

Dynamic Modeling and Very Short-term Prediction of Wind Power Output Using Box-Cox Transformation

Kengo Urata^{1,a}, Masaki Inoue^{1,b}, Dai Murayama^{2,c} and Shuichi Adachi^{1,d}

¹ Faculty of Science and Technology, Keio University,
3-14-1 Hiyoshi, Kohoku-ku, Yokohama, 223-8522, Japan.

² Toshiba Corporation Power Systems Company,
2-4 Suehiro-cho, Tsurumi-ku, Yokohama, 230-0045, Japan

E-mail: ^akengourata@keio.jp, ^bminoue@appi.keio.ac.jp,
^cdai.murayama@toshiba.co.jp, ^dadachi@appi.keio.ac.jp

Abstract. We propose a statistical modeling method of wind power output for very short-term prediction. The modeling method with a nonlinear model has cascade structure composed of two parts. One is a linear dynamic part that is driven by a Gaussian white noise and described by an autoregressive model. The other is a nonlinear static part that is driven by the output of the linear part. This nonlinear part is designed for output distribution matching: we shape the distribution of the model output to match with that of the wind power output. The constructed model is utilized for one-step ahead prediction of the wind power output. Furthermore, we study the relation between the prediction accuracy and the prediction horizon.

1. Introduction

Recently, renewable energies, such as solar power and wind power energies, are installed to the existing power networks that is composed of the conventional generators, such as thermal and nuclear energies. The aim of the installation is to reduce environment load and to prevent global warming. Such renewable energies contribute to CO₂ reduction and realize a safer power system. However, since their energy outputs uncertainly fluctuate depending on a weather condition, direct installation of the renewable energies inevitably causes demand supply imbalance in the entire power system.

An approach to this problem is equipping the renewable energies with battery storages to smooth the output fluctuations. In some cases, battery storages can be substituted by solar-thermal energies that are connected to renewable energies and compensate the energy outputs, e.g., [1]. The efficient smoothing operation in such storage-equipped energy systems, mainly relies on accurate prediction of the renewable energy outputs. In this paper, we develop a prediction method of renewable energy outputs, in particular, model-based prediction of the wind power output. Many methods of predicting the wind power output based on statistical models have been proposed [2][3]. For example, an autoregressive (AR) model and a Just-In-Time (JIT) model are designed for performing short-term prediction of the wind power output, for 1–10 minutes-ahead prediction. In this paper, we design a statistical model for very-short term prediction of the wind power output, for 1–10 seconds-ahead prediction. The predicted output can be utilized for effective governor-free operation, which controls outputs of the conventional generators to maintain a constant frequency.

To improve the prediction accuracy, it is necessary to precisely match the time series behavior of real output with that of the model output. In addition, it is also important to match the distribution [4] of



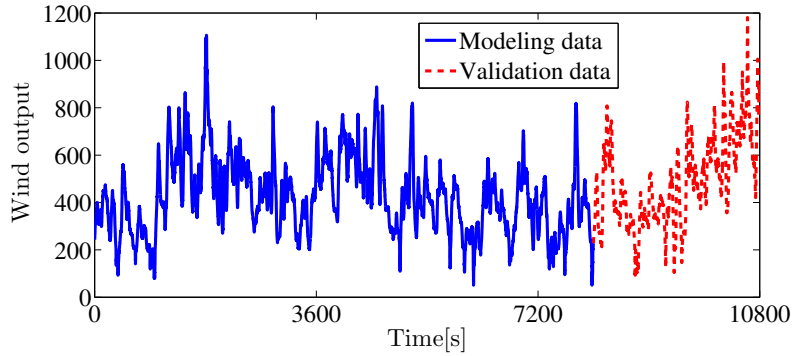


Fig. 1. Wind power output data $y(k)$

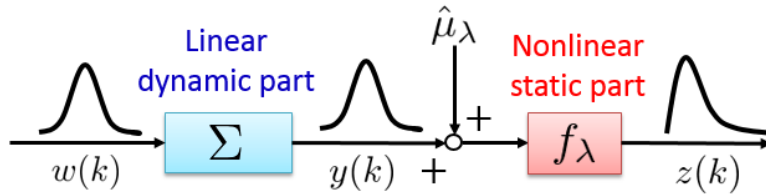


Fig. 2. Model structure

the model output with that of the real output. To this end, we consider a model structure that combines an AR model with a nonlinear transformation for the output distribution matching. Then, we propose the modeling method for this model. We utilize the Box-Cox transformation ^[5], which transforms the time series with a non-Gaussian distribution into those with a Gaussian distribution. Finally, we apply the proposed method to measurement data of wind power output. In the prediction result, we show the effectiveness of the proposed distribution matching. Furthermore, we study the relation between the prediction accuracy and the prediction horizon.

