Wind Speed Forecasting by Using Time Series

Feyza TURAN
Department of Electrical Engineering, Yildiz Technical University, Turkey
Email: feyzaozcan@gmail.com

Bedri KEKEZOĞLU
Department of Electrical Engineering, Yildiz Technical University, Turkey
Email: bkekez@yildiz.edu.tr

Abstract: As a result of the rapid development of technology and consequently increasing energy demand, the world's underground resources have started to run out. Global warming, along with the damage to the environment, has become a bigger problem. This has increased the interest in renewable energy sources, called clean energy. However, some problems have arisen in the integration of renewable energy sources into the power system. The biggest problem encountered in this regard is that the power outputs of these systems cannot be known with absolute accuracy. This is because renewable energy generation systems are directly connected to the environmental parameters (pressure, temperature, humidity, wind direction, ghosting effect, etc.). There are many methods and many researches that have been developed in order to estimate how much renewable energy sources will produce within the desired time. In this study, information about time series methods, AR, MA, ARMA and ARIMA methods were examined. In order to estimate the wind speed, 3 separate regions, 3 different estimation time scales and 3 different time zones were selected. The wind speed was analyzed in detail, and the time series were evaluated to be efficient and not efficient.

Keyword: ARIMA; Time Series Analysis; Wind Energy; Wind Speed Forecast

1. INTRODUCTION

Nowadays, electrical energy has become a constituent of modern life. The technologies of electricity generation have diversified by the time. Conventional electricity generation systems source on fossil fuels. Considering the depletion rate of fossil fuels, it is possible to encounter a serious crisis in the near future. In addition, carbon emissions and high cost are the disadvantages of traditional energy production [1]. This has revealed the necessity of finding an alternative source for electrical energy. For this reason, alternative electrical energy sources have been the most important research and application area of recent years. The so-called renewable sources of energy have become an alternative to traditional energy production. Especially wind power has come to the forefront with its advantages such as easy applicability, sufficient technology and environmental benefits. According to the records of 2018, approximately %7.7 of the installed capacity of Turkey is comprised of wind power plants [2].

Although there are many positive aspects of renewable energy generation, some problems arise in case of integration with power system. The first of these problems is the uncertain energy generation, depending on the environmental factors. Supply and demand balance in electricity energy system; pricing, planning, energy quality etc. reasons are very important [3].

Energy production uncertainty of renewable energy sources also makes energy management difficult. The solution to this problem is possible with an accurate estimate of the power that can be obtained from the wind. And wind speed forecasting is necessary to obtain a share in other energy production methods [4]. Some of the benefits of estimating wind speed are: short-term forecasting is used energy management, and the medium-term forecasting is used maintenance and repair time can be decided [5].

In the literature, many methods related to wind speed estimation are mentioned. Some of these are artificial neural networks, fuzzy logic, time series analysis.

Box-Jenkins, which makes time series analyzes more applicable for all disciplines, has been introduced to the literature with the methodology of estimating with non-stationary series [6]. Box and Jenkins interpreted the time series methods in an understandable way. Therefore, time series analysis is also known as the Box-Jenkins method. And then Gujarati in his book, helped scientists by explaining all the issues related to time series [7]. In the reference [8] information about the estimation methods are given with the time scale classification used in estimation and the estimation methods comparison table. The
reference [9] asserts the estimation method should be selected depending on the region where the wind data belongs. In [10], the study examined the differences in wind speed according to seasonality and geographical location. The study given in [11] reports that the short-term wind speed of the MA provides better results when evaluated in different time scales. The reference [12] has chosen ARIMA model as the best fit model to make a wind forecast for Karabük in Turkey. In [Jiang] estimated the decline in coal use by using time series because of the increased share of wind energy in electricity generation. In [13], ARIMA, ANFIS and artificial neural network models is compared and very short term estimation of ARIMA method is found to be successful with two other methods. On the other hand, it has been shown that the success rate decreases as the time interval increases. The reference [14] used the large wind speed data set by reducing and applied ARIMA and ANN models. This study said that different ARIMA models are very similar but they differ quite a bit with ANN. In [15] Naive model and ARIMA has been compared respect to error rates on power system management estimations and ARIMA model has given lower error rates. According to the results of the study [16], the hybrid use of artificial neural networks and ARIMA model gives much better results than the use of these two methods individually.

