Wind Power Forecasting Algorithm Based on Similarity of Multivariate Time Series

Hui-Ying JIN¹, Yong-Qiang YANG², Zhan-Feng WANG¹, Wei-Jun MA¹, Yong SU¹, and Yun-Peng PAN³

¹Nanjing Goldwater Shang Yang Information Technology Co., Ltd, Nanjing, 211101, China
²Yunnan Electric Power Dispatch and Control Center, Kunming, 650000, China
³Nanjing Sancha River Estuary Sluice Management Office, Nanjing, 210014, China

Corresponding author

Keywords: Wind power, Time series, Similarity, Prediction, Meteorological factors.

Abstract. An accurate prediction of wind power plants has great significance to the power leveling and stability of a power grid. While the electricity market gradually introducing competition mechanism, it could maximize the growth of economic benefits for electricity companies. From a complex system’s point of view, this paper introduces the forecasting of wind power by considering the power generation of wind plants as a black box system, using weighted average and analysis of the historical time series similarity between the meteorological forecast data and the relevant factors affecting wind power generation. The experiment has proof this algorithm simpler, more accurate and maneuverable compared than other wind power algorithms.

1. Introduction

With the continuous embedding of the green energy strategy in China, wind energy has been paid more and more attention as a renewable clean energy. In recent years, wind power generator price has gradually decreased and a large scale of commercial application has become a reality with the constant improvement in manufacturing wind power equipment. It is estimated that the wind power generation in China will reach approximately 15% of the global power grid around 2030. Due to the fluctuation of wind power, the rapid development of wind power provides not only a clean energy to the national economy, but also a great challenge to the dispatch management and difficulties in regulation control of power grids.

As the increasing practicability of artificial intelligence and big data, the accuracy of power prediction has also been improved. There are a lot of wind-power prediction algorithms, some of them use meteorological forecast data or not. Relatively speaking, the algorithm with usage of meteorological forecast data is more accurate and the forecast time is longer. In general, the accuracy of the wind-power prediction algorithms is impacted by various factors such as terrain, wind, location, air density, temperature, pressure, etc. In order to overcome these non-linear factors, researchers have applied the artificial neural networks to wind-power prediction. However, due to the uncertainty of the parameters, the prediction length is limited.
Over years, a study based on the similarity of time series via various measurements has received attentions[2,3]. If the wind-power plant is treated as a black box, its power could actually be considered as a response under a variety of input elements. Obviously if the input is the same, the output will also be the same. Based on this principle, this paper presents a wind-power prediction algorithm based on multiple time series similarity (WPFSMTS, Wind power forecasting algorithm based on similarity of multivariate time series). Its basic idea is to use the predicted time-series of meteorological elements as input, find similar sequences in historically observed data, and use the weighted average of the historical power time as the predicted power. Through the simulation of the actual data, it is found that this algorithm has the advantages of simple operation and high prediction precision, compared with the ARMA algorithm and the prediction algorithm of wavelet neural network.

This paper is structured as follows: Section 2 discusses the related work of wind-power prediction; Section 3 discusses the characteristics of wind power plant; Section 4 explains the main steps of WPFSMTS algorithm; Section 5 compares and explains the effect of algorithm by experiments; the last section is a summary.

2. Related Work

There are many wind power prediction algorithms, which can be classified by different taxonomy methods. For example, it can be categorized into long-term prediction algorithm, medium-term prediction algorithm, short-term prediction algorithm and ultra-short-term prediction algorithm. According to whether the physical mechanism is used, it could also be categorized into physical-cause method. In this paper, all the algorithms are categorized into forecasting methods based on using numerical weather-forecasting or not.

The method of wind-power prediction based on historical data only uses the continuity of time series to predict power, which is accurate in short time but with increasing prediction errors in long term. For instance, Kalman filter[4], persistence algorithm[5], ARMA algorithm[6,7], linear regression model[8], and adaptive fuzzy-logic algorithm.

