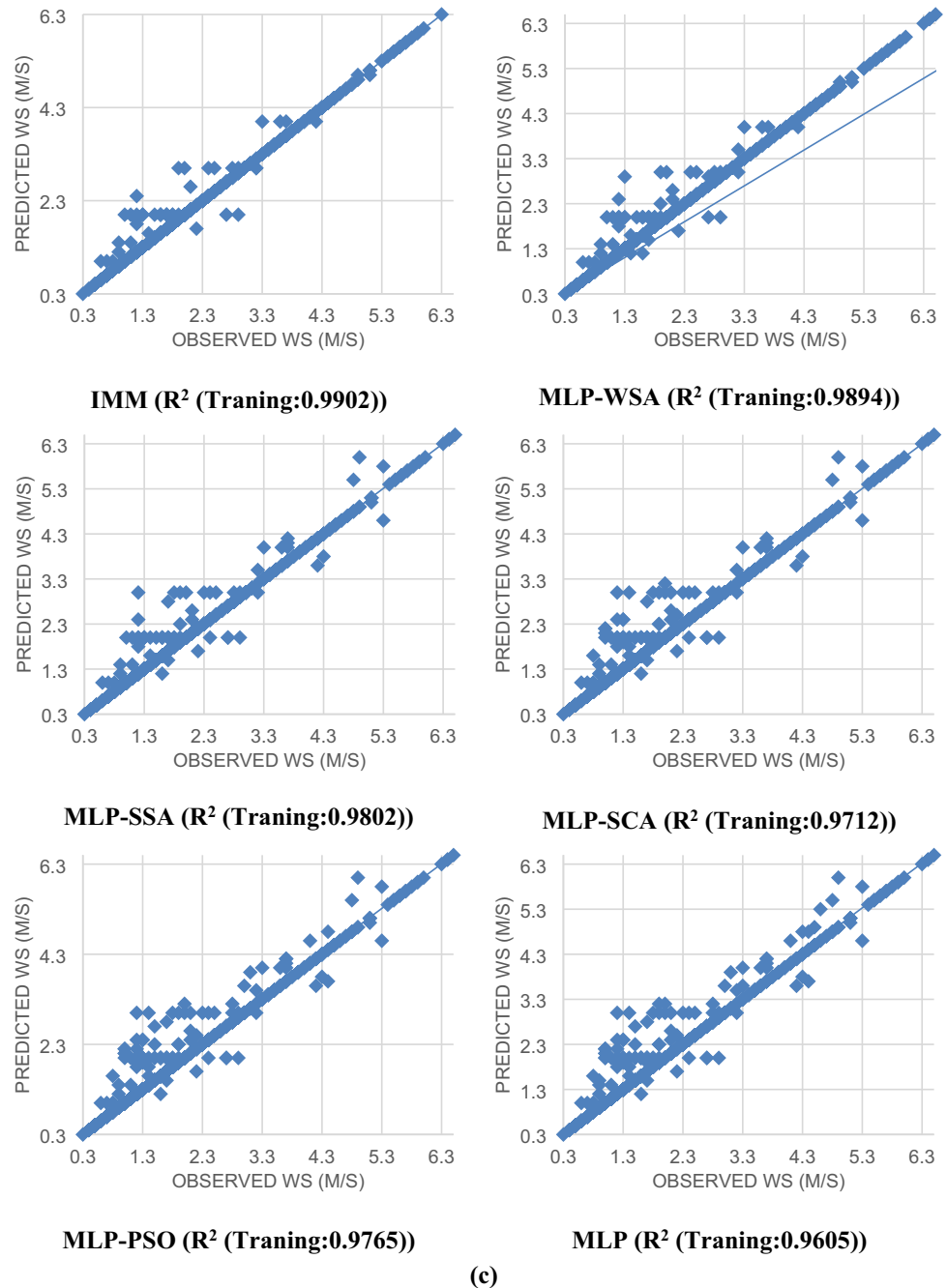
Fig. 8 continued

PSO, and MLP models recorded lower  $R^2$  values. Figure 8d depicts the testing  $R^2$  values of the IMM model (0.9923). In Fig. 8e, the IMM and MLP-WSA models also exhibited the best testing  $R^2$ .

Figure 9 displays the Scatter Index (S.I.) values of the models under evaluation. The S.I. value for testing levels was determined to be 0.09, 0.11, 0.17, 0.21, 0.24, and 0.25 for the IMM, MLP-WSA, MLP-SSA, MLP-SCA, MLP-

PSO, and MLP models, respectively, in the AS station. Based on these findings, the IMM, MLP-WSA, and MLP-SSA models performed well, achieving excellent, good, and good accuracy ratings, respectively. The MLP-SCA, MLP-PSO, and MLP models achieved fair accuracy, indicating room for improvement. In the B.L. station, the accuracy ratings for the IMM, MLP-WSA, MLP-SSA, MLP-SCA, MLP-PSO, and MLP models were excellent,

