



Estimating the short-term and long-term wind speeds: implementing hybrid models through coupling machine learning and linear time series models

Saeid Mehdizadeh¹ · Ali Kozekalani Sales² · Mir Jafar Sadegh Safari³

Received: 4 February 2020 / Accepted: 28 April 2020 / Published online: 9 May 2020
© Springer Nature Switzerland AG 2020

Abstract

Wind speed data are of particular importance in the design and management of wind power projects. In the current study, three types of linear time series models including autoregressive (AR), moving average (MA), and autoregressive moving average (ARMA) were employed to estimate short-term (i.e., daily) and long-term (i.e., monthly) wind speeds. The required data were gathered, respectively, from the Tabriz and Zahedan stations in the northwest and southeast of Iran. The MA models outperformed the AR and ARMA on the both daily and monthly scales. Daily and monthly wind speed values, as a function of lagged wind speed data, were then estimated using two machine learning models of random forests (RF) and multivariate adaptive regression splines (MARS). It was found that the RF and MARS provided similar results; however, RF performed slightly better than the MARS. Finally, the stand-alone time series and machine learning models were coupled to improve the accuracy of the wind speed estimation. Accordingly, the hybrid RF-AR, RF-MA, RF-ARMA, MARS-AR, MARS-MA, and MARS-ARMA models were implemented. It was concluded that, the hybrid models outperformed the stand-alone RF and MARS for both short- and long-term wind speed estimations where, the RF-AR and MARS-AR hybrid models provided the best performances. The hybrid models tested in the present study could be effective alternatives to the stand-alone machine learning-based RF and MARS models for the estimation of wind speed time series.

Keywords Wind speed · Estimation · Time series models · Machine learning models · Stand-alone and hybrid models

1 Introduction

Wind speed affects decisions related to agriculture, maritime and air transport, urban air pollution management, and many other important domains. Recently, wind energy as a renewable energy source has received considerable attention worldwide, as the uncertainties inherent in our dependence on fossil fuels are increasingly recognized [1, 2]. Wind energy plays a vital role in a sustainable future as an important and widely applicable green energy alternative [3]. The analysis of wind characteristics, such as speed,

is essential for the design and implementation of wind energy projects; these characteristics directly affect the energy production potential of a wind power plant. Robust techniques must be used for the interpretation and estimation of wind speeds due to the challenges presented by the random/stochastic, indeterminant, discontinuous, and fluctuating nature of the wind speed time series [4–8]. Accurate forecasting and estimation of wind speeds are essential if wind energy is to reach its full potential as an alternative energy in a more sustainable future.

✉ Mir Jafar Sadegh Safari, jafar.safari@yasar.edu.tr; Saeid Mehdizadeh, saeid.mehdizadeh@gmail.com; Ali Kozekalani Sales, ali.kozekalani1368@gmail.com | ¹Department of Water Engineering, Urmia University, Urmia, Iran. ²Department of Civil Engineering, Elm-o-Fan University College of Science and Technology, Urmia, Iran. ³Department of Civil Engineering, Yaşar University, Izmir, Turkey.



Wind speed can be measured directly by anemometers installed at the desired site, or indirectly, using wind speed estimation models. These can be classified as either physical or statistical models [9]. Physical models employ physical and mathematical formulas to estimate wind speed information [10]; however, these models can be complex and time-consuming to employ. Statistical models, including the time series models and machine learning-based techniques, have the potential to estimate wind speed using historical data and other related parameters [2]. In another classification, wind speed estimation models can be categorized in the stand-alone and hybrid paradigms. In fact, the hybrid models are developed to improve the performance of stand-alone models when estimating the intended time series like wind speed studied in the present study. Different approaches could be taken into consideration to implement the hybrid models. One of these methods is the use of data preprocessing techniques including the wavelet analysis (WA), empirical mode decomposition (EMD), etc. Hybridizing the time series models and machine learning ones could also be used as alternatives to the stand-alone models with a dependable level of precision. Finally, bioinspired optimization algorithms are capable of optimizing the parameters of machine learning models to increase the accuracy of stand-alone models.

Wind speed time series data, like other meteorological data, consist of both stochastic and deterministic components [11, 12]. In this context, time series and machine learning models are able to capture the stochastic and deterministic terms, respectively. Stand-alone time series and machine learning models focus on estimating one of these components; however, a truly accurate wind speed estimation technique should consider both terms. Hybrid machine learning-time series models have the ability to capture both the stochastic and deterministic terms and could have great potential as tools for estimating meteorological data.

In recent decades, numerous studies have explored wind speed estimation applying various techniques, including the stand-alone and hybrid models. Liu et al. [13] proposed two hybrid models to predict wind speed by combining the autoregressive integrated moving average (ARIMA) with artificial neural networks (ANNs) and a Kalman filter (KF). The obtained results showed that the hybrid models had superior outcomes than the classical ARIMA model. A comparative study by Li and Shi [14] evaluated the accuracy of three types of ANN, namely back-propagation (BP), radial basis function (RBF), and adaptive linear element (ALR). The RBF was found to provide better estimates of wind speed time series. Hu and Chen [15] applied the differential evolution algorithm to improve the performance of long short-term memory (LSTM) models in wind speed forecasting. The accuracy of the developed

hybrid model was improved compared to the stand-alone one. Liu et al. [16] improved the efficiency of a support vector machine (SVM) model by coupling it with genetic algorithms (GA) for estimating wind speed time series. In addition, particle swarm optimization (PSO) was used by Kong et al. [17] to increase the accuracy of wind speed estimation via the reduced SVM (RSVM). Jiang et al. [18] applied two optimization algorithms, including the PSO and cuckoo search (CS), to optimize the parameters of the SVM technique. The hybrid SVM-CS demonstrated better results than the SVM-PSO. Khosravi et al. [19] compared the performance of machine learning models in wind speed estimation. The results suggested that the group method of data handling outperformed the other models used. Liu et al. [20] introduced three types of hybrid models by coupling the extreme learning machine (ELM) with the wavelet packet decomposition (WPD) and empirical mode decomposition (EMD) for wind speed forecasting. The performance of the proposed models was compared with ARIMA, SVM, and ELM, and the hybrid WPD-EMD-ELM model surpassed the others. Nikolic et al. [21] compared the accuracy of ANN, SVM, ELM, and genetic programming (GP) for wind speed estimation. They reported that the ELM showed the best performance. Mohandes et al. [22] estimated wind speed profiles at various heights with the application of an adaptive neuro-fuzzy inference system (ANFIS). Their results suggested that the ANFIS is a reliable approach for estimating wind speed. Meng et al. [23] developed a hybrid model by coupling ANN with WPD and a crisscross algorithm. The proposed hybrid model was found to present superior results compared to the single ANN for wind speed forecasting. Shukur and Lee [24] introduced a hybrid model, KF-ANN, utilizing ARIMA. While the single KF and ANN models outperformed the single ARIMA, the hybrid model yielded improved estimates of wind speed time series compared to the stand-alone models. Torres et al. [25] examined the performance of the autoregressive moving average (ARMA) time series model and confirmed its potential for wind speed modeling. The performance of ANN and seasonal ARIMA (SARIMA) was evaluated by Bivona et al. [26] in estimating the wind speed data. Both techniques produced similar results. Each of these studies confirmed the applicability of time series and machine learning models for the estimation of wind speeds. Wu et al. [27] hybridized the ELM, complete ensemble EMD (CEEMD), and multi-objective gray wolf optimization (MOGWO) and reported that such a hybrid model can be appropriately employed for wind speed estimation. Bilgili and Sahin [28] predicted wind speed using comparative analysis of regression, namely linear regression (LR), nonlinear regression (NLR), and artificial neural network (ANN) methods. In this study, three measuring stations, including Antakya, Mersin, and Samandag, were used.

Findings illustrated that ANN method had better results than the two used models such as LR and NLR. Ghorbani et al. [29] predicted short-term wind speed with machine learning techniques of ANNs and genetic expression programming (GEP) comparing with MLR method. The results indicated that both the GEP and ANN were trusted selections as well as MLR method. Time series analysis of wind speed with time-varying turbulence is investigated by Ewing et al. [30] to obtain wind energy production, operational and financial hedging strategies of wind-related risk applying the ARMA-GARCH-in-mean framework. In general, the findings revealed that the results have practical implications for wind plant operations.

