DOI 10.24425/aee.2021.138274

# A combined method for wind power generation forecasting

TUAN-HO LE

*Faculty of Engineering and Technology, Quy Nhon University*
*Quy Nhon, Binh Dinh Province, 820000, Vietnam*
*e-mail: tuanhole@qnu.edu.vn*

(Received: 31.03.2021, revised: 12.08.2021)

**Abstract:** Most of the existing statistical forecasting methods utilize the historical values of wind power to provide wind power generation prediction. However, several factors including wind speed, nacelle position, pitch angle, and ambient temperature can also be used to predict wind power generation. In this study, a wind farm including 6 turbines (capacity of 3.5 MW per turbine) with a height of 114 meters, 132-meter rotor diameter is considered. The time-series data is collected at 10-minute intervals from the SCADA system. One period from January 04$^{th}$, 2021 to January 08$^{th}$, 2021 measured from the wind turbine generator 06 is investigated. One period from January 01$^{st}$, 2021 to January 31$^{st}$, 2021 collected from the wind turbine generator 02 is investigated. Therefore, the primary objective of this paper is to propose a combined method for wind power generation forecasting. Firstly, response surface methodology is proposed as an alternative wind power forecasting method. This methodology can provide wind power prediction by considering the relationship between wind power and input factors. Secondly, the conventional statistical forecasting methods consisting of autoregressive integrated moving average and exponential smoothing methods are used to predict wind power time series. Thirdly, response surface methodology is combined with autoregressive integrated moving average or exponential smoothing methods in wind power forecasting. Finally, the two above periods are performed in order to demonstrate the efficiency of the combined methods in terms of mean absolute percent error and directional statistics in this study.

**Key words:** autoregressive integrated moving average, exponential smoothing method, forecasting, response surface methodology, wind power

## Nomenclature

ARIMA is the autoregressive integrated moving average, RSM is the response surface methodology, EXP is the exponential smoothing method, $x_1$ is the wind speed, $x_2$ is the nacelle position,

$x_3$ is the pitch angle, $x_4$ is the ambient temperature, MAPE is the mean absolute percent error, $D_{\text{stat}}$ is the directional statistics, $y_t$ is the wind power generation at time period $t$, $\hat{y}_t$ is the estimated value (forecasted value) of $y$ at time period $t$.

## 1. Literature review

Because of the environmental issues, especially the exhaustion of fossil fuels, wind power sources have been highly integrated into electrical power systems in recent decades. Power from wind produces no air pollution or carbon dioxide emissions, no water consumption, and it is an inexhaustible renewable energy source [1]. Global wind report 2019 of GWEC (global wind energy council) showed that the 60.4 GW of new installations brings global cumulative wind power capacity up to 651 GW. New installations in the onshore wind market reached 54.2 GW, while the offshore wind market passed a milestone of 6 GW, making up 10% of the global new installation in 2019 the highest level to date. GWEC Market Intelligence expects that over 355 GW of new capacity will be added. That is nearly 71 GW of new installations each year until 2024 [2]. The rapid growth of this kind of energy has provided an alternative energy source in the electrical power systems. However, the high penetration of this energy source into the grid has caused several challenges for electrical power system operators. In addition, wind power is intermittent availability because it is weather-dependent. Therefore, accurate wind power forecasts are necessary for the efficient operation and the process of grid integration of wind power. As wind energy makes significant penetration into the electricity grid, the need for accurate predictions of wind power generation becomes critical and urgent [3, 4].

Numerous models and methods were proposed to forecast wind power generation. Based on a time-scale framework, wind power forecasting methods are normally divided into four different types, namely immediate short-term, short-term, medium-term, and long-term predictions [5]. For time-series data, various statistical methods and models were suggested for wind power forecasting and estimation. Statistical models are easy to use and cheaper to develop compared to other models. Basically, statistical methods use the previous history of wind data to perform a forecast over the next few hours; they are suitable for short periods of time [6]. An exponential smoothing approach was used for time-series wind power short-horizon prediction [7, 8]. This approach was also used for time-series wind power prediction in real-time electricity markets [9]. The methods based on autoregressive (AR), moving average (MA), autoregressive moving average (ARMA), or ARIMA models were employed to predict time-series wind power generation as in [10–14]. Recently, the artificial intelligence-based forecasting methods have been used to predict time-series data. Among these methods, models of multilayer neural networks have been developed to make the wind energy forecasts as in [6, 15, 17–19]. In this line, the recurrent neural networks models are also used to forecast wind power generation [20, 21]. The performance of the neural network models has been improved by using a deep learning approach in wind power forecasting [22, 23]. In the artificial intelligence-based forecasting methods, the support vector machine is also integrated into wind power predictions [24–26]. Obtaining the wind power forecasting results by using the artificial intelligence-based forecasting methods is time consuming since the performance of these methods must be trained. In addition, the training algorithms for the artificial intelligence methods are complicated. Another direction in non-stationary time series investigation is the deterministic chaos theory. Based on chaos theory, numerous efforts have been made to develop forecasting models to predict short-term wind power generation [27–29]. The accuracy of wind power prediction is highly addressed in these studies. Wind power forecasting errors for different methods by using different evaluation criteria were investigated in these existing studies.

Numerous efforts have been made in order to improve the forecasting accuracy and robustness of these methods. These models were improved or combined with other forecasting methods for better time-series wind power forecasting results as in [30–34]. In these studies, a statistical forecasting method is combined with another statistical forecasting method or an artificial intelligence-based forecasting method, and vice versa. However, useful information such as air pressure, air temperature, and climatic information, etc can be also employed for the forecast of wind power by using physical models as in [35]. Literature reviews of wind power forecasting methods can be found in [36, 37].

## 2. Introduction

In the existing studies in the literature, most of the wind power forecasting methods deal with time-series data problem. Unfortunately, in the practical operation of wind power plants, several factors can affect the power output. These operational parameters including wind speed, nacelle position, pitch angle, ambient temperature, etc can be used to predict wind power generation. In this case study, a wind farm, located on the south-central coast, Vietnam, including 6 turbines (capacity of 3.5 MW per turbine) with a height of 114 meters, 132-meter rotor diameter is considered. Given the topology of the land, the wind speed, nacelle position, pitch angle, ambient temperature, and active power of a wind turbine generator in this wind farm are measured based on the SCADA system. In order to provide the appropriate operation schedules for the wind farm, the wind power generation forecasting issue is necessary. Normally, the existing wind power generation forecasting methods are used for time-series data prediction. This study also investigates the effects of the operational parameters on the wind power forecasting results. In this case, RSM is used to investigate the relationship between the operational parameters of wind power plants and their power output. RSM is a collection of mathematical and statistical techniques that was introduced in the early 1950s by Box and Wilson [38]. A comprehensive presentation of RSM and its application can be found in [39]. In addition, the RSM-based forecasting method can be combined with other statistical forecasting models which are used to predict wind power time series for better prediction results in this case study. Therefore, the main motivation of this paper is to propose a combined method for wind power forecasting. RSM is proposed to predict wind power generation based on the operational parameters including wind speed, nacelle position, pitch angle, ambient temperature. Statistical models consisting of ARIMA and EXP are used to forecast wind power time series. The combined forecasting method is proposed by using RSM and ARIMA model/EXP. Finally, two different periods from practical wind power data is performed to illustrate the efficiency of the proposed methods. The data from the wind turbine generator 06 is measured in the period from January 04$^{th}$, 2021 to January 08$^{th}$, 2021. The data from the wind turbine generator 02 is measured in the period from January 01$^{st}$, 2021 to January 31$^{st}$, 2021. An overview of the proposed hybrid method for a wind power generation forecasting method in this study is presented in Fig. 1.

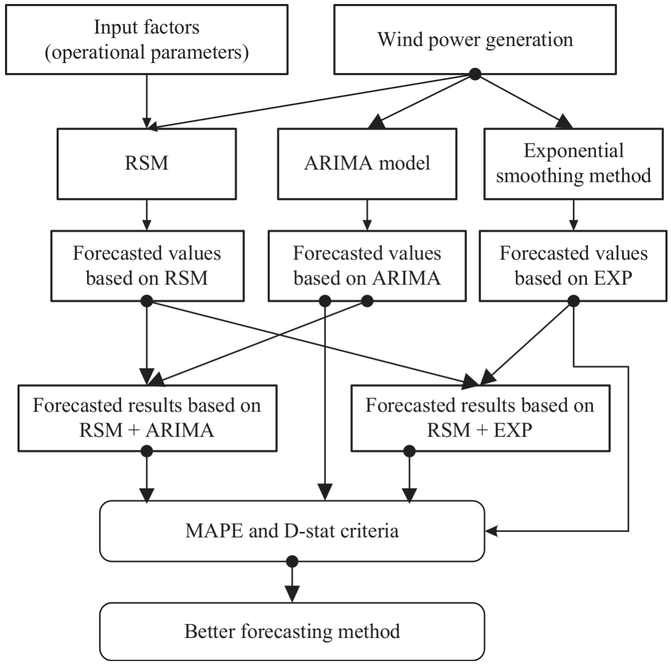


Fig. 1. Overview of the proposed hybrid method for wind power generation forecasting method

The remainder of this paper is organized as follows. The next section will discuss the proposed wind power forecasting methods used in this paper. Section 4 will investigate the two periods of wind power generation. The final section will be the conclusions and further studies.

## 3. The proposed methods

### 3.1. Response surface methodology

RSM is usually used to analyze and estimate the functional relationship between the input variables and the corresponding output response. The general form of RSM can be represented as follows:

$$\boldsymbol{y} = \boldsymbol{X}\boldsymbol{\beta} + \boldsymbol{\varepsilon}, \tag{1}$$

where: $\boldsymbol{y}$ is the vector of the observations, $\boldsymbol{X}$ is the matrix consisting of the levels of the input variables, $\boldsymbol{\beta}$ is the vector of the regression coefficients, and $\boldsymbol{\varepsilon}$ is the vector of random errors. By using the least-squares method, the estimator of $\boldsymbol{\beta}$ is

$$\hat{\boldsymbol{\beta}} = (\boldsymbol{X}'\boldsymbol{X})^{-1}\boldsymbol{X}'\boldsymbol{y}. \tag{2}$$

Therefore, the fitted regression model of e.g. (1) becomes

$$\hat{\boldsymbol{y}} = \boldsymbol{X}\hat{\boldsymbol{\beta}}. \tag{3}$$

In scalar form, the estimated second-order regression function of e.g. (3) can be shown as

$$\hat{y}_{\text{sec}} = \beta_0 + \sum_{i=1}^{p} \beta_i x_i + \sum_{i=1}^{p} \beta_i (x_i)^2 + \sum_{i=1}^{p} \sum_{j \neq i} \beta_{ij} x_i x_j \,. \tag{4}$$

In scalar form, the estimated third-order regression function of e.g. (3) can be shown as

$$\begin{aligned}\hat{y}_{\text{third}} = \beta_0 + \sum_{i=1}^{p} \beta_i x_i + \sum_{i=1}^{p} \beta_i (x_i)^2 + \sum_{i=1}^{p} \sum_{j \neq i} \beta_{ij} x_i x_j + \sum_{i=1}^{p} \beta_i (x_i)^3 + \\ + \sum_{i=1}^{p} \sum_{j \neq i} \beta_{ij} x_i (x_j)^2. \end{aligned} \tag{5}$$

The degree of fit of RSM is expressed by the coefficient of determination ($R^2$). In this study, the third-order regression function of RSM is proposed as an alternative wind power forecasting method. The four different input factors including wind speed, nacelle position, pitch angle, ambient temperature, are used to estimate the wind power generation. The proposed RSM – based wind power forecasting procedure can be shown in Fig. 2.

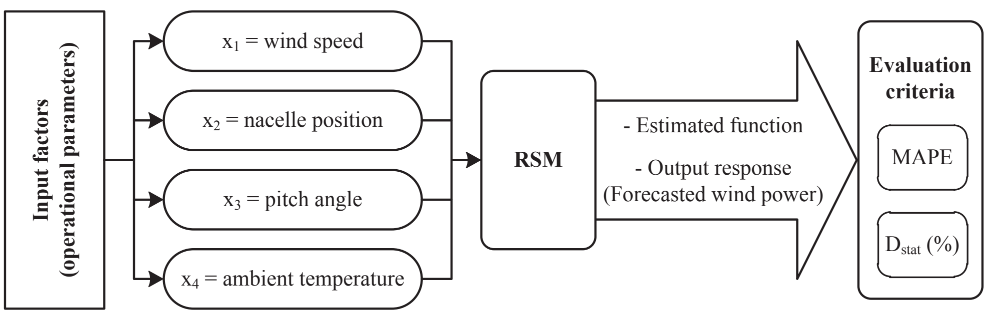


Fig. 2. The proposed RSM – based wind power forecasting procedure

### 3.2. Autoregressive integrated moving average

The ARIMA model first proposed by Box and Jenkins (1970) is a class of statistical models for time series data forecast and analysis. In order to predict the time series data, the ARIMA model uses three basic sequence processes including autoregressive, integrated, and moving average. The ARIMA model uses the past time-series data plus an error in order to predict future values. The application of the ARIMA model for time series forecasting includes the following steps: identification, estimation, diagnostic checking, and model's use. The general forecasting ARIMA model for time series is represented as

$$\hat{y}_t = c + \phi_1 y_{t-1} + \ldots + \phi_p y_{t-p} + \theta_1 \varepsilon_{t-1} + \ldots + \theta_q \varepsilon_{t-q} \,, \tag{6}$$

where $c$ is the constant; $\hat{y}_t$ is the estimated value of $y$ at time period $t$; $y_{t-1}, \ldots, y_{t-p}$ are the lagged values of $y$ at different time periods $t-1, \ldots, t-p$; $\varepsilon_{t-1}, \ldots, \varepsilon_{t-q}$ are the lagged errors at different time periods $t-1, \ldots, t-q$; $\phi_i$ $(i = 1, \ldots, p)$ and $\theta_i$ $(i = 1, \ldots, q)$ are the model parameters. This model is named the ARIMA $(p, d, q)$ model where $p$, $d$, and $q$ are non-negative integers. In the ARIMA model, $p$, $d$, and $q$ represent the order (number of time lags) of the autoregressive model, the degree of differencing, and the order of the moving-average model, respectively.

### 3.3. Exponential smoothing method

The EXP method proposed in the late 1950s by Brown (1959) is another class of statistical models for time-series data forecast and analysis. The principle of the EXP is to assign exponentially decreasing weights over time to make forecasts. Three different main types of exponential smoothing time-series forecasting methods are represented as follows:

***Simple exponential smoothing method***: This method is used for forecasting data without a trend or seasonality. For any time period $t$, the forecasted value $\hat{y}_t$ is

$$\hat{y}_t = \alpha y_{t-1} + (1 - \alpha)\hat{y}_{t-1}, \tag{7}$$

where the smoothing $\alpha$ is a constant between 0 and 1.

***Double exponential smoothing method***: This method is used for forecasting data with a trend. For any time period $t$, the forecasted value $\hat{y}_t$ can be calculated as

$$\begin{aligned} L_t &= \alpha y_t + (1 - \alpha)\left(L_{t-1} + T_{t-1}\right) \\ T_t &= \beta\left(L_t - L_{t-1}\right) + (1 - \beta)T_{t-1}, \\ \hat{y}_t &= L_{t-1} + T_{t-1} \end{aligned} \tag{8}$$

where $L_t$ and $T_t$ are the level and the trend at time $t$, respectively. $\beta$ is the weight (or smoothing constant) for the trend.

***Triple exponential smoothing method***: This method is used for forecasting data with a trend and seasonality. For any time period $t$, the forecasted value $\hat{y}_t$ can be calculated as

$$\begin{aligned} L_t &= \alpha\left(y_t - S_{t-p}\right) + (1 - \alpha)\left(L_{t-1} + T_{t-1}\right) \\ T_t &= \beta\left(L_t - L_{t-1}\right) + (1 - \beta)T_{t-1} \\ S_t &= \delta\left(y_t - L_t\right) + (1 - \delta)S_{t-p} \\ \hat{y}_t &= L_{t-1} + T_{t-1} + S_{t-p} \end{aligned}, \tag{9}$$

where $S_t$ is the seasonal component at time $t$, $\delta$ is the weight (or smoothing constant) for the seasonal component, and $p$ is the seasonal period.

### 3.4. The combined method

Different forecasting methods may provide different prediction results. However, forecasts from a given method may provide some useful information that is not conveyed in forecasts from other methods. Thus, instead of choosing a single forecasting method, it seems reasonable to consider aggregating information by generating forecasts from several methods and then combining these forecasts [40]. In this paper, the forecasts obtained from RSM are combined with the forecasts obtained from the ARIMA model or EXP. Therefore, the combined forecast at time $t$, $C_t$, is given by

$$C_t = w_1 \hat{y}_{1,t} + w_2 \hat{y}_{2,t}\,, \tag{10}$$

where $\hat{y}_{1,t}$ and $\hat{y}_{2,t}$ are the forecasts at time $t$ from the first set and the second set, respectively. By using the variance-covariance method, the corresponding weights in e.g. (10) can be calculated as

$$w_1 = \frac{\sum e_{2,t}^2 - \sum e_{1,t} e_{2,t}}{\sum e_{1,t}^2 + \sum e_{2,t}^2 + 2 \sum e_{1,t} e_{2,t}}\,, \tag{11}$$

$$w_2 = \frac{\sum e_{1,t}^2 - \sum e_{1,t} e_{2,t}}{\sum e_{1,t}^2 + \sum e_{2,t}^2 + 2 \sum e_{1,t} e_{2,t}}\,, \tag{12}$$

where $e_{1,t}$ and $e_{2,t}$ represent the individual forecast errors at time $t$.

### 3.5. Evaluation criteria

In order to evaluate the performance of the forecasting method, several criteria were proposed. In this paper, two primary evaluation criteria are used as follows:

MAPE is used to evaluate the prediction accuracy.

$$\mathrm{MAPE} = \frac{\sum\limits_{t=1}^{n} \dfrac{|e_t|}{y_t}}{n} = \frac{\sum\limits_{t=1}^{n} \dfrac{|y_t - \hat{y}_t|}{y_t}}{n}\,, \tag{13}$$

where $n$ is the number of observations.

$D_{\mathrm{stat}}$ is used to evaluate the ability of direction prediction.

$$D_{\mathrm{stat}} = \frac{1}{n} \sum_{t=1}^{n} d_t\,, \tag{14}$$

where $d_t = 1$ if $(\hat{y}_{t+1} - y_t)\,(y_{t+1} - y_t) \geq 0$, and $d_t = 0$ otherwise.

## 4. The investigated periods

### 4.1. Period 1

In this period, the time-series data from the wind turbine generator 06 is collected at 10-minute intervals in the period from January 04th, 2021 to January 08th, 2021. The total number of observations is 720. The wind speed, nacelle position, pitch angle, ambient temperature, and active power of the wind turbine generator in the investigated period are illustrated in Figs. 3–7, respectively.

By using the proposed RSM–based wind power forecasting procedure as shown in Fig. 2 with a third-order regression function, the functional relationship between the active power ($y$) and the operational parameters including wind speed ($x_1$), nacelle position ($x_2$), pitch angle ($x_3$), and

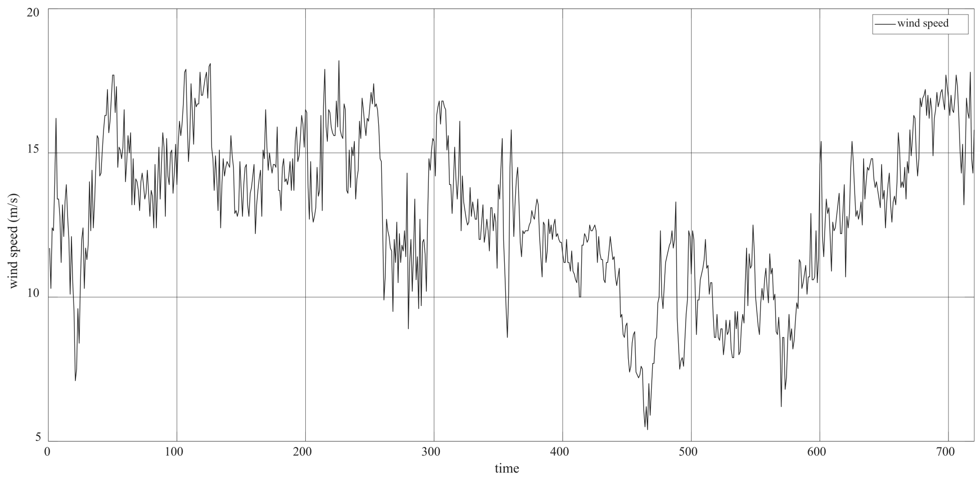


Fig. 3. Wind speed

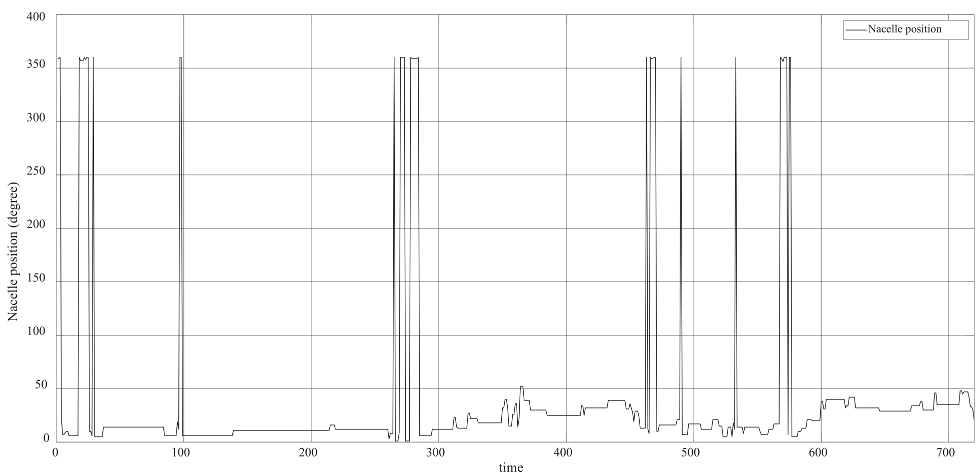


Fig. 4. Nacelle position

ambient temperature ($x_4$) is represented in e.g. (15). The corresponding analysis of a variance (ANOVA) table is shown in Table 1.

$$
\begin{aligned}
y = {} & -26389 - 10628x_1 + 1185x_2 - 3668x_3 + 10465x_4 - 124x_1^2 - 0.506x_2^2 - \\
& - 48.3x_3^2 - 783x_4^2 - 7.27x_1x_2 + 114x_1x_3 + 1070x_1x_4 - 3.43x_2x_3 - 86.8x_2x_4 + \\
& + 245x_3x_4 - 8.85x_1^3 - 0.000044x_2^3 - 0.412x_3^3 + 16.3x_4^3 + 0.0965x_1^2x_2 + \\
& + 7.17x_1^2x_3 + 13.86x_1^2x_4 + 0.00629x_1x_2^2 - 0.0437x_1x_2x_3 + 0.133x_1x_2x_4 + \\
& + 2.44x_1x_3^2 - 11.19x_1x_3x_4 - 27.9x_1x_4^2 - 0.00033x_2^2x_3 + 0.01797x_2^2x_4 + \\
& + 0.0737x_2x_3^2 + 0.130x_2x_3x_4 + 1.632x_2x_4^2 - 0.22x_3^2x_4 - 1.9x_3x_4^2\,.
\end{aligned}
\tag{15}
$$