Fig. 8 continued



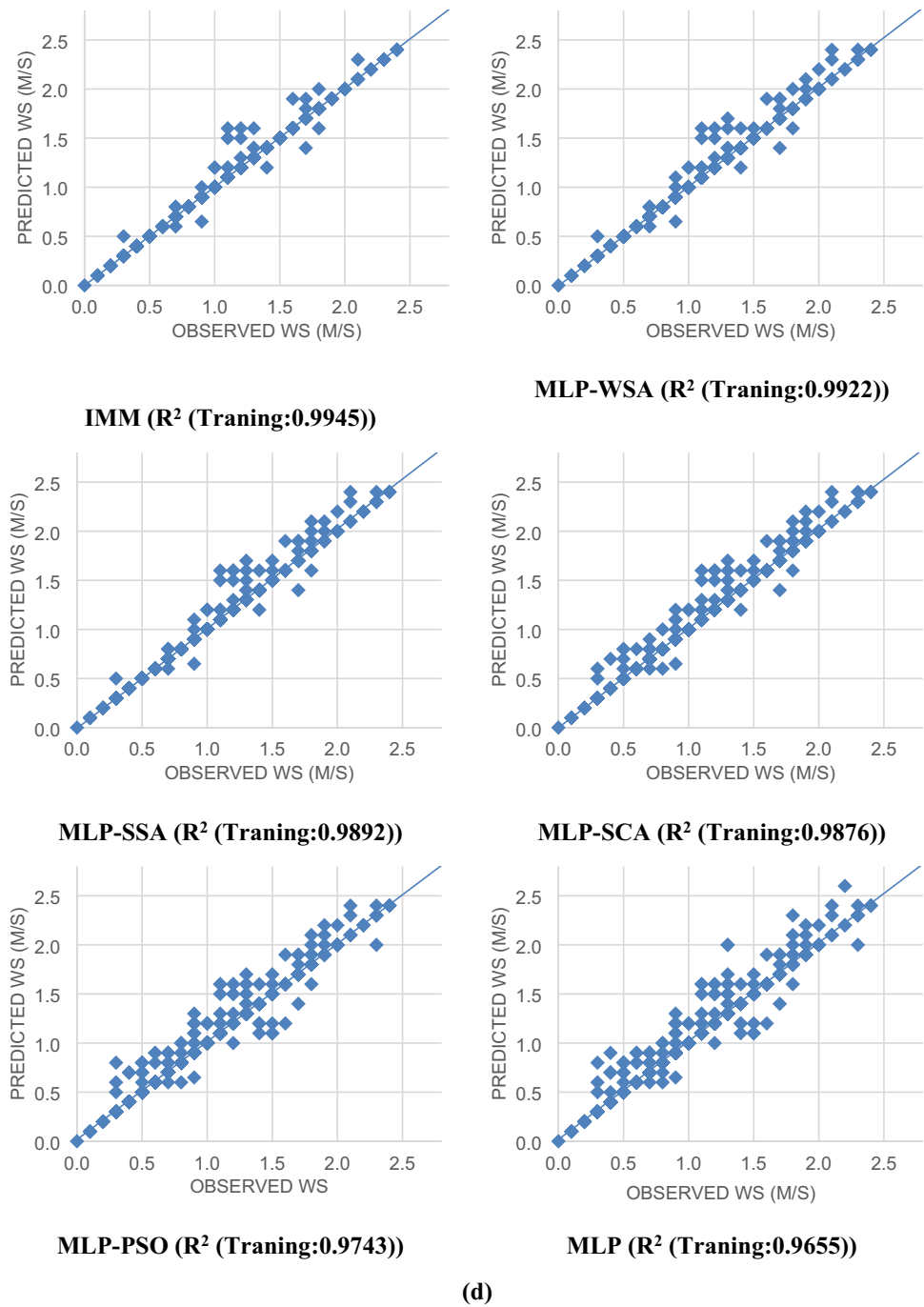
good, good, fair, fair, and fair, respectively. Furthermore, in the C.H., I.P., and K.B. stations, the IMM model's performance was deemed excellent, while the MLP models achieved fair accuracy.

Figure 10 presents the heat maps depicting the relative error of various models. The variation of relative error for the IMM model in all stations ranged from 0 to 5%, whereas the range of relative errors in MLP model was between 20 to 25%. The findings revealed that the relative error of MLP-WSA ranged from 0 to 10, 0 to 10, 5 to 10, 0

to 10, and 0 to 10 for the AS, B.L., CH, I.P., and K.B. stations, respectively.

The CPU time was calculated for various models. The results indicate that for the AS station, the CPU time for the IMM model was 230 s and 260 s without and with fuzzy reasoning, respectively. Similarly, for the B.L. station, the MLP-WSA model had a CPU time of 250 s and 282 s without and with fuzzy reasoning, respectively. However, the results across stations showed that by utilizing the fuzzy reasoning, the CPU time became lower than time calculated with fuzzy reasoning.

