ICIC Express Letters
Part B: Applications
Volume **17**, Number **5**, May **2026**

ICIC International ©2026 ISSN 2185-2766
pp. **449–458**

ESTIMATION PARAMETERS OF WIND SPEED AND DIRECTION FOR WIND TURBINE APPLICATIONS USING SARIMA METHOD

Syafaruddin[1,2,*], Yusri Syam Akil[1], Putri Aisyah Kurnia[1], Agusman[1]
Sri Mawar Said[1] and Anuar Mohamad[3]

[1]Department of Electrical Engineering
[2]Department of Informatics
Universitas Hasanuddin
Jl. Poros Malino Km. 6 Bontomarannu, Gowa 92127, Sulawesi Selatan, Indonesia
{ yusakil; srimawarsaid }@unhas.ac.id; { kurniapa23d; agusman17d }@student.unhas.ac.id
*Corresponding author: syafaruddin@unhas.ac.id

[3]Electrical Engineering Studies
Universiti Teknologi Mara, Cawangan
Pulau Pinang, Permatang Pauh 13500, Penang, Malaysia
anuarmoh@uitm.edu.my

Received July 2025; accepted September 2025

Abstract. *Forecasting is a statistical method that plays a crucial role in informed decision-making. This study forecasted wind speed and direction variables for wind turbine applications. The method used is one of the time series methods, namely SARIMA (Seasonal Autoregressive Integrated Moving Average), which uses past and present values of the studied variable to produce accurate short-term forecasts. The non-seasonal ARIMA method encountered a problem where the forecast results tended to be flat and ignore the seasonal component of the time series data. With the SARIMA method, there are five different historical data ranges forecasted to identify the optimal length of historical data in the SARIMA method, which contributes less accuracy in predicting wind speed and direction with long historical data. This method can only accommodate one seasonal pattern, whereas many patterns will be formed if there is too much historical data. This makes it difficult to determine which pattern is the most influential and may ignore other patterns affecting the forecasting results. In this study, the trials of the SARIMA method are applied to wind power locations in Jeneponto and Sidrap. Based on the forecasting performance metrics (MSE, RMSE, MAD, and MAPE), the accuracy of the SARIMA method is highly dependent on the data's length and the data's variation for predictions. In this respect, it is obtained that the less historical data used, the more accurate the forecasting results with the SARIMA method.*

Keywords: Statistical method, Wind speed, Wind direction, SARIMA method, Estimation parameters

1. **Introduction.** Wind energy is one of the renewable energy sources that can significantly contribute to domestic electrical energy needs, especially in remote areas. The use of wind energy is popular in certain countries such as Germany, China, the United States of America, Spain, and India. The country utilizes wind energy because of its favourable geographical location and wide coastline. The wind power plant is a pollution-free energy source where the wind source is available anywhere; therefore, the plant can solve the environmental problems and the lack of electricity. Nowadays, the prospect of developing wind power technology is still very high. However, one of the obstacles to the operation of wind power is the erratic changes in wind speed and direction [1]. Therefore, it becomes another challenge and problem in wind plant operation.

DOI: 10.24507/icicelb.17.05.449

Forecasting tasks play a crucial role in informed decision-making. Forecasting predicts what will happen in the future based on relevant data and past data information with a mathematical model. One of the methods used in forecasting is the time series-based data method, based on past information from variables or past errors [2]. Forecasting can also be a subjective intuition prediction or a combination of mathematical models adjusted with good judgment [3]. In this respect, output power forecasting in wind power generation is an important component of an overall electrical energy management system. Proper forecasting helps improve power system reliability, reduce generation costs, and schedule maintenance plans [4].

Wind turbines, as one of the energy-producing tools, have several characteristics highly dependent on the wind speed. According to the basic equation of mechanical output power i.e., $P_0 = \frac{1}{2}\rho A\nu^3$, it is shown that wind energy output is the cubic multiplication of the wind speed. The faster the wind speed, the greater the power generated until saturation occurs at a certain wind speed, where the output power is forced to be constant to avoid mechanical and electrical failures in the wind turbine. The faster the wind speed, the greater the rotational speed of the rotor produced for the excellent energy output. Wind with a speed of 4 m/s is considered the minimum wind speed for rotating a turbine with a small power generation capacity. In comparison, 25 m/s is the maximum wind speed that does not interfere with the strength of the wind turbine structure.

The quality of wind speed is measured with the Beaufort scale, an empirical measure related to wind speed for observing land or sea conditions. This scale was invented by Francis Beaufort in 1805. Beaufort measures wind speed by describing its effect on the speed of ships and ocean waves. The Beaufort scale uses numbers and symbols. The larger the number on the Beaufort scale, the stronger the wind will blow, and it can even be more destructive. The Beaufort scale starts from 1 for the calmest wind gusts to 12 for the most destructive gusts. The Beaufort scale remains useful for wind scale today [5]. This measurement approach is also suitable for finding the optimal location of a wind farm based on the quality of wind speed and direction.

Some previous works on wind speed forecasting from different countries can be found in the literature. For example, in [6], wind speed in Malaysia is predicted using ARIMA and Simple Exponential Smoothing (SES) models. The hybrid model provides a positive outcome for predicting wind speed compared to the single model of ARIMA and SES. The model performs better at capturing the dynamic behaviour of the studied meteorological variable. In [7], the authors propose a cluster-based technique for wind speed forecasting. The results, using 12 different datasets from the NREL site, C-ARIMA, and C-GAS (models with a clustering technique), show better performance than the classical ARIMA and GAS models. In [8], a seasonal ARIMA model is proposed for short-term wind speed forecasting. The data was collected from a wind farm in Morocco for three different months in 2018. The model gives good results for an hour-ahead forecast with a prediction error below 16%. [9] compares ARIMA and NARX models for wind speed forecasting using measurement data from two sites in Mexico. It is mentioned that even though both models give satisfactory results, the NARX model shows better performance, as indicated by the applied error measures.

Another study in [10] uses the fractional-ARIMA (f-ARIMA) to forecast wind speed up to 48 hours ahead using data from four sites in North Dakota. It is noted that as compared to the persistence model, the f-ARIMA gives better results as it can improve the forecasting accuracy by around 42%. In [11], the ARIMA model is applied for wind speed and direction forecasting using 7 days of data from one site in Western Australia. Obtained forecasting errors for wind speed and direction are less than 5% and 16%, respectively. Recently, in [12], a time series model including SARIMA is proposed for wind speed forecasting of Ras-Gharib in Egypt using daily and monthly data. The authors apply MAPE, MAE, RMSE, and PR-value as error metrics to select the best models. Next, [13]

gives a comparative study on several methods (statistical, traditional machine learning, and deep learning models) of wind speed forecasting using fifteen minutes of data from two wind farms in Gansu province. It is mentioned that deep learning models, particularly Convolutional Neural Networks, give the best prediction.

Many studies mainly focus on wind speed forecasting [11], such as in [6-10,12,13]. Moreover, variations in data and observed locations affect the applicability and performance of wind speed forecasting using certain methods [13]. In line with [14], located around the equator, Indonesia experiences daily and seasonal variations in wind speed. This study considers the estimation of variables that are the basis for controlling wind turbines in the form of wind speed and direction, so that the wind turbine can work optimally to achieve a higher rotor speed and optimal electricity energy production. As the unpredictable nature of wind speed and direction, the wind speed and direction are then forecasted using the Seasonal Autoregressive Integrated Moving Average (SARIMA) method. The simulation results indicate that the grouping data can improve forecasting accuracy by accommodating only one seasonal pattern.

The article is organized as follows. Section 1 presents literature and critical reviews about the utilization of different models of ARIMA method for wind speed forecasting, including finding research gaps why this model extends to seasonal ARIMA to obtain more accurate results. The section follows with a mathematical model of ARIMA and SARIMA methods, including the performance index equation for validation. Section 3 shows the testing scenario without data grouping with conventional ARIMA and data grouping with a SARIMA method then followed by Section 4 with a discussion of results. The conclusion in Section 5 includes the future approach for this research.