In this study, the wind speed variations of three different points in Marmara region, in Turkey, is estimated by using time series, via ARIMA method. The results obtained by three different approaches, which are short-term, medium-term and long-term, are presented and interpreted. The original values of the work can be listed as follows:

- Wind data of real regions are used within the study.
- The behavior of the ARIMA method against different wind speed variations is demonstrated over the error rates.
- The behavior of the ARIMA method against different prediction intervals was demonstrated by error rates.
- The effects of seasonal changes on estimated quantities are proved.

In the 2nd part of the study, time series are defined, mathematical analyzes and results obtained in section 3 are shared. The study is concluded in Section 4.

2. TIME SERIES

Time series is a set of ordered measurements of a size [17]. Time series are used to analyze the process, to better understand, to define and then to model. Thus, it is tried to estimate the future forecast of the series. If there is no systematic change in the mean and variance of time series, this time series is called stationary. However, the series encountered in real life are generally not stationary. The series has a trend and varies with the average of variance.

In these cases it is possible to analyze the series by ‘stationarizing’ [18]. Time series models are examined under 4 headings. Depending on whether the series is stationary and the character’s analysis, it is decided which method to use. The aim is to determine the linear stochastic (coincidental) model that best matches the series and simplifies them in the simplest way [12]. Time series are modeled with different structures in terms of the relationships between the data. These structures; AR, MA, ARMA, ARIMA.

2.1 AR (Autoregressive Model)

This model is called an autoregressive model if the future values can be estimated by adding the error term to the historical data of the time series. The AR model is shown in Equation 1.

\[ Y_t = \alpha_1 Y_{t-1} + \alpha_2 Y_{t-2} + \ldots + \alpha_p Y_{t-p} + \varepsilon_t \]  

In the equation, \( Y_t \) is wind speed values, \( \alpha \) is autoregressive parameter and \( \varepsilon_t \) is error term. Additionally, \( p \) represents the degree of the AR model.

2.2 MA (Moving Average)

If the future value of the series can be obtained by adding a constant term to the weighted average of the errors of historical values, this process is called the Moving Average. The mathematical expression of the MA model is presented in Equation 2.

\[ Y_t = \mu + \varepsilon_t + \theta_1 \varepsilon_{t-1} + \theta_2 \varepsilon_{t-2} + \ldots + \theta_q \varepsilon_{t-q} \]  

In this equation, \( \mu \) is fixed term, \( \varepsilon_t \) is error term and \( \theta \) is unknown parameters about MA process. Also \( q \) represents the degree of the MA model.

2.3 ARMA (Autoregressive Moving Average)

The time series is modeled with serial ARMA in this case if it contains both autoregressive and moving average processes for its relationships [19]. Mathematical expression is shown in Equation 3.

\[ Y_t = \theta_1 Y_{t-1} + \theta_2 Y_{t-2} + \ldots + \theta_p Y_{t-p} + \varepsilon_t + \rho_1 \varepsilon_{t-1} + \rho_2 \varepsilon_{t-2} + \ldots + \rho_q \varepsilon_{t-q} \]  

According to Equation 3, \( \theta \) is fixed term, \( \alpha \) and \( \beta \) are unknown parameters and \( \varepsilon_t \) is error term.

To model the series, look at whether it is stationary or not. Stationarizing should be performed if it is not stationary.

2.4 ARIMA (Autoregressive Integrated Moving Average)

If the non-stationary series can be stationarized by
taking the difference, it is modeled by the ARIMA process. In ARIMA (p, d, q), p specifies the number of the AR process, d, the difference, and the coefficient of the q, MA process. It is tried to determine the coefficients that reflect the series in the most successful way. The most appropriate p, d, q will give the best estimate result. Determined coefficients are written in the equation instead of the future value is reached.

3. WIND SPEED FORECAST WITH ARIMA

This study seasonal wind speed estimates 3 zones in the Marmara region of Turkey was held. In this study, ARIMA was preferred as the prediction method. The data set consists of real hourly values of wind speeds. Using the first 3-week data on the monthly data considered for each season in the forecasting process, the remaining data were estimated and compared with the actual data.

![Figure 1 Monthly forecast results for Zone 1](image1)

![Figure 2 Monthly forecast results for Zone 2](image2)
The 3-week data consists of 504 values in total. The results of the estimations performed within 3 different time scales are shown in Figure 1, Figure 2 and Figure 3, respectively.