The present study aims to (1) evaluate the efficiency of three linear time series models including autoregressive (AR), moving average (MA), and ARMA, for short- and long-term wind speed estimation; (2) investigate the accuracy of the machine learning-based random forest (RF) and multivariate adaptive regression spline (MARS) models for wind speed estimation; (3) develop and implement six hybrid models via coupling the noted linear time series and machine learning models; and (4) compare the performance of all of the developed stand-alone and hybrid models. To achieve these objectives, two stations, the Tabriz station northwest of Iran and the Zahedan station southeast of Iran, were selected as the study region.

It is worthy to note that despite widespread applications of machine learning and time series models, specifically machine learning models, the hybrid machine learning-time series models have received less attention for estimating the wind speed.

2 Materials and method

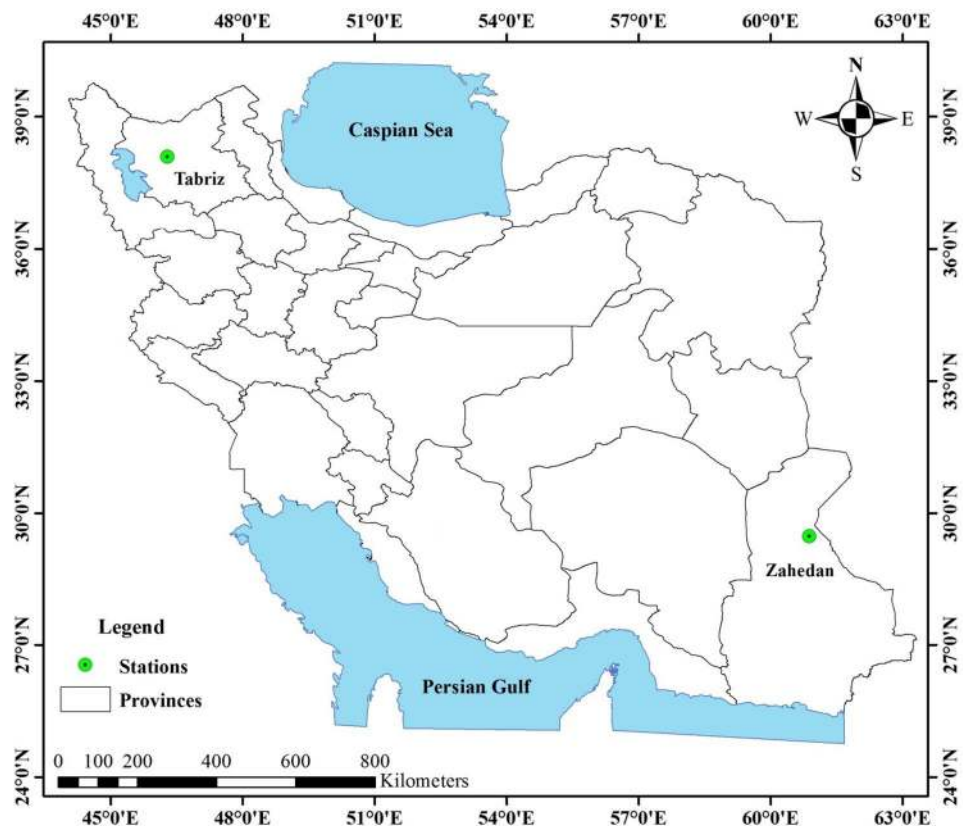
2.1 Study location

In the current study, the Tabriz and Zahedan stations located in the northwest and southeast of Iran, respectively, were selected as the case study sites (as shown in Fig. 1). The Tabriz station, with a latitude of $38^{\circ} 05' N$, a longitude of $46^{\circ} 17' E$, and an elevation of 1361.0 m above the sea level, has a semiarid climate. The Zahedan station has an arid climate and is located at a latitude of $29^{\circ} 28' N$, a longitude of $60^{\circ} 53' E$, and an elevation of 1370.0 m above the sea level.

2.2 Data

Raw data for the present study, including daily and monthly wind speed data from the selected stations, were collected from the Iranian Meteorological Organization

Fig. 1 Geographical position of the studied locations in Iran



(IMO) during a 25-year period from January 1, 1990, to December 31, 2014. The collected data were divided into two separate datasets, the training and testing datasets. For both of the studied time scales, the first 20-year data (i.e., 80% of the entire data from 1990 to 2009) were applied as the training dataset, while the last 5-year data, between 2010 and 2014 (i.e., 20% of the entire data), were used as the testing dataset. Figure 2 shows the time series plots for the observed daily and monthly wind speeds of the studied locations during both the training and testing periods.

In addition, statistical parameters of the wind speed data for the study period, including the minimum (X_{min}), maximum (X_{max}), average (X_{mean}), and standard deviation (X_{SD}), are listed in Table 1. The daily wind speed data ranged from 0.00 to 11.25 m s⁻¹ (training phase), and 0.00 to 10.63 m s⁻¹ (testing phase) at the Tabriz station, as well as between 0.00 and 14.61 m s⁻¹ (training phase), 0.38 and 15.25 m s⁻¹ (testing phase) at the Zahedan station. Moreover, the standard deviations of the daily wind speed data were higher than the monthly data at both stations.

Daily and monthly wind speed datasets for the studied locations were standardized utilizing the following equation:

$$WS_s = \frac{WS_o - \overline{WS_o}}{\sigma_{WS_o}} \tag{1}$$

where WS_s , WS_o , $\overline{WS_o}$, and σ_{WS_o} denote the standardized wind speed, the observed wind speed, the average of the observed wind speed data, and the standard deviation of the observed wind speed data, respectively.

2.3 Time series models employed

The idea of time series was initially proposed by Thomas and Fiering [31] and then developed by Box and Jenkins [32]. The time series models used in the current study include autoregressive (AR), moving average (MA), and autoregressive moving average (ARMA); they are classified as linear time series models. The AR model is developed according to the Markov chain theory, i.e., any time series data recorded at the current time t are related to its values at a former time (i.e., $t - 1$, $t - 2$, ...). To achieve the stochastic term in the AR, a simple linear regression is performed on the past values of observed data (e.g., wind speed) in the standardized form; however, in the MA model, it is done on the past values of the stochastic term. On the other hand, this regression is performed simultaneously in the ARMA on the past values of the standardized observational data and stochastic term. The AR, MA, and ARMA models can be, respectively, written as:

