A new hybridization filtering-based linear-nonlinear models for time series forecasting

Mehrnaz Ahmadi and Mehdi Khashei*

Department of Industrial and Systems Engineering, Isfahan University of Technology (IUT), Isfahan, Iran.

*Corresponding Author: Khashei@cc.iut.ac.ir

ABSTRACT

The prediction is one of the most influential factors in management and efficient utilization in various sciences as well as economic planning. Since there is a direct relationship between the accuracy of predictions and the quality of the decisions made, today, despite the numerous prediction methods and the achievement of accurate predictions, most researchers still try to combine different methods in order to obtain more accurate results. In order to improve the accuracy of predictions, the present study introduces a new parallel hybrid methodology based on trend-residual data preprocessing to time series predict. In the proposed model, different patterns and structures of trend and residual as well as linear and nonlinear, are simultaneously modeled. In the first stage of the proposed method, the data is analyzed by the Kalman filter (KF) method and divided into two groups of trend and residual patterns. Then, the trend patterns from the previous step, with the original data, are simultaneously considered as input of the autoregressive integrated moving average with explanatory variable (ARIMAX) and multilayer perceptron (MLP), as linear and nonlinear forecasting models, respectively. Then, this step is repeated for residual patterns. In this way, the proposed model can model four types of patterns, including linear-trend, linear-residual, nonlinear-trend, and nonlinear-residual. Finally, the results of these patterns, along with the Kalman trend, are combined together in a parallel hybridization process, and final forecasts of the proposed model are obtained. Numerical results of 2 wind power and speed benchmarks indicate that the proposed model can approximately improve 41.73% and 42.14% the performance of its single linear and nonlinear components, respectively. Furthermore, the proposed model can yield more accurate results than traditional series-based components combination hybrid models, parallel-based components combination hybrid models,

Pre-processing-based linear models, and preprocessing-based nonlinear models by the same components. The proposed method can roughly improve 16.35%, 24.39%, 36.09%, and 14.98% the performance of these hybrid models, respectively. Therefore, the proposed model may be a suitable alternative for single as well as hybrid models for wind power forecasting, especially when more accurate results and/or more quality decisions are required.

Keywords: Hybrid structures, Trend-residual preprocessing, Parallel hybridization, Autoregressive Integrated Moving Average with Explanatory Variable (ARIMAX), Multi-Layer Perceptrons (MLPs), Kalman filter.

INTRODUCTION

In recent years, the idea of using a mathematical model to describe the behavior of physical phenomena has been very much considered. Specifically, a definitive model, based on physical laws, enables researchers to calculate the number of time dependencies precisely at any moment in time. However, in the real world, we often face time-dependent phenomena with many unknown or unavailable factors (Lindley, 2010; Roulston et al., 2003). In this case, when it is not possible to achieve a definite model, the prediction methods are wide-used, especially when the past

observations of a variable and primary relationships between specific observations are available. Forecasting methods that are used in different fields of science can be categorized based on various aspects. For example, the prediction methods used in the field of wind energy can be divided into four categories of 1) ultra short term (several seconds to four hours), 2) short term (4 to 24 hours), 3) medium-term (1 to 7 days), and 4) long term (more than 7 days) (Zack, 2003; Soman et al., 2010). Also, the structure of forecasting methods can be divided into two types of 1) single methods and 2) hybrid methods. Each of these categories can also be subdivided into smaller subgroups.

For example, single methods can be divided into three subclasses of 1) physical methods, 2) statistical methods, and 3) intelligent methods; and hybrid methods can be divided into four subcategories of 1) data preprocessing based approaches, 2) parameter optimization-based approaches, 3) postprocessing based approaches, and 4) component-based in series or parallel approaches. The physical methods are based on the numerical weather prediction models and using some data such as surface roughness, orography, obstacles, pressure, and temperature to estimate the local wind power, speed, and direction (Lange et al., 2008). The physical approaches use more accurate physical descriptions for modeling in comparison to statistical approaches (Kariniotakis et al., 2004; Lange et al., 2006). Marjanovic *et al.* (Marjanovic et al. 2014) have introduced a weather research and forecasting (WRF) model for 48h-ahead wind power forecasting. The simulation results show a significant improvement. Traditional statistical methods such as quantile regression models (Wang et al., 2019; Nathaniel et al., 2018; Wang et al., 2018; Lahouar et al., 2017; Zhang et al., 2015; Bessa et al., 2012), time series models (Jakob et al., 2018; Robles–Rodriguez & Dochain, 2018; T. Filik & U. Filik, 2017; Ziel et al., 2016; Lydia et al., 2016; Guo et al., 2014), grey models (Kou et al., 2014; Kou et al., 2013; Ramirez–Rosado et al., 2009), stochastic differential equations models (Xydas et al., 2017; Iversen et al., 2016), and Markov regime-switching models (Carpinone et al., 2015; Song et al., 2014; D'Amicoa et al., 2014) are among the most important and more popular methods, which extensively used to wind power and speed forecasting.

The SARIMA model for 24 hours of wind speed forecasting is provided (Bivona et al., 2011). In this study, wind speed time series in two regions of Italy are used to test the proposed model, whose numerical results show the efficiency of the SARIMA model. The Hammerstein Auto-Regressive model (HAR) for 1-24 hours wind speed forecasting is used (Maatallah et al., 2015). The proposed model is applied to real data from two different sites. The results show that the proposed HAR model is better than ARIMA and MLP models in terms of different indicators such as RMSE, MAE, and MAPE. Despite all the statistical model's advantages, they have several disadvantages and limitations, such as the inability to model nonlinear, complex, and multiple structures. These limitations encourage researchers to develop intelligent methods to address the defects of statistical models. Many researchers have developed intelligent methods due to their unique features in modeling complex nonlinear patterns in underlying data. Intelligence approaches, such as artificial neural networks (Wang et al., 2018; Yu et al., 2018; Ahmed & Khalid, 2018; He & Li, 2017; Wang et al., 2017) fuzzy sets and systems (Dong et al., 2017; Taslimi Renani et al., 2016), and support vector machines (Hu et al., 2014; Yang & Zhao, 2012), have been developed and frequently applied for wind power, speed, and direction forecasting.

The LRNN model for short-term wind power and speed forecasting is used (Olaofe et al., 2014). Empirical results indicate that the proposed model can yield more accurate results than others. The NESN-P and NESN-MP models for medium-term wind speed and direction forecasting are used (Chitsazan et al., 2019). To demonstrate the efficiency of the proposed methods, they have compared with ESN and adaptive neuro-fuzzy inference system (ANFIS). Intelligent methods are a universal approximation to modeling and analyzing systems with appropriate accuracy to lifting the linear limitation of statistical approaches. However, many studies indicate that artificial neural networks may be incompatible and inaccurate in some specific situations (Khashei & Bijari, 2010). Generally, both statistical and intelligent single models, despite their unique features and specifications, have a critical limitation to modeling and providing accurate results.

In other words, by using a single model, a specific part of the patterns and relationships in the raw data is modeled. Therefore, using a single model may increase the risk of using an inappropriate model for modeling (Khashei & Bijari, 2011). Recently, researchers have more focused on hybrid models in order to yield more comprehensiveness and consequently achieve more accuracy and low risk in time-series forecasting. In other words, the main idea of hybrid methods is to take advantage of single models simultaneously for more comprehensive, accurate, and less risk modeling. Hybrid methods can be categorized into different categories the 1) data preprocessing based hybrid models (DAH), 2) parameter optimization based hybrid models (PAH), 3) postprocessing based hybrid models (POH), and 4) components combination based hybrid models (CCH). Data preprocessing-based hybrid models decompose wind

data into smooth and regular parts that are easily identifiable, and then each decomposed segment provides an appropriate prediction before combining with the predicted models. A new data preprocessing-based hybrid technique based on VMD, and MKRR model for 10 min, 30 min, and one-hour wind power and speed forecasting is used (Naik et al., 2018). In this paper, the original wind speed and power data have been decomposed by VMD decomposed model; then, decomposed patterns have been used as input for the MKRR model. Parameter optimization-based hybrid models, as well as postprocessing-based hybrid models, have a significant contribution in increasing the accuracy of wind forecasts.

A new parameter optimization-based hybrid technique, HAP-ACO-PSO, for ultra-short-term wind power forecasting is used (Rahmani et al., 2013). The hourly wind power of the Binaloud wind farm has been collected and used to train and test the developed model. A postprocessing-based hybrid model with an error feedback scheme (IRBFNN-EF) for short-term wind power and speed forecasting is provided (Chang et al., 2017). In this study, a selected day in 2014 is used to test the proposed model and four other compared models. Another widely-used hybrid approaches are models that combine two or more single models or components in series or parallel structures to achieve higher accuracy in prediction (Khashei et al., 2013). A series-based hybrid model, based on integrating machine learning techniques and physical knowledge modeling (grey-box) for short-term wind speed forecasting is provided (Vaccaro et al., 2012). A parallel-based hybrid model, the KF-ANN model for daily wind speed forecasting is proposed (Shukur & Lee, 2015). In this study, daily wind speed data from Iraq and Malaysia is used. Some other hybrid methods for wind power and speed forecasting are stated in Table (1).

As mentioned previously, hybrid structures can be categorized into different categories. The literature demonstrates that the data preprocessing-based hybrid models are the most popular and widely-used method among other hybrid models (Liu & Chen, 2019). Because after the preprocessing techniques, the high-frequency sub-signals are extracted in order to smooth the raw data, so the prediction accuracy increases. However, the difference of the preprocessing-based hybrid model in this article with other data preprocessing-based hybrid models in the literature is that both trend and residual patterns are simultaneously used. On the other hand, the main idea of the proposed hybrid model is that the residual data that has been considered as noise in other hybrid models and not used in the modeling is considered as input data. The reason is that it may enclose beneficial information and patterns for modeling; so, by eliminating this information, the accuracy of modeling may be reduced.

Also, a parallel structure is a comprehensive modeling and can produce more accurate and low-risk forecasting models. The literature shows that the combination of components in the parallel form is generally more usual and more accurate than the combination in series (Hajirahimi & Khashei, 2019). Accordingly, the key goal of this paper is to propose a hybrid approach that can simultaneously cover linear/nonlinear, trend/residual, and different hybridization structures; and also can yield accurate forecasts. For this reason, in the proposed model, the underlying raw data sets are considered to be composed of trend and residual patterns that can be decomposed by the Kalman filter.

Table (1). Some recent proposed hybrid models for wind power and speed forecasting.

[Ref.]	Year	Domain	Time-scale	Applied Model(s)	Technique(s)
Jiang et al.	2019	Wind power	Ultra short term	PICP, LSSVM	PAH
Chen et al.	2019	Wind speed	Ultra short term	QPSO, LSSVM	PAH
Zhang et al.	2019	Wind speed	Ultra short term	OSO, RELM-C	PAH
Jiang et al.	2019	Wind speed	Ultra short term	PSO, MODE-FTS	PAH
Jiang et al.	2018	Wind speed	Ultra short term	CS, V-SVM	PAH
Liu et al.	2018	Wind speed	Ultra short term	GWO, RELM	PAH
Liu et al.	2018	Wind speed	Ultra short term	WF, BFGS	POH
Wang et al.	2018	Wind speed	Ultra short term	ARIMA, ELM	POH
Li et al.	2018	Wind speed	Ultra short term	RELM, LSTM	POH
Jiang et al.	2018	Wind speed	Medium-term	NWP, GPR	CCH
Akçay & Filik	2018	Wind speed	Ultra short term	LSSVM, GARCH	CCH
Liu et al.	2018	Wind power	Short term	ARIMAX, PR	CCH
Hu & Chen	2018	Wind speed	Ultra short term	LSTM, ELM	CCH
Yu et al.	2018	Wind speed	Ultra short term	LSTMDE, HELM	CCH
Jiang & Li	2018	Wind speed	Ultra short term	RNN, SVM, LSTM	CCH

Sun et al.	201 7	Wind speed	Short term	KPCA, CVR, COR	DAH
Wang et al.	201 7	Wind speed	Ultra short term	CEEMD, ENN	DAH
Wang et al.	201 7	Wind speed	Ultra short term	VMD, GAWNN	DAH
Ma et al.	201 7	Wind speed	Ultra short term	SSA, GDFNN	DAH
Chang et al.	201 7	Wind Power	short term	EF, IRBFNN	POH
Cassol a & Burlando	201 7	Wind speed	short term	ARMA, KF	CCH
Aghaja ni et al.	201 6	Wind power	Short term	WT, HNN	DAH
Azimi et al.	201 6	Wind power	short term	DWT, HANTS, MLPNN	DAH
Shao & Deng	201 6	Wind power	Ultra short term	MADF, LLE	DAH
Lynch et al.	201 4	Wind Speed	Ultra short term	KF,YR. NO NWP	POH
Zhao et al.	201 2	Wind speed	Long term	NWP, KF	CCH

THE PROPOSED PARALLEL TREND-RESIDUAL (PTR) HYBRID MODEL

The prediction of time series based on combined modeling has been one of the most important research areas in recent years. Hybrid models have been used to improve forecasting accuracy by combining the benefits of individual forecasting models, as well as omitting their disadvantages in forecasting. Combined models in the subject literature, basically, include the combination of different prediction models under series or parallel structures and can also be the combination of various prediction models with preprocessing and/or postprocessing techniques. The data preprocessing method is the widely-used approach in modeling. The reason for the extensive applications of the preprocessing methods is that the raw data is analyzed for irrelevant and redundant information, noisy and unreliable data, outlier data, etc., before entering into the forecasting process. Because of this, the use of preprocessing techniques can increase the accuracy of predictions. But in the subject literature, residual data in the preprocessing processes is completely excluded and is considered as noise in the modeling and forecasting procedure. Consequently, the main idea of our proposed hybrid model is to use 1) the residual data in the proposed model, 2) different hybridization structures, 3) linear and nonlinear modeling in order to increase the comprehensiveness of modeling and prediction accuracy. In general, the difference between our proposed model and the subject literature can be summarized as follows:

1. In classical structures based on the data preprocessing technique, only the trend data from the data preprocessing process is entered into the prediction model. However, in the proposed method, trend data and residual data are simultaneously considered as input variables in the prediction procedure. Accordingly, if

the data obtained from the studied system involves dynamic multiple-trend structures, the proposed model can better model them.

2. In the classical structures in the subject literature, raw data, or the only data from the data preprocessing process, is used in the prediction model. However, in the proposed method, the raw data, trend data, and residual data and their lags are simultaneously entered into the prediction model. Accordingly, if the preprocessed data of the studied system still contains specific modelable patterns, the proposed model can model more patterns.

3. In the classical structures based on the data preprocessing, the preprocessed data do not enter the final model separately. However, in the proposed method, the preprocessed data are also used as input of the prediction model.

4. In the classical hybrid approaches, only the parallel combination of different forecasting models with the same inputs, in general, are used. However, in the proposed method, in addition to using three different components, including the Kalman filter, the multilayer perceptron, and autoregressive integrated moving average with explanatory variable models, the similar models with different inputs are also combined in the parallel form. By doing so, more emphasis can be placed on the fundamental patterns and relationships in the modeling process.

In this way, considering the autoregressive integrated moving average with explanatory variable, the multilayer perceptron as prediction models, and the Kalman filter as a data preprocessing method, the overall structure of the proposed method is presented in Fig. (1). The steps of the proposed method can be summarized as follows:

Stage I: *Data Preprocessing technique:* In the first stage of the proposed method, the raw data is decomposed by the Kalman filter technique. The Kalman filter could be described as an approach consisting of two decomposing stages. The decomposing patterns are given by Eq. (1).

$$Y_t = Y_t^T + Y_t^R \tag{1}$$

where, Y_t is the output of the Kalman filter, Y_t^T is the decomposing trend patterns, and Y_t^R is the residual decomposing patterns.

Stage II: *Linear Trend/Residual Modeling:* In the second stage of the proposed model, at first, the raw data and decomposed trend patterns from the previous stage are included in the autoregressive integrated moving average with explanatory variable (ARIMAX) model in order to model the linear trend correlation structures in the underlying data sets. Then, the raw data and decomposed residual data are used in the ARIMAX model in order to model the remained modelable linear correlation structures in the underlying data that are left as noise in the residuals.
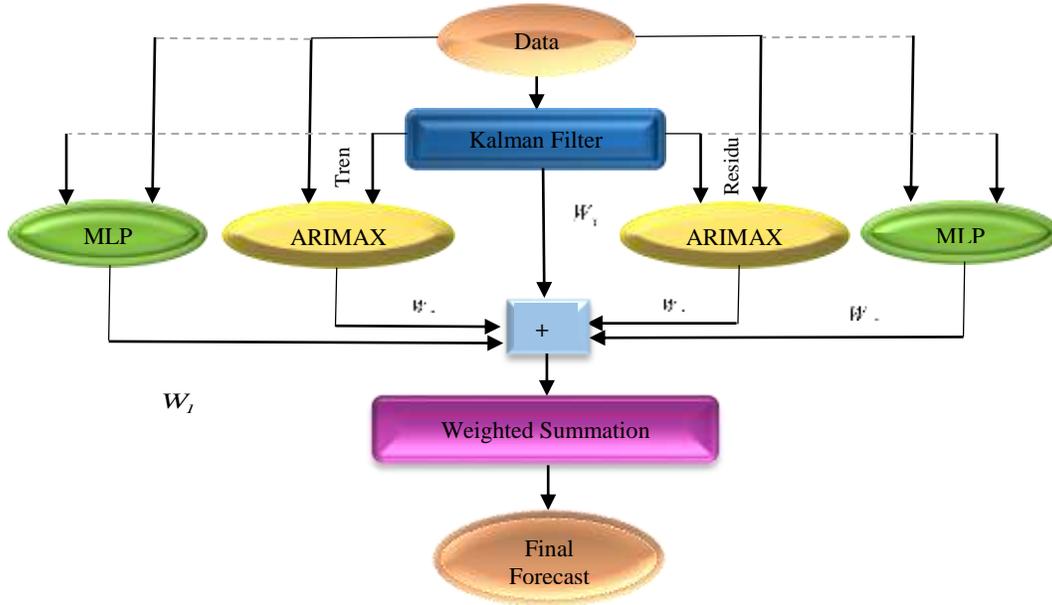
Fig 1. The overall structure of the proposed model.

The ARIMAX model is similar to a multivariate regression model but also allows us to take advantage of autocorrelation that may be present in residuals of the regression to improve the accuracy of a forecast. Or, as another viewpoint, The ARIMAX model is an Auto-Regressive Integrated Moving Average (ARIMA) model that also consists of the exogenous entrance, besides its autoregressive (AR) and moving average (MA) parameters. In this way, the ARIMAX model can be explained as the combination of the Autoregressive $(AR(p))$, Integrated (d), Moving Average $(MA(q))$, and Exogenous $X(r)$ variables, which is often represented by $ARIMAX(p,d,Q,r)$. A simplified form to represent this model is described in Eq. (2).

$$Z_t = \beta + \sum_{i=1}^{p} \phi_i z_{t-i} + \sum_{j=1}^{q} \theta_j \varepsilon_{t-j} + \sum_{j=1}^{m} \xi_j \varphi_j + \varepsilon_t \qquad (2)$$

where Z_t is a dependent variable at the time t, β is a constant; Z_{t-1} is a dependent variable (lagged by the time steps, i; ϕ_i is a coefficient of Z_{t-1}; p is the maximum number of time intervals; φ_j represents the exogenous variables (in this case, trend data, wind power data); ξ_j represents the coefficients of the exogenous variables; q is the maximum number of exogenous variables; θ_j is the coefficient of the term ε_{t-j}, which represents the error in the time t lagged from j. Finally, ε_t is the error component of the model with $\varepsilon_t \in N(0,\sigma^2)$ The coefficients of the models are estimated by regression (Khashei et al., 2009).

Stage III: *Nonlinear Trend/Residual Modeling:* Since the $ARIMAX$ model cannot model the nonlinear correlation structures, in the next stage of the proposed model, a multilayer perceptron (MLP) is used in order to model the nonlinear, trend and residual correlation structures in the underlying data sets as follows (Zendehboudi, 2016):

$$Y_{i,t} = \sum_{j=0}^{P} W_j \cdot g \left(\sum_{k=0}^{q} W_{kj} Y_{i,t-k} \right) + \varepsilon_t \qquad (i = 1,2,...,N, \quad t = 1,2,...,T) \qquad (3)$$

where, w_j and w_{kj} are connection weights, q is the number of input nodes and p is the number of hidden nodes.

Stage IV: Parallel combination: In the last step of the proposed model, the final prediction of the proposed model is calculated from the results of all components and their weights in the parallel structure by Eq. (4).

$$y_{Total} = W_1\, y_{TMLP}(t) + W_2\, y_{TARIMAX}(t) + W_3\, y_{KF}(t) + W_4\, y_{RARIMAX}(t) + W_5\, y_{RMLP}(t) \tag{4}$$

where, $y_{TMLP}(t)$ and $y_{TARIMAX}(t)$ are the prediction of the *MLP* and *ARIMAX* models for trend data at the time t, respectively. In the similar, the $y_{RARIMAX}(t)$ and $y_{RMLP}(t)$ are the forecasting of the *ARIMAX* and *MLP* models for residual data at the time t, respectively. The $y_{KF}(t)$ is the Kalman filter model output value at the time t and $w_i\,(i = 1, 2, ..., 5)$ are the weights of components. In this paper, these weights are estimated by the ordinary least squares method (Chahkoutahi & Khashei, 2017).

APPLYING THE PROPOSED METHOD TO PREDICT WIND POWER

DATA SETS

In this section, in order to evaluate the accuracy of the proposed model and calculate its performance against other methods, 2 wind power and speed data sets are used. The first data set is hourly wind power data from a wind power plant located in Spain has been used. This hourly data set consists of 168 points that show in Fig. (2). The 85% of data (e.g., 144 observations) is applied as the training sample, and 15% of the remaining data (e.g., 24 observations) is used as the test sample in order to evaluate the performance of different models. The second data set is the wind speed data set, which is gathered every 5 min in the Colorado State on 09/02/2012 from 00:00 to 14:00, and totally consists of 169 points. Approximately, the first 80% of data (e.g. 133 observations) is applied as the training sample, and 20% of the remaining data (e.g., 36 observations) is used as the test sample. These data sets are among the most widely used benchmark data sets in the field of wind forecasting, which has been frequently used in the literature (Azimi et al. 2016).
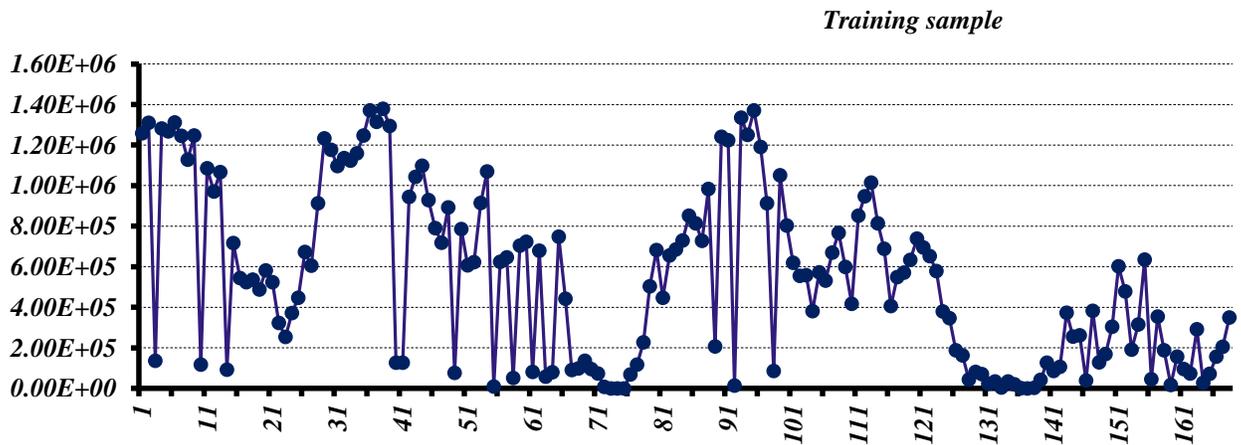
Fig 2. The wind power data sets.

PERFORMANCE EVALUATION CRITERIA

The results presented in this paper are based on some evaluation criteria related to the outputs of 1) Kalman filter model, 2) trend ARIMAX model, 3) trend MLP model, 4) residual ARIMAX model, 5) residual MLP model, 6) and the proposed hybrid model. In this paper, different indicators such as MAE (Mean Absolute Error) and RMSE (Root Mean Squared Error) are used, which are calculated as follows:

$$MAE = \frac{1}{N}\sum_{t=1}^{N}|e_t| \tag{5}$$

$$RMSE = \sqrt{\frac{1}{N}\sum_{t=1}^{N}(e_t)^2} \tag{6}$$

where, e_t and A_t $(t=1,2,...,m)$ are forecast error and forecasted value at time t, respectively, and N is the number of total underlying data.

RESULTS OF THE PROPOSED MODEL

In this section, the first wind power data set is used to design different models as well as the proposed model. After designing each model using training data, the performance of the models is calculated using the criteria mentioned above based on the test data, and ultimately a comprehensive assessment of the methods' performance compared with each other in the training and the test data has been made.

RESULTS OF THE DATA PREPROCESSING TECHNIQUE

Stage I: Kalman Filter

Following the process of the proposed method, the Kalman filter model is first used for preprocessing input data. The trend and residual pattern obtained from the Kalman filter are shown in contrast to the actual data in Fig. (3).
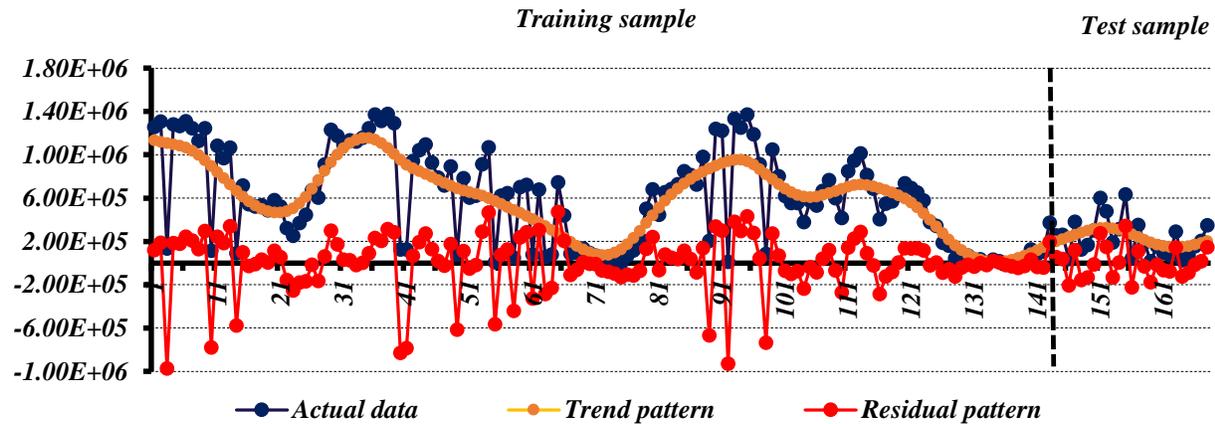
Fig 3. Trend and residual pattern in contrast to the actual data.

RESULTS OF THE TREND/RESIDUAL LINEAR MODELING

Stage II: Trend/Residual and Autoregressive Integrated Moving Average with Explanatory Variable models

i) The trend patterns in the data and the relevant lags, along with the raw data, are entered into the autoregressive integrated moving average with explanatory variable (TARIMAX) model. The results of this model are obtained using the Eviews package software. The estimation of the TARIMAX model using evaluation characteristics (absolute error (AE) and squared error (SE)) are illustrated in Fig. (4). The performance of the TARIMAX model is extensively presented in Table (2).

TARIMAX(AE)
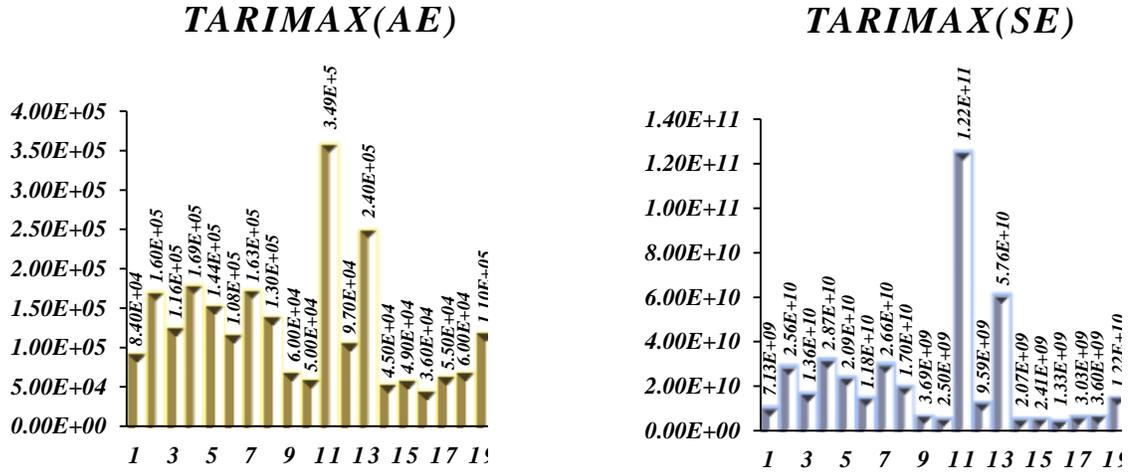
TARIMAX(SE)

Fig 4. The performance of the TARIMAX model for test data.

Table 2. The performance of the TARIMAX model results in the test day.