The zone with the highest speed on the points considered within the scope of the study was identified as Zone 2. In this Zone, the maximum value of the wind speed exceeds 20m/s, while the maximum wind speed in Zone 1 is 10m/s and in Zone 3 it is 6.5m/s. The Box-Jenkins method was used in the estimation phase. Accordingly, the data sets were analyzed first when analyzing the data set, we first plotted the series using MATLAB or Eviews programs. And the partial autocorrelation coefficients of the autocorrelation were examined and it was decided whether the series was stationary or not.

![Figure 3 Monthly forecast results for Zone 3](image)

It was decided that the wind speed was not stationary. An attempt was made to stationarizing the series using the EViews program. Then, ARIMA coefficients were determined for each data set and the estimation was performed. And comparison graphs were drawn, where all the predicted results for each case were combined. As a result of the analyzes performed, it has been observed that the accuracy of the estimation accuracy increases as the time scale of the estimated size decreases. Some statistical error measurement equations are used to compare the results [20]. The average relative error rates of the different time scales for the regions examined are shown in Table 1, Table 2 and Table 3, respectively.

**TABLE I PERCENTAGE RELATIVE ERROR VALUES FOR FEBRUARY PERIOD**

<table>
<thead>
<tr>
<th>Time interval</th>
<th>Zone 1</th>
<th>Zone 2</th>
<th>Zone 3</th>
</tr>
</thead>
<tbody>
<tr>
<td>10 days</td>
<td>57.64</td>
<td>62.53</td>
<td>45.14</td>
</tr>
<tr>
<td>1 days</td>
<td>52.74</td>
<td>51.68</td>
<td>41.8</td>
</tr>
<tr>
<td>1 hours</td>
<td>36.42</td>
<td>25.69</td>
<td>32.67</td>
</tr>
</tbody>
</table>

**TABLE II PERCENTAGE RELATIVE ERROR VALUES FOR MARCH PERIOD**

<table>
<thead>
<tr>
<th>Time interval</th>
<th>Zone 1</th>
<th>Zone 2</th>
<th>Zone 3</th>
</tr>
</thead>
<tbody>
<tr>
<td>10 days</td>
<td>71.69</td>
<td>47.77</td>
<td>47.51</td>
</tr>
<tr>
<td>1 days</td>
<td>59.33</td>
<td>39.33</td>
<td>49.75</td>
</tr>
<tr>
<td>1 hours</td>
<td>36.82</td>
<td>16.76</td>
<td>33.71</td>
</tr>
</tbody>
</table>

The biggest error rates in the study are seen in Zone 1. On the other hand, the lowest error rates were obtained as a result of the estimations for Zone 2. It is found that the very variable of the wind profile of zone 2 affects the estimation accuracy. Again, as the time scale of the prediction accuracy decreases, the error rates are increased by decreasing the hourly estimates. The most accurate estimation on the examined regions was observed in the estimates of Zone 2 for March with a 16.76% error rate. On the other hand, the worst estimation scenario with 107.15% error rate was observed in the Zone 1 forecasts for June.
4. CONCLUSION

Wind speed estimation is extremely important for power generation estimation in terms of the quality-stability of the electricity market and the electrical power system. In many countries, incentive premiums are given to companies that are successful in the forecast point.

For this reason, it is very important for both the manufacturer and the energy supplier to estimate the wind speeds and to obtain the hourly output power results of the wind turbines. The uncertainty of how much power will be transferred to the grid is eliminated by wind speed estimation and a more secure system is achieved.

Within the scope of this study, different data sets were analyzed with different selected estimation intervals. ARIMA models were applied for each case. As a result of the study, it was observed that the efficiency increased in the short-term wind speed estimation. As the time interval increases, the error rates have increased.

In this study, it was determined that time series estimation for regions with stable wind speeds may be sufficient, but hybrid estimation methods should be used in order to increase the accuracy.

REFERENCES


Authors Biography

Feyza Turan received the B.Sc., the M.Sc. degree in electrical engineering Yıldız Technical University (YTU) in 2015 and 2019 respectively. Her research interests are renewable energy, alternative energy sources.
Bedri Kekezoğlu received the B.Sc., the M.Sc. and Ph.D. degree in electrical engineering Yıldız Technical University (YTU) in 2005, 2007, and 2013, respectively. He is currently working as associated professor in Electrical Engineering Department of YTU. His current research interests include energy distribution, renewable energy, power system analysis, railway systems.

Cite this paper: