Hybrid Short-Term Wind Power Prediction Based on Markov Chain

Liangsong Zhou¹, Xiaotian Zhou², Hao Liang², Mutao Huang¹* and Yi Li³

¹School of Electrical and Electronic Engineering, Huazhong University of Science and Technology, Wuhan, China, ²Department of Electrical and Computer Engineering, University of Alberta, Edmonton, AB, Canada, ³College of Engineering, University of Washington, Seattle, WA, United States

This article proposes a combined prediction method based on the Markov chain to realize precise short-term wind power predictions. First, three chaotic models are proposed for the prediction of chaotic time series, which can master physical principles in wind power processes and guide long-term prediction. Then, considering a mechanism switching between different physical models via a Markov chain, a combined model is constructed. Finally, the industrial data from a Chinese wind farm were taken as a study case, and the results validated the feasibility and superiority of the proposed prediction method.

Keywords: wind power prediction, combined model, Markov chain, chaotic time series, data-driven

1 INTRODUCTION

To solve the global energy crisis and environmental pollution around the world, wind energy has become one of the most potential sources due to its clean and renewable properties (Brouwer et al., 2016; Yang et al., 2021a). In China, America, and many European countries, more and more wind power generation has been highly concentrated to the power system in the past decades (Huber et al., 2014; Yang et al., 2021b; Yang et al., 2022a), while in this case, the intermittency and variability of wind bring great challenges to the safety of power systems (Tang et al., 2020; Yang et al., 2022b). Therefore, an accurate wind power prediction system is required by system operators to mitigate the undesirable effects. Especially as the installed capacity of wind farms increases, this situation is more urgent than ever.

Currently, the major wind power prediction system related to power system schedule is based on short-term predictions. The existing short-term wind power prediction methods are mainly data-driven models, also known as statistical models. These methods usually utilize the statistical features of historical data and machine learning algorithms to train and build models, for example, auto-regressive and moving average models (ARMA), KNN, neural networks (NNs), support vector machines (SVMs), extreme learning machines (ELMs), and so on (Valipour et al., 2013; Doucoure et al., 2016; Tang et al., 2021). For instance, in Kanna and Singh (2012), several data-driven models were introduced for short-term wind power prediction. In Liu et al. (2012), an ELM-based wind power prediction model was also applied to the real wind farm data. All of these models could achieve high accuracy in short-term prediction, but their shortage is also obvious since residuals are cumulated as prediction is going on.

To improve the performance of wind power predictions, there are several ways for model improvement. This first type is to consider the combination of physical and statistical factors of data together. In this way, the multiple step prediction could be improved with the assistance of physical trends. For example, one useful model considering wind power data’s physical factors is based on chaotic time series prediction, since the process of wind was validated to obey the variance of the chaotic system (Lange and Focken 2006). In Lei et al. (2007), wind power data were verified as a chaotic time series and reconstructed for wind power prediction. The second type is to consider...
different patterns in time series data and to build a hybrid prediction model. For example, some short-term prediction models utilize signal decomposition to generate different patterns, for example, wavelet decomposition and empirical mode decomposition (EMD) (An et al., 2011; Li et al., 2021a), and then combine predictions of different patterns for the final prediction. In this way, the prediction can be improved with more robustness. The more direct way is to use different models to reflect different patterns in prediction. For example, in Tascikaraoglu and Uzunoglu (2014), a linear combined model was proposed to average the prediction results of different methods. In Ouyang et al. (2016), a Markov chain was regarded as a switching regime to select different parameters in a prediction model.

Considering the aforementioned descriptions, this article proposed a hybrid model which considers the aforementioned two factors to realize precise short-term wind power prediction. The advantages of the proposed model can be summarized as follows: First, considering the physical characteristics of wind power time series, different chaotic time series models are used to form the model library. Since chaotic time series models are physics-based models, they are more reliable than data-driven models, especially in long-term trend prediction. Second, considering the mechanism of hybrid modeling, this study considers using the Markov chain to estimate wind patterns and to guide the selection of suitable physical models in prediction. Finally, combining these contributions to realize a high-precision wind power prediction, industrial wind power data are studied to validate the effectiveness of the proposed method.