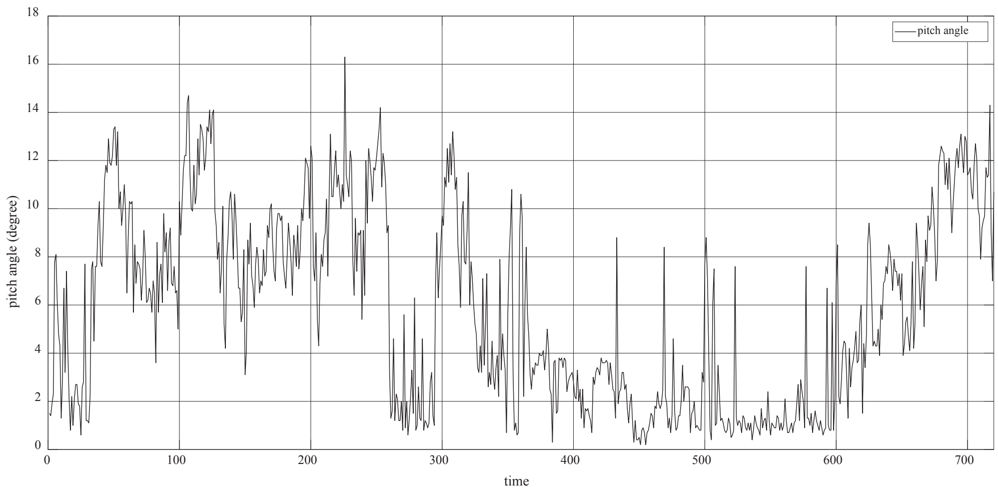


Fig. 5. Pitch angle

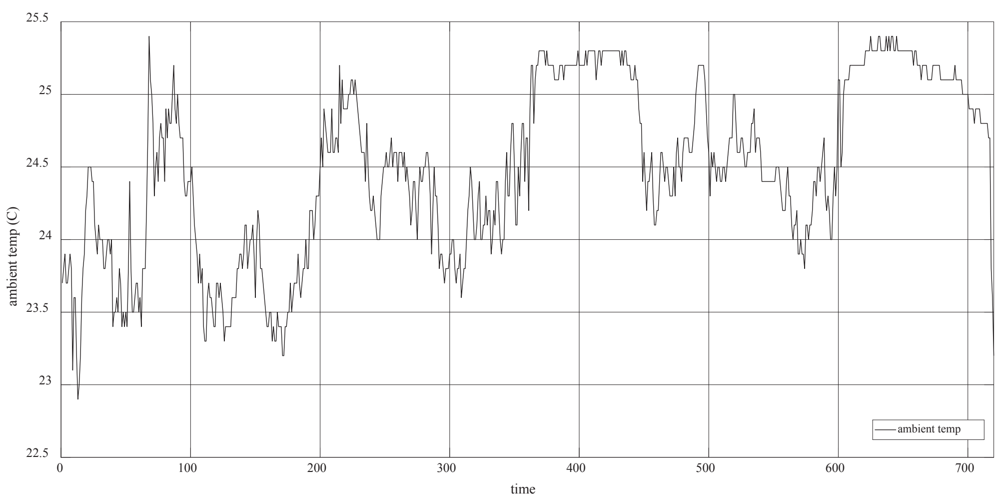


Fig. 6. Ambient temperature

According to Table 1, the p-value of the regression model is 0.000 (less than 0.05). This indicates that the regression is statistically significant. In addition, the coefficient of determination (R-sq equals 97.10%) indicates that the regression predictions almost perfectly fit the data. Furthermore, the p-values of $x_2$, $x_2 * x_2$, $x_1 * x_2$, $x_2 * x_4$, $x_1 * x_1 * x_1$, $x_1 * x_1 * x_2$, $x_1 * x_1 * x_3$, $x_1 * x_1 * x_4$, $x_1 * x_2 * x_2$, $x_2 * x_2 * x_4$, $x_2 * x_3 * x_3$ and $x_2 * x_4 * x_4$ are less than 0.05. This indicates these input parameters have a high relationship with the output response ($y$). In this period, the ARIMA (1, 1, 1) model was used to make the wind power forecast. In addition, a simple EXP with the smoothing $\alpha = 0.7$ was employed to predict the wind power. Based on the variance-covariance method, the proposed RSM–based wind power forecasting procedure is

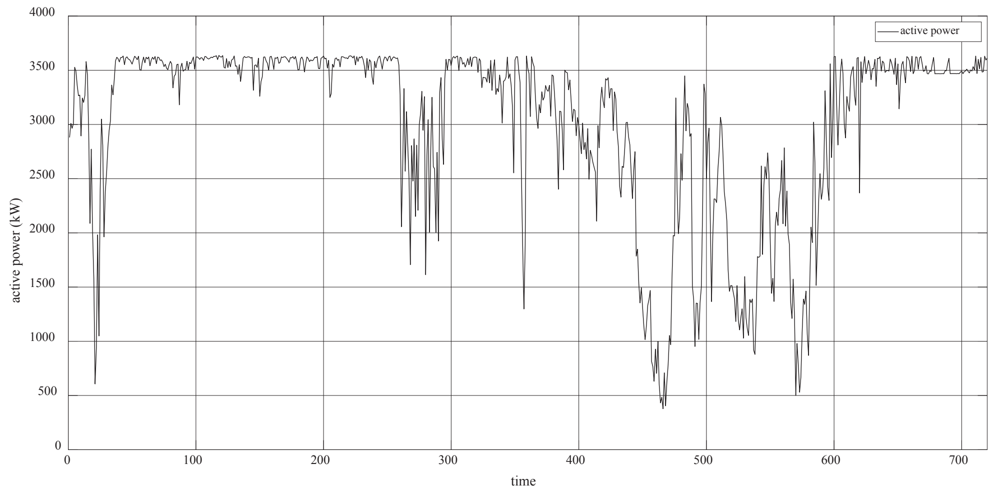


Fig. 7. Active power

combined with the ARIMA (1, 1, 1) model and EXP to make better forecasting results. The five forecasting results from the RSM, ARIMA (1, 1, 1) model, EXP, RSM+ARIMA, and RSM+EXP with the actual wind power are represented in Fig. 8.

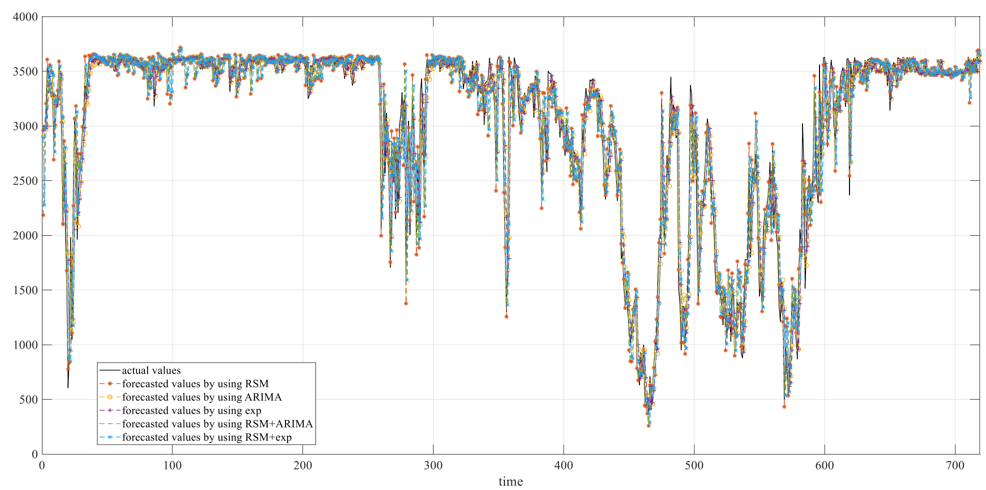


Fig. 8. The forecasted results of wind power generation

In this figure, the actual wind power values are demonstrated by the solid line (black color). The forecasted values by using the RSM, ARIMA model, EXP, RSM and ARIMA model, as well as RSM and EXP, are represented by the dash lines with a * marker (red color), o marker (yellow color), + marker (purple color), no marker (light green color), and a pentagram marker (light blue color), respectively. Most of the forecasted lines are approximate to the actual value

Table 1. ANOVA table

| **Source** | **DF** | **Adj SS** | **Adj MS** | **F-Value** | **P-Value** |
|---|---|---|---|---|---|
| Regression | 34 | 472386982 | 13893735 | 675.40 | 0.000 |
| $x_1$ | 1 | 11954 | 11954 | 0.58 | 0.446 |
| $x_2$ | 1 | 152629 | 152629 | 7.42 | 0.007 |
| $x_3$ | 1 | 4342 | 4342 | 0.21 | 0.646 |
| $x_4$ | 1 | 622 | 622 | 0.03 | 0.862 |
| $x_1 * x_1$ | 1 | 13120 | 13120 | 0.64 | 0.425 |
| $x_2 * x_2$ | 1 | 184637 | 184637 | 8.98 | 0.003 |
| $x_3 * x_3$ | 1 | 8770 | 8770 | 0.43 | 0.514 |
| $x_4 * x_4$ | 1 | 2196 | 2196 | 0.11 | 0.744 |
| $x_1 * x_2$ | 1 | 85593 | 85593 | 4.16 | 0.042 |
| $x_1 * x_3$ | 1 | 7367 | 7367 | 0.36 | 0.550 |
| $x_1 * x_4$ | 1 | 19682 | 19682 | 0.96 | 0.328 |
| $x_2 * x_3$ | 1 | 4840 | 4840 | 0.24 | 0.628 |
| $x_2 * x_4$ | 1 | 130259 | 130259 | 6.33 | 0.012 |
| $x_3 * x_4$ | 1 | 3084 | 3084 | 0.15 | 0.699 |
| $x_1 * x_1 * x_1$ | 1 | 1215264 | 1215264 | 59.08 | 0.000 |
| $x_2 * x_2 * x_2$ | 1 | 2132 | 2132 | 0.10 | 0.748 |
| $x_3 * x_3 * x_3$ | 1 | 10809 | 10809 | 0.53 | 0.469 |
| $x_4 * x_4 * x_4$ | 1 | 5343 | 5343 | 0.26 | 0.610 |
| $x_1 * x_1 * x_2$ | 1 | 359302 | 359302 | 17.47 | 0.000 |
| $x_1 * x_1 * x_3$ | 1 | 262385 | 262385 | 12.75 | 0.000 |
| $x_1 * x_1 * x_4$ | 1 | 107105 | 107105 | 5.21 | 0.023 |
| $x_1 * x_2 * x_2$ | 1 | 158012 | 158012 | 7.68 | 0.006 |
| $x_1 * x_2 * x_3$ | 1 | 69612 | 69612 | 3.38 | 0.066 |
| $x_1 * x_2 * x_4$ | 1 | 19210 | 19210 | 0.93 | 0.334 |
| $x_1 * x_3 * x_3$ | 1 | 42083 | 42083 | 2.05 | 0.153 |
| $x_1 * x_3 * x_4$ | 1 | 45122 | 45122 | 2.19 | 0.139 |
| $x_1 * x_4 * x_4$ | 1 | 34002 | 34002 | 1.65 | 0.199 |
| $x_2 * x_2 * x_3$ | 1 | 924 | 924 | 0.04 | 0.832 |
| $x_2 * x_2 * x_4$ | 1 | 129127 | 129127 | 6.28 | 0.012 |
| $x_2 * x_3 * x_3$ | 1 | 169348 | 169348 | 8.23 | 0.004 |
| $x_2 * x_3 * x_4$ | 1 | 4214 | 4214 | 0.20 | 0.651 |
| $x_2 * x_4 * x_4$ | 1 | 115288 | 115288 | 5.60 | 0.018 |
| $x_3 * x_3 * x_4$ | 1 | 112 | 112 | 0.01 | 0.941 |
| $x_3 * x_4 * x_4$ | 1 | 486 | 486 | 0.02 | 0.878 |
| Error | 685 | 14091244 | 20571 | R-sq | R-sq(adj) |
| Total | 719 | 486478226 | | 97.10% | 96.96% |

line. This indicates the efficiency of the forecasting models and methods. The evaluation criteria of the RSM, ARIMA model, EXP, RSM and ARIMA model, as well as RSM and EXP in wind power forecasting are illustrated in Table 2.

Table 2. Evaluation criteria of the proposed forecasting models and methods

| Evaluation criteria | RSM | ARIMA | Exp | RSM+ARIMA | RSM+exp |
|---|---|---|---|---|---|
| MAPE (%) | 4.3129 | 9.2140 | 9.1015 | 4.1971 | 4.1964 |
| $D_{\text{stat}}$ (%) | 76.4951 | 62.0306 | 61.4743 | 76.7733 | 76.9124 |

Based on the results in Table 2, the proposed RSM-based forecasting method can provide better prediction results (with MAPE = 4.3129% and $D_{\text{stat}}$ = 76.4951%) compared to the traditional statistical forecasting methods consisting of the ARIMA model (with MAPE = 9.2140% and $D_{\text{stat}}$ = 62.0306%) and EXP (with MAPE = 9.1015% and $D_{\text{stat}}$ = 61.4743%) in this period. Furthermore, the combined forecasting methods based on the variance-covariance method consisting of RSM and ARIMA (with MAPE = 4.1971% and $D_{\text{stat}}$ = 76.7733%) as well as RSM and EXP (with MAPE = 4.1964% and $D_{\text{stat}}$ = 76.9124%) can provide more accurate forecasting results compared to the others. Obviously, the proposed forecasting method consisting of RSM and EXP is the best wind power forecasting regarding MAPE and $D_{\text{stat}}$ in this period.

### 4.2. Period 2

In this period, the time-series data from the wind turbine generator 02 is collected at 10-minute intervals in the period from January 01st, 2021 to January 31st, 2021. The total number of observations is 4461. The wind speed, nacelle position, pitch angle, ambient temperature, and active power of the wind turbine generator in the investigated period are illustrated in Figs. 9– 12, and Fig. 13, respectively.

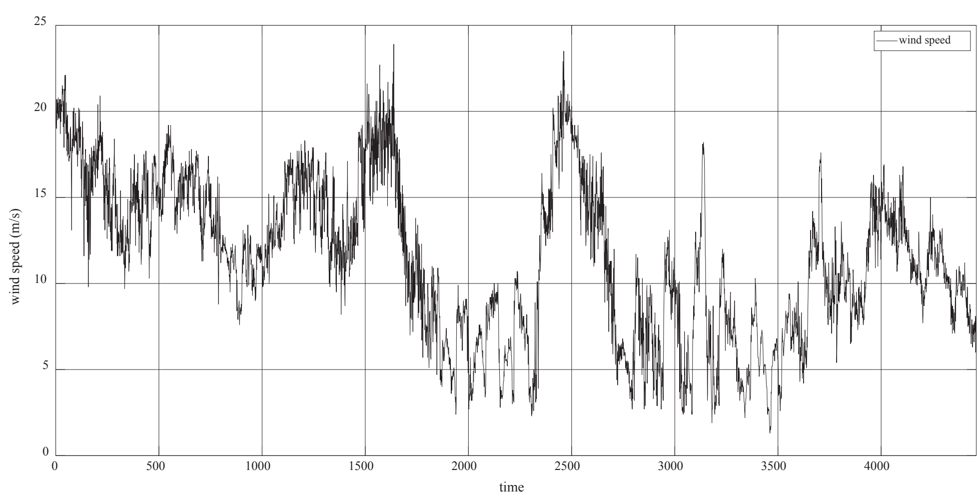


Fig. 9. Wind speed

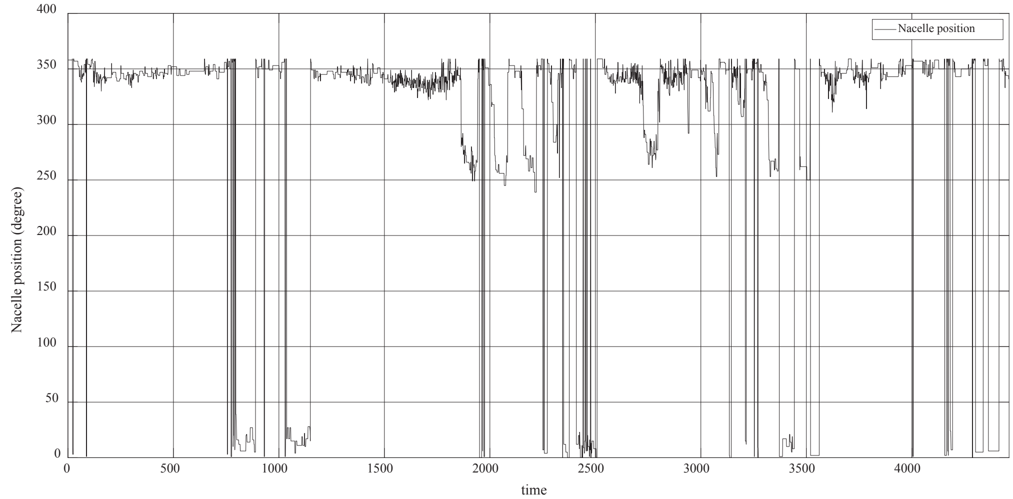


Fig. 10. Nacelle position

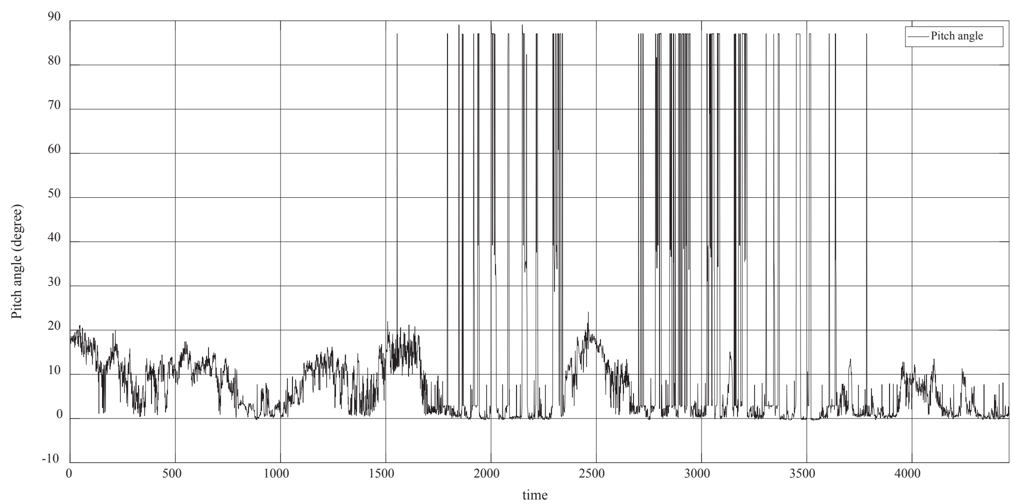


Fig. 11. Pitch angle

By using the proposed RSM–based wind power forecasting procedure as shown in Fig. 2 with third-order regression function, the functional relationship between the active power ($y$) and the operational parameters including wind speed ($x_1$), nacelle position ($x_2$), pitch angle ($x_3$), and ambient temperature ($x_4$) is represented in e.g. (16). The corresponding analysis of the variance

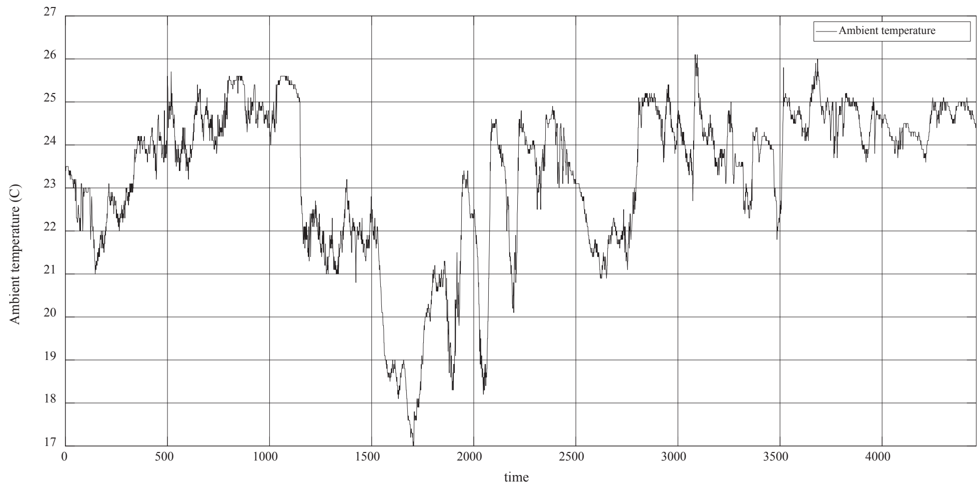


Fig. 12. Ambient temperature

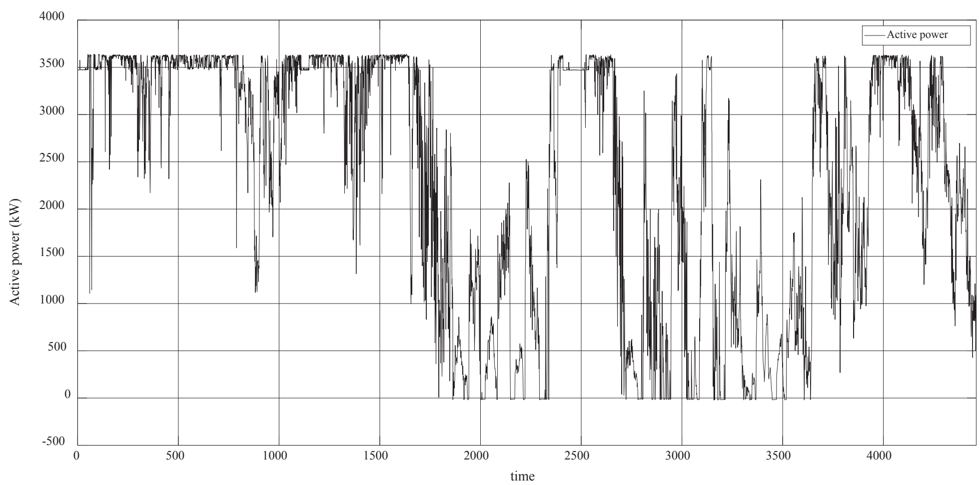


Fig. 13. Active power

(ANOVA) table is shown in Table 3.

$$\begin{aligned} y = {} & -38681 - 558x_1 - 63.2x_2 + 69.6x_3 - 4371x_4 + 78.55124x_1^2 + 0.0085x_2^2 - \\ & - 3.348x_3^2 + 156.2x_4^2 - 1.124x_1x_2 - 26.65x_1x_3 + 23.5x_1x_4 - 0.012x_2x_3 5.98x_2x_4 + \\ & + 5.34x_3x_4 - 0.6786x_1^3 + 0.000046x_2^3 + 0.03159x_3^3 - 1.819x_4^3 - 0.0356x_1^2x_2 + \\ & + 0.1725x_1^2x_3 - 2.104x_1^2x_4 + 0.003105x_1x_2^2 + 0.00016x_1x_2x_3 + 0.0314x_1x_2x_4 + \\ & + 0.18468x_1x_3^2 + 0.198x_1x_3x_4 + 0.238x_1x_4^2 + 0.000231x_2^2x_3 - 0.00281x_2^2x_4 - \\ & - 0.001756x_2x_3^2 + 0.00407x_2x_3x_4 - 0.1193x_2x_4^2 - 0.0403x_3^2x_4 - 0.079x_3x_4^2 . \end{aligned} \tag{16}$$

Table 3. ANOVA table

| **Source** | **DF** | **Adj SS** | **Adj MS** | **F-Value** | **P-Value** |
|---|---|---|---|---|---|
| Regression | 34 | 7560324430 | 222362483 | 3633.48 | 0.000 |
| $x_1$ | 1 | 241975 | 241975 | 3.95 | 0.047 |
| $x_2$ | 1 | 259476 | 259476 | 4.24 | 0.040 |
| $x_3$ | 1 | 55585 | 55585 | 0.91 | 0.341 |
| $x_4$ | 1 | 1137328 | 1137328 | 18.58 | 0.000 |
| $x_1 * x_1$ | 1 | 8856330 | 8856330 | 144.72 | 0.000 |
| $x_2 * x_2$ | 1 | 5960 | 5960 | 0.10 | 0.755 |
| $x_3 * x_3$ | 1 | 7002379 | 7002379 | 114.42 | 0.000 |
| $x_4 * x_4$ | 1 | 988395 | 988395 | 16.15 | 0.000 |
| $x_1 * x_2$ | 1 | 553924 | 553924 | 9.05 | 0.003 |
| $x_1 * x_3$ | 1 | 5233892 | 5233892 | 85.52 | 0.000 |
| $x_1 * x_4$ | 1 | 94720 | 94720 | 1.55 | 0.214 |
| $x_2 * x_3$ | 1 | 439 | 439 | 0.01 | 0.932 |
| $x_2 * x_4$ | 1 | 462637 | 462637 | 7.56 | 0.006 |
| $x_3 * x_4$ | 1 | 67741 | 67741 | 1.11 | 0.293 |
| $x_1 * x_1 * x_1$ | 1 | 5697813 | 5697813 | 93.10 | 0.000 |
| $x_2 * x_2 * x_2$ | 1 | 553104 | 553104 | 9.04 | 0.003 |
| $x_3 * x_3 * x_3$ | 1 | 29262017 | 29262017 | 478.15 | 0.000 |
| $x_4 * x_4 * x_4$ | 1 | 655842 | 655842 | 10.72 | 0.001 |
| $x_1 * x_1 * x_2$ | 1 | 3053028 | 3053028 | 49.89 | 0.000 |
| $x_1 * x_1 * x_3$ | 1 | 2012297 | 2012297 | 32.88 | 0.000 |
| $x_1 * x_1 * x_4$ | 1 | 5669566 | 5669566 | 92.64 | 0.000 |
| $x_1 * x_2 * x_2$ | 1 | 7192092 | 7192092 | 117.52 | 0.000 |
| $x_1 * x_2 * x_3$ | 1 | 185 | 185 | 0.00 | 0.956 |
| $x_1 * x_2 * x_4$ | 1 | 287497 | 287497 | 4.70 | 0.030 |
| $x_1 * x_3 * x_3$ | 1 | 60474200 | 60474200 | 988.17 | 0.000 |
| $x_1 * x_3 * x_4$ | 1 | 232859 | 232859 | 3.80 | 0.051 |
| $x_1 * x_4 * x_4$ | 1 | 29956 | 29956 | 0.49 | 0.484 |
| $x_2 * x_2 * x_3$ | 1 | 2800122 | 2800122 | 45.75 | 0.000 |
| $x_2 * x_2 * x_4$ | 1 | 453271 | 453271 | 7.41 | 0.007 |
| $x_2 * x_3 * x_3$ | 1 | 1879712 | 1879712 | 30.72 | 0.000 |
| $x_2 * x_3 * x_4$ | 1 | 32641 | 32641 | 0.53 | 0.465 |
| $x_2 * x_4 * x_4$ | 1 | 582699 | 582699 | 9.52 | 0.002 |
| $x_3 * x_3 * x_4$ | 1 | 920336 | 920336 | 15.04 | 0.000 |
| $x_3 * x_4 * x_4$ | 1 | 31655 | 31655 | 0.52 | 0.472 |
| Error | 4426 | 270863662 | 61198 | R-sq | R-sq(adj) |
| Total | 4460 | 7831188091 | | 96.54% | 96.22% |

According to Table 3, the p-value of the regression model is 0.000 (less than 0.05). This indicates that the regression is statistically significant. In addition, the coefficient of determination (R-sq equals 96.54%) indicates that the regression predictions almost perfectly fit the data. Furthermore, the p-values of $x_1$, $x_2$, $x_4$, $x_1 * x_1$, $x_2 * x_2$, $x_3 * x_3$, $x_4 * x_4$, $x_1 * x_2$, $x_1 * x_3$, $x_2 * x_4$, $x_1 * x_1 * x_1$, $x_2 * x_2 * x_2$, $x_3 * x_3 * x_3$, $x_4 * x_4 * x_4$, $x_1 * x_1 * x_2$, $x_1 * x_1 * x_3$, $x_1 * x_1 * x_4$, $x_1 * x_2 * x_2$, $x_1 * x_2 * x_4$, $x_1 * x_3 * x_3$, $x_2 * x_2 * x_3$, $x_2 * x_2 * x_4$, $x_2 * x_3 * x_3$, $x_2 * x_4 * x_4$, and $x_3 * x_3 * x_4$ are less than 0.05. This indicates these input parameters have a high relationship with the output response ($y$). In this period, the ARIMA (1, 1, 1) model was used to make the wind power forecast. In addition, a simple EXP with the smoothing $\alpha = 0.6$ was employed to predict the wind power. Based on the variance-covariance method, the proposed RSM–based wind power forecasting is combined with the ARIMA model and EXP to make better forecasting results. The five forecasting results from the RSM, ARIMA model, EXP, RSM+ARIMA, and RSM+EXP with the actual wind power are represented in Fig. 14.

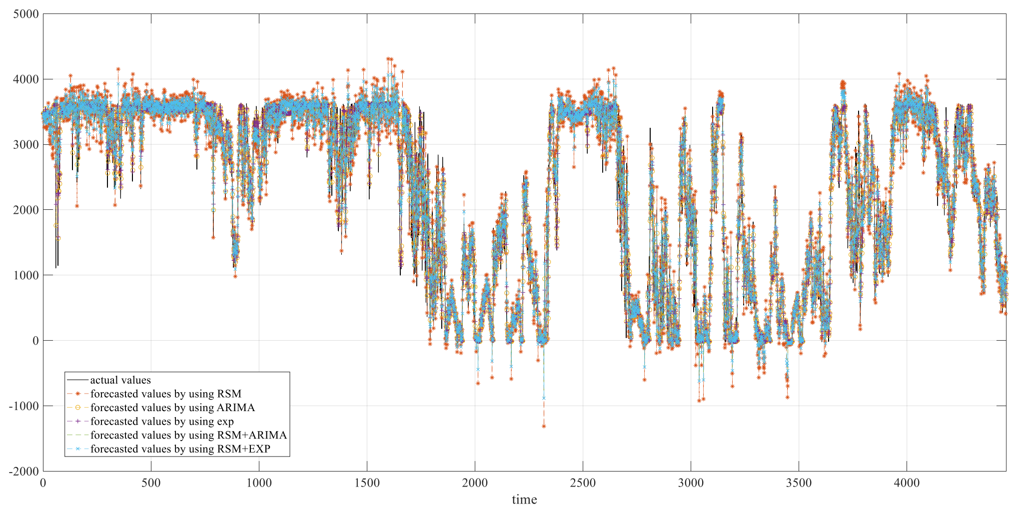


Fig. 14. The forecasted results of wind power generation

The actual wind power values are demonstrated by the solid line (black color). The forecasted values by using the RSM, ARIMA model, EXP, RSM and ARIMA model, as well as RSM and EXP are represented by the dash lines with a * maker (red color), o marker (yellow color), + marker (purple color), no marker (light green color), and a pentagram marker (light blue color), respectively. Most of the forecasted lines are approximate to the actual value line. The evaluation criteria of the RSM, ARIMA model, EXP, RSM+ARIMA, and RSM+EXP in wind power forecasting are illustrated in Table 4.

Based on the results in Table 4, the proposed RSM-based forecasting method can provide better prediction results (with MAPE = 83.3192% and $D_{\text{stat}}$ = 70.4709%) compared to the traditional statistical forecasting methods consisting of the ARIMA model (with MAPE = 88.2990% and $D_{\text{stat}}$ = 59.2825%) and EXP (with MAPE = 85.8814% and $D_{\text{stat}}$ = 59.0807%) in this period. Furthermore, the combined forecasting methods based on variance-covariance method consisting

Table 4. Evaluation criteria of the proposed forecasting models and methods

| **Evaluation criteria** | **RSM** | **ARIMA** | **Exp** | **RSM+ARIMA** | **RSM+exp** |
|---|---|---|---|---|---|
| MAPE (%) | 83.3192 | 88.2990 | 85.8814 | 75.3274 | 75.1407 |
| $D_{\text{stat}}$ (%) | 70.4709 | 59.2825 | 59.0807 | 70.1794 | 70.1570 |

of RSM and ARIMA (with MAPE = 75.3274% and $D_{\text{stat}}$ = 70.1794%) and RSM and EXP (with MAPE = 75.1407% and $D_{\text{stat}}$ = 70.1570%) can provide more accurate forecasting results compared to the others. Obviously, the proposed forecasting method consisting of RSM and ARIMA is the best wind power forecasting regarding MAPE and $D_{\text{stat}}$ in this period.

## 5. Conclusions

In this paper, RSM is proposed as an alternative forecasting method for wind power generation. Wind power is forecasted by considering the relational relationship between operational parameters (wind speed, nacelle position, pitch angle, and ambient temperature) and wind power, instead of time-series data, in this method. In addition, two popular statistical forecasting methods consisting of the ARIMA model and EXP are used to make the time-series wind power data. By using the variance-covariance method, RSM is combined with the ARIMA model and EXP, respectively. The combined forecasting methods consisting of RSM+ARIMA and RSM+ the exponential smoothing method show more accurate prediction results in terms of MAPE and $D_{\text{stat}}$ compared to the only RSM or ARIMA or exponential smoothing method. For further studies, all of the factors that affect wind power generation can be used such as air density, turbine swept areas, etc. In addition, the forecasting results can be improved by using the neural network, support vector machine, and deep machine learning-based forecasting methods with a suitable period dividing algorithm. Moreover, the missing value issues in wind power forecasting can be investigated. Other evaluation criteria can also be used to identify better forecasting methods.

**Acknowledgements**

This research is conducted within the framework of science and technology projects at the institutional level of Quy Nhon University under the project code T2021.713.16.

**References**

[1] Ren G., Liu J., Wan J., Guo Y., Yu D., *Overview of wind power intermittency: Impacts, measurements, and mitigation solutions*, Applied Energy, vol. 204, pp. 47–65 (2017), DOI: 10.1016/j.apenergy.2017.06.098.

[2] https://gwec.net/global-wind-report-2019/, accessed March 2021.

[3] Lerner J., Grundmeyer M., Garvert M., *The importance of wind forecasting*, Renewable Energy Focus, vol. 10, no. 2, pp. 64–6 (2009).

[4] Li G., Shi J., *On comparing three artificial neural networks for wind speed forecasting*, Applied Energy, vol. 87, pp. 2313–2320 (2010).

[5] Chang W.Y., *A literature review of wind forecasting methods*, Journal of Power and Energy Engineering, vol. 2, no. 04, pp. 161–168 (2014), DOI: 10.4236/jpee.2014.24023.

[6] Barbosa de Alencar D., de Mattos Affonso C., Limão de Oliveira R.C., Moya Rodriguez J.L., Leite J.C., Reston Filho J.C., *Different models for forecasting wind power generation: Case study*, Energies, vol. 10, no. 12, 1976 (2017), DOI: 10.3390/en10121976.

[7] Kusiak A., Zhang Z., *Short-horizon prediction of wind power: A data-driven approach*, IEEE Transactions on Energy Conversion, vol. 25, no. 4, pp. 1112–1122 (2010), DOI: 10.1109/TEC.2010.2043436.

[8] Zhu X., Genton M.G., *Short-term wind speed forecasting for power system operations*, International Statistical Review, vol. 80, no. 1, pp. 2–23 (2012).

[9] Jónsson T., Pinson P., Nielsen H.A., Madsen H., *Exponential smoothing approaches for prediction in real-time electricity markets*, Energies, vol. 7, no. 6, pp. 3710–3732 (2014), DOI: 10.3390/en7063710.

[10] Hodge B.M., Zeiler A., Brooks D., Blau G., Pekny J., Reklatis G., *Improved wind power forecasting with ARIMA models*, Computer Aided Chemical Engineering, vol. 29, pp. 1789–1793 (2011), DOI: 10.1016/B978-0-444-54298-4.50136-7.

[11] Chen P., Pedersen T., Bak-Jensen B., Chen Z., *ARIMA-based time series model of stochastic wind power generation*, IEEE Transactions on Power Systems, vol. 25, no. 2, pp. 667–676 (2009).

[12] Kavasseri R.G., Seetharaman K., *Day-ahead wind speed forecasting using f-ARIMA models*, Renewable Energy, vol. 34, no. 5, pp. 1388–1393 (2009), DOI: 10.1016/j.renene.2008.09.006.

[13] Torres J.L., Garcia A., De Blas M., De Francisco A., *Forecast of hourly average wind speed with ARMA models in Navarre (Spain)*, Solar energy, vol. 79, no. 1, pp. 65–77 (2005), DOI: 10.1016/j.solener.2004.09.013.

[14] Sfetsos A., *A novel approach for the forecasting of mean hourly wind speed time series*, Renewable energy, vol. 27, no. 2, pp. 163–174 (2002), DOI: 10.1016/S0960-1481(01)00193-8.

[15] Dumitru C.D., Gligor A., *Daily average wind energy forecasting using artificial neural networks*, Procedia Engineering, vol. 181, pp. 829–836 (2017), DOI: 10.1016/j.proeng.2017.02.474.

[16] Peiris A.T., Jayasinghe J., Rathnayake U., *Forecasting Wind Power Generation Using Artificial Neural Network:"Pawan Danawi"—A Case Study from Sri Lanka*, Journal of Electrical and Computer Engineering (2021), DOI: 10.1155/2021/5577547.

[17] Liu Y., Zhang H., *An empirical study on machine learning models for wind power predictions*, In 2016 15th IEEE International Conference on Machine Learning and Applications (ICMLA), pp. 758–763 (2016), DOI: 10.1109/ICMLA.2016.0135.

[18] Ranganayaki V., Deepa S.N., *An intelligent ensemble neural network model for wind speed prediction in renewable energy systems*, The Scientific World Journal (2016), DOI: doi.org/10.1155/2016/9293529.

[19] Sapronova A., Meissner C., Mana M., *Short time ahead wind power production forecast*, Journal of Physics: Conference Series, vol. 749, no. 1, 012006 (2016), DOI: 10.1088/1742-6596/749/1/012006.

[20] López E., Valle C., Allende-Cid H., Allende H., *Comparison of recurrent neural networks for wind power forecasting*, In Mexican Conference on Pattern Recognition, pp. 25–34, Springer (2020), DOI: 10.1007/978-3-030-49076-8_3.

[21] Manero J., Béjar J., Cortés U., *Predicting wind energy generation with recurrent neural networks*, In International Conference on Intelligent Data Engineering and Automated Learning, pp. 89–98, Springer (2018), DOI: 10.1007/978-3-030-03493-1_10.

[22] Liu B., Zhao S., Yu X., Zhang L., Wang Q., *A novel deep learning approach for wind power forecasting based on WD-LSTM model*, Energies, vol. 13, no. 18, pp. 4964 (2020), DOI: 10.3390/en13184964.

[23] Dong D., Sheng Z., Yang T., *Wind power prediction based on recurrent neural network with long short-term memory units*, In 2018 International Conference on Renewable Energy and Power Engineering (REPE), pp. 34–38 (2018), DOI: 10.1109/REPE.2018.8657666.

[24] Zeng J., Qiao W., *Support vector machine-based short-term wind power forecasting*, In 2011 IEEE/PES Power Systems Conference and Exposition, pp. 1–8 (2011), DOI: 10.1109/PSCE.2011.5772573.

[25] Wang J., Sun J., Zhang H., *Short-term wind power forecasting based on support vector machine*, In 2013 5th International Conference on Power Electronics Systems and Applications (PESA), pp. 1–5 (2013), DOI: 10.1109/PESA.2013.6828211.

[26] Li L.L., Zhao X., Tseng M.L., Tan R.R., *Short-term wind power forecasting based on support vector machine with improved dragonfly algorithm*, Journal of Cleaner Production, vol. 242, 118447 (2020), DOI: 10.1016/j.jclepro.2019.118447.

[27] Popławski T., Szeląg P., Bartnik R., *Adaptation of models from determined chaos theory to short-term power forecasts for wind farms*, Bulletin of the Polish Academy of Sciences: Technical Sciences, vol. 68, no. 6, pp. 1491–1501 (2020), DOI: 10.24425/bpasts.2020.135400.

[28] Zhu L., Shi H., Ding M., Gao T., Jiang Z., *Wind power prediction based on the chaos theory and the GABP neural network*, In 2019 IEEE Innovative Smart Grid Technologies-Asia (ISGT Asia), pp. 4221–4224 (2019), DOI: 10.1109/ISGT-Asia.2019.8881549.

[29] Wang C., Zhang H., Fan W., Fan X., *A new wind power prediction method based on chaotic theory and Bernstein Neural Network*, Energy, vol. 117, pp. 259–271 (2016), DOI: 10.1016/j.energy.2016.10.041.

[30] Yuan X., Tan Q., Lei X., Yuan Y., Wu X., *Wind power prediction using hybrid autoregressive fractionally integrated moving average and least square support vector machine*, Energy, vol. 129, pp. 122–137 (2017), DOI: 10.1016/j.energy.2017.04.094.

[31] Lau A., McSharry P., *Approaches for multi-step density forecasts with application to aggregated wind power*, The Annals of Applied Statistics, vol. 4, no. 3, pp. 1311–1341 (2010), DOI: 10.1214/09-AOAS320.

[32] Hwang M.Y., Jin C.H., Lee Y.K., Kim K.D., Shin J.H., Ryu K.H., *Prediction of wind power generation and power ramp rate with time series analysis*, In 2011 3rd International Conference on Awareness Science and Technology (iCAST), pp. 512–515 (2011), DOI: 10.1109/ICAwST.2011.6163182.

[33] Reddy V., Verma S.M., Verma K., Kumar R., *Hybrid approach for short term wind power forecasting*, In 2018 9th International Conference on Computing, Communication and Networking Technologies (ICCCNT), pp. 1–5 (2018), DOI: 10.1109/ICCCNT.2018.8494034.

[34] Jiang Y., Xingying C.H.E.N., Kun Y.U., Yingchen L.I.A.O., *Short-term wind power forecasting using hybrid method based on enhanced boosting algorithm*, Journal of Modern Power Systems and Clean Energy, vol. 5, no. 1, pp. 126–133 (2017), DOI: 10.1007/s40565-015-0171-6.

[35] Landberg L., *A mathematical look at a physical power prediction model*, Wind Energy, vol. 1, no. 1, pp. 23–28 (1998).

[36] Chang W.Y., *A literature review of wind forecasting methods*, Journal of Power and Energy Engineering, vol. 2, no. 04, pp. 161–168 (2014), DOI: 10.4236/jpee.2014.24023.

[37] Hanifi S., Liu X., Lin Z., Lotfian S., *A critical review of wind power forecasting methods—past, present and future*, Energies, vol. 13, no. 15, 3764 (2020), DOI: 10.3390/en13153764.

[38] Box G.E., Wilson K.B., *On the experimental attainment of optimum conditions*, Journal of the Royal Statistical Society. Series B. Methodological, vol. 13, pp. 1–45 (1951).

[39] Myers R.H., Montgomery D.C., Anderson-Cook C.M., *Response surface methodology: process and product optimization using designed experiments*, John Wiley & Sons (2016).

[40] Makridakis S., Winkler R.L., *Averages of forecasts: Some empirical results*, Management science, vol. 29, no. 9, pp. 987–996 (1983).