Fig. 8 continued



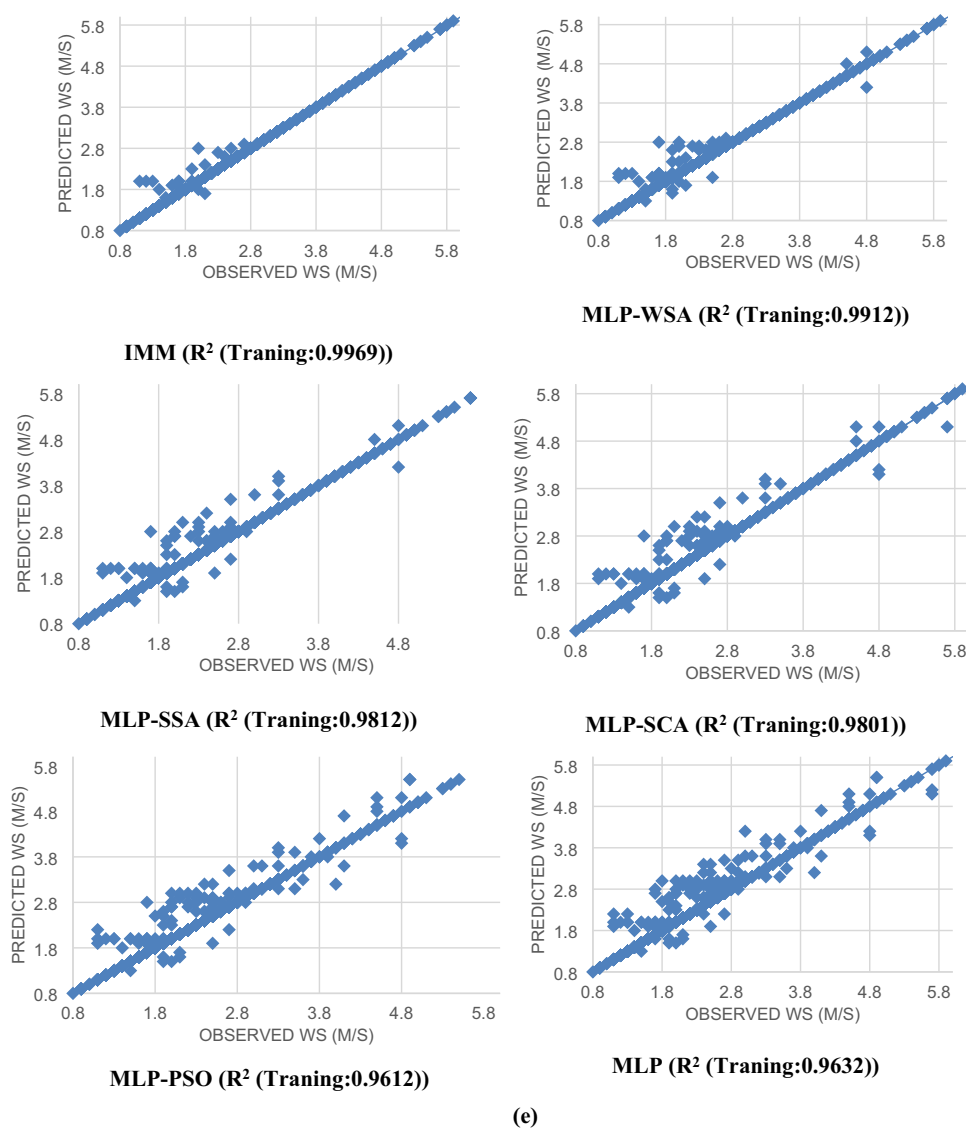
#### 4.4 Concluding discussion

By analyzing the results, it is worthy notable that the WSA optimization algorithm helped enhance the accuracy of MLP and resulted in the MLP-WSA outperforming other MLP models, such as MLP-SSA, MLP-SCA, MLP-PSO, and MLP. Hence, the developed MLP model can be regarded as an effective means of predicting various climate and hydrological variables. Additionally, the IMM model has demonstrated its ability to enhance the accuracy

of MLP models by aggregating data from multiple MLP models.

The study's outcomes support the findings of earlier research studies. Liu et al. (2013) found that utilizing optimization algorithms like particle swarm optimization and genetic algorithms could enhance the MLP model's precision for forecasting W.S. Liu et al. (2015) combined fast ensemble decomposition and optimization algorithms with the MLP to predict W.S, resulting in hybrid MLP models outperforming standalone MLP models. Moreover,

Fig. 8 continued



Samadianfard et al. (2020) merged optimization algorithms with the MLP model, indicating that the whale optimization algorithm boosted the MLP models' accuracy by utilizing advanced operators.

Future research could investigate the use of the WSA in combination with other soft computing models such as the radial basis function neural network and SVM models to forecast W.S. Furthermore, further research can examine the impact of uncertainty on the models' outputs caused by the uncertainty of model parameters and inputs. While the MLP-WSA showed superior performance in this study, future research can utilize multi-criteria decision-making methods to determine the most appropriate model based on different analyses.

Future studies could explore the possibility of defining multiple objective functions for tuning the MLP parameters, which would allow the identification of the best input

combination and MLP parameters simultaneously. This approach does not require additional preprocessing methods like the Gamma test to identify the optimal inputs. To achieve this, two objective functions could be defined. The first objective function would focus on identifying the optimal MLP parameters, while the second objective function would concentrate on finding the best input combination. Therefore, it would be necessary to modify the WSA into a multi-objective optimization algorithm capable of solving such problems.

In situations where climate data are unavailable, alternative input combinations like latitude, longitude, and the number of available data points can still provide useful insights into predicting W.S. This approach can be particularly useful for scenarios where the availability of data is limited. Although fuzzy reasoning can reduce computational time, it is worth noting that optimization algorithms