2 PREDICTION MODEL FOR CHAOTIC TIME SERIES

Generally, if time series data are from a chaotic system, it is possible to reproduce the dynamics information of its related system. To realize this idea, time series data are usually embedded into a new diffeomorphism space where data have the same dynamic behaviors as the original space and better representation ability (Rand and Young 1988; Chen et al., 2019). Therefore, data in the reconstructed phase space would be better for modeling and data mining. Here, referring to Packard et al. (1980), wind power time series could be reconstructed as chaotic time series according to Takens’ theory (Rand and Young 1988), expressed in the following form:

\[ \mathbf{x}_n = (x_{n1}, x_{n2}, \ldots, x_{n(m-1)t}) \in \mathbb{R}^m, \quad n = 1, 2, \ldots, N \quad n = N - (m - 1)t, \]

where \( \mathbf{x}_n \) represents the series of wind power data; \( x_n \) is the reconstructed data, which is a vector \( \mathbb{R}^m \); and \( r \) and \( m \) are delay time and embedding dimension parameters, which can be calculated by using the mutual information (MI) method (Fraser and Swinney 1986) and false nearest neighbors (FNN) method (Xiong et al., 2017), respectively. Then, based on the reconstructed chaotic time series in eq. 1, several prediction models could be built, for example, local prediction models, global prediction methods, prediction methods based on the Lyapunov exponent and Volterra prediction models (Zhang and Liang 2012; Jinquan et al., 2016). Considering the precision, the global prediction method is rarely used in wind power prediction, so three selected models are introduced as follows.

2.1 Local Linear Model With First Order

Local prediction models have different orders, but the first-order linear model is used the most. Assuming a given point \( x(t) \) in the phase space, \( q \) nearest points \( x_i(t), i = 1, 2, \ldots, q \), are taken as reference in prediction. Then the prediction model is described in eq. 2:

\[ \begin{bmatrix} x_1(t + k) \\ x_2(t + k) \\ \vdots \\ x_q(t + k) \end{bmatrix} = A + Bx(t), \]

where \( A \) and \( B \) are the coefficient matrices of the prediction model, which can be trained by the least square method, and \( k \) is the prediction time interval, but its value cannot be too large since the evolution over a long period is not linear for chaotic systems.

Then the prediction is calculated as \( x(t + k) = A + Bx(t) \).

2.2 Prediction Method Based on the Lyapunov Exponent

In the chaotic system, it is known that two nearby tracks are linear in the short term and have an exponent trend in the longer term. This feature is utilized in the prediction method based on the Lyapunov exponent. Then the prediction model is expressed in eq. 3:

\[ L_{t+k} = L_t \cdot e^{\lambda_k}, \]

where \( k \) is the prediction time interval, \( L_t \) is the distance between \( x(t) \) and its nearest neighbor \( x(t',k) \), \( L_{t+k} \) is the distance between \( x(t+k) \) and \( x(t'+k) \), \( \lambda_k \) is the Lyapunov exponent related to the interval \( k \), and it is calculated when \( T = k \).

2.3 Volterra Prediction Model

The Volterra algorithm is a non-linear filter and can be expressed in eq. 4:

\[ x(n+1) = h_0 + \sum_{t=0}^{n} h_1(t)x(n-t) + \sum_{t_1=0} h_2(t_1,t_2)x(n-t_1)x(n-t_2) + \sum_{t_1=0}^{n} \sum_{t_2=0}^{n} \sum_{t_3=0}^{n} h_3(t_1,t_2,t_3)x(n-t_1)x(n-t_2)x(n-t_3) + \ldots, \]

where \( x(n+1) \) is the predicted value, \( x(n-t) \) are historical values, \( p \) is the order of a Volterra filter, and \( h \) is the coefficient corresponding to different elements. It is seen that a Volterra filter with a high order is complex, so a second-order model is usually used for predicting wind power time series, as described in eq. 5:
\[ x(t + k) = h_0 + \sum_{\nu=0}^{m-1} h_1(\nu)x(t - \nu\tau) + \sum_{\nu=0}^{m-1} \sum_{\omega=0}^{m-1} h_2(\nu, \omega)x(t - \nu\tau)x(t - \omega\tau), \] (5)

where \( m \) is the embedded dimension, \( \tau \) is the delay time, \( k \) is the prediction time interval, and \( u, v, \) and \( w \) represent the index of the selected data.