Time (h)	Evaluation Criteria		Time (h)	Evaluation Criteria	
	AE	SE		AE	SE
1:00	8.40×10^4	7.13×10^9	**13:00**	2.40×10^5	5.76×10^{10}
2:00	1.60×10^5	2.56×10^{10}	**14:00**	4.50×10^4	2.07×10^9
3:00	1.16×10^5	1.36×10^{10}	**15:00**	4.90×10^4	2.41×10^9
4:00	1.69×10^5	2.87×10^{10}	**16:00**	3.60×10^4	1.33×10^9
5:00	1.44×10^5	2.09×10^{10}	**17:00**	5.50×10^4	3.03×10^9
6:00	1.08×10^5	1.18×10^{10}	**18:00**	6.00×10^4	3.60×10^9
7:00	1.63×10^5	2.66×10^{10}	**19:00**	1.10×10^5	1.22×10^{10}
8:00	1.30×10^5	1.70×10^{10}	**20:00**	1.20×10^5	1.45×10^{10}
9:00	6.00×10^4	3.69×10^9	**21:00**	8.05×10^4	6.49×10^9

10:00	5.00×10^4	2.50×10^9	**22:0**	7.76×10^4	6.03×10^9
11:00	3.49×10^5	1.22×10^{11}	**23:0**	4.73×10^4	2.24×10^9
12:00	9.70×10^4	9.59×10^9	**24:0**	8.90×10^4	8.00×10^9
MAE	1.10×10^5		**RMSE**	1.31×10^5	

Results of Table (2) indicate that the TARIMAX model can achieve 1.10×10^5 and 1.31×10^5 in the MAE and RMSE in the whole test day, respectively. The errors of the TARIMA model for training and test data sets are also shown in Fig. (5).
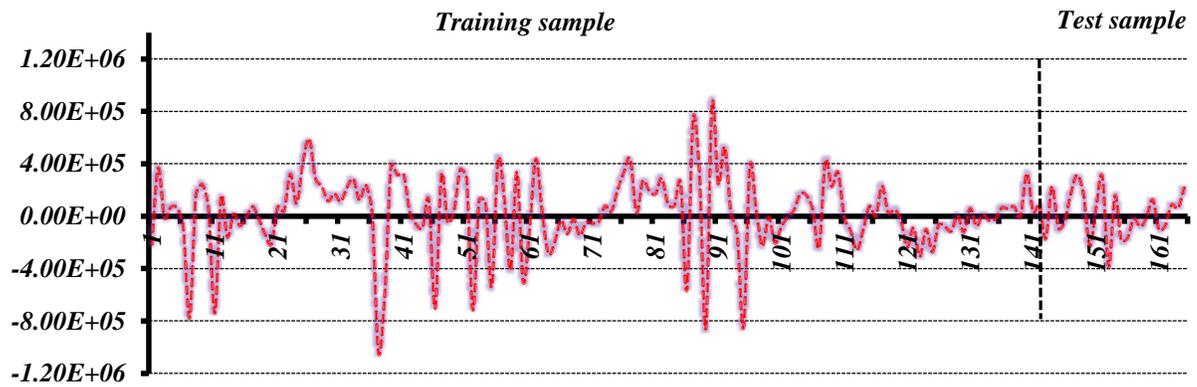
Fig 5. The errors of the TARIMAX model for training and test data sets.

ii) In the next stage of step 2, the residual patterns in the data and the relevant lags, along with the raw data, are entered into the Autoregressive Integrated Moving Average with the Explanatory Variable (RARIMAX) model. The performance indicators of the designed RARIMAX model are shown in Fig. (6). The performance of the RARIMAX model is extensively presented in Table (3).
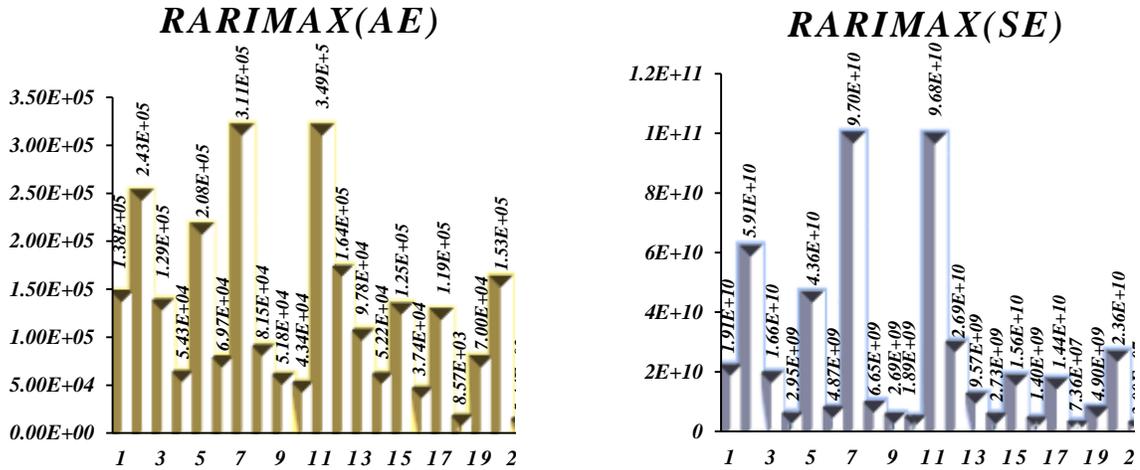
Fig 6. The performance of the RARIMAX model for test data.

Table 3. The performance of the RARIMAX model results in the test day.

Time (h)	Statistical Characteristics		Time (h)	Statistical Characteristics	
	AE	SE		AE	SE
1:00	1.38×10^5	1.91×10^{10}	**13:00**	1.64×10^5	9.57×10^9
2:00	2.43×10^5	5.91×10^{10}	**14:00**	9.78×10^4	2.73×10^9
3:00	1.29×10^5	1.66×10^{10}	**15:00**	5.22×10^4	1.56×10^{10}
4:00	5.43×10^4	2.95×10^9	**16:00**	1.25×10^5	1.40×10^9
5:00	2.08×10^5	4.36×10^{10}	**17:00**	3.74×10^4	1.44×10^{10}
6:00	6.97×10^4	4.87×10^9	**18:00**	1.19×10^5	7.00×10^7
7:00	3.11×10^5	9.70×10^{10}	**19:00**	8.57×10^3	4.91×10^9
8:00	8.15×10^4	6.65×10^9	**20:00**	7.00×10^4	2.36×10^{10}
9:00	5.18×10^4	2.69×10^9	**21:00**	1.53×10^5	2.00×10^7

10:00	4.34×10^{4}	1.89×10^{9}	**22:00**	5.44×10^{3}	8.20×10^{9}
11:00	3.11×10^{5}	9.68×10^{10}	**23:00**	9.05×10^{4}	1.86×10^{9}
12:00	1.38×10^{5}	2.69×10^{10}	**24:00**	4.30×10^{4}	4.29×10^{10}
MAE	**1.17×10^{5}**		**RMSE**	**1.45×10^{5}**	

Results of Table (3) indicate that the RARIMAX model can achieve 1.17×10^{5} and 1.45×10^{5} in MAE, and RMSE in the whole test day, respectively. The errors of the RARIMAX model for training and test data sets are also shown in Fig. (7).

Training sample
Test sample
1.00E+06
5.00E+05
0.00E+00
-5.00E+05
-1.00E+06
1
11
21
31
41
51
61
71
81
91
101
111
121
131
141
151
161
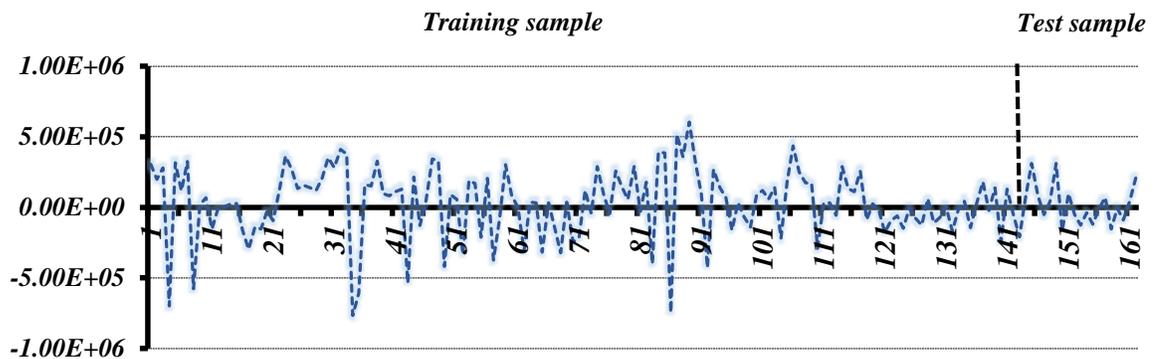
Fig 7. The errors of the RARIMAX model for training and test data sets.

RESULTS OF THE TREND/RESIDUAL NONLINEAR MODELING

Stage III: Trend/Residual and Multilayer Perceptron models

i) The MLP is the nonlinear model that is sketched in the proposed model. Similar to the previous stage, the trend pattern in the data and the relevant lags, along with the raw data, are entered into the Multilayer Perceptron model. The results of this model are obtained using the Matlab package software. The estimation of the TMLP model is presented in Fig. (8). The results of the TMLP model are expansively reported in Table (4).

TMLP(AE)

TMLP(SE)
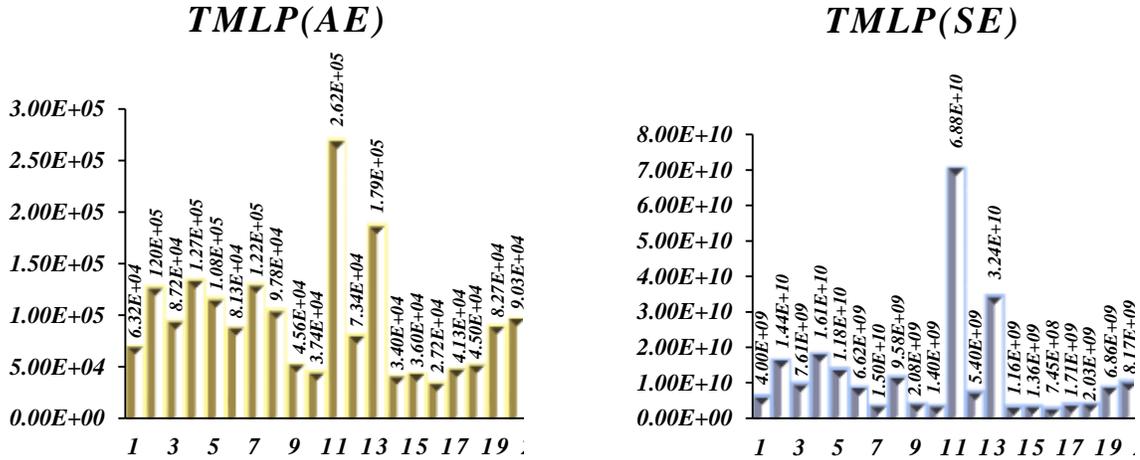
Fig 8. The performance of the TMLP model for test data.