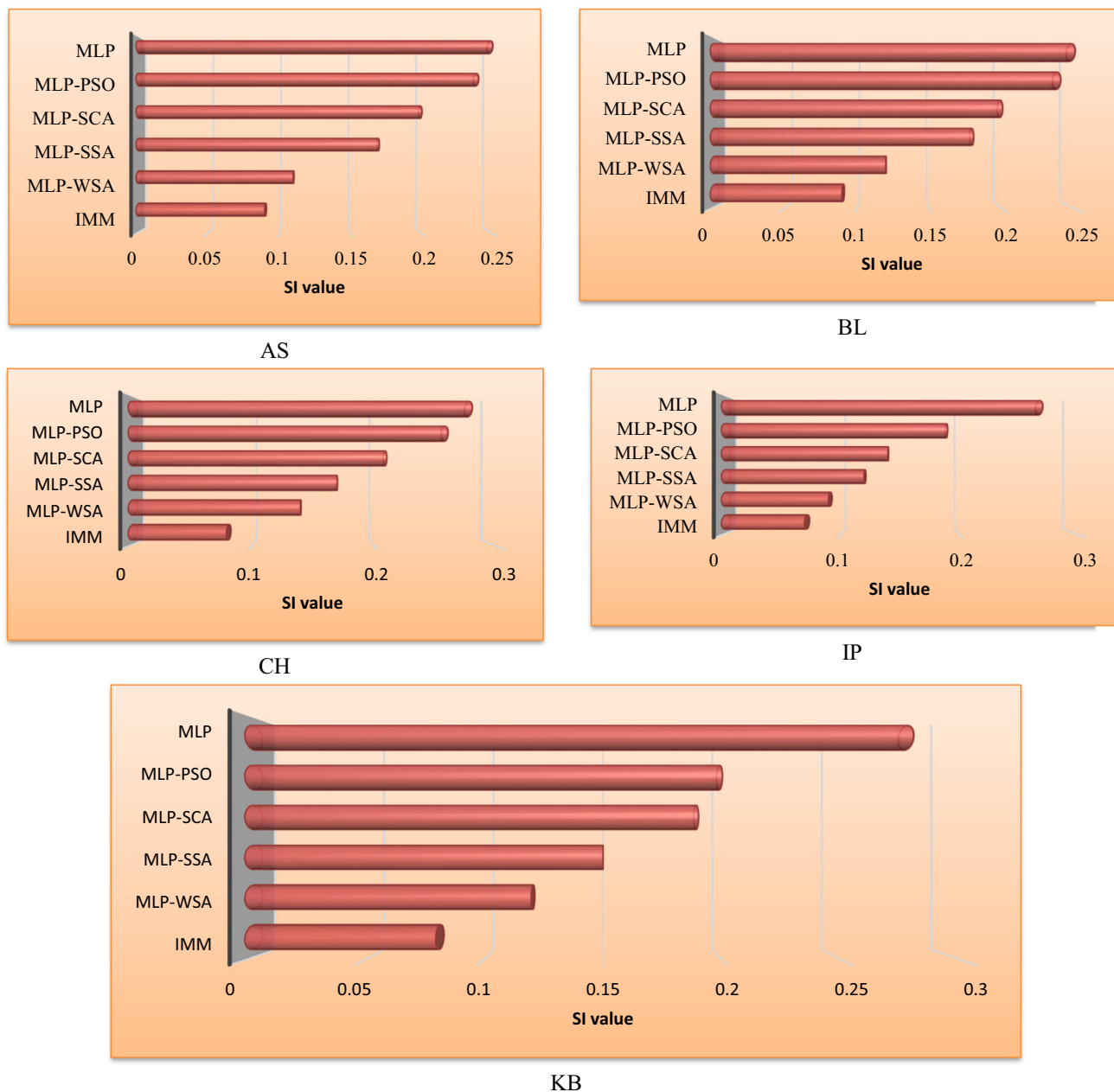


Fig. 9 The SI values for different models in the different stations

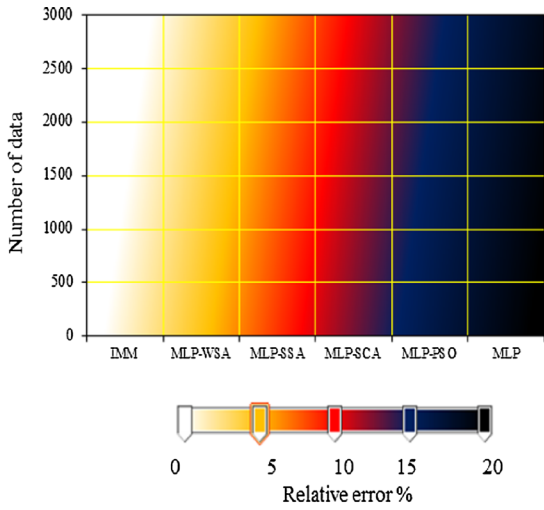
can also be effective at reducing the computational time required. These algorithms may be especially beneficial when they converge more quickly.

### 5 Conclusion

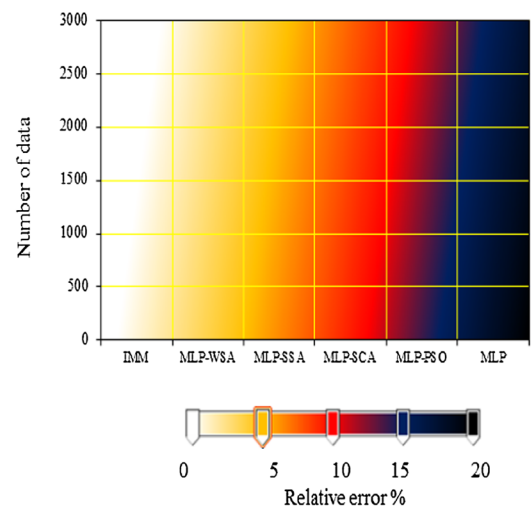
Wind energy aims to mitigate the environmental pollution resulted from the consumption of fossil fuels. Accurately predicting wind speed is essential in managing energy and generating power. This work utilized an optimization algorithm called the WSA to train an MLP in five stations

in Malaysia. Additionally, the outputs of several MLP models, including the MLP-WSA, MLP-SSA, MLP-SCA, MLP-PSO, and MLP, were applied to the IMM. To find the best input combination, a Gamma test was conducted. The results showed that the MLP-WSA outperformed other MLP models, with the lowest RMSE of 3.95 m/s. In terms of accuracy, the IMM model had the highest NSE in the B.L. station, whereas the MLP model had the lowest NSE in the same station. During testing, the MAE of the IMM model was recorded at 2.55 m/s, which was significantly lower compared to the MAE of the MLP-WSA, MLP-SSA, MLP-SCA, MLP-PSO, and MLP models, which were 2.55,