2. **Mathematical Models of ARIMA Method.**

The ARIMA model consists of three basic steps, namely the identification stage, the assessment and testing stage, and the diagnostic examination. Furthermore, the ARIMA model can be used for forecasting if the model obtained is adequate. The purpose of this model is to determine a good statistical relationship between the predicted variable and the historical value of the variable so that forecasting can be done with the model presented in [15].

ARIMA is also known as the Box-Jenkins time series method. ARIMA has very good accuracy for short-term forecasting, while for long-term forecasting, the accuracy of forecasting is not good. Usually, it will tend to be flat (flat/constant) for a fairly long period. The Autoregressive Integrated Moving Average (ARIMA) model completely ignores the other variables when making forecasts. ARIMA uses past and present values of the studied variable to produce accurate short-term forecasts. ARIMA is suitable if the observations from the time series data are statistically related.

It is worth noting that most time series are non-stationary and that the AR and MA aspects of the ARIMA model only deal with stationary time series. Stationarity means that there is no growth or decline in the data. The data should be roughly horizontal along the time axis. In other words, fluctuations in the data are around a constant mean value, independent of time, and the variance of these fluctuations remains essentially constant over time. A time series that is not stationary must be converted into stationary data by differencing. What is meant by differencing is calculating changes or differences in the value of observations. The difference value obtained is checked again to determine whether it is stationary. If it is not stationary, then differencing is done again. If the variance is not stationary, then a logarithmic transformation is performed.

2.1. **Classification of ARIMA models.**

The acronym ARIMA stands for Autoregressive Integrated Moving Average. Lags of the stationarized series in the forecasting equation are called "autoregressive" terms (p), lags of the forecast errors are called "moving average" terms (q), and a time series which needs to be differenced to be made stationary

is said to be an "integrated" version of a stationary series (d). A non-seasonal ARIMA model is classified as an "ARIMA (p, d, q)" model, where

- p is the number of autoregressive terms,
- d is the number of nonseasonal differences needed for stationarity, and
- q is the number of lagged forecast errors in the prediction equation.

The general form of an *autoregressive model* with order p or ARIMA model $(p, 0, 0)$ is stated as follows:

$$Y_t = \mu + \phi_1 Y_{t-1} + \phi_2 Y_{t-2} + \cdots + \phi_p Y_{t-p} \tag{1}$$

Meanwhile, the general form of the *moving average model of* order q or ARIMA $(0, 0, q)$ is stated as follows:

$$Y_t = e_t - \theta_1 e_{t-1} - \theta_2 e_{t-2} - \cdots - \theta_q e_{t-q} \tag{2}$$

where

Y_t = time series value at time t
$Y_{t-1}, \ldots, Y_{t-p}$ = actual value at time t, $lag\,1, \ldots, lag\,p$
μ = constant parameter
e_t = noise error term at time t
$\phi_1, \ldots, \phi_p$ = autoregressive parameter $1^{th}, \ldots, p^{th}$
$\theta_1, \ldots, \theta_q$ = moving average parameter $1^{th}, \ldots, q^{th}$
$e_{t-1}, \ldots, e_{t-q}$ = error at time $t - q$, $lag\,1, \ldots, lag\,q$

If the non-seasonal difference for stationary is ignored, for example, the ARIMA $(1, 0, 1)$ can be considered as follows.

$$Y_t = \phi_1 Y_{t-1} + e_t - \theta_1 e_{t-1} \tag{3}$$

where

$$AR(1) = (1 - \phi_1 B) Y_t \tag{4}$$

$$MA(1) = \mu + (1 - \theta_1 B) e_t \tag{5}$$

In comparison, if the non-seasonal difference for stationary is added, for instance, the ARIMA $(1, 1, 1)$ can be considered as follows.

$$(1 - B)(1 - \phi_1 B) Y_t = \mu + (1 - \theta_1 B) e_t \tag{6}$$

where B is the backshift operator; $\left(B = \frac{Y_{t-1}}{Y_t}\right)$.

2.2. **Seasonal ARIMA models.** Seasonality is defined as a pattern that repeats itself at regular intervals. For *stationary data*, seasonal factors can be determined by identifying the autocorrelation coefficients at two or three *time lags* significantly different from zero. Autocorrelation that is significantly different from zero indicates the presence of a pattern in the data. One must look at the high autocorrelation to recognize the presence of seasonal factors.