Table 4. The performance of the TMLP results in the test day.

Time (h)	Evaluation Criteria		Time (h)	Evaluation Criteria	
	AE	SE		AE	SE
1:00	6.32×10^4	4.00×10^9	**13:00**	1.79×10^5	3.24×10^{10}
2:00	1.20×10^5	1.44×10^{10}	**14:00**	3.40×10^4	1.16×10^9
3:00	8.72×10^4	7.61×10^9	**15:00**	3.60×10^4	1.36×10^9
4:00	1.27×10^5	1.61×10^{10}	**16:00**	2.72×10^4	7.45×10^8
5:00	1.08×10^5	1.18×10^{10}	**17:00**	4.13×10^4	1.71×10^9
6:00	8.13×10^4	6.62×10^9	**18:00**	4.50×10^4	2.03×10^9
7:00	1.22×10^5	1.50×10^{10}	**19:00**	8.27×10^4	6.86×10^9
8:00	9.78×10^4	9.58×10^9	**20:00**	9.03×10^4	8.17×10^9
9:00	4.56×10^4	2.08×10^9	**21:00**	6.04×10^4	3.66×10^9
10:00	3.74×10^4	1.40×10^9	**22:00**	5.80×10^4	3.39×10^9

11:00	2.62×10^5	6.88×10^{10}	**23:00**	3.50×10^4	1.27×10^9
12:00	7.34×10^4	5.40×10^9	**24:00**	6.69×10^4	4.49×10^9
MAE	8.27×10^4		**RMSE**	9.79×10^4	

Results of Table (4) indicate that the TMLP model can achieve 8.27×10^4 and 9.79×10^4 in MAE, and RMSE in the whole test day, respectively. The errors of the TMLP model for training and test data sets are also shown in Fig. (9). These results indicate that the MLP model can achieve more accurate and more reliable results than ARIMAX. It demonstrates that the underlying data set consists of the linear patterns and structures, in addition to nonlinear ones. Thus, the classic ARIMAX model is inadequate for modeling and cannot comprehensively and appropriately model all existing patterns and structures in the data.
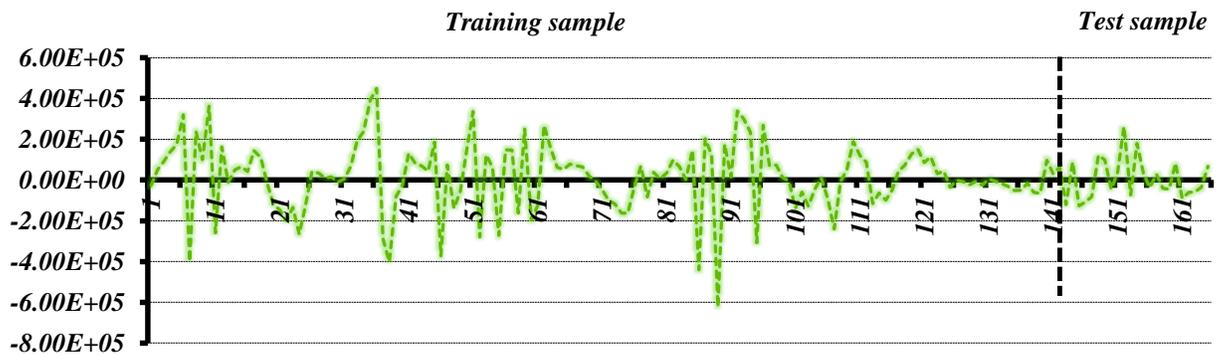
Fig 9. The errors of the TMLP model for training and test data sets.

ii) In a similar process, the residual patterns in the data and the relevant lags, along with the raw data, are entered into the MLP model. The performance indicators of the designed TMLP model are shown in Fig. (10). The performance of the TMLP model is extensively presented in Table (5).
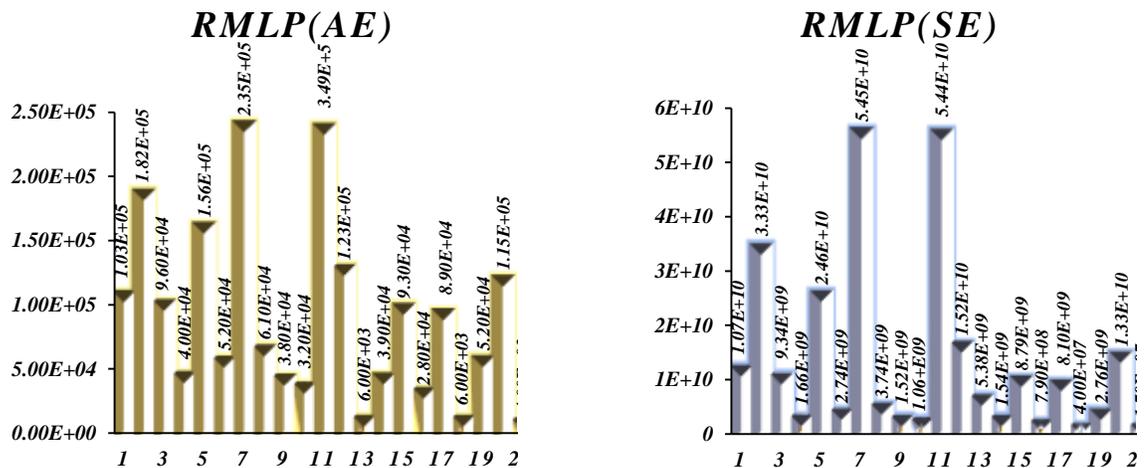
Fig 10. The performance of the RMLP model for test data

385

Table 5. The performance of the RMLP results in the test day.

Time (h)	Statistical Characteristics		Time (h)	Statistical Characteristics	
	AE	SE		AE	SE
1:00	1.03×10^5	1.07×10^{10}	**13:00**	6.00×10^3	5.38×10^9
2:00	1.82×10^5	3.33×10^{10}	**14:00**	3.90×10^4	1.54×10^9
3:00	9.60×10^4	9.34×10^9	**15:00**	9.30×10^4	8.79×10^9
4:00	4.00×10^4	1.66×10^9	**16:00**	2.80×10^4	7.90×10^8
5:00	1.56×10^5	2.64×10^{10}	**17:00**	8.90×10^4	8.10×10^9
6:00	5.20×10^4	2.74×10^9	**18:00**	6.00×10^3	4.10×10^7
7:00	2.35×10^5	5.45×10^{10}	**19:00**	5.20×10^4	2.76×10^9
8:00	6.10×10^4	3.74×10^9	**20:00**	1.15×10^5	1.33×10^{10}
9:00	3.80×10^4	1.52×10^9	**21:00**	4.00×10^3	1.60×10^7
10:00	3.20×10^4	1.06×10^9	**22:00**	6.70×10^4	4.62×10^9
11:00	2.33×10^5	5.44×10^{10}	**23:00**	3.20×10^4	1.04×10^9
12:00	1.23×10^5	1.52×10^{10}	**24:00**	1.55×10^5	2.41×10^{10}
MAE	8.81×10^4		**RMSE**	1.09×10^5	

Results of Table (5) show that the RMLP model can achieve 8.81×10^4 and 1.09×10^5 in MAE, and RMSE in the whole test day, respectively. The errors of the RMLP model for training and test data sets are also shown in Fig. (11).
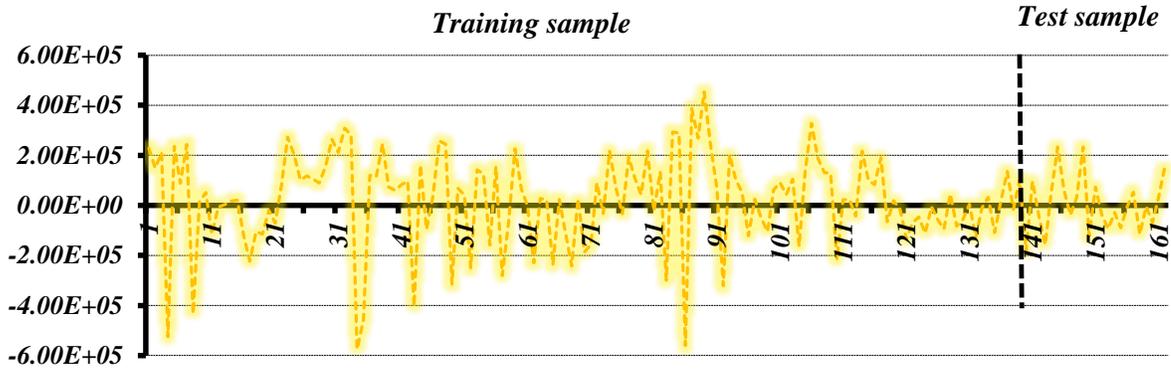
Fig 11. The errors of the RMLP model for training and test data sets.