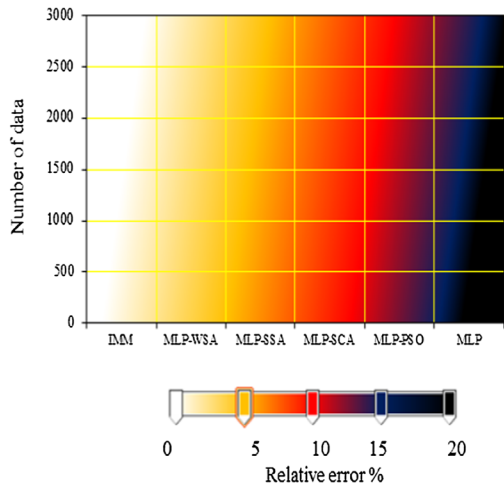
AS



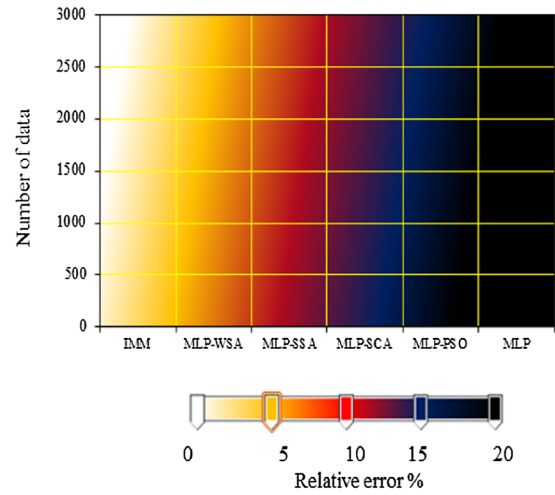
BL



CH



IP



KB

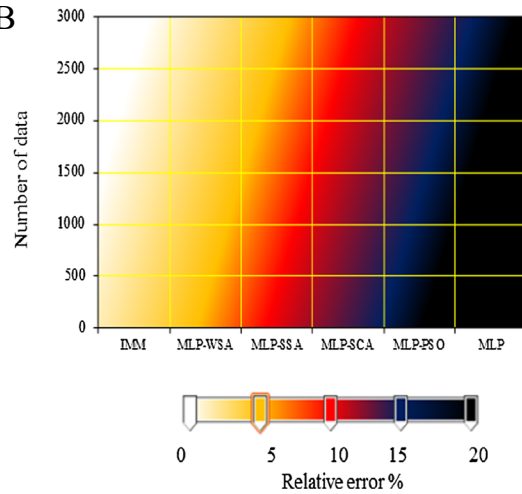


Fig. 10 The relative error for the different models

2.98, 3.44, 3.98, and 4.12 m/s, respectively. Similarly, the MAE of the IMM model was found to be significantly lower than others in the C.H. station. Specifically, it was 6%, 22%, 25%, 39%, and 41% lower compared to the MAE of the MLP-WA, MLP-SCA, MLP-SSA, MLP-PSO, and MLP models, respectively. The IMM and MLP-WSA also performed better in terms of NSE in the I.P. station, and its accuracy was superior to that of other models in the B.K. station. Furthermore, it was observed that incorporating fuzzy reasoning in the modeling process reduced the CPU time required for analysis. Overall, combining whale search algorithm with the MLP model and using hybrid Gamma test based on fuzzy reasoning provided the most accurate prediction of wind speed.

This paper focused on MLP particularly and combined it with optimization algorithms. This limitation of targeting only MLP can be addressed in future. This work opens a door to explore other neural network architectures, such as CNNs, LSTMs, and transformers to be trained with the various optimization algorithms to improve the prediction performance.

List of acronyms

Abbreviation	Definition
SGD	Stochastic gradient decent
MSE	Mean squared error
W.S	Wind speed
MLP	Multilayer perceptron
WSA	Water strider algorithm
MLP-SCA	MLP-sine cosine
MLP-SSA	MLP-salp swarm
MLP-PSO	MLP-particle swarm optimization
IMM	Inclusive multiple model
ANN	Artificial neural network
SVM	Support vector machine
SCMs	Soft computing models
GPR	Gaussian Process Regression
GRNN	Generalized Regression Neural Network
ANFIS	Adaptive neuro-fuzzy interface system
CLSTMN	Convolutional Long Short-Term Memory Network
RNN	Recurrent Neural Network
RBFNN	Radial Basis Function Neural Network
WOA	Whale Optimization Algorithm
MLKSSVM	Multi-Kernel Least Square Support Vector Machine
GSA	Gravitational Search Algorithm
A.S	Alor Setar
B.L	Bayan Lepas
C.H	Cameron Highlands
I.P	Ipoh
K.B	Kota Bharu
RMSE	Root-mean-square error

Abbreviation	Definition
S.I	Scatter Index
NSE	Nash–Sutcliffe efficiency
U95	Uncertainty with 95% confidence level
MAE	Mean absolute error

**Acknowledgements** The author would like to thank the Malaysian Meteorological Department (MMD) for providing this study with the data.

**Author contributions** Mohammad Ehteram: methodology, formal analysis, visualization, and writing—review and editing; Fatemeh Panahi: methodology, formal analysis, visualization, and writing—review and editing, Nouar AlDahoul: methodology, formal analysis, visualization, and writing—review and editing; Ali Najah Ahmed: methodology and writing—review and editing, Yuk Feng Huang: data curation and writing—review and editing, Ahmed Elshafie: writing—review and editing and supervision.