To deal with seasonality, the short common notation is

$$ARIMA\,(p, d, q)(P, D, Q)^S \tag{7}$$

where

(p, d, q) = non-seasonal part of the model
(P, D, Q) = seasonal part of the model
S = number of periods per season

The notation used in seasonal ARIMA is easy and common, as extended from Equation (6). For example, the model of ARIMA $(0, 1, 1)(0, 1, 1)^{12}$ is described as follows:

$$(1 - B)\left(1 - \phi_1 B^{12}\right) Y_t = (1 - \theta_1 B)\left(1 - \phi_1 B^{12}\right) e_t \tag{8}$$

However, to use it in forecasting requires an elaboration of the equation and making it a more general regression equation. For the above model, the form is

$$Y_t = Y_{t-1} + Y_{t-12} - Y_{t-13} - \theta_1 e_{t-1} - \phi_1 e_{t-12} + \theta_1 \phi_1 e_{t-13} \tag{9}$$

To predict one period in the future, namely $\hat{Y}_t$ or Y_{t+1}, then as in the following equation:

$$\hat{Y}_t = Y_{t+1} = Y_t + Y_{t-11} - Y_{t-12} + e_{t+1} - \theta_1 e_t - \phi_1 e_{t-11} + \theta_1 \phi_1 e_{t-12} \tag{10}$$

The value of e_{t+1} will not be known because the expected value for the random error in the future must be set equal to zero. However, from the *fitted model*, the values of e_t, e_{t-11}, e_{t-12} can be replaced with their empirically determined values (obtained after the Marquardt algorithm's last iteration). Of course, if the forecast is far ahead, it will not get an empirical value for "e" after some time, and therefore, its expected value will be entirely zero.

The identification process of the seasonal model depends on statistical tools in the form of autocorrelation and partial autocorrelation, as well as knowledge of the system (or process) being studied. There are two basic ways to get these parameters:

a) By *trial and error*, testing several different values and selecting one (or a set of values if there is more than one parameter to be estimated) minimizes the *sum of squared residuals*.
b) Repair iteratively, selecting an initial estimate and then letting the computer program refine iteratively.

The model is considered good if the *error value* is random, meaning it no longer has a certain *pattern*. In other words, the obtained model effectively captures the existing data patterns. To examine the randomness of the *error value*, a test is conducted on the value of the autocorrelation coefficient of *error* using one of the following two statistical equations.

1) *Q Box* and *Pierce* test

$$Q = n' \sum_{k=1}^{m} r_k^2 \tag{11}$$

2) *Ljung-Box* test:

$$Q = n'(n'+2) \sum_{k=1}^{m} \frac{r_k^2}{(n'-k)} \tag{12}$$

where

$n' = n - (d + SD)$
d = order of difference, not seasonal factor
D = order of differentiation of seasonal factors
S = number of periods per season
m = maximum time *lag*
r_k = autocorrelation for *time lag* $1, 2, 3, 4, \ldots, k$
Testing Criteria:
If $Q \leq X^2(\alpha, df)$, the *error value* is random (the model is acceptable)
If $Q > X^2(\alpha, df)$, the *error value* is not random (the model cannot be accepted)
where
X^2 = *Chi-Square*
α = a given significance level
df = degrees of freedom

2.3. **Best model selection.** The best model can be determined by the average value of the percentage of forecasting error over a determined period, such as Mean Absolute Percentage Error (MAPE) by the following equation.

$$MAPE = \frac{\sum_{t=1}^{T} \frac{\left|Y_t - \hat{Y}_t\right|}{Y_t}}{T} \times 100\% \tag{13}$$

The other performance indexes, such as Mean Squared Error (MSE), Root Mean Squared Error (RMSE), and Mean Absolute Deviation (MAD) to measure the accuracy of forecasting results, are shown in equations as follows.

$$MSE = \frac{1}{n} + \sum_{t=1}^{n} \left(Y_t - \hat{Y}_t\right)^2 \tag{14}$$

$$RMSE = \sqrt{\frac{1}{n} + \sum_{t=1}^{n} \left(Y_t - \hat{Y}_t\right)^2} \tag{15}$$

$$MAD = \frac{1}{n} \sum_{t=1}^{n} \left|Y_t - \hat{Y}_t\right| \tag{16}$$

where

Y_t = actual value at time t
$\hat{Y}_t$ = estimated value at time t
T = number of forecasting/estimating periods
n = number data of observations.

3. **Scenarios of Simulation Results.** Simulation planning is carried out for the input parameters of the wind turbine, namely, simulation calculations to estimate wind speed and direction parameters using the SARIMA (Box-Jenkins) method based on the wind conditions around the wind turbine. The data scenario is the wind speed data every hour at an altitude of 150 m above sea level, which consists of 8760 wind speed and direction data points every hour in one year. This simulation aims to obtain information regarding the prediction of maximum wind turbine output power, or according to the turbine generator rating, based on environmental factors around the turbine, so that the yaw angle setting on the wind turbine becomes more controlled.

The simulation scenario carried out with the first step is to determine the stationarity of the wind speed and direction data every hour for 2 and 6 days, forecasting for 1 week and 1-month historical data, respectively, where this becomes the basis for determining the forecasting model to be used. The approach of forecasting simulation is divided into five, namely

- Using 11 months/8017-hour period data to forecast wind conditions for the next 2 days (48 data)
- Using 3 months/2160-hour period data to forecast wind conditions for the next 2 days (48 data)
- Using 1 month/719-hour period data to forecast wind conditions for the next 2 days (48 data)
- Using 2 weeks/336-hour period data to forecast wind conditions for the next 6 days (144 data)
- Using 1-week data/168-hour period data to forecast wind conditions for the next 6 days (144 data)

It can be seen in Table 1 that the SARIMA method is less accurate in predicting wind speed and direction with long historical data, as shown by the values of the used evaluation metrics, which are lowest for one week of historical data (excluding MAD). Conversely, this method is more accurate for less historical data. This is because the SARIMA method is designed for short-term forecasting (according to the theory). This method can only accommodate one seasonal pattern, whereas many patterns will be formed if there is too much historical data. As in this study, there are many patterns, such as every 4 hours, 6 hours, 12 hours, 24 hours, 3 days, weekly, monthly, and every 3 months. If the number

Table 1. Simulation results for various timescales

Performance metrics	Forecasting data			
	1-week historical data		1-month historical data	
	Speed (m/s)	Direction (Degree)	Speed (m/s)	Direction (Degree)
MSE	2.531	2.403	3.761	3.225
RMSE	1.591	1.550	1.939	1.796
MAD	2.236	2.265	3.124	3.256
MAPE	2.659	2.456	3.778	3.347
Performance metrics	3-months historical data		11-months historical data	
	Speed (m/s)	Direction (Degree)	Speed (m/s)	Direction (Degree)
MSE	4.765	4.256	5.734	5.217
RMSE	2.183	2.063	2.395	2.284
MAD	4.078	4.025	5.026	5.265
MAPE	4.987	4.875	5.879	6.327

of historical data increases, it is difficult to determine the pattern as the most influential data and to ignore other patterns that also affect the forecasting results.

4. **Discussion.** The application of the SARIMA model is based on the data collected from Windy.App data whose coordinates are set in the wind farm locations in Jeneponto and Sidrap. The data obtained is hourly data, which is sampled every 3 hours. This data is the latest recorded wind direction and speed data on 8-9 May 2021. In this simulation, the forecasting approach is made for a shorter time in an hour, than the initial approach of day interval prediction. This is due to the intention to foresee the SARIMA prediction capability for quick response at time intervals.

In this simulation, wind direction and speed data will be predicted for the next three hours using available historical data, which is a week. This data has met the requirements for use in the SARIMA method, classified as no data grouping and data grouping as follows. The results for no data grouping can be seen in Table 2. Wind characteristics that can be utilized in wind turbines are winds that flow regularly between 4 m/s to 25 m/s. Wind with a speed of 4 m/s is considered the minimum wind speed to rotate the turbine with a small power generation capacity, while 25 m/s is the maximum wind speed that does not interfere with the strength of the wind turbine structure. Based on the forecast results, the wind turbine can only operate before 14.00 on May 9, 2021, provided that the wind blows at a speed above 4 m/s continuously. Based on the wind turbine forecast results, the wind turbine will not operate for the next three hours because the wind speed is below 4 m/s.

The second simulation approach is conducted by grouping day data and night data, which will be forecasted separately in the hope of better forecasting results. This is done to reduce variations in data patterns, an obstacle for the SARIMA method, which can accommodate only one seasonal pattern. Meanwhile, in this data grouping approach, the before noon (A.M.) and afternoon (P.M.) are grouped, and forecasting will be carried out separately in the expectation of better results. This is conducted to reduce variations in data patterns, an obstacle for the SARIMA method, which can accommodate only one seasonal pattern. Table 3 compares the Actual (A) and Forecasted (F) results, along with data grouping, for the Jeneponto and Sidrap wind farm locations. There are very close values between actual and forecasted data, with and without data grouping (Table 2). Better forecasting results, indicated by lower error rates in performance indicators, such as MAPE indicator (less than 1.0), are obtained when data grouping is conducted using the SARIMA method, as summarized in Table 4. In addition, the lowest value of RMSE, which is the root of MSE indicates the high accuracy of SARIMA model. In

TABLE 2. Actual (A) and Forecasted (F) results for Jeneponto and Sidrap windfarm locations with no data grouping

Locations	Wind farm of Jeneponto					Wind farm of Sidrap				
Parameters	Wind speed (m/s)		Direction (Degree)		Output power	Wind speed (m/s)		Direction (Degree)		Output power
Time	A	F	A	F	(MW)	A	F	A	F	(MW)
08/05/2021 02:00	5.1	5.64	355	132	0.74	5.6	3.67	40	26.38	263.86
08/05/2021 05:00	5.5	5.49	25	44	8.99	7.9	7.44	295	239	239.56
08/05/2021 08:00	6.8	6.96	35	28	2.39	8.1	8.97	295	298	298.91
08/05/2021 11:00	6.0	6.06	135	156	3.27	7.6	7.17	195	187	187.45
08/05/2021 14:00	7.8	7.27	130	227	5.60	6.4	6.26	145	138	138.24
08/05/2021 17:00	8.6	8.02	135	176	5.83	7.4	7.13	95	145	145.89
08/05/2021 20:00	4.1	4.30	140	254	5.98	8.0	8.53	105	105	105.51
08/05/2021 23:00	7.7	7.70	165	109	1.49	6.1	6.90	265	219	219.76
09/05/2021 02:00	8.3	8.59	30	25	0.62	6.3	6.20	260	187	187.41
09/05/2021 05:00	8.0	8.05	310	325	5.75	7.5	7.84	275	174	174.32
09/05/2021 08:00	8.9	8.89	270	208	1.94	7.1	5.72	265	195	195.65
09/05/2021 11:00	8.4	8.50	255	149	9.12	8.1	8.78	255	167	167.84
09/05/2021 14:00	3.8	2.48	275	174	8.84	2.6	2.25	240	124	124.20
09/05/2021 17:00	3.0	2.18	265	1530	2.83	3.2	2.68	235	123	123.36
09/05/2015 20:00	3.2	3.76	265	209	1.43	2.0	2.38	270	205	105.32
09/05/2015 23:00	3.8	4.26	265	280	2.09	2.4	2.22	275	265	165.72

TABLE 3. Actual (A) and Forecasted (F) results for Jeneponto and Sidrap windfarm locations with data grouping