### 3 HYBRID MODELING WITH THE MECHANISM OF MARKOV CHAIN

The previously mentioned three models were proposed to apply in chaotic time series prediction in many references. The three methods perform differently in different scenarios and even have different accuracy in the same period in same scenario. Therefore, a suitable solution to achieve relatively better performance is to construct a combined model as described as follows:

\[ \hat{y} = \sum w_i \cdot \hat{y}_i, \] (6)

where \( \hat{y} \) represents the output of an individual prediction model, \( \hat{y}_i \) is the final output of the combined model, and \( w_i \) is the weight coefficient for the combination of different models.

In Ouyang et al. (2016), the Markov chain was used as a regime switching different models to improve the accuracy of short-term wind power prediction. The process is decided by the transition probability matrix of a Markov chain. In this article, a new combined model based on the transition probability and prediction models of chaotic time series is proposed to predict wind power.

By assuming these three chaotic time series models as three different patterns, as \( \{S_1, S_2, S_3\} \), the idea of the Markov process is to suppose that the states \( s_{t+1} \) are related to the states \( s_t \), described in eq. 7:

\[ P_{ij}(t, t+1) = P(s_{t+1} = S_j | s_t = S_i), \] (7)

where \( P_{ij}(t, t+1) \) is the transition probability on the condition of satisfying the current pattern \( S_i \) and the next pattern \( S_j \). Assuming the wind power time series as \( \Omega_t \) and the \( t^{th} \) value is \( y_t \), then the forward probability is defined in eq. 8:
\[ p(j_t) = P(s_t = j|\Omega_t) = \frac{P(y_t, s_t = j|\Omega_{t-1})}{P(y_t|\Omega_{t-1})} \]  

where \( p(j_t) \) is the probability of \( s_t = j \) with a given series \( \Omega_t \), and the numerator and denominator are calculated by eq. 9:

\[
\begin{align*}
P(y_t, s_t = j|\Omega_{t-1}) &= p(j)_{t} \cdot P(y_t|s_t = j, \Omega_{t-1}), \\
 p(j)_{t-1} &= P(s_t = j|\Omega_{t-1}), \\
 P(y_t|\Omega_{t-1}) &= \sum_{j=1}^{N} P(y_t, s_t = j|\Omega_{t-1}).
\end{align*}
\]  

where \( N = 3 \), since there are only three available prediction models. Conversely, assuming that the whole wind power time series is given as \( \Omega_T \) (where \( T \) is the length of the given time series), the backward probability is calculated in eq. 10:

\[ p(j_t) = P(s_t = j|\Omega_T). \]  

Combining the forward probability and backward probability, the probability of states at each time is calculated by using the maximum likelihood estimation (MLE) method. Then the weights of the combined model in eq. 6 are decided by the calculated probability, and the final prediction model is expressed as eqn 11:

\[ \hat{y}_t = p_1 \cdot \hat{y}_1 + p_2 \cdot \hat{y}_2 + p_3 \cdot \hat{y}_3, \]  

where \( p_i \) is the probability of the \( ith \) model.

4 EXPERIMENTS AND DISCUSSION

4.1 Prediction Based on Individual Models

In this study, a real-world wind power dataset from Chinese wind farms is taken for the case study. This dataset collected samples for 2 years with a sampling period of 15 min. There are a total of 70,176 samples. By taking 70% of the data (viz., historical data from the last 17 months) as the training set, the data from the rest of the months are used for validation and testing. First, the phase space of the data is reconstructed based on eq. 1 by applying the reconstruction parameters with \( \tau = 7 \) and \( m = 9 \) based on MI and

![Figure 2: Probability of three states (three models) at each time.](image)
FNN. Then the reconstructed chaotic time series are used to train and predict wind power based on three models introduced in Section 2.

Figure 1 depicts the prediction results of three models: $S_1$, $S_2$, and $S_3$ (representing local linear prediction models with first-order, prediction models based on the Lyapunov exponent, and Volterra prediction models, respectively). Figures 1A, B present the predicted wind power and prediction errors, respectively. It is qualitatively found that these three models have good performance in predicting wind power in Figure 1A. However, the prediction errors in Figure 1B illustrate that each model has its best prediction performance at different local times.

Moreover, to evaluate the time series prediction performance quantitatively, some useful error metrics could be applied (He et al., 2017). For example, the commonly used mean absolute error (MAE), mean squared error (MSE), root-mean-square error (RMSE), and the standard deviation of absolute error (SDofAE) are possible options. These metrics have advantages in evaluating the error on prediction amplitude and are commonly used in prediction research. Moreover, a kind of transverse error metric could also be used for evaluation, such as correlation coefficient (CC), which aims to evaluate the prediction model’s performance in time delay. In this study, four selected metrics are given to evaluate the performance of wind power time series predictions:

$$\begin{align*}
\text{MAE} &= \frac{1}{n} \sum_{i=1}^{n} |y_i - \hat{y}_i|/n, \\
\text{RMSE} &= \sqrt{\frac{1}{n} \sum_{i=1}^{n} (y_i - \hat{y}_i)^2/n}, \\
\text{SDofAE} &= \sqrt{\frac{1}{n} \sum_{i=1}^{n} [(y_i - \hat{y}_i) - \text{MAE}]^2}/n, \\
\text{CC} &= \frac{\text{cov}(y, \hat{y})}{\sqrt{\text{var}(y) \sqrt{\text{var}(\hat{y})}}},
\end{align*}$$

(12)

where $y_i$ and $\hat{y}_i$ are predicted and observed values of wind power time series, respectively, and $n$ is the number of testing data points. According to the definition of error metrics before, we can see that the former three (MAE, RMSE, and SDofAE) are required to be small when the model performs well. However, on the other hand, the value of CC needs to be as close to 1 as possible.

Table 1 shows the values of four error indicators in wind power prediction. $S_1$, $S_2$, and $S_3$ represent the three prediction models introduced in Section 2, respectively. It is seen that the difference between the former three models is not large and that $S_1$ has the worst prediction performance, while $S_3$ has the best prediction performance according to the values of the error indicators.

4.2 Prediction Based on the Proposed Hybrid Model

According to the description of the proposed method, a combined model could effectively improve the shortages of individual models at local prediction. Therefore, a combined model is proposed in eq. 11, where its weights are selected as the transition probability of the Markov chain. Taking the three given prediction models as three states, the transition probabilities of states at each time are calculated based on eqs 7–10.

Figure 2 depicts the transition probability of three states in the same period as Figure 1. It is seen that the wind power is predicted by one model with a given probability at each time. Therefore, if the probability of each state is taken as the weight of the combined model in eq. 11, then wind power for the same period with Figure 1 can be predicted.

Then, also making use of the given error metrics in evaluation, the performance of the proposed method is presented in the following table. For the consideration of the comparison study, several other prediction methods are also evaluated, such as two combined models, namely, a linear combined model of $S_1$, $S_2$, and $S_1$ and a Markov-switching-autoregression model (MSAR), three generally used data-driven models based on neural networks (NNs), support vector machine (SVM) (Li et al., 2021b), and extreme learning machine (ELM) (Shen et al., 2020). All of their results are presented in Table 2.

Table 2 shows the error metrics of wind power prediction based on different models, including individual data-driven models and combined models. Through the comparison, some phenomena could be found. First, the combined models generally outperform all of these individual models, including the common statistical models and the chaotic time series models in Table 1, which validates the effectiveness of combined models in improving wind power prediction performance. Second, through comparison between individual and combined chaotic models, it is seen that considering the combination or switching of different physical models is helpful to improve prediction accuracy. Moreover, the switching mechanism with consideration of the Markov chain can outperform the simple linear combination model. Finally, the proposed method combining both the chaotic time series model and the Markov chain could further improve the prediction performance. In the comparison study, it is reasonable that for individual chaotic time series, models cannot achieve obvious superiority than traditional data-driven models which can achieve high-precision in short-term prediction. However, chaotic time series models are more reliable than data-driven models since they consider physical principles of the wind development process. Therefore, chaotic models would be

| TABLE 2 | Wind power prediction using various methods. |
|---|---|---|---|---|
| MAE | RMSE | SDofAE | CC |
| Combined method | Proposed | 16.398 | 17.887 | 14.211 | 0.9850 |
| | Linear | 16.660 | 22.541 | 16.831 | 0.9884 |
| | MSAR | 16.547 | 19.538 | 15.482 | 0.9788 |
| Data-driven | SVM | 18.610 | 26.789 | 19.290 | 0.9276 |
| | ELM | 18.651 | 27.885 | 19.797 | 0.9359 |
| | NN | 17.741 | 24.154 | 16.402 | 0.9506 |

Figure 1A, B depicts the prediction results of three models: $S_1$, $S_2$, and $S_3$ (representing local linear prediction models with first-order, prediction models based on the Lyapunov exponent, and Volterra prediction models, respectively). Figures 1A, B present the predicted wind power and prediction errors, respectively. It is qualitatively found that these three models have good performance in predicting wind power in Figure 1A. However, the prediction errors in Figure 1B illustrate that each model has its best prediction performance at different local times.

Moreover, to evaluate the time series prediction performance quantitatively, some useful error metrics could be applied (He et al., 2017). For example, the commonly used mean absolute error (MAE), mean squared error (MSE), root-mean-square error (RMSE), and the standard deviation of absolute error (SDofAE) are possible options. These metrics have advantages in evaluating the error on prediction amplitude and are commonly used in prediction research. Moreover, a kind of transverse error metric could also be used for evaluation, such as correlation coefficient (CC), which aims to evaluate the prediction model’s performance in time delay. In this study, four selected metrics are given to evaluate the performance of wind power time series predictions:

$$\begin{align*}
\text{MAE} &= \frac{1}{n} \sum_{i=1}^{n} |y_i - \hat{y}_i|/n, \\
\text{RMSE} &= \sqrt{\frac{1}{n} \sum_{i=1}^{n} (y_i - \hat{y}_i)^2/n}, \\
\text{SDofAE} &= \sqrt{\frac{1}{n} \sum_{i=1}^{n} [(y_i - \hat{y}_i) - \text{MAE}]^2}/n, \\
\text{CC} &= \frac{\text{cov}(y, \hat{y})}{\sqrt{\text{var}(y) \sqrt{\text{var}(\hat{y})}}},
\end{align*}$$

(12)
helpful to capture wind trends, and guarantee good accuracy in the long-term prediction in the future study.

5 CONCLUSION

To realize precise short-term wind power prediction, this article proposed a more reliable combined model considering both chaotic time series modeling and the Markov chain mechanism. First, by reconstructing the wind power data into a new space with the consideration of the wind’s chaotic physical features, three chaotic time series models consist of the primary model library. Second, by taking different models as the states of the Markov chain, and a combined model utilizing the transition probability of the Markov chain as the weight is constructed. The results of Table 2 verify that the combined model based on the Markov chain is feasible to predict wind power. Moreover, through the comparison study, the proposed method is validated to improve the precision of wind power prediction, which will be helpful for direct scheduling and planning in the power system in the future.

DATA AVAILABILITY STATEMENT

The original contributions presented in the study are included in the article/Supplementary Materials, further inquiries can be directed to the corresponding author.

AUTHOR CONTRIBUTIONS

LZ put forward the main research points; MH completed manuscript writing and revision; XZ, YL, and HL completed simulation research; LZ collected relevant background information; and MH revised grammar and expression. All authors contributed to manuscript revision, read, and approved the submitted version.
Yang, N., Dong, Z., Wu, L., Zhang, L., Shen, X., Chen, D., et al. (2021). A Comprehensive Review of Security-Constrained Unit Commitment. *J. Mod. Power Syst. Clean Energ.*

Yang, N., Qin, T., Wu, L., Huang, Y., Huang, Y., Xing, C., et al. (2022). A Multi-Agent Game Based Joint Planning Approach for Electricity-Gas Integrated Energy Systems Considering Wind Power Uncertainty. *Electric Power Syst. Res.* 204, 107673. doi:10.1016/j.epsr.2021.107673

Yang, N., Yang, C., Wu, L., Shen, X., Jia, J., Li, Z., et al. (2021). Intelligent Data-Driven Decision-Making Method for Dynamic Multi-Sequence: An E-Seq2Seq Based SCUC Expert System. *IEEE Trans. Ind. Inform.* 18, 3126. doi:10.1109/TII.2021.3107406

Yang, N., Yang, C., Xing, C., Ye, D., Jia, J., Chen, D., et al. (2022). Deep Learning-based SCUC Decision-making: An Intelligent Data-driven Approach with Self-learning Capabilities. *IET Generation Trans. Dist.* 16 (4), 629–640. doi:10.1049/gtd2.12315

Zhang, X., and Liang, J. (2012). Chaotic Characteristics Analysis and Prediction Model Study on Wind Power Time Series. *Acta Phys. Sin.* 61 (19), 70–81. doi:10.7498/aps.61.190507

**Conflict of Interest:** The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

**Publisher’s Note:** All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Copyright © 2022 Zhou, Zhou, Liang, Huang and Li. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.