**Funding** Not applicable.

**Data availability** Data are available from the authors upon reasonable request.

**Declarations**

**Conflict of interest** The authors declare that there is no conflict of interest.

**Ethical approval** Not applicable.

**Informed consent** Not applicable.

**References**

Abba SI, Pham QB, Saini G, Linh NTT, Ahmed AN, Mohajane M et al (2020) Implementation of data intelligence models coupled with ensemble machine learning for prediction of water quality index. *Environ Sci Pollut Res*. <https://doi.org/10.1007/s11356-020-09689-x>

Abedi M, Tan X, Klausner JF, Bénard A (2023a) Solar desalination chimneys: investigation on the feasibility of integrating solar chimneys with humidification–dehumidification systems. *Renewable Energy* 202:88–102

Abedi M, Tan X, Klausner JF, Murillo MS, Benard A (2023b) A comparison of the performance of a data-driven surrogate model of a dehumidifier with mathematical model of humidification–dehumidification system. In: *AIAA SCITECH 2023 forum*, p 2329

Ahmed K, Ewees AA, El Aziz MA, Hassanien AE, Gaber T, Tsai P-W et al (2016) A hybrid krill-ANFIS model for wind speed forecasting. *Adv Intell Syst Comput*. [https://doi.org/10.1007/978-3-319-48308-5\\_35](https://doi.org/10.1007/978-3-319-48308-5_35)

AlDahoul N, Momo MA, Chong KL et al (2023) Streamflow classification by employing various machine learning models for peninsular Malaysia. *Sci Rep* 13:14574. <https://doi.org/10.1038/s41598-023-41735-9>