The the numerical weather prediction(NWP) based models imploy the weather forecast data is used as the input to predict the future power. The general resolution of NWPs is tens of square kilometers, and usually further interpolation is needed to improve the prediction accuracy. This method can be broadly divided into two categories[11]: one is statistical model, and the other is physical model. The statistical models use a linear or nonlinear mapping between the numerical weather forecast, the wind farm observation data and the power of the wind farm, including autoregressive technology[10], neural network[9], and so on. The advantages of the statistical method includes the following aspects: they does not need have the knowleded of the specific location of the power plant and the wind power generator layout. The disadvantage is that the long-term observation data and algorithm training are needed, and they are easy to be carried out. The shortages of these the algorithms are not universal, and it need a lot trainning data. The essence of the physical model method is to improve the resolution of the model of the numerical weather prediction model so that it can accurately predict the weather conditions of each typhoon machine. However, it needs to know the relative positions and geographical coordinates of all wind driven generators. This method requires more parameters and does not have strong operability in practice.
3. Wind Power Prediction

The difference between the wind power plants and wind power generators lays on that wind power plant prediction need the parameters, such as the topology of the wind power generators, the power of the wind power plant is running state, wind, wind direction, air density, topography and other wind turbine wake and so on, as shown in Eq. 1.

\[ P = F(\text{state}, \text{wind}, \text{direction}, \text{density}, \text{geo}, \text{topology}) \] (1)

The interaction between these factors is nonlinear and cannot be described by accurate mathematical models. In order to improve the prediction accuracy, many nonlinear models are used to model the wind power, and the neural network is widely used. The black box is adopted in the neural network, and the influence between each fan and various factors is not analyzed. The relation between the power and the parameters of the power plant is illustrated by adjusting the parameters of the implicit network node. The algorithm is not very robust and requires special training to select parameters.

Many wind power forecasting systems predict power by predicting wind power, and actual wind power is the result of many comprehensive factors. In order to verify this hypothesis, we take a test for a wind plant located in Yunnan, by analyzing the correlation of wind power. The method is as follows: the similarity test on the strength of the time series of wind power plant by day, and then the similarity of the corresponding power of the test, and the correlation between the two inferred. In order to ensure the preciseness of the logic, we first establish the inverse proposition that all wind power systems are determined by wind force, so we can deduce that the power curve is approximately the same if the wind time curve is approximately the same. On the other hand, if the wind time series are similar and the power is different, the problem of multiple factors should be considered.

For this reason, in the historical data of a wind power plant in Yunnan, the operation data of one month were selected, and the wind speed curve and power curve similar to time series were selected. The similarity of sequences has many kinds of measurements, such as Minkowski distance, dynamic time, bend distance, mode distance and complex distance\(^{12}\), etc. In this paper, the Mahalanobis distance is used to calculate the power and wind speed curve in order to eliminate the influence of dimension. As can be seen from the Figure 1, the power curves are very similar in the selected two time periods, while the wind speed curves are very similar. Further observation shows that the peak power tends to be later than the peak of wind speed, which is due to the inertia of the wind power generation system, but this relationship is established. Therefore, it is feasible to predict wind power with time series similarity.
4. Wind Power Forecasting Algorithm Based on Similarity of Multivariate Time Series

Based on the above findings, this paper proposes a multivariate time series similarity based wind power prediction algorithm, the basic idea is as follows: firstly, for prediction of meteorological data, search for the similar curve in the history of meteorological observation data in the same month, choose the most similar days, then more similar differences using wind speed forecast for power. This algorithm considers that wind power plant is a complex system and receives various factors, so the algorithm can be divided into the following steps:

4.1 Searching of Similar Sequences

For each wind plant, historical data is selected to calculate, and the similarity of wind speed time series is calculated for all historical observation data and weather forecast data. The wind speed time series in D is \( V_0^o \), \( V_1^o \), ..., \( V_{95}^o \), the predicted wind speed time series is \( V_p = <V_0^p, V_1^p, ..., V_{95}^p> \), the similarity between them is:

\[
Sim_d = \left( \sum_{i=0}^{95} \left| V_i^d - V_i^p \right|^2 \right)^{1/2}
\] (1)

In actual wind power plant, there are many wind speed data collected, including wind tower data and wind speed observed on wind power generator. The wind tower data generally includes wind speed of 10 meters, 30 meters, 50 meters, 70 meters, 90 meters, and the wind speed of wind power generator is the head wind speed. In actual operation, it is better to predict the wind speed near the fan center, that is, the nose wind speed or the wind speed of 50 meters and 70 meters.