Time	Speed (m/s)		Direction (Degree)		Time	Speed (m/s)		Direction (Degree)	
	A	F	A	F		A	F	A	F
	Wind farm of Jeneponto					Wind farm of Sidrap			
01/05/2021 02:00	4	4.01	260	256	01/05/2021 14:00	5.3	5.29	290	289
01/05/2021 05:00	3.7	3.68	265	276	01/05/2021 17:00	6.3	6.28	285	284
01/05/2021 08:00	3.8	3.76	260	261	01/05/2021 20:00	5.9	6	285	284
01/05/2021 11:00	4.4	4.56	280	278	01/05/2021 23:00	5.7	5.67	265	266
02/05/2021 02:00	5.2	5.1	265	264	02/05/2021 14:00	7.1	7.09	280	278
02/05/2021 05:00	5.5	5.4	260	259	02/05/2021 17:00	7.4	7.38	290	289
02/05/2021 08:00	5.6	5.7	255	254	02/05/2021 20:00	7.6	7.75	285	287
02/05/2021 11:00	6.3	6.28	275	276	02/05/2021 23:00	6.7	6.81	270	270
03/05/2021 02:00	5.5	5.49	255	255	03/05/2021 14:00	7.9	7.82	295	297
03/05/2021 05:00	5.2	5.18	255	254	03/05/2021 17:00	8	7.99	290	289
03/05/2021 08:00	4.7	4.75	265	267	03/05/2021 20:00	8.2	8.12	280	278
03/05/2021 11:00	6	6.01	290	287	03/05/2021 23:00	7.1	7.12	265	266
04/05/2021 02:00	6.1	6.02	255	254	04/05/2021 14:00	7.3	7.21	285	278
04/05/2021 05:00	5.4	5.38	255	254	04/05/2021 17:00	8.1	8.09	275	274
04/05/2021 08:00	5.1	5.08	255	253	04/05/2021 20:00	6.9	6.78	275	274
04/05/2021 11:00	6.8	6.87	280	278	04/05/2021 23:00	6.9	6.88	270	268
05/05/2021 02:00	6.9	7.01	85	86	05/05/2021 14:00	7.5	7.66	95	94
05/05/2021 05:00	6.4	6.38	90	91	05/05/2021 17:00	7.4	7.45	95	95
05/05/2021 08:00	6.5	6.46	95	94	05/05/2021 20:00	6.1	6.08	95	94
05/05/2021 11:00	7.9	7.89	105	103	05/05/2021 23:00	6.2	6.17	95	95
06/05/2021 02:00	7.5	7.46	95	94	06/05/2021 14:00	6	6	95	94
06/05/2021 05:00	7.4	7.38	85	84	06/05/2021 17:00	5.3	5.23	105	106
06/05/2021 08:00	6.5	6.41	105	104	06/05/2021 20:00	4.2	4.18	80	81
06/05/2021 11:00	5.1	5.08	95	94	06/05/2021 23:00	4.3	4.27	85	83
07/05/2021 02:00	1.3	1.27	60	60	07/05/2021 14:00	1.9	1.89	145	144
07/05/2021 05:00	0.9	0.89	315	316	07/05/2021 17:00	1.3	1.23	135	140
07/05/2021 08:00	1.2	1.128	70	71	07/05/2021 20:00	0.5	0.44	10	8
07/05/2021 11:00	1.4	1.51	135	134	07/05/2021 23:00	1.9	1.89	355	356

TABLE 4. Performance indexes of forecasting results with and without data grouping

Performance metrics	No data grouping (Table 2)		With data grouping (Table 3)	
	Speed (m/s)	Direction (Degree)	Speed (m/s)	Direction (Degree)
MSE	3.721	4.203	0.756	0.926
RMSE	1.939	2.050	0.869	0.962
MAD	3.026	4.924	0.631	0.735
MAPE	4.487	4.433	0.971	0.302

comparison, the MAD values show more stability in terms of outliers than MSE results due to non-quadratic error performances.

5. **Conclusions.** One of the obstacles in operating wind turbines is the difficulty of predicting wind speed and direction. The SARIMA method is a short-term forecasting method. The prediction accuracy of the SARIMA method is highly dependent on the length and variation of the data. The SARIMA method can only accommodate one seasonal pattern, so it is still possible for a pattern to be neglected. Forecasting wind speed and direction for wind turbine applications using the SARIMA method is performed in two wind farm locations on Sulawesi Island. The simulation results indicate that the grouping data can improve forecasting accuracy. This can improve the performance of SARIMA by accommodating only one seasonal pattern. As future research, the statistical frequency of wind speed by Weibull, Nakagami, and Rayleigh distributions, is conducted to obtain more accurate wind speed data requirements for long-term utilization of wind power.

Acknowledgment. This collaborative research is conducted following the financial support from the research scheme of Thematic Research Group (TRG) 2025 with the contract No. 00518UN4.22/PT.01.03/2025 supported by *Lembaga Penelitian dan Pengabdian pada Masyarakat* (LPPM) of Universitas Hasanuddin.

REFERENCES

[1] A. P. Sari, H. Suzuki, T. Kitajima, T. Yasuno, D. A. Prasetya and A. Rabi, Deep convolutional long short-term memory for forecasting wind speed and direction, *SICE Journal of Control, Measurement, and System Integration*, vol.14, no.2, pp.30-38, 2021.

[2] H. F. J. Katuuk, Syafaruddin and Y. S. Akil, Wind speed calculation utilizing the Hellmann coefficient method and roughness length for wind turbine application, *The 7th International Seminar on Research of Information Technology and Intelligent Systems (ISRITI)*, pp.1030-1035, 2024.

[3] S. Alshamali, D. Alawadhi and E. Aljuwaiser, Nonlinear control design for maximum power point tracking of a PMSG-based wind turbine, *International Journal of Innovative Computing, Information and Control*, vol.20, no.1, pp.313-328, 2024.

[4] B. M. Ali et al., Sustainable strategies for preventive maintenance and replacement in photovoltaic power systems: Enhancing reliability, efficiency, and system economy, *Unconventional Resources*, vol.6, pp.1-13, 2025.

[5] L. Hasse, *Beaufort Wind Scale in Encyclopedia of Atmospheric Sciences*, Academic Press, 2003.

[6] N. A. B. Kamisan, M. H. Lee, S. F. Hassan, S. M. Norrulashikin, M. E. Nor and N. H. A. Rahman, Forecasting wind speed data by using a combination of ARIMA model with single exponential smoothing, *Mathematical Modelling of Engineering Problems*, vol.8, no.2, pp.207-212, 2021.

[7] V. Kushwah, R. Wadhvani and A. K. Kushwah, Trend-based time series data clustering for wind speed forecasting, *Wind Engineering*, vol.45, pp.992-1001, 2020.

[8] I. Tyass, A. Bellat, A. Raihani, K. Mansouri and T. Khalili, Wind speed prediction based on seasonal ARIMA model, *E3S Web of Conferences*, vol.336, no.00034, pp.1-6, 2022.

[9] E. Cadenas, W. Rivera, R. Campos-Amezcua and C. Heard, Wind speed prediction using a univariate ARIMA model and a multivariate NARX model, *Energies*, vol.9, no.109, pp.1-15, 2016.

[10] R. G. Kavasseri and K. Seetharaman, Day-ahead wind speed forecasting using f-ARIMA models, *Renewable Energy*, vol.34, pp.1388-1393, 2009.

[11] E. Yatiyana, S. Rajakaruna and A. Ghosh, Wind speed and direction forecasting for wind power generation using ARIMA model, *IEEE 2017 Australasian Universities Power Engineering Conference (AUPEC)*, pp.1-6, 2018.

[12] O. A. El-Kashty, A. A. Daoud and E. E. El-Araby, Forecasting of short-term and long-term wind speed of Ras-Gharib using time series analysis, *International Journal of Renewable Energy Research*, vol.13, no.1, pp.258-272, 2023.

[13] X. Li, K. Li, S. Shen and Y. Tian, Exploring time series models for wind speed forecasting: A comparative analysis, *Energies*, vol.16, no.7785, pp.1-22, 2023.

[14] H. Suyono, D. O. Prabawanti, M. Shidiq, R. N. Hasanah, U. Wibawa and A. Hasibuan, Forecasting of wind speed in Malang city of Indonesia using adaptive neuro-fuzzy inference system and autoregressive integrated moving average methods, *IEEE 2020 International Conference on Technology and Policy in Energy and Electric Power (ICT-PEP)*, pp.131-136, 2020.

[15] Central Bureau of Statistics, *daps.bps.go.id*, https://daps.bps.go.id/file_article/77/arima.pdf, Accessed on March 9, 2021.