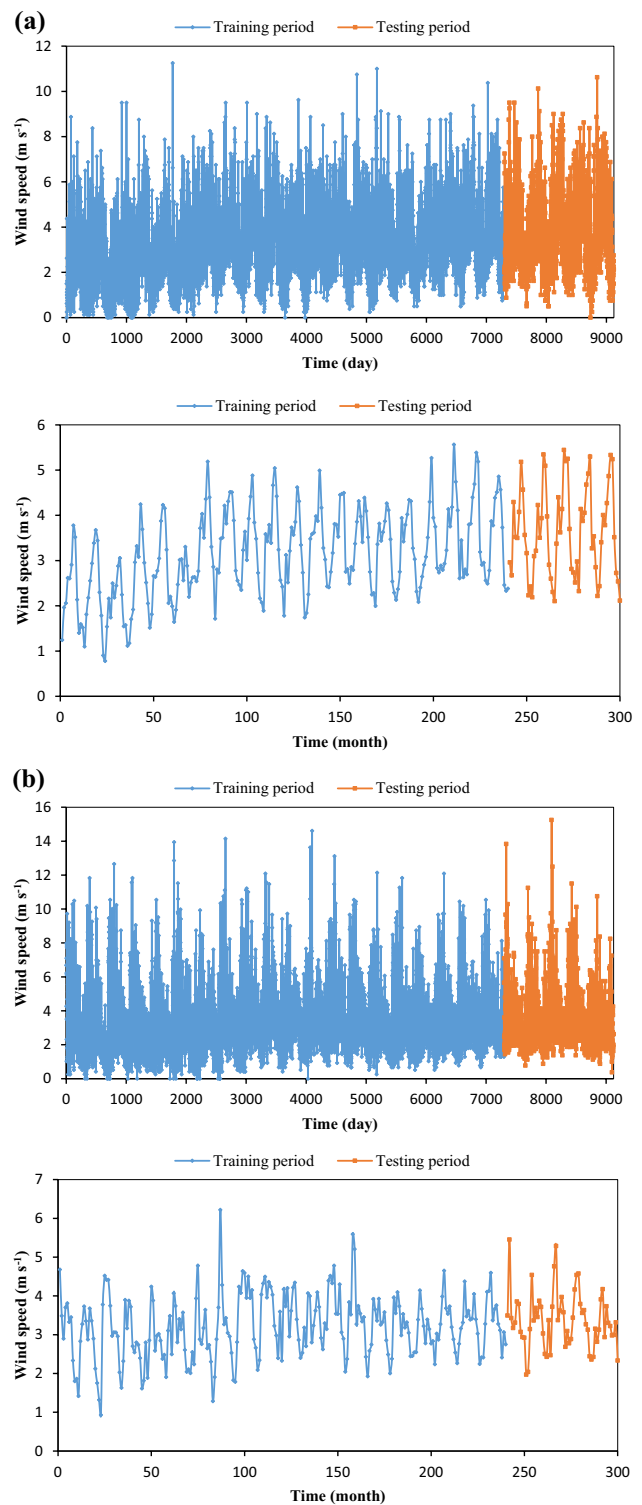


Fig. 2 Time series of the observed daily and monthly wind speeds at the studied locations during the 1990–2014: **a** Tabriz station; **b** Zahedan station

Table 1 Statistical parameters of the daily and monthly wind speed data at the studied locations

Station	Time scale	Dataset	X_{min}	X_{max}	X_{mean}	X_{SD}
Tabriz	Daily	Training	0.00	11.25	3.13	1.61
		Testing	0.00	10.63	3.64	1.68
	Monthly	Training	0.78	5.56	3.13	0.95
		Testing	2.10	5.45	3.64	1.00
Zahedan	Daily	Training	0.00	14.61	3.21	1.87
		Testing	0.38	15.25	3.36	1.56
	Monthly	Training	0.93	6.22	3.22	0.82
		Testing	1.97	5.45	3.37	0.72

$$Z_t = \sum_{i=1}^p (\varphi_i \cdot Z_{t-i}) + \varepsilon_t \tag{2}$$

$$Z_t = \sum_{j=1}^q (\theta_j \cdot \varepsilon'_{t-j}) + \varepsilon'_t \tag{3}$$

$$Z_t = \sum_{i=1}^p (\varphi_i \cdot Z_{t-i}) - \sum_{j=1}^q (\theta_j \cdot \varepsilon'_{t-j}) + \varepsilon'_t \tag{4}$$

where Z_t indicates the observed time series in the standardized form, p and q represent the order of AR and MA, respectively, φ_i and θ_j , respectively, denote the coefficients of AR and MA, and ε_t , ε'_t , and ε''_t represent the stochastic series of the observed data estimated via the AR, MA, and ARMA models, respectively.

Estimation procedure by the time series models comprises three steps: (1) standardizing the observed data, (2) specifying the parameters of the used models (i.e., φ_i and θ_j in Eqs. 2–4), and (3) reporting the optimal AR, MA, and ARMA models, specifically, identifying the models that result in the lowest errors according to the chosen evaluation metrics.

2.4 Machine learning models employed

2.4.1 Random forest

Random forest (RF), a model proposed by Breiman [33], is a relatively new supervised data classification approach. Furthermore, RF is a decision tree-based method in which a number of decision trees are combined to create the final RF. Simple training and tuning of parameters, generation of accurate results, and the ability to fit nonlinear models are among the advantages of the RF method [33].

In general, a decision tree can provide good results for training datasets; however, it may produce poor results in the testing phase (i.e., over-fitting) due to the existence of a tree containing many splits. For such a situation, tree pruning is essential to reduce the number of trees and

splits [34, 35]. The pruning process often leads to improvements in the performance of a decision tree model. However, decision trees in the RF technique are grouped without pruning. Instead, randomization through bagging and random feature selection is used to combat over-fitting and generate accurate results [33, 36]. Random samples are selected as training sets using the bagging method. In general, two-thirds of the training samples, called out-of-bag samples, are applied toward tree growth, and the remaining third is used for cross-validation.

The number of trees (ntree) and predictors on each node (mtry) are two user-defined parameters affecting the performance of RF. Breiman [33] stated that ntree = 500 are enough to obtain good results. While using more than 500 trees does not improve considerably model performance, it also increases the required time to create the final RF algorithm structure. In terms of nodes, mtry = 1 was used in the current study since only one input (i.e., 1-day or 1-month lagged wind speed data) was used as an estimator in RF.

2.4.2 Multivariate adaptive regression splines

Multivariate adaptive regression spline (MARS) was initially developed by Friedman [37] and is arranged as nonparametric regression models. Nonparametric term implies that the relationships among the input and output parameters are not already known to the user. This is similar to the multiple linear regression (MLR) approach. However, in MLR, a unique slope is assigned for each predictor by the MLR, while in MARS, various slopes are given for the different ranges of input variables. This leads to the generation of more flexible regression models for estimating the target parameter. The solution space in MARS is divided into some intervals of input variables, and splines are fitted to each interval. Specific spline functions are developed for each interval, and the endpoints of each interval are called knots. The X predictor is mapped to another variable (i.e., Y) via the basis functions. The basis functions can be formulated as either of the following equations:

$$Y = \max(0; C - X) \tag{5}$$

$$Y = \max(0; X - C) \tag{6}$$

where C illustrates the knot (or a threshold value). The basis functions are combined together to produce the final structure of the MARS as follows [38]:

$$f(x) = \beta_0 + \sum_{m=1}^M \beta_m \cdot BF_m(x) \tag{7}$$

where β_0 is an initial constant, β_m is the coefficient of the m th basis function, and BF_m is the m th basis function. The β constants are estimated through the least squares regression technique.

Estimating the target parameter using MARS involves a forward and backward step. During the forward step, overfitting often occurs, along with a number of surplus knots; the MARS model produced in this first step seemingly offers poor results. The backward step involves a pruning technique to remove the extra knots and improve model performance. The optimal model is that with the smallest generalized cross-validation (GCV) value.

2.5 Models development

2.5.1 Time series models development

When developing the stand-alone time series models used in the current research, different orders of p in AR, q in MA, and p and q in ARMA were examined, and the best-performing models were then selected and reported. The AR, MA, and ARMA time series models were fitted to the standardized wind speed data, and the best-performing models from this step were selected.

2.5.2 Machine learning models development

Two types of machine learning models, i.e., RF and MARS, were employed to estimate daily and monthly wind speed time series. Here, lagged wind speed data (i.e., 1 day and 1 month lagged data) were utilized as inputs into the RF and MARS models to estimate the wind speed of the current day or month.

2.5.3 Hybrid models development

The main objective of current study is to improve the estimation performance of stand-alone RF and MARS models through coupling the machine learning models with the time series models that are capable of capturing the stochastic term of wind speed data (i.e., AR, MA, and ARMA). To this end, six hybrid models were developed and

evaluated, including RF-AR, RF-MA, RF-ARMA, MARS-AR, MARS-MA, and MARS-ARMA. The hybrid machine learning-time series models were developed using the following equation:

$$WS_H = WS_S + WS_D \tag{8}$$

where WS_H , WS_S , and WS_D denote the wind speed data estimated via the hybrid machine learning-time series models, the estimated wind speed (i.e., stochastic term) produced by the stand-alone time series models, and the estimated wind speed (i.e., deterministic term) found using the stand-alone machine learning models, respectively.