- Allawi MF, Othman FB, Afan HA, Ahmed AN, Hossain MS, Fai CM et al (2019) Reservoir evaporation prediction modeling based on artificial intelligence methods. *Water (Switz)*. <https://doi.org/10.3390/w11061226>
- Banadkooki FB, Ehteram M, Ahmed AN, Teo FY, Ebrahimi M, Fai CM et al (2020) Correction to: Suspended sediment load prediction using artificial neural network and ant lion optimization algorithm (*Environmental Science and Pollution Research*, (2020), 27, 30, (38094–38116), <https://doi.org/10.1007/s11356-020-09876-w>). *Environ Sci Pollut Res* 27:38117–38119. <https://doi.org/10.1007/s11356-020-10139-x>
- Chen H, Fan DL, Fang L, Huang W, Huang J, Cao C et al (2020) Particle swarm optimization algorithm with mutation operator for particle filter noise reduction in mechanical fault diagnosis. *Int J Pattern Recognit Artif Intell* 34:2058012. <https://doi.org/10.1142/s0218001420580124>
- Daoui A, Yamni M, Karmouni H, Sayyouri M, Qjidaa H (2021) Biomedical signals reconstruction and zero-watermarking using separable fractional order Charlier–Krawtchouk transformation and Sine Cosine Algorithm. *Signal Process* 180:107854. <https://doi.org/10.1016/j.sigpro.2020.107854>
- Darwish A, Ezzat D, Hassanien AE (2020) An optimized model based on convolutional neural networks and orthogonal learning particle swarm optimization algorithm for plant diseases diagnosis. *Swarm Evol Comput* 52:100616. <https://doi.org/10.1016/j.swevo.2019.100616>
- Dhiman G (2021) SSC: a hybrid nature-inspired meta-heuristic optimization algorithm for engineering applications. *Knowl Based Syst* 222:106926. <https://doi.org/10.1016/j.knosys.2021.106926>
- Ehteram M, Ahmed AN, Latif SD, Huang YF, Alizamir M, Kisi O et al (2020) Design of a hybrid ANN multi-objective whale algorithm for suspended sediment load prediction. *Environ Sci Pollut Res*. <https://doi.org/10.1007/s11356-020-10421-y>
- Ehteram M, Ferdowsi A, Faramarzpour M, Al-Janabi AMS, Al-Ansari N, Bokde ND et al (2021) Hybridization of artificial intelligence models with nature inspired optimization algorithms for lake water level prediction and uncertainty analysis. *Alex Eng J* 60:2193–2208. <https://doi.org/10.1016/j.aej.2020.12.034>
- El-Shafie A, Alsulami HM, Jahanbani H, Najah A (2013) Multi-lead ahead prediction model of reference evapotranspiration utilizing ANN with ensemble procedure. *Stoch Environ Res Risk Assess* 27:1423–1440. <https://doi.org/10.1007/s00477-012-0678-6>
- Ewees AA, Abdelaziz M, Al-Qaness MAA, Khalil HA, Kim S (2020) Improved artificial bee colony using sine-cosine algorithm for multi-level thresholding image segmentation. *IEEE Access* 8:26304–26315. <https://doi.org/10.1109/access.2020.2971249>
- GnanaSheela K, Deepa SN (2013) Neural network based hybrid computing model for wind speed prediction. *Neurocomputing* 122:425–429. <https://doi.org/10.1016/j.neucom.2013.06.008>
- Hwang YK, Ibrahim MZ, Ahmed AN, Albani A (2019) An optimized ANN measure-correlate-predict method for long-term wind prediction in Malaysia. In: *Proceedings of the conference on industrial and commercial use of energy, ICUE*, vol 2018. <https://doi.org/10.23919/ICUE-GESD.2018.8635790>
- Jumin E, Zaini N, Ahmed AN, Abdullah S, Ismail M, Sherif M et al (2020) Machine learning versus linear regression modelling approach for accurate ozone concentrations prediction. *Eng Appl Comput Fluid Mech* 14:713–725. <https://doi.org/10.1080/19942060.2020.1758792>
- Jumin E, Basaruddin FB, Yusoff YBM, Latif SD, Ahmed AN (2021) Solar radiation prediction using boosted decision tree regression model: a case study in Malaysia. *Environ Sci Pollut Res*. <https://doi.org/10.1007/s11356-021-12435-6>
- Kaveh A, Dadras EA (2020) Water strider algorithm: a new metaheuristic and applications. *Structures* 25:520–541. <https://doi.org/10.1016/j.istruc.2020.03.033>
- Khatibi R, Ghorbani MA, Pourhosseini FA (2017) Stream flow predictions using nature-inspired Firefly Algorithms and a Multiple Model strategy—directions of innovation towards next generation practices. *Adv Eng Inform* 34:80–89. <https://doi.org/10.1016/j.aei.2017.10.002>
- Kong X, Liu X, Shi R, Lee KY (2015) Wind speed prediction using reduced support vector machines with feature selection. *Neurocomputing* 169:449–456. <https://doi.org/10.1016/j.neucom.2014.09.090>
- Kumar M (2020) Social, economic, and environmental impacts of renewable energy resources. *Wind Sol Hybrid Renew Energy Syst*. <https://doi.org/10.5772/intechopen.89494>
- Kumar G, Malik H (2016) Generalized regression neural network based wind speed prediction model for western region of India. *Procedia Comput Sci* 93:26–32. <https://doi.org/10.1016/j.procs.2016.07.177>
- Lei H, Lei T, Yuenian T (2021) Sports image detection based on particle swarm optimization algorithm. *Microprocess Microsyst* 80:103345. <https://doi.org/10.1016/j.micpro.2020.103345>
- Li E, Zhou J, Shi X, JahedArmaghani D, Yu Z, Chen X et al (2020) Developing a hybrid model of salp swarm algorithm-based support vector machine to predict the strength of fiber-reinforced cemented paste backfill. *Eng Comput*. <https://doi.org/10.1007/s00366-020-01014-x>
- Liu H, Tian H, Li Y (2012) Comparison of two new ARIMA-ANN and ARIMA-Kalman hybrid methods for wind speed prediction. *Appl Energy* 98:415–424. <https://doi.org/10.1016/j.apenergy.2012.04.001>
- Liu H, Tian H, Chen C, Li Y (2013) An experimental investigation of two Wavelet-MLP hybrid frameworks for wind speed prediction using GA and PSO optimization. *Int J Electr Power Energy Syst* 52:161–173. <https://doi.org/10.1016/j.ijepes.2013.03.034>
- Liu H, Tian H, Liang X, Li Y (2015) New wind speed forecasting approaches using fast ensemble empirical model decomposition, genetic algorithm, Mind Evolutionary Algorithm and Artificial Neural Networks. *Renew Energy* 83:1066–1075. <https://doi.org/10.1016/j.renene.2015.06.004>
- Liu H, Mi X, Li Y (2018) Smart deep learning based wind speed prediction model using wavelet packet decomposition, convolutional neural network and convolutional long short term memory network. *Energy Convers Manag* 166:120–131. <https://doi.org/10.1016/j.enconman.2018.04.021>
- Liu Y, Qin H, Zhang Z, Pei S, Jiang Z, Feng Z et al (2020) Probabilistic spatiotemporal wind speed forecasting based on a variational Bayesian deep learning model. *Appl Energy* 260:114259. <https://doi.org/10.1016/j.apenergy.2019.114259>
- Mohamadi S, Sammen SS, Panahi F, Ehteram M, Kisi O, Mosavi A et al (2020) Zoning map for drought prediction using integrated machine learning models with a nomadic people optimization algorithm. *Nat Hazards* 104:537–579. <https://doi.org/10.1007/s11069-020-04180-9>
- Muslim TO, Ahmed AN, Malek MA, AbdulmohsinAfan H, Khaleel Ibrahim R, El-Shafie A et al (2020a) Investigating the influence of meteorological parameters on the accuracy of sea-level prediction models in Sabah, Malaysia. *Sustainability* 12:1193. <https://doi.org/10.3390/su12031193>
- Muslim TO, Ahmed AN, Malek MA, Afan HA, Ibrahim RK, El-Shafie A et al (2020b) Investigating the influence of meteorological parameters on the accuracy of sea-level prediction models in Sabah, Malaysia. *Sustain*. <https://doi.org/10.3390/su12031193>
- Najah Ahmed A, Binti Othman F, AbdulmohsinAfan H, Khaleel Ibrahim R, Ming Fai C, Shabbir Hossain M et al (2019) Machine

- learning methods for better water quality prediction. *J Hydrol* 578:124084. <https://doi.org/10.1016/j.jhydrol.2019.124084>
- Navas RKB, Prakash S, Sasipraba T (2020) Artificial Neural Network based computing model for wind speed prediction: a case study of Coimbatore, Tamil Nadu, India. *Phys A Stat Mech Appl* 542:123383. <https://doi.org/10.1016/j.physa.2019.123383>
- Neggaz N, Houssein EH, Hussain K (2020) An efficient henry gas solubility optimization for feature selection. *Expert Syst Appl* 113364
- Osman AIA, Ahmed AN, Chow MF, Huang YF, El-Shafie A (2021) Extreme gradient boosting (Xgboost) model to predict the groundwater levels in Selangor Malaysia. *Ain Shams Eng J*
- Paikray HK, Das PK, Panda S (2021) Optimal multi-robot path planning using particle swarm optimization algorithm improved by sine and cosine algorithms. *Arab J Sci Eng* 46:3357–3381. <https://doi.org/10.1007/s13369-020-05046-9>
- Ramasamy P, Chandel SS, Yadav AK (2015) Wind speed prediction in the mountainous region of India using an artificial neural network model. *Renew Energy* 80:338–347. <https://doi.org/10.1016/j.renene.2015.02.034>
- Salgotra R, Singh U, Singh S, Singh G, Mittal N (2021) Self-adaptive salp swarm algorithm for engineering optimization problems. *Appl Math Model* 89:188–207. <https://doi.org/10.1016/j.apm.2020.08.014>
- Samadianfard S, Hashemi S, Kargar K, Izadyar M, Mostafaeipour A, Mosavi A et al (2020) Wind speed prediction using a hybrid model of the multi-layer perceptron and whale optimization algorithm. <https://doi.org/10.20944/preprints202002.0233.v1>
- Shabani E, Hayati B, Pishbahar E, Ghorbani MA, Ghahremanzadeh M (2021) A novel approach to predict CO<sub>2</sub> emission in the agriculture sector of Iran based on Inclusive Multiple Model. *J Clean Prod* 279:123708. <https://doi.org/10.1016/j.jclepro.2020.123708>
- Sharafati A, Asadollah SBHS, Neshat A (2020) A new artificial intelligence strategy for predicting the groundwater level over the Rafsanjan aquifer in Iran. *J Hydrol* 591:125468. <https://doi.org/10.1016/j.jhydrol.2020.125468>
- Sun S, Fu J, Li A, Zhang P (2020) A new compound wind speed forecasting structure combining multi-kernel LSSVM with two-stage decomposition technique. *Soft Comput* 25:1479–1500. <https://doi.org/10.1007/s00500-020-05233-8>
- Tubishat M, Ja'afar S, Alswaiti M, Mirjalili S, Idris N, Ismail MA et al (2021) Dynamic Salp swarm algorithm for feature selection. *Expert Syst Appl* 164:113873. <https://doi.org/10.1016/j.eswa.2020.113873>
- Xin-gang Z, Ji L, Jin M, Ying Z (2020) An improved quantum particle swarm optimization algorithm for environmental economic dispatch. *Expert Syst Appl* 152:113370. <https://doi.org/10.1016/j.eswa.2020.113370>
- Xu W, Liu P, Cheng L, Zhou Y, Xia Q, Gong Y et al (2021) Multi-step wind speed prediction by combining a WRF simulation and an error correction strategy. *Renew Energy* 163:772–782. <https://doi.org/10.1016/j.renene.2020.09.032>
- Yafouz A, Ahmed AN, Zaini N, El-Shafie A (2021) Ozone concentration forecasting based on artificial intelligence techniques: a systematic review. *Water Air Soil Pollut* 232:79. <https://doi.org/10.1007/s11270-021-04989-5>
- Yu C, Li Y, Bao Y, Tang H, Zhai G (2018) A novel framework for wind speed prediction based on recurrent neural networks and support vector machine. *Energy Convers Manag* 178:137–145. <https://doi.org/10.1016/j.enconman.2018.10.008>
- Zhang C, Wei H, Zhao X, Liu T, Zhang K (2016) A Gaussian process regression based hybrid approach for short-term wind speed prediction. *Energy Convers Manag* 126:1084–1092. <https://doi.org/10.1016/j.enconman.2016.08.086>
- Zhang Y, Pan G, Chen B, Han J, Zhao Y, Zhang C (2020) Short-term wind speed prediction model based on GA-ANN improved by VMD. *Renew Energy* 156:1373–1388. <https://doi.org/10.1016/j.renene.2019.12.047>

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

Springer Nature or its licensor (e.g. a society or other partner) holds exclusive rights to this article under a publishing agreement with the author(s) or other rightsholder(s); author self-archiving of the accepted manuscript version of this article is solely governed by the terms of such publishing agreement and applicable law.

## Authors and Affiliations

Mohammad Ehteram<sup>1</sup> · Fatemeh Panahi<sup>2</sup> · Nouar AlDahoul<sup>3</sup> · Ali Najah Ahmed<sup>4</sup>  · Yuk Feng Huang<sup>5</sup> · Ahmed Elshafie<sup>6,7</sup>

✉ Ali Najah Ahmed  
mahfoodh@sunway.edu.my

Mohammad Ehteram  
mohammdehteram@semnan.ac.ir

Fatemeh Panahi  
fpanahi@kashanu.ac.ir

Nouar AlDahoul  
naa9497@nyu.edu

Yuk Feng Huang  
huangyf@utar.edu.my

Ahmed Elshafie  
elshafie@um.edu.my

<sup>1</sup> Department of Water Engineering and Hydraulic Structures, Faculty of Civil Engineering, Semnan University, Semnan, Iran

<sup>2</sup> Faculty of Natural Resources and Earth Sciences, University of Kashan, Kashan, Iran

<sup>3</sup> Computer Science, New York University, Abu Dhabi, United Arab Emirates

<sup>4</sup> Research Centre For Human-Machine Collaboration (HUMAC), School of Engineering and Technology, Sunway University, No. 5, Jalan Universiti Bandar Sunway, 43000 Darul Ehsan, Selangor, Malaysia

<sup>5</sup> Department of Civil Engineering, Lee Kong Chian Faculty of Engineering and Science, Universiti Tunku Abdul Rahman, Kajang, Selangor, Malaysia

<sup>6</sup> Department of Civil Engineering, Faculty of Engineering, University of Malaya (U.M.), 50603 Kuala Lumpur, Malaysia

<sup>7</sup> National Water and Energy Center, United Arab Emirate University, P.O. Box 15551, Al Ain, United Arab Emirates