4.2 Weighting Coefficient Determination

Through the Sequence comparing, we can find the closest several power curves in the historical data. In fact, all of these curves can be used to predict the power. In order to ensure the
stability of the algorithm, the closest $M$ time series are chosen to predict the power. In order to ensure the accuracy of prediction, the weights are weighted according to the similarity of the time series, and the weights of different sequences are determined by Eq. (2), $Sim_1$, $Sim_2$, $\ldots$, $Sim_M$.

$$w_i = \frac{Sim_i}{\sum_{n=1}^{M} Sim_n}$$

(2)

### 4.3 Prediction of $T$ Moment Power

Given there are $M$ similar power time series, the power at $t$ is $P_t^m$, and the measured time power is $P_t$, which is determined by the observed wind speed, predicted wind speed and similarity coefficient, as shown in Eq.3

$$P_t = \sum_{i=1}^{M} w_d P_t^o \frac{V_t^i}{V_t}$$

(3)

The pseudo code description of the WPFSMTS algorithm is given below.
Algorithm: WPFSMTS
Input:
\{V_t^O\} // Historical power of wind power plant
\{V_p\} // Predicted wind speed series
\{P_t^O\} // Historical power series
\{P_t\} // Predicted power series

Begin
1: Initialize \{V_t^O\},\{V_p\},\{P_t^O\},\{P_t\} //Initializing parameters
2: \{ Sim_i \},\{w_i\};
3: For i to N
4: \{Sim_i\}←ComputeSimilarity(V_p, V_t^O) //Computing similarity
5: End
7: For i to M
8: w_i=ComputerWeight(Sim_i);
9: End
10: For t to 96 // Predicting wind power
11: \{P_t\}←PredictPower(P_t^O,V_p, V_t^O)
12: End
End

Figure 2. The pseudo codes of WPFSMTS algorithm.

5. Experiment Analysis

In order to illustrate the effectiveness of the algorithm, the meteorological observation data and power data from a wind power plant in Yunnan from 2013 to 2016 were simulated. Meanwhile, we choose the most common ARMA model and neural network algorithm in wind power forecasting, and compare the prediction accuracy.

According to the needs of power dispatching business, the power value is predicted every 15 minutes, and 96 values are predicted in one day. In order to ensure the fairness of prediction algorithm, the selection of parameters in ARMA is based on the AIC and BIC criteria\[^{11}\], the initial parameters are 15, and after 225 iterations, select q=3, p=5. At the same time, the neural network algorithm selects the best parameters after 5000 iterations. Figure 4 shows the relative errors of the three algorithms for predicting the next day power. As can be seen from the curve, the wavelet neural network prediction algorithm error is larger, not very stable, ARMA and WPFSMTS are relatively close.
In order to explain the difference of prediction accuracy between the three algorithms, the mean and variance of the absolute value of the prediction error and the running time of the three algorithms are calculated. In the three algorithms, the average error of WAVENN algorithm is minimum, but it is very close to that of WPFSMTS, but its prediction accuracy fluctuates most. From the point of view of algorithm time efficiency, WPFSMTS is two times higher than that of WPFSMTS, so it is acceptable to consider both time efficiency and precision.

<table>
<thead>
<tr>
<th>Algorithms</th>
<th>Running time(Second)</th>
<th>Average error</th>
<th>Variance of error</th>
</tr>
</thead>
<tbody>
<tr>
<td>ARMA</td>
<td>48.7</td>
<td>14.7%</td>
<td>8.7%</td>
</tr>
<tr>
<td>WPFSMTS</td>
<td>1.4</td>
<td>10.6%</td>
<td>8.41%</td>
</tr>
<tr>
<td>WAVENN</td>
<td>190.5</td>
<td>9.9%</td>
<td>11.6%</td>
</tr>
</tbody>
</table>

6. Conclusion

The proportion of power generation in the new energy power plant is increasing in the proportion of electricity generation, and the power forecasting of wind farm has more and more influence on the regulation of power grid dispatching. The former prediction algorithm based on neural network and time series does not consider the time series of mutual influence. From the big data prospect, the proposed similarity based on the time series of wind power prediction algorithm to predict the object as an object from the whole, and for a specific time, specific effects and considering the wind speed, thus to improve the prediction accuracy. On the same experimental data set, through comparing with ARMA algorithm and neural network algorithm, it is found that the algorithm has higher prediction accuracy. In the future, the power
dispatching system will be popularized and applied, which will contribute to the promotion of power dispatching.

References