RESULTS OF THE COMBINING

Stage IV

According to the process of the proposed model and after modeling components, the weight of each component, i.e., trend ARIMA, residual ARIMA, trend MLP, residual MLP, and the Kalman filter, is calculated. After that, final forecasts of the proposed model are calculated by Eq. (11). These forecasting results using absolute and squared error values are presented in Fig. (12). The actual and fitted values of the proposed model for train and test day are shown in Fig. (13). Also, the evaluation metrics for the proposed model for test day are reported in Table (6).
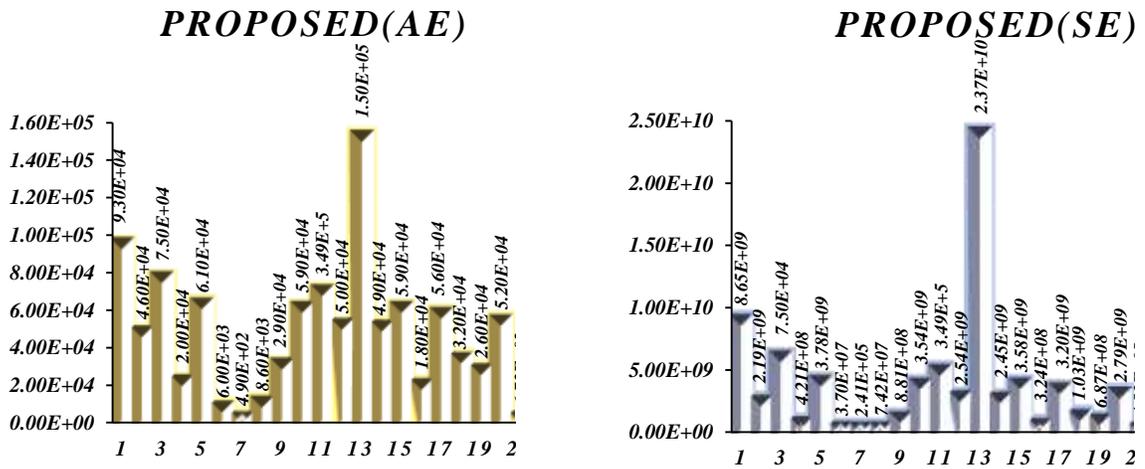
Fig 12. The performance of the proposed model
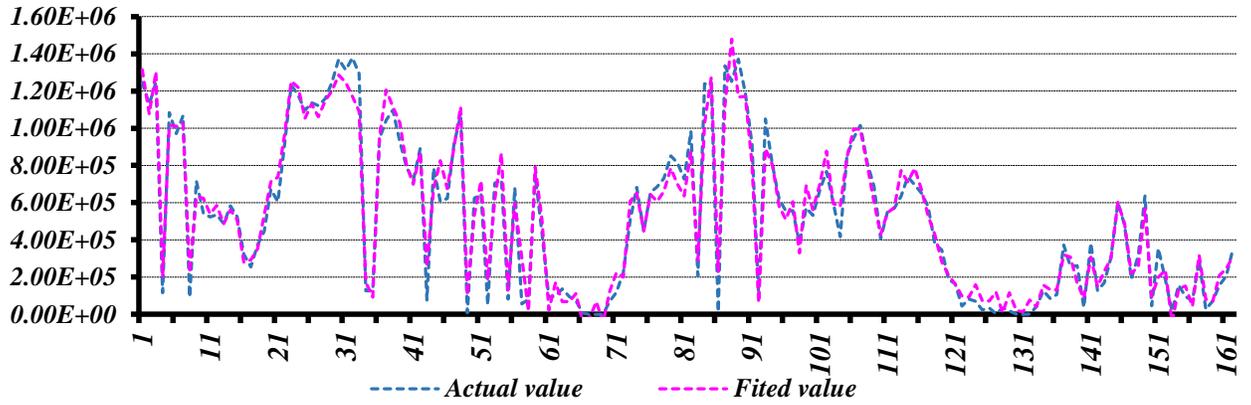
Fig 13. Actual and fitted values of the proposed model for train and test data sets.

Table 6. Results of the proposed model on a test day.

Time (h)	Statistical Characteristics		Time (h)	Statistical Characteristics	
	AE	SE		AE	SE
1:00	9.30×10^4	8.65×10^9	13:00	1.50×10^5	2.37×10^{10}
2:00	4.60×10^4	2.19×10^9	14:00	4.90×10^4	2.45×10^9
3:00	7.50×10^4	5.74×10^9	15:00	5.90×10^4	3.58×10^9
4:00	2.00×10^4	4.21×10^8	16:00	1.80×10^4	3.24×10^8
5:00	6.10×10^4	3.78×10^9	17:00	5.60×10^4	3.20×10^9
6:00	6.00×10^3	3.60×10^7	18:00	3.20×10^4	1.03×10^9
7:00	4.90×10^2	2.00×10^5	19:00	2.60×10^4	6.87×10^8
8:00	8.60×10^3	7.40×10^7	20:00	5.20×10^4	2.79×10^9
9:00	2.90×10^4	8.81×10^8	21:00	6.72×10^2	4.51×10^5
10:00	5.90×10^4	3.54×10^9	22:00	5.20×10^4	2.76×10^9

11:00	6.80×10^4	4.65×10^9	**23:00**	3.90×10^4	1.56×10^9
12:00	5.00×10^4	2.54×10^9	**24:00**	3.80×10^4	1.50×10^9
MAE	4.59×10^4		**RMSE**	5.36×10^4	

Results of Table (6) show that the proposed model can achieve 4.59×10^4, and 5.36×10^4 in MAE, and RMSE in the whole test day, respectively. The error values of the proposed model for the train and test data sets are also shown in Fig. (14).
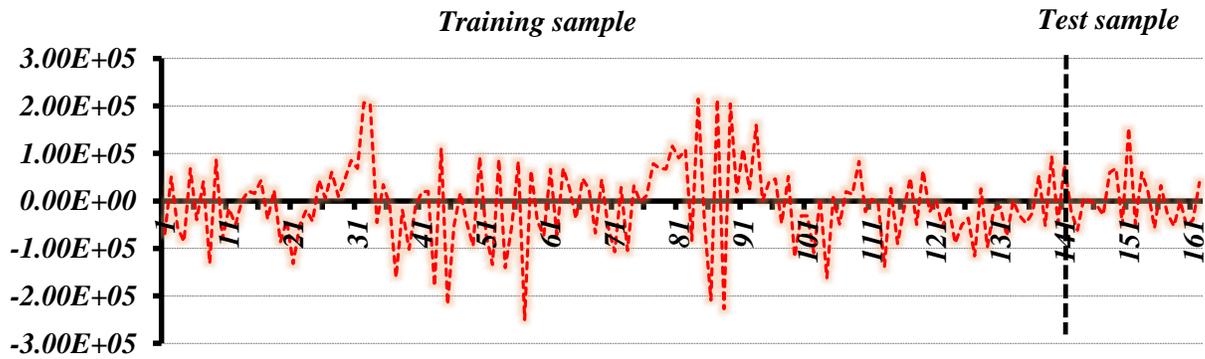
Fig 14. The errors of the proposed model for train and test data sets.

COMPARISON WITH OTHER THOSE HYBRID AND INDIVIDUAL MODELS

In this section, the performance of the proposed model is compared with some other individual and hybrid models using the 2 wind power and speed data sets. Evaluation criteria, e.g., MAE, and RMSE of the models in training and test data sets, are summarized in Table (7). Besides, the improvement percentage of the proposed model in comparison with those models is given in Table (8). In addition, Table (9) summarizes the improvement percentage of the proposed model against other classes of models as follows:

1)   SLM: Single linear model(s) (e.g., ARIMA),

2)   SNM: Single nonlinear model(s) (e.g., MLP),

3)   STM: Single trend model(s) (e.g., TARIMAX, and TMLP),

4)   SRM: Single residual model(s) (e.g., RARIMAX, and RMLP),

5)   PLM: Preprocessing-based linear model(s) (e.g., KARIMAX),

6)   PNM: Preprocessing-based nonlinear model(s) (e.g., KMLP),

7)   SCM: Series-based components combination hybrid model(s) (e.g., ARIMA-MLP),

8)   PCM: Parallel-based components combination hybrid model(s) (e.g., ARIMA/MLP).

Numerical results show that the proposed model in overall can improve averagely, 37.34 and 27.00 percentages in MAE, and 34.18 and 29.81 percentages in RMSE the performance of other individual and hybrid models in train and test data sets, respectively. The biggest value of these improvements is related to the ARIMA model that only can model the linear patterns and uses none of the trend and residual patterns of the preprocessing. The proposed model can achieve 46.73% in MAE and 46.37 % in RMSE improvement than the ARIMA.

Table 7. Forecasting performance of different models in train and test data sets.

Model	Data Sets	MAE		RMSE	
		Train	Test	Train	Test
ARIMA	*Sotavento Data Set*	2.11×10^5	1.47×10^5	2.99×10^5	1.75×10^5
MLP		1.58×10^5	1.10×10^5	2.24×10^5	1.31×10^5
TARIMAX		1.64×10^5	1.10×10^5	2.23×10^5	1.31×10^5
RARIMAX		1.92×10^5	1.17×10^5	2.50×10^5	1.45×10^5
TMLP		1.22×10^5	8.27×10^4	1.67×10^5	9.79×10^4
RMLP		1.44×10^5	8.81×10^4	1.87×10^5	1.09×10^5
ARIMA-MLP		8.34×10^4	5.78×10^4	1.05×10^5	6.93×10^4
ARIMA/MLP		1.48×10^5	6.83×10^4	1.85×10^5	7.79×10^4
KARIMAX		1.44×10^5	1.06×10^5	2.01×10^5	1.34×10^5
KMLP		9.27×10^4	5.77×10^4	1.20×10^5	7.34×10^4
Proposed model		$\mathbf{6.64 \times 10^4}$	$\mathbf{4.95 \times 10^4}$	$\mathbf{8.64 \times 10^4}$	$\mathbf{5.63 \times 10^4}$
ARIMA	*Colorado Data Set*	1.034	1.016	1.307	1.228
MLP		0.825	0.991	1.105	1.219

Model				
TARIMAX	0.957	1.002	1.115	1.213
RARIMAX	1.018	1.007	1.223	1.220
TMLP	0.768	0.926	1.033	1.162
RMLP	0.816	0.988	1.091	1.211
ARIMA-MLP	0.947	0.987	1.077	1.189
ARIMA/MLP	0.753	0.912	1.001	1.112
KARIMAX	0.849	0.978	1.081	1.193
KMLP	0.713	0.901	0.983	1.136
Proposed model	**0.651**	**0.863**	**0.912**	**1.027**

The second and third places are also related to the ARIMA model with the use of the preprocessed residual and trend data, respectively. The proposed model can reach 43.36% in MAE and 41.96% in RMSE improvement than the RARIMAX. It demonstrates that the idea of using residual patterns instead of ignoring them can be a more efficient strategy. The proposed model can also obtain 40.08% in MAE and 37.95% in RMSE improvement than the TARIMAX. It clearly proves the reason for using trend patterns instead of residual patterns in the literature. The use of the trend patterns instead of residual patterns averagely improves 11.72% and 11.40% in MAE and RMSE the performance of the ARIMA model.

Table 8. Improvement percentages of the proposed model against other models.

Model	Data Sets	MAE		RMSE	
		Train	Test	Train	Test
ARIMA	*Sotavento Data Set*	68.53	66.32	71.10	67.82
MLP		57.97	55.00	61.42	57.02
TARIMAX		59.51	55.00	61.25	57.02

X	RARIMA		65.41	57.69	65.44	61.17
	TMLP		45.57	40.14	48.26	42.49
	RMLP		53.88	43.81	53.79	48.34
MLP	ARIMA-		20.38	14.36	17.71	18.76
MLP	ARIMA/		55.14	27.53	53.30	27.73
X	KARIMA		53.70	56.54	56.99	57.85
	KMLP		28.37	14.21	28.00	23.29
E	**AVERAGE**		**50.85**	**43.06**	**51.73**	**46.15**
	ARIMA		37.04	15.05	30.22	16.36
	MLP		21.09	12.91	17.46	15.75
X	TARIMA		31.97	13.87	18.20	15.33
X	RARIMA	*Colorado Data Set*	36.05	14.29	25.42	15.81
	TMLP		15.23	6.80	11.71	11.61
	RMLP		20.22	12.65	16.40	15.19
MLP	ARIMA-		31.25	12.56	15.32	13.62
MLP	ARIMA/		13.54	5.37	8.89	7.64
X	KARIMA		23.32	11.75	15.63	13.91
	KMLP		8.69	4.21	7.22	9.59

AVERAGE	**23.84**	**10.946**	**16.647**	**13.481**

The results of the MLP model also indicate almost the same conclusions; while using the MLP model and modeling nonlinear patterns, performances are improved. The use of the MLP instead of the ARIMA model alone can averagely improve the obtained results by more than 33% in both criteria.

Table 9. Improvement percentages of the proposed model against other classes of models.

Classes	Data Sets	MAE		RMSE	
		Train	Test	Train	Test
1- SLM models		68.53	66.32	71.10	67.82
2- SNM models		56.85	53.92	60.42	55.99
3- STM models		49.27	44.08	51.60	46.32
4- SRM models	*Sotavento Data Set*	53.33	43.14	53.35	47.65
5- PLM models		53.70	56.54	56.99	57.85
6- PNM models		24.61	10.05	24.00	19.12
7- SCM models		20.38	14.36	17.71	18.76
8- PCM models		55.14	27.53	53.30	27.73
AVERAGE		**47.73**	**39.49**	**48.56**	**42.66**
1- SLM models		37.04	15.05	30.22	16.36
2- SNM models	*Colorado Data Set*	21.09	12.91	17.46	15.75
3- STM models		23.60	10.33	14.95	13.47

393

4- SRM models	28.135	13.47	20.91	15.5
5- PLM models	23.32	11.75	15.63	13.91
6- PNM models	8.69	4.21	7.22	9.59
7- SCM models	31.25	12.56	15.32	13.62
8- PCM models	13.54	5.37	8.89	7.64
AVERAGE	**23.33**	**10.70**	**16.32**	**13.23**

Roughly, these results are also repeated for its preprocessed versions (e.g., the TMLP instead of the TARIMAX and RMLP instead of the RARIMAX model). At the same time, the proposed model averagely can reach 29.78% in MAE and 30.97% in RMSE improvement than the TMLP and RMLP. It can be concluded from these results that at first, the use of the nonlinear modeling approach can be a better solution than linear ones for improving the forecasting accuracy. Second, the proposed idea of using residual as well as trend patterns in both linear and nonlinear environments can be an appropriate way to yield more accurate results by a more comprehensiveness modeling process. Furthermore, the proposed model outperforms both series and parallel component combination based hybrid models in both MAE and RMSE criteria in train and test data sets. The proposed model can reach 25.81% and 13.46% in MAE and 16.51% and 16.19% in RMSE improvement than the series combination based hybrid model in train and test data sets, respectively. The proposed model can obtain 34.34% and 16.45% in MAE and 31.09% and 17.68% in RMSE improvement than the parallel combination based hybrid model in train and test data sets, respectively.

CONCLUSION

Advanced modeling to predict time series has become a leading subject in much recent research in the literature. On the other hand, improving the accuracy of predictions, especially in complex environments and nonlinear systems, has become one of the most challenging and the hardest issues for researchers. Hybridization is one of the most common ways of improving forecasting accuracy as well as reducing the risk of using inappropriate single models that have extensively and successfully been used in recent investigations. Wind energy is one of the most important clean energy and suitable alternative for fossil fuels. Due to the importance of wind power forecasting for planning purposes, several methods have been developed and used to predict this energy source more precisely. However, increasing the accuracy of wind power and speed forecasting models is a crucial challenge for researchers, due to specific characteristics, as well as the existence of sophisticated linear and nonlinear patterns in the data. In this paper, a new filter-based hybrid model is proposed in which these different structures and patterns can simultaneously be modeled. The main idea of the proposed model, which differs from other existing hybrid models, is that in the proposed model, the residual patterns are not set aside and are not considered as noise, due to the fact that there is no logical evidence supporting it. Empirical results indicate that the proposed model, due to simultaneously using linear/nonlinear, trend/residual, models can model more different patterns and structures in underlying data, can achieve more accurate results.

REFERENCES

Lindley, D. 2010. The energy storage problem. Nature 463: 18–20.

Roulston, MS., Kaplan, DT., Hardenberg, J., & Smith, LA. 2003. Using medium-range weather forecasts to improve the value of wind energy production. Renewable Energy 28: 585–602.

Zack, J. 2003. Overview of wind energy generation forecasting. Albany, NY: TrueWind Solutions, LLC & AWS Scientific, Inc.

Soman, SS., Zareipour, H., Malik, O., & Mandal, P. 2010. A review of wind power and wind speed forecasting methods with different time horizons. In: Proceedings of the 2010 north american power symposium (NAPS): 1-8.

Lange, M., & Focken, U. 2008. New developments in wind energy forecasting," IEEE Power and Energy Society General Meeting 2008 - Conversion and Delivery of Electrical Energy in the 21st Century: 20-4.

Kariniotakis, G., Pinson, P., Siebert, N., Giebel, G., & Barthelmie, R. 2004. The state of the art in short-term prediction of wind power from an offshore perspective. 2004 Sea Tech Week.

Lange, M., & Focken, U. 2006. Physical approach to short-term wind power prediction. New York, NY: Springer.

Marjanovic, N., Wharton, S., & Chow, F. 2014. Investigation of model parameters for high–resolution wind energy forecasting: Case studies over simple and complex terrain. Journal of Wind Engineering and Industrial Aerodynamics 134: 10-24.

Wang, Y., Wang, H., Srinivasan, D., & Hu, Q. 2019. Robust functional regression for wind speed forecasting based on Sparse Bayesian learning. Renewable Energy 132: 43–60.

Nathaniel, S. P., & Lukas, G. S. 2018. Statistical approach for improved wind speed forecasting for wind power production. Sustainable Energy Technologies and Assessments 27: 180–91.

Wang, Y., Xie, Z., Hu, Q., & Xiongc, Sh. 2018. Short–term wind speed prediction using an extreme learning machine model with error correction. Energy Conversion and Management 163: 384–406.

Lahouar A., & Ben Hadj Slama, J. 2017. Hour–ahead wind power forecast based on random forests. Renewable Energy 109: 529–41.

Zhang, Y., Wang, J., & Luo, X. 2015. Probabilistic wind power forecasting based on logarithmic transformation and boundary kernel. Energy Conversion and Management 96: 440–51.

Bessa, R. J., Miranda, V., Botterud, A., Zhou, Z., & Wang, J. 2012. Time–adaptive quantile–copula for wind power probabilistic forecasting. Renewable Energy 40: 29–39.

Messner, J. W., & Pinson, P. 2019. Online adaptive lasso estimation in vector autoregressive models for high dimensional wind power forecasting. International Journal of Forecasting 35(4): 1485-98.

Robles–Rodriguez, C.E., & Dochain, D. 2018. Decomposed Threshold ARMAX Models for short to medium term wind power forecasting. IFAC Papers On Line 51(13): 49–54.

Filik, T., & Filik, U.B. 2017. Very Short Term Wind Speed Forecasting Using Multivariable Dense Data with WLS–MARMA Model. Energy Procedia 107: 259–63.

Ziel, F., Croonenbroeck, C., & Ambach, D. 2016. Forecasting wind power – Modeling periodic and non–linear effects under conditional heteroscedasticitya. Applied Energy 177: 285–97.

Lydia, M., Suresh Kumar, S., Selvakumar, A. I., & Kumar, G. E. P. 2016. Linear and non–linear autoregressive models for short–term wind speed forecasting. Energy Conversion and Management 112: 115–24.

Guo, Zh., Chi, D., Wu, J.,& Zhang, W. 2014. A new wind speed forecasting strategy based on the chaotic time Series Modelling technique and the Apriori algorithm. Energy Conversion and Management 84: 140–51.

Kou, P., Liang, D., Gao, F., & Gao, L. 2014. Probabilistic wind power forecasting with online model selection and warped gaussian process. Energy Conversion and Management 84: 649–63.

Kou, P., Gao, F., & Guan, X. 2013. Sparse online warped Gaussian process for wind power probabilistic forecasting. Applied Energy 108: 410–28.

Ramirez–Rosado, I. J., Fernandez–Jimenez, L. A., Monteiro, C., Sousa, Jo., & Bessa, R. 2009. Comparison of two new short–term wind–power forecasting systems. Renewable Energy 34: 1848–54.

Xydas, E., Qadrdan, M., Marmaras, Ch., Cipcigan, L., Jenkins, N., & Ameli, H. 2017. Probabilistic wind power forecasting and its application in the scheduling of gas–fired generators. Applied Energy 192: 382–94.

Iversen, E. B., Morales, J. M., Møller, J. K., & Madsen, H. 2016. Short–term probabilistic forecasting of wind speed using stochastic differential equations. International Journal of Forecasting 32: 980–90.

Carpinone, A., Giorgio, M., Langella, & R., Testa, A. 2015. Markov chain modeling for very–short–term wind power forecasting. Electric Power Systems Research 122:152–8.

Song, Zh., Jiang, Y., & Zhang, Z. 2014. Short–term wind speed forecasting with Markov–switching model. Applied Energy 130: 103–12.

D'Amicoa, G., Petroni, F., & Prattico, F. 2014. Wind speed and energy forecasting at different time scales: A nonparametric approach. Sustainable Energy Technologies and Assessments. Physica A 406: 59–66.

Bivona, S., Bonanno, G., Burlon, R., Gurrera, D., & Leone, C. 2011. Stochastic models for wind speed forecasting. Energy Conversion and Management 52: 1157–65.

Maatallah, O. A., Achuthan, A., Janoyan, K., & Marzocca, P. 2015. Recursive wind speed forecasting based on Hammerstein Auto–Regressive model. Applied Energy 145: 191–7.

Wang, K., Qi, X., Liu, H., & Song, J. 2018. Deep belief network based k–means cluster approach for short–term wind power forecasting. Energy 165: 840–52.

Yu, R., Gao, J., Yu, M., Lu, W., Xu, T., Zhao, M., Zhang, J., Zhang, R., & Zhang, Zh. 2018. LSTM–EFG for wind power forecasting based on sequential correlation features.Future Generation Computer Systems 93: 33–42.

Ahmed, A., & Khalid, M. 2018. An intelligent framework for short–term multi–step wind speed forecasting based on Functional Networks. Applied Energy 225: 902–11.

He, Y., & Li, H. 2017. Probability density forecasting of wind power using quantile regression neural network and kernel density estimation. Energy Conversion and Management 150: 108–21.

Wang, H.zh., Li, G.q., Wang, G.b., Peng, J.ch., Jiang, H., & Liu, Y.t. 2017. Deep learning based ensemble approach for probabilistic wind power forecasting. Applied Energy188: 56–70.

Dong, Q., Sun, Y., & Li, P. 2017. A novel forecasting model based on a hybrid processing strategy and an optimized local linear fuzzy neural network to make wind power forecasting: A case study of wind farms in China. Renewable Energy 102: 241–57.

Taslimi Renani, E., Elias, M. F., Rahim, & N. Abd. 2016. Using data–driven approach for wind power prediction: A comparative study. Energy Conversion and Management 118: 193–203.

Hu, Q., Zhang, Sh., Xie, Z., Mi, J., & Wan, J. 2014. Noise model based v–support vector regression with its application to short–term wind speed forecasting. Neural Networks 57: 1–11.

Yang, Y., & Zhao, Y. 2012. Prevailing Wind Direction Forecasting for Natural Ventilation Adjustment in Greenhouses Based on LE–SVM. Renewable Energy 16: 252-58.

Zaccheus O. O. 2014. A 5–day wind speed & power forecasts using a layer recurrent neural network (LRNN). Sustainable Energy Technologies and Assessments 6: 1–24.

Chitsazan, M. A., Sami Fadali, M., & Trzynadlowski, A.j. 2019. Wind speed and wind direction forecasting using echo state network with nonlinear functions. Renewable Energy 131: 879–89.

Khashei, M., & Bijari, M. 2010. An artificial neural network (p, d, q) model for time series forecasting. Expert Systems with Applications. 37: 479-89.

Khashei, M., & Bijari, M. 2011. A novel hybridization of artificial neural networks and ARIMA models for time series forecasting. Applied Soft Computing. 11: 2664-75.

Naik, J., Bisoi, R., & Dash, P.K. 2018. Prediction interval forecasting of wind speed and wind power using modes decomposition based low rank multi–kernel ridge regression. Renewable Energy 129: 357-83.

Rahmani, R., Yusof, R., Seyedmahmoudian, M.M., & Mekhilef, S. 2013. Hybrid technique of ant colony and particle swarm optimization for short term wind energy forecasting. Journal of Wind Engineering and Industrial Aerodynamics 123: 163-70.

Chang, G.W., Lu, H.J., Chang, Y.R., & Lee, Y.D. 2017. Short–term wind speed and power forecasting using an ensemble of mixture density neural networks. Renewable Energy105: 301-11.

Khashei, M., Bijari, M., & Mokhatab, F. 2013. Hybrid Fuzzy Auto-Regressive Integrated Moving Average (FARIMAH) model for Forecasting the Foreign Exchange Markets. International Journal of Computational Intelligence Systems 6: 954-68.

Vaccaro, A., Bontempi, G., Ben Taieb, S., & Villacci, D. 2012. Adaptive local learning techniques for multiple–step–ahead wind speed forecasting. Electric Power Systems Research 83: 129–35.

Shukur, O. B., & Lee, M. H. 2015. Daily wind speed forecasting through hybrid KF–ANN model based on ARIMA", Renewable Energy76: 637-47.

Sun, Sh., Qiao, H., Wei, Y., & Wang, Sh. 2017. A new dynamic integrated approach for wind speed forecasting. Applied Energy 197:151-62.

Wang, J., Du, P., Niu, T., & Yang, W. 2017. A novel hybrid system based on a new proposed algorithm—Multi–Objective Whale Optimization Algorithm for wind speed forecasting. Applied Energy 208: 344–60.

Wang, D., Luo, H., Grunder, O., & Lin, Y. 2017. Multi–step ahead wind speed forecasting using an improved wavelet neural network combining variational mode decomposition and phase space reconstruction. Renewable Energy 113:1345-58.

Ma, X., Jin, Y., & Dong, Q. 2017. A generalized dynamic fuzzy neural network based on singular spectrum analysis optimized by brain storm optimization forshort–term wind speed forecasting. Applied Soft Computing 54: 296–312.

Aghajani, A., Kazemzadeh, R., & Ebrahimi, A. 2016. A novel hybrid approach for predicting wind farm power production based on wavelet transform, hybrid neural networks and imperialist competitive algorithm. Energy Conversion and Management 121: 232-40.

Azimi, R., Ghofrani, M., & Ghayekhloo, M. 2016. A hybrid wind power forecasting model based on data mining and wavelets analysis. Energy Conversion and Management 127: 208-25.

Shao, H., & Deng, X. 2016. Short–term wind power forecasting using model structure selection and data fusion techniques. Electrical Power and Energy Systems 83: 79–86.

Jiang, P., Li, R., & Li, H. 2019. Multi–objective algorithm for the design of prediction intervals for wind power forecasting model. Applied Mathematical Modelling 67: 101-22.

Chen, Y., Heb, Zh., Shang, Zh., Li, C., Li, L., & Xu, M. 2019. A novel combined model based on echo state network for multi–step ahead wind speed forecasting: A case study of NREL. Energy Conversion and Management 179: 13–29.

Zhang, D., Peng, X., Pan, K., & Liu, Y. 2019. A novel wind speed forecasting based on hybrid decomposition and online Sequential outlier robust extreme learning machine. Energy Conversion and Management 180: 338–57.

Jiang, P., Wang, Y., & Wang, J. 2018. Short–term wind speed forecasting using a hybrid model. Energy 119: 561–77.

Jiang, P., Yang, H., & Heng, J. 2019. A hybrid forecasting system based on fuzzy time Series and multi–objective optimization for wind speed forecasting. Applied Energy 235:786–801.

Liu, H., Wu, H., & Li, Y. 2018. Smart wind speed forecasting using EWT decomposition, GWO evolutionary optimization, RELM learning and IEWT reconstruction. Energy Conversion and Management 161: 266–83.

Liu, H., Duan, Zh., Han, F. Z., & Li, Y. F. 2018. Big multi–step wind speed forecasting model based on secondary Decomposition, ensemble method and error correction algorithm. Energy Conversion and Management 156: 525–41.

Wang, L., Lid, X., & Bai, Y. 2018. Short–term wind speed prediction using an extreme learning machine model with error correction. Energy Conversion and Management 162: 239–50.

Li, Y., Wu, H., & Liu, H. 2018. Multi–step wind speed forecasting using EWT decomposition, LSTM principal computing, RELM subordinate computing and IEWT reconstruction. Energy Conversion and Management 167: 203–19.

Chang, G.W., Lu, H.J., Chang, Y.R., & Lee, Y.D. 2017. Short–term wind speed and power forecasting using an ensemble of mixture density neural networks. Renewable Energy 105: 301–11.

Lynch, C., OMahony, M. J., & Scully, T. 2014. Simplified method to derive the Kalman Filter covariance matrices to predict wind speeds from a NWP model. Energy Procedia 62: 676–85.

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. Journal of Wind Engineering & Industrial Aerodynamics 174: 28–38.

Akçay, H., & Filik, T. 2017. Short–term wind speed forecasting by spectral analysis from long–term observations with missing values. Applied Energy 191: 653–62.

Cassola, F., & Burlando, M. 2012. Wind speed and wind energy forecast through Kalman filtering of Numerical Weather Prediction model output. Applied Energy 99: 154-66.

Zhao, P., Wang, J., Xia, J., Dai, Y., Sheng, Y., & Yue, J. 2012. Performance evaluation and accuracy enhancement of a day–ahead wind power forecasting system in China. Renewable Energy 43: 234–41.

Liu, H., Mi, X. W., & Li, Y.F. 2018. Wind speed forecasting method based on deep learning strategy using empirical wavelet transform, long short term memory neural network and Elman neural network. Energy Conversion and Management 156: 498–514.

Hu, Y.L., & Chen, L. 2018. A nonlinear hybrid wind speed forecasting model using LSTM network, hysteretic ELM and Differential Evolution algorithm. Energy Conversion and Management 173: 123–42.

Yu, Ch., 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 Conversion and Management 178:137–45.

Jiang, P., & Li, Ch. 2018. Research and application of an innovative combined model based on a modified optimization algorithm for wind speed forecasting. Measurement 124: 395–412.

Liu, H., & Chen, Ch. 2019. Data processing strategies in wind energy forecasting models and applications: A comprehensive review. Applied Energy 2491: 392-408.

Hajirahimi, Z., & Khashei, M. 2019. Weighted sequential hybrid approaches for time series forecasting. Physica A: Statistical Mechanics and its Applications 531:121717.

Khashei, M., Bijari, M., & Raissi, G. A. 2009. Improvement of Auto–Regressive Integrated Moving Average Models Using Fuzzy Logic and Artificial Neural Networks. Neurocomputing 72: 956–67.

Zendehboudi, A. 2016. Implementation of GA-LSSVM modelling approach for estimating the performance of solid desiccant wheels. Energy Conversion and Management 127: 245-55.

Chahkoutahi, F., & Khashei, M. 2017. A seasonal direct optimal hybrid model of computational intelligence and soft computing techniques for electricity load forecasting. Energy 140: 988–1004.