3 Application and results

Five statistical metrics, including root-mean-square error (RMSE), mean absolute error (MAE), relative RMSE (RRMSE), coefficient of determination (R^2), and Nash–Sutcliffe efficiency (NSE), were applied to evaluate the performance of the developed models. These metrics were calculated through the following formulas:

$$RMSE = \sqrt{\frac{\sum_{i=1}^N (WS_{o,i} - WS_{e,i})^2}{N}} \tag{9}$$

$$MAE = \frac{\sum_{i=1}^N |WS_{o,i} - WS_{e,i}|}{N} \tag{10}$$

$$RRMSE = \frac{\sqrt{\frac{\sum_{i=1}^N (WS_{o,i} - WS_{e,i})^2}{N}}}{WS_o} \tag{11}$$

$$R^2 = \left[\frac{\sum_{i=1}^N (WS_{o,i} - \overline{WS_o}) \cdot (WS_{e,i} - \overline{WS_e})}{\sqrt{\sum_{i=1}^N (WS_{o,i} - \overline{WS_o})^2 \cdot \sum_{i=1}^N (WS_{e,i} - \overline{WS_e})^2}} \right]^2 \tag{12}$$

$$NSE = 1 - \frac{\sum_{i=1}^N (WS_{o,i} - WS_{e,i})^2}{\sum_{i=1}^N (WS_{o,i} - \overline{WS_o})^2} \tag{13}$$

where $WS_{o,i}$ and $WS_{e,i}$ denote the i th observed and estimated wind speed, $\overline{WS_o}$ and $\overline{WS_e}$ indicate the mean of the observed and estimated wind speed values, and N is the total number of observations.

To measure the performance of all the developed models qualitatively, the RRMSE values were categorized as follows: excellent ($0\% < RRMSE < 10\%$), good

(10% < RRMSE < 20%), fair (20% < RRMSE < 30%), and poor (RRMSE > 30%).

Scatter and time series plots were also used to visually qualify the performance of the machine learning and hybrid models.

Various AR, MA, and ARMA time series models were firstly fitted to the standardized wind speed data, and then, the optimal models that yielded the best performance were selected. Table 2 presents the equations of

the optimal time series models for the study locations on both daily and monthly scales.

The values of RMSE, MAE, RRMSE, R^2 , and NSE for each stand-alone time series models for both stations are summarized in Tables 3 and 4. A comparative evaluation of the stand-alone time series models developed at the study sites clearly shows that the MA models, i.e., MA(1) on a daily scale and MA(3) on a monthly scale at the Tabriz station, and MA(1) on both the daily and monthly scales

Table 2 Mathematical equations developed from the optimal time series models at the study sites

Station	Time scale	Models	Equations
Tabriz	Daily	AR(1)	$Z_t = (0.446344 \times Z_{t-1}) + \epsilon_t$
		MA(1)	$Z_t = (0.380492 \times \epsilon'_{t-1}) + \epsilon'_t$
		ARMA(1,1)	$Z_t = (0.650312 \times Z_{t-1}) - (0.27029 \times \epsilon''_{t-1}) + \epsilon''_t$
	Monthly	AR(1)	$Z_t = (0.697019 \times Z_{t-1}) + \epsilon_t$
		MA(3)	$Z_t = (0.186448 \times \epsilon'_{t-1}) + (0.230132 \times \epsilon'_{t-2}) + (0.124544 \times \epsilon'_{t-3}) + \epsilon'_t$
		ARMA(4,1)	$Z_t = (0.426033 \times Z_{t-1}) + (0.207713 \times Z_{t-2}) + (0.04959 \times Z_{t-3}) + (0.146769 \times Z_{t-4}) - (0.18781 \times \epsilon''_{t-1}) + \epsilon''_t$
Zahedan	Daily	AR(1)	$Z_t = (0.457735 \times Z_{t-1}) + \epsilon_t$
		MA(1)	$Z_t = (0.464683 \times \epsilon'_{t-1}) + \epsilon'_t$
		ARMA(1,1)	$Z_t = (0.437226 \times Z_{t-1}) + (0.026778 \times \epsilon''_{t-1}) + \epsilon''_t$
	Monthly	AR(1)	$Z_t = (0.398564 \times Z_{t-1}) + \epsilon_t$
		MA(1)	$Z_t = (0.289439 \times \epsilon'_{t-1}) + \epsilon'_t$
		ARMA(1,1)	$Z_t = (0.577657 \times Z_{t-1}) - (0.30129 \times \epsilon''_{t-1}) + \epsilon''_t$

Table 3 Statistical results of the AR, MA, and ARMA time series models implemented at the Tabriz station

Time scale	Models	Training					Testing				
		RMSE (m s ⁻¹)	MAE (m s ⁻¹)	RRMSE (%)	R ²	NSE	RMSE (m s ⁻¹)	MAE (m s ⁻¹)	RRMSE (%)	R ²	NSE
Daily	AR(1)	0.625	0.481	19.934	0.851	0.850	0.631	0.460	17.329	0.870	0.859
	MA(1)	0.486	0.373	15.521	0.910	0.909	0.491	0.357	13.466	0.920	0.915
	ARMA(1,1)	1.188	0.915	37.907	0.537	0.458	1.200	0.874	32.928	0.571	0.491
Monthly	AR(1)	0.421	0.325	13.456	0.807	0.802	0.372	0.308	10.225	0.936	0.860
	MA(3)	0.203	0.153	6.477	0.958	0.954	0.165	0.139	4.528	0.991	0.972
	ARMA(4,1)	0.476	0.362	15.204	0.754	0.747	0.406	0.347	11.165	0.946	0.833

Table 4 Statistical results of the AR, MA, and ARMA time series models implemented at the Zahedan station

Time scale	Models	Training					Testing				
		RMSE (m s ⁻¹)	MAE (m s ⁻¹)	RRMSE (%)	R ²	NSE	RMSE (m s ⁻¹)	MAE (m s ⁻¹)	RRMSE (%)	R ²	NSE
Daily	AR(1)	0.795	0.593	24.739	0.820	0.819	0.668	0.470	19.863	0.818	0.816
	MA(1)	0.721	0.532	22.458	0.852	0.851	0.603	0.425	17.953	0.851	0.850
	ARMA(1,1)	0.721	0.538	22.466	0.852	0.851	0.607	0.426	18.048	0.850	0.848
Monthly	AR(1)	0.213	0.168	6.638	0.932	0.931	0.217	0.164	6.438	0.924	0.907
	MA(1)	0.144	0.111	4.477	0.969	0.969	0.150	0.116	4.453	0.960	0.956
	ARMA(1,1)	0.440	0.341	13.693	0.733	0.708	0.455	0.352	13.514	0.721	0.592

at the Zahedan station, outperformed the AR and ARMA models. The statistical metric values obtained for the best-performing MA models on a monthly scale during the testing period are $RMSE = 0.165 \text{ m s}^{-1}$, $MAE = 0.139 \text{ m s}^{-1}$, $RRMSE = 4.528\%$, $R^2 = 0.991$, $NSE = 0.972$ for MA(3) at the Tabriz station, and $RMSE = 0.150 \text{ m s}^{-1}$, $MAE = 0.116 \text{ m s}^{-1}$, $RRMSE = 4.453\%$, $R^2 = 0.960$, $NSE = 0.956$ for MA(1) at the Zahedan station. The developed ARMA models generally display the weakest performance, except for the model applied to the daily data from the Zahedan station, where the ARMA models demonstrate slightly better results than AR(1).

The AR(1) and MA(1) models at the Tabriz station on a daily scale can be classified as good ($10\% < RRMSE < 20\%$), while the ARMA(1,1) model displays poor performance with $RRMSE > 30\%$. For the Zahedan station on a daily scale, the performance of all the time series models is within the fair and good classes during the training and testing periods, respectively. On a monthly scale, the performance of the MA(3) model at the Tabriz and the results of the AR(1) and MA(1) models at the Zahedan station are excellent. In terms of precision, the performance of the AR(1) and ARMA(4,1) models at the Tabriz station and ARMA(1,1) at the Zahedan station is classified as good.

The evaluation metric values calculated for the stand-alone RF and MARS models at the Tabriz and Zahedan stations are given in Tables 5 and 6, respectively. The stand-alone MARS and RF models performed poorly in the estimation of daily wind speed, yielding $RRMSE$ values greater than 30%. Conversely, when using monthly wind speed data, both the MARS and RF models produced good results according to the $RRMSE$ criterion classification ($10\% < RRMSE < 20\%$).

In total, six hybrid models were developed and evaluated: i.e., RF-AR, RF-MA, RF-ARMA, MARS-AR, MARS-MA, and MARS-ARMA. The statistic metrics computed for the hybrid models are listed in Tables 5 and 6.

Based on the $RRMSE$ index, the performance of all hybrid models was excellent ($RRMSE < 10\%$) for data from both stations on a monthly scale. On a daily scale at the Tabriz station, the accuracy of hybrid models generated via coupling the RF and MARS models with the AR and MA models was excellent, while the performance of the RF-ARMA and MARS-ARMA models generally belonged to the good class. For the Zahedan station on a daily scale, the accuracy of the hybrid models was classified as excellent, except for the RF-MA and MARS-MA models during the training period, which displayed good performance.

The MA and ARMA time series models, respectively, demonstrate the best and worst wind speed estimation ability at both locations on daily and monthly scales. The stand-alone RF and MARS models generally illustrate similar results on both time scales; however, the RF model

shows slightly improved statistics compared to the MARS. Comparing the accuracy of stand-alone time series and machine learning models (i.e., AR, MA, ARMA, RF, and MARS) at the study sites reveals that the stand-alone AR, MA, and ARMA time series models yield better estimates of wind speed than the stand-alone RF and MARS on a daily scale. This result is also observed on a monthly scale for the Zahedan station during both training and testing periods; however, the RF and MARS show better results than the AR and ARMA at the Tabriz station during the training period.

The developed hybrid models outperform the stand-alone RF and MARS, yielding lower $RMSE$, MAE , and $RRMSE$ values, as well as higher R^2 and NSE values (see Tables 5 and 6). For example, the error statistics of best-performing hybrid model at the Tabriz station on a daily scale during the testing period (i.e., MARS-AR(1)) are $RMSE = 0.081 \text{ m s}^{-1}$, $MAE = 0.061 \text{ m s}^{-1}$, $RRMSE = 2.220\%$, $R^2 = 0.998$, and $NSE = 0.998$, while the same statistics for the corresponding stand-alone model (i.e., MARS) are $RMSE = 1.279 \text{ m s}^{-1}$, $MAE = 0.936 \text{ m s}^{-1}$, $RRMSE = 35.094\%$, $R^2 = 0.441$, and $NSE = 0.422$. Similarly, the values of $RMSE$, MAE , $RRMSE$, R^2 , and NSE at the Zahedan station improved from 1.249 m s^{-1} , 0.892 m s^{-1} , 37.152% , 0.358 , and 0.357 in the stand-alone MARS to 0.045 m s^{-1} , 0.023 m s^{-1} , 1.330% , 0.999 , and 0.999 in the hybrid MARS-AR(1) during the testing period.

To graphically evaluate and compare the performance of the proposed hybrid models with the stand-alone machine learning models, the best-performing hybrid models for both daily and monthly scales (presented with the error statistics in bold in Tables 5 and 6) and the corresponding stand-alone machine learning models (i.e., MARS) were selected. The respective scatter and time series graphs were plotted, as shown in Figs. 3 and 4. The scatter plots show significant dispersion around the 1:1 line for the stand-alone MARS models, especially on a daily scale, while fewer dispersions are apparent for the hybrid models, indicating improved performance on the part of the hybrid models compared to the stand-alone models. The results of time series plots suggest the ability of the proposed hybrid models to capture the observed wind speed time series. In contrast, the stand-alone MARS model illustrates poor performance. In terms of estimating the high/peak and low daily and monthly wind speed values, the stand-alone MARS model tends to overestimate the low values and underestimate the peak values, particularly on a daily scale. However, the hybrid models offer acceptable performances in estimating the high and low values of the wind speed time series. This trend was also observed for all the other stand-alone and hybrid models.

Among the hybrid models, the RF-AR and MARS-AR generally performed the best on both the daily and monthly scales, specifically on a daily scale. The hybrid

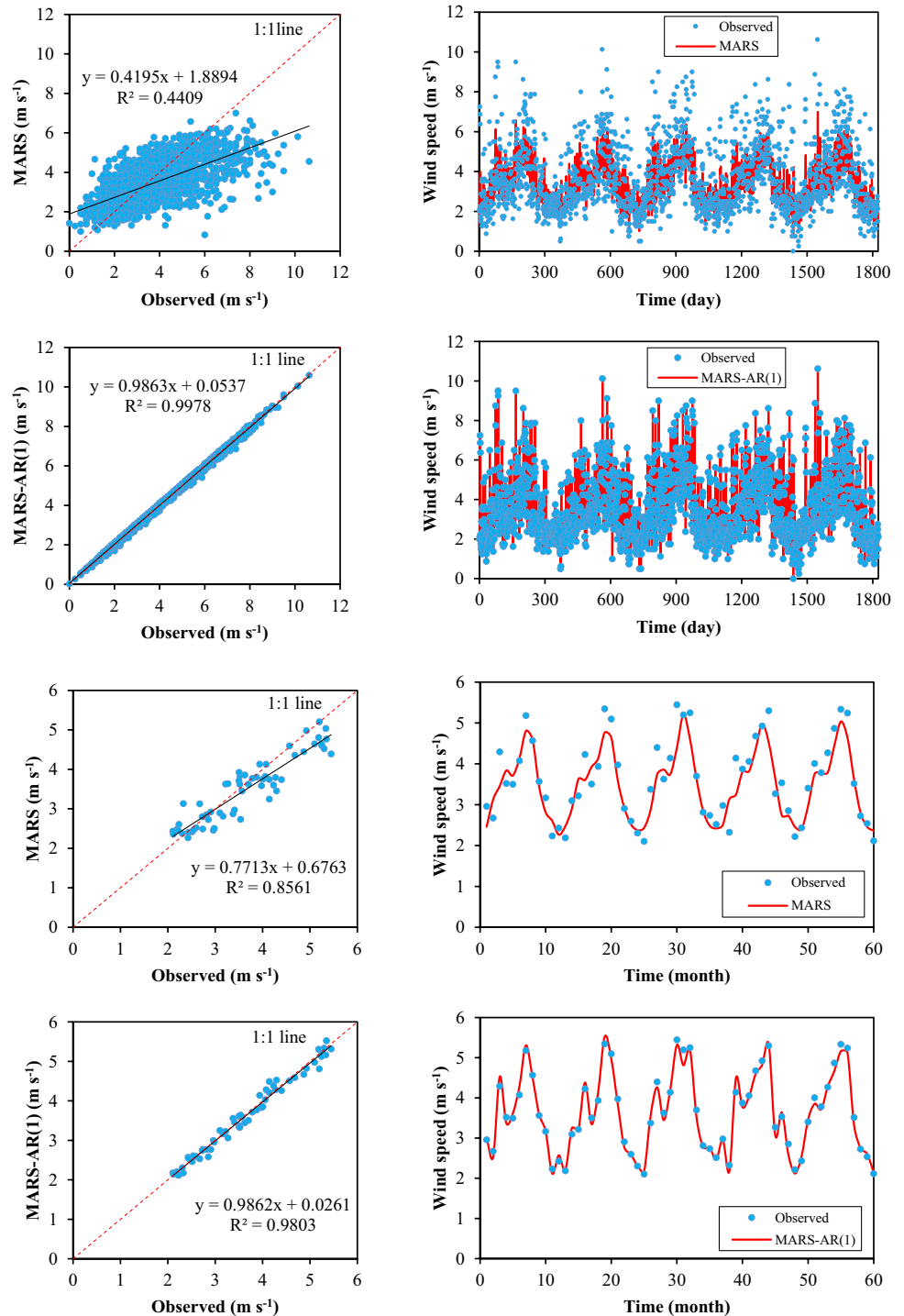
Table 5 Statistical results of the stand-alone machine learning models and hybrid machine learning-time series models implemented at the Tabriz station

Time scale	Models	Training					Testing				
		RMSE (m s ⁻¹)	MAE (m s ⁻¹)	RRMSE (%)	R ²	NSE	RMSE (m s ⁻¹)	MAE (m s ⁻¹)	RRMSE (%)	R ²	NSE
Daily	RF	1.240	0.940	39.568	0.411	0.410	1.273	0.933	34.940	0.448	0.427
	RF-AR(1)	0.130	0.095	4.159	0.994	0.993	0.132	0.093	3.614	0.994	0.994
	RF-MA(1)	0.258	0.193	8.243	0.979	0.974	0.222	0.167	6.099	0.985	0.983
	RF-ARMA(1,1)	0.641	0.475	20.448	0.844	0.842	0.684	0.487	18.767	0.840	0.835
	MARS	1.250	0.950	39.866	0.402	0.401	1.279	0.936	35.094	0.441	0.422
	MARS-AR(1)	0.075	0.057	2.395	0.998	0.998	0.081	0.061	2.220	0.998	0.998
Monthly	MARS-MA(1)	0.243	0.186	7.741	0.981	0.977	0.220	0.164	6.051	0.986	0.983
	MARS-ARMA(1,1)	0.621	0.462	19.798	0.853	0.852	0.658	0.472	18.054	0.851	0.847
	RF	0.353	0.273	11.279	0.863	0.861	0.403	0.324	11.082	0.866	0.835
	RF-AR(1)	0.186	0.148	5.929	0.971	0.962	0.175	0.135	4.823	0.969	0.969
	RF-MA(3)	0.329	0.238	10.499	0.933	0.879	0.166	0.139	4.562	0.988	0.972
	RF-ARMA(4,1)	0.157	0.122	5.013	0.974	0.972	0.181	0.132	4.966	0.973	0.967
	MARS	0.369	0.284	11.789	0.851	0.848	0.418	0.333	11.476	0.856	0.823
	MARS-AR(1)	0.176	0.142	5.624	0.971	0.965	0.141	0.116	3.887	0.980	0.980
	MARS-MA(3)	0.306	0.217	9.766	0.937	0.895	0.148	0.125	4.067	0.991	0.978
	MARS-ARMA(4,1)	0.166	0.131	5.287	0.971	0.969	0.153	0.122	4.203	0.984	0.976

Table 6 Statistical results of the stand-alone machine learning models and hybrid machine learning-time series models implemented at the Zahedan station

Time scale	Models	Training					Testing				
		RMSE (m s ⁻¹)	MAE (m s ⁻¹)	RRMSE (%)	R ²	NSE	RMSE (m s ⁻¹)	MAE (m s ⁻¹)	RRMSE (%)	R ²	NSE
Daily	RF	1.515	1.113	47.188	0.344	0.343	1.240	0.885	36.890	0.368	0.366
	RF-AR(1)	0.129	0.096	4.009	0.995	0.995	0.128	0.096	3.798	0.994	0.993
	RF-MA(1)	0.349	0.258	10.864	0.968	0.965	0.304	0.222	9.049	0.967	0.962
	RF-ARMA(1,1)	0.142	0.109	4.412	0.995	0.994	0.141	0.108	4.208	0.993	0.992
	MARS	1.526	1.121	47.505	0.335	0.334	1.249	0.892	37.152	0.358	0.357
	MARS-AR(1)	0.052	0.027	1.630	0.999	0.999	0.045	0.023	1.330	0.999	0.999
	MARS-MA(1)	0.334	0.246	10.391	0.971	0.968	0.276	0.195	8.216	0.972	0.969
	MARS-ARMA(1,1)	0.086	0.067	2.668	0.998	0.998	0.073	0.055	2.186	0.998	0.998
Monthly	RF	0.487	0.381	15.143	0.643	0.643	0.467	0.373	13.865	0.585	0.570
	RF-AR(1)	0.125	0.094	3.893	0.978	0.976	0.116	0.078	3.437	0.976	0.974
	RF-MA(1)	0.142	0.109	4.423	0.975	0.970	0.113	0.076	3.352	0.976	0.975
	RF-ARMA(1,1)	0.301	0.213	9.349	0.872	0.864	0.320	0.231	9.486	0.837	0.799
	MARS	0.511	0.397	15.892	0.608	0.607	0.466	0.378	13.841	0.583	0.572
	MARS-AR(1)	0.122	0.095	3.800	0.978	0.978	0.119	0.087	3.535	0.974	0.972
	MARS-MA(1)	0.139	0.107	4.329	0.974	0.971	0.112	0.080	3.311	0.977	0.975
	MARS-ARMA(1,1)	0.296	0.213	9.214	0.873	0.868	0.319	0.233	9.460	0.835	0.800

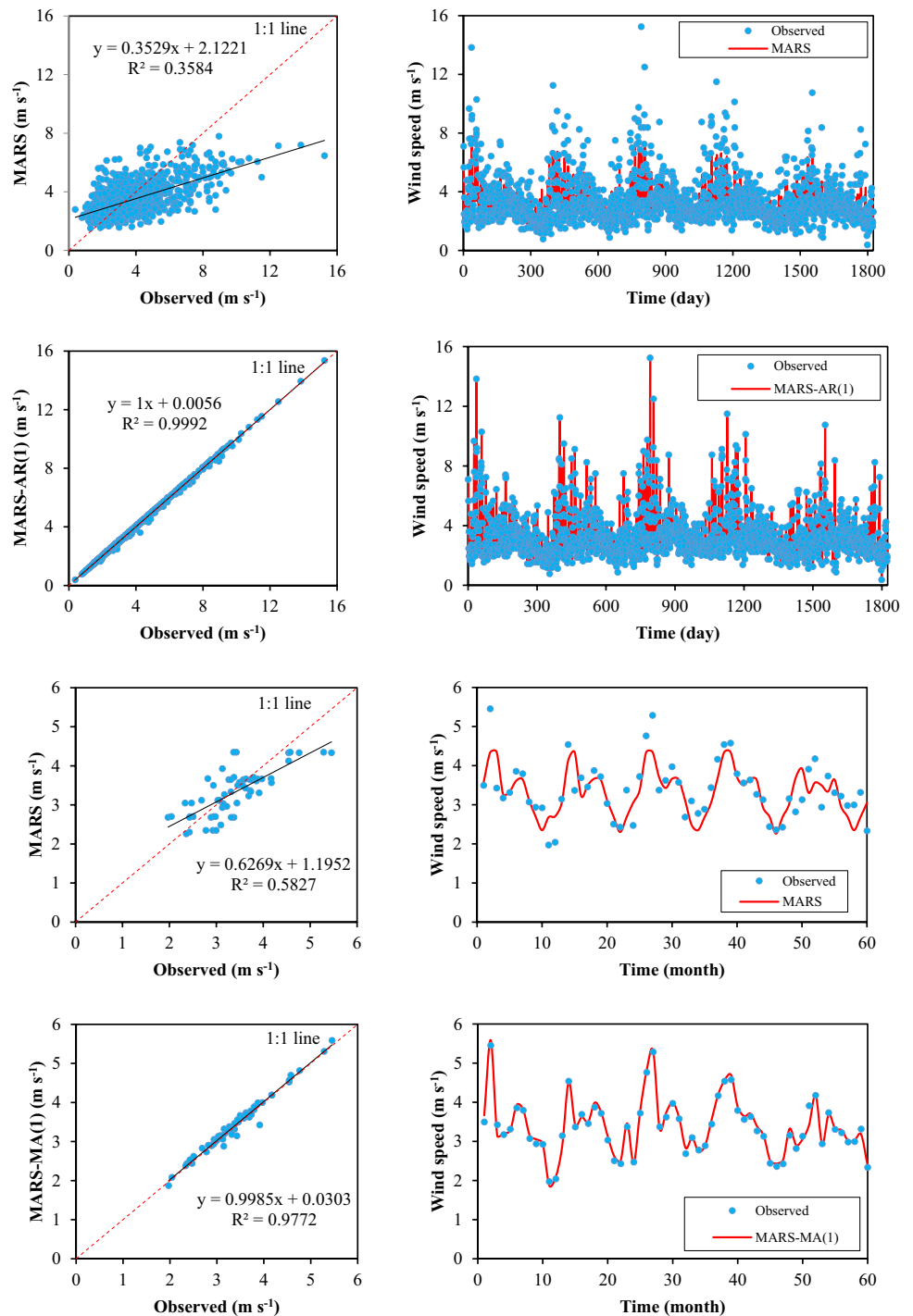
Fig. 3 Scatter and time series plots of the observed daily and monthly wind speeds via the best hybrid models and the corresponding stand-alone machine learning models at Tabriz station



models created through coupling the MARS and time series models (i.e., MARS-AR, MARS-MA, and MARS-ARMA) led to more accurate estimates of the daily and monthly wind speed in comparison with the hybrid RF-AR, RF-MA, and RF-ARMA models. Improved performance on the part of the hybrid models was observed for the Tabriz station on the monthly time scale, except for the RF-MA(3) and MARS-MA(3) models during the

training period, where the performance of MA(3) is better than the mentioned hybrid models. At the Zahedan station on a monthly scale, the AR(1) and MA(1) models were found to be superior to the hybrid RF-ARMA(1,1) and MARS-ARMA(1,1) models.

Fig. 4 Scatter and time series plots of the observed daily and monthly wind speeds via the best hybrid models and the corresponding stand-alone machine learning models at Zahedan station



4 Discussions

Wind speed depicts one of the most plentiful sources of renewable energy. Moreover, wind speed is clearly a vital factor in the performance of buildings and structures from an engineering point of view. To this end, the short-term and long-term wind speeds for two stations

of Tabriz and Zahedan located at northwest and south-east of Iran, respectively, are investigated.

As concluded, the hybrid models created through hybridizing the machine learning and time series models performed better than the stand-alone ones in estimating wind speeds. It can be justified considering the fact that the hybrid models employ both the stochastic and deterministic terms of the wind speed data in the modeling

procedure, which could be considered as an advantage of the hybrid models. Contrary, the stand-alone machine learning and time series models focus only on modeling the deterministic and stochastic terms, respectively. Indeed, it is a major drawback of the stand-alone models.

The results achieved in the present work (i.e., a much better performance of the hybrid machine learning-time series models compared to the stand-alone machine learning models) are in agreement with the outcomes of previous works [11, 12, 39–44]. The different hybrid models were developed by the authors via coupling the various machine learning and time series models to estimate the hydrological and meteorological data and reported the higher accuracies of hybrid models than the stand-alone machine learning models.

In addition, various hybrid models have been implemented in the literature [13, 24, 45–48] to improve the performance of wind speed modeling by coupling the machine learning and time series models. In these studies, various types of the time series models including the linear and nonlinear ones were hybridized with the machine learning models. The outcomes have verified the superiority of hybrid models compared to the stand-alone machine learning ones, which support the results achieved in the present work.

5 Conclusions

The current research explores the estimation of daily and monthly wind speed data at two stations in Iran, namely the Tabriz and Zahedan stations. To this end, three stand-alone time series models, including AR, MA, ARMA, and two machine learning models consisting of RF and MARS, were developed and applied to estimate daily and monthly wind speed data at the studied stations. The error metrics, RMSE, MAE, RRMSE, R^2 , and NSE, were utilized to compare the performance of all the developed models. Of the stand-alone time series models, the MA models performed better than the AR and ARMA models on both the daily and monthly time scales. For the case of machine learning models, the RF and MARS models performed at a similar level, yielding poor results when applied to daily data (RRMSE > 30%), and good results ($10\% < \text{RRMSE} < 20\%$) when applied to monthly data.

To improve wind speed estimation accuracy, hybrid models were developed by coupling AR, MA, and ARMA with the RF and MARS modeling approaches. The results showed a much better performance for the hybrid machine learning-time series models compared to the stand-alone machine learning models. Of the six hybrid models (i.e., RF-AR, RF-MA, RF-ARMA, MARS-AR, MARS-MA, and MARS-ARMA), the RF-AR and MARS-AR models

outperformed the others in terms of estimation accuracy. Generally, the performance of the hybrid models developed in the current study ranked as excellent according to the RRMSE classification criterion, for both daily and monthly time scales.

The present study explored the performance of six types of hybrid models developed through coupling the linear AR, MA, and ARMA models with the RF and MARS machine learning models. Further investigation into alternative hybrid models is suggested, through integrating diverse time series models, including the nonlinear autoregressive conditional heteroscedasticity (ARCH) model and self-exciting threshold autoregressive (SETAR) model, with machine learning models such as ANN, ANFIS, gene expression programming (GEP), convolutional neural networks, and deep learning. In addition, the authors recommend that the methodology presented in the current study could be employed to estimate solar radiation time series, toward a deeper understanding of another source of renewable and sustainable energy.

Compliance with ethical standards

Conflict of interest The authors declare that they have no conflict of interest.

References

1. Zhang Z, Qin H, Liu Y, Yao L, Yu X, Lu J, Jiang Z, Feng Z (2019) Wind speed forecasting based on quantile regression minimal gated memory network and Kernel density estimation. *Energy Convers Manag* 196:1395–1409. <https://doi.org/10.1016/j.enconman.2019.06.024>
2. Zhang Z, Qin H, Liu Y, Wang Y, Yao L, Li Q, Li J, Pei S (2019) Long short-term memory network based on neighborhood gates for processing complex causality in wind speed prediction. *Energy Convers Manag* 192:37–51. <https://doi.org/10.1016/j.enconman.2019.04.006>
3. Georgilakis PS (2008) Technical challenges associated with the integration of wind power into power systems. *Renew Sustain Energy Rev* 12(3):852–863. <https://doi.org/10.1016/j.rser.2006.10.007>
4. Li R, Jin Y (2018) A wind speed interval prediction system based on multi-objective optimization for machine learning method. *Appl Energy* 228:2207–2220. <https://doi.org/10.1016/j.apenergy.2018.07.032>
5. Duan Z, Liu H (2019) An evolution-dependent multi-objective ensemble model of vanishing moment with adversarial auto-encoder for short-term wind speed forecasting in Xinjiang wind farm, China. *Energy Convers Manage* 198:111914. <https://doi.org/10.1016/j.enconman.2019.111914>
6. Jiang P, Li R, Li H (2019) Multi-objective algorithm for the design of prediction intervals for wind power forecasting model. *Appl Math Model* 67:101–122. <https://doi.org/10.1016/j.apm.2018.10.019>
7. Asim Singh SN, Mohapatra A (2019) Repeated wavelet transform based ARIMA model for very short-term wind speed forecasting.

- Renew Energy 136:758–768. <https://doi.org/10.1016/j.renene.2019.01.031>
8. Jiang P, Liu Z (2019) Variable weights combined model based on multi-objective optimization for short-term wind speed forecasting. *Appl Soft Comput* 82:105587. <https://doi.org/10.1016/j.asoc.2019.105587>
 9. Yuan X, Tan Q, Lei X, Yuan Y, Wu X (2017) Wind power prediction using hybrid autoregressive fractionally integrated moving average and least square support vector machine. *Energy* 129:122–137. <https://doi.org/10.1016/j.energy.2017.04.094>
 10. Allen DJ, Tomlin AS, Bale CSE, Skea A, Vosper S, Gallani ML (2017) A boundary layer scaling technique for estimating near-surface wind energy using numerical weather prediction and wind map data. *Appl Energy* 208:1246–1257. <https://doi.org/10.1016/j.apenergy.2017.09.029>
 11. Mehdizadeh S, Behmanesh J, Khalili K (2017) A comparison of monthly precipitation point estimates at 6 locations in Iran using integration of soft computing methods and GARCH time series model. *J Hydrol* 554:721–742. <https://doi.org/10.1016/j.jhydrol.2017.09.056>
 12. Mehdizadeh S, Fathian F, Adamowski JF (2019) Novel hybrid artificial intelligence-time series models for monthly streamflow modeling. *Appl Soft Comput* 80:873–887. <https://doi.org/10.1016/j.asoc.2019.03.046>
 13. Liu H, Tian HQ, Li YF (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>
 14. Li G, Shi J (2010) On comparing three artificial neural networks for wind speed forecasting. *Appl Energy* 87(7):2313–2320. <https://doi.org/10.1016/j.apenergy.2009.12.013>
 15. Hu YL, Chen L (2018) A nonlinear hybrid wind speed forecasting model using LSTM network, hysteretic ELM and differential evolution algorithm. *Energy Convers Manag* 173:123–142. <https://doi.org/10.1016/j.enconman.2018.07.070>
 16. Liu D, Niu D, Wang H, Fan L (2014) Short-term wind speed forecasting using wavelet transform and support vector machines optimized by genetic algorithm. *Renew Energy* 62:592–597. <https://doi.org/10.1016/j.renene.2013.08.011>
 17. Kong X, Liu X, Shi R, Li KW (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>
 18. Jiang P, Wang Y, Wang J (2017) Short-term wind speed forecasting using a hybrid model. *Energy* 119:561–577. <https://doi.org/10.1016/j.energy.2016.10.040>
 19. Khosravi A, Machado L, Nunes RO (2018) Time-series prediction of wind speed using machine learning algorithms: a case study Osorio wind farm, Brazil. *Appl Energy* 224:550–566. <https://doi.org/10.1016/j.apenergy.2018.05.043>
 20. Liu H, Mi X, Li Y (2018) An experimental investigation of three new hybrid wind speed forecasting models using multi-decomposing strategy and ELM algorithm. *Renew Energy* 123:694–705. <https://doi.org/10.1016/j.renene.2018.02.092>
 21. Nikolic V, Motamedi S, Shamshirband S, Petković D, Ch S, Arif M (2016) Extreme learning machine approach for sensorless wind speed estimation. *Mechatronics* 34:78–83. <https://doi.org/10.1016/j.mechatronics.2015.04.007>
 22. Mohandes M, Rehman S, Rahman SM (2011) Estimation of wind speed profile using adaptive neuro-fuzzy inference system (ANFIS). *Appl Energy* 88(11):4024–4032. <https://doi.org/10.1016/j.apenergy.2011.04.015>
 23. Meng A, Ge J, Yin H, Chen S (2016) Wind speed forecasting based on wavelet packet decomposition and artificial neural networks trained by crisscross optimization algorithm. *Energy Convers Manag* 114:75–88. <https://doi.org/10.1016/j.enconman.2016.02.013>
 24. Shukur OB, Lee MH (2015) Daily wind speed forecasting through hybrid KF-ANN model based on ARIMA. *Renew Energy* 76:637–647. <https://doi.org/10.1016/j.renene.2014.11.084>
 25. Torres JL, Garcia A, De Blas M, De Francisco A (2005) Forecast of hourly average wind speed with ARMA models in Navarre (Spain). *Sol Energy* 79(1):65–77. <https://doi.org/10.1016/j.solener.2004.09.013>
 26. Bivona S, Bonanno G, Burlon R, Gurrera D, Leone C (2011) Stochastic models for wind speed forecasting. *Energy Convers Manag* 52(2):1157–1165. <https://doi.org/10.1016/j.enconman.2010.09.010>
 27. Wu C, Wang J, Chen X, Du P, Yang W (2020) A novel hybrid system based on multi-objective optimization for wind speed forecasting. *Renew Energy* 146:149–165. <https://doi.org/10.1016/j.renene.2019.04.157>
 28. Bilgili M, Sahin B (2010) Comparative analysis of regression and artificial neural network models for wind speed prediction. *Meteorol Atmos Phys* 109:61–72. <https://doi.org/10.1007/s00703-010-0093-9>
 29. Ghorbani MA, Khatibi R, Fazelifard MH, Naghipour L, Makarynsky O (2016) Short-term wind speed predictions with machine learning techniques. *Meteorol Atmos Phys* 128:57–72. <https://doi.org/10.1007/s00703-015-0398-9>
 30. Ewing BT, Kruse JB, Schroeder JL (2006) Time series analysis of wind speed with time-varying turbulence. *Environmetrics* 17:119–127. <https://doi.org/10.1002/env.754>
 31. Thomas HA, Fiering MB (1962) Mathematical synthesis of stream flow sequences for the analysis of river basin by simulation. Harvard University Press, Cambridge, p 751 p
 32. Box GE, Jenkins GM (1976) Time series analysis. Forecasting and control. Holden-Day, San Francisco
 33. Breiman L (2001) Random forests. *Mach Learn* 45:5–32
 34. Breiman L, Friedman J, Stone CJ, Olshen RA (1984) Classification and regression trees. Chapman & Hall/CRC, New York
 35. Casella G, Fienberg S, Olkin I (2006) An introduction to statistical learning. In: James G, Witten D, Hastie T, Tibshirani R (eds) Design, vol 102. Springer, New York
 36. Breiman L (1996) Bagging predictors. *Mach Learn* 24:123–140
 37. Friedman JH (1991) Multivariate adaptive regression splines. *Ann Stat* 19(1):1–67
 38. Mehdizadeh S (2018) Assessing the potential of data-driven models for estimation of long-term monthly temperatures. *Comput Electron Agric* 144:114–125. <https://doi.org/10.1016/j.compag.2017.11.038>
 39. Mehdizadeh S (2018) Estimation of daily reference evapotranspiration (ET₀) using artificial intelligence methods: offering a new approach for lagged ET₀ data-based modeling. *J Hydrol* 559:794–812. <https://doi.org/10.1016/j.jhydrol.2018.02.060>
 40. Mehdizadeh S, Behmanesh J, Khalili K (2018) New approaches for estimation of monthly rainfall based on GEP-ARCH and ANN-ARCH hybrid models. *Water Resour Manag* 32(2):527–545. <https://doi.org/10.1007/s11269-017-1825-0>
 41. Mehdizadeh S, Kozekalani Sales A (2018) A comparative study of autoregressive, autoregressive moving average, gene expression programming and Bayesian networks for estimating monthly streamflow. *Water Resour Manag* 32(9):3001–3022. <https://doi.org/10.1007/s11269-018-1970-0>
 42. Fathian F, Mehdizadeh S, Kozekalani Sales A, Safari MJS (2019) Hybrid models to improve the monthly river flow prediction: integrating artificial intelligence and non-linear time series model. *J Hydrol* 575:1200–1213. <https://doi.org/10.1016/j.jhydrol.2019.06.025>

43. Mehdizadeh S, Fathian F, Safari MJS, Khosravi A (2020) Developing novel hybrid models for estimation of daily soil temperature at various depths. *Soil Tillage Res* 197:104513. <https://doi.org/10.1016/j.still.2019.104513>
44. Mehdizadeh S (2020) Using AR, MA, and ARMA time series models to improve the performance of MARS and KNN approaches in monthly precipitation modeling under limited climatic data. *Water Resour Manag* 34(1):263–282. <https://doi.org/10.1007/s11269-019-02442-1>
45. Cadenas E, Rivera W (2010) Wind speed forecasting in three different regions of Mexico, using a hybrid ARIMA–ANN model. *Renew Energy* 35(12):2732–2738. <https://doi.org/10.1016/j.renene.2010.04.022>
46. Han Q, Meng F, Hu T, Chu F (2017) Non-parametric hybrid models for wind speed forecasting. *Energy Convers Manag* 148:554–568. <https://doi.org/10.1016/j.enconman.2017.06.021>
47. Jiang Y, Huang G, Peng X, Li Y, Yang Q (2018) A novel wind speed prediction method: hybrid of correlation-aided DWT, LSSVM and GARCH. *J Wind Eng Ind Aerodyn* 174:28–38. <https://doi.org/10.1016/j.jweia.2017.12.019>
48. Zhang Y, Zhao Y, Kong C, Chen B (2020) A new prediction method based on VMD-PRBF-ARMA-E model considering wind speed characteristic. *Energy Convers Manag* 203:112254. <https://doi.org/10.1016/j.enconman.2019.112254>

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