2. Proposed method

A typical time series data of the wind power output data is shown in Fig. 1. We design a dynamic model that expresses the wind power output denoted by $\{y(k)\}$. The proposed model includes the nonlinear transformation as illustrated in Fig. 2. In this figure, the model has cascade structure composed of two parts: One is a linear dynamic part that is driven by a Gaussian white noise and described by an AR model. The other is a nonlinear static part that is driven by the output of the linear part. This nonlinear part is designed such that the distribution of its output matches with that of the wind power output. In this section, we explain the proposed model in more detail. Then, we propose a design procedure of the nonlinear function and the AR model from given wind power output data.

2.1. Model structure

In Fig. 2, the linear dynamic part G represents an AR model described by

$$G : z(k) = \frac{1}{1 + a_1 q^{-1} + \dots + a_n q^{-n}} w(k), \quad (1)$$

where $z(k)$ and $w(k)$ represent the output of G and a zero-mean Gaussian white noise, respectively. The symbol n is the model order, $a_i, i \in \{1, \dots, n\}$ is the constant coefficient, and q is the shift operator.

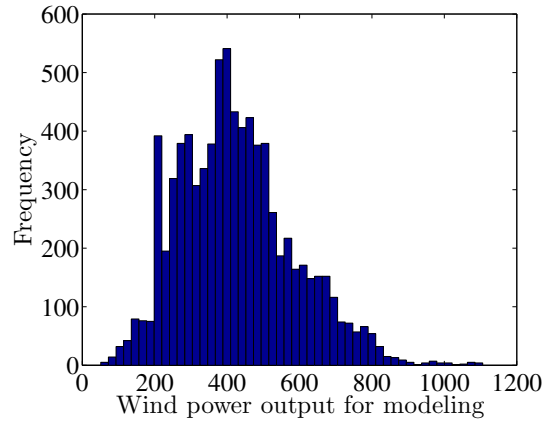


Fig. 3. Histogram of wind power data at three hours period where sampling time is one second

General time series data is conventionally often modeled by the AR model under the implicit assumption that the data follows a Gaussian distribution [5]. However, as illustrated in Fig. 3, the histogram of the wind power output does not follow any Gaussian distribution. This means that the AR model cannot completely express the wind power output. In order to transform the given wind power output data into time series data with the Gaussian distribution, the nonlinear static part f_λ is connected to the output of the AR model as illustrated in Fig. 2.

We define the inverse function of f_λ and illustrate the role of f_λ . We consider the following Box-Cox transformation, which shapes the time series $\{y(k)\}$ into the time series $\{z(k)\}$ with the Gaussian distribution. For any positive value y , the Box-Cox transformation is defined by

$$z := g_\lambda(y) = \begin{cases} \lambda^{-1}(y^\lambda - 1), & \lambda \neq 0 \\ \log y, & \lambda = 0 \end{cases}, \quad (2)$$

where the constant λ is the design parameter. The inverse function of the Box-Cox transformation (2) is also defined as

$$\begin{aligned} y = g_\lambda^{-1}(z) &= \begin{cases} (\lambda z + 1)^{\frac{1}{\lambda}}, & \lambda \neq 0 \\ e^z, & \lambda = 0 \end{cases} \\ &:= f_\lambda(z), \end{aligned}$$

and this inverse function represents the nonlinear static part $f_\lambda(z)$.

2.2. Modeling method

First, we find the optimal parameter λ^* which minimizes the gap in the distribution of the time series data and the Gaussian distribution. Note that g_λ in (2) depends on λ . For a fixed λ , we obtain the mean $\hat{\mu}_\lambda$ and variance $\hat{\sigma}_\lambda$ of the transformed data $\{z(k)\}$ by applying the maximum likelihood estimation. The optimal parameter λ^* is given by

$$\lambda^* = \arg \min_{\lambda} \sum_{i=1}^N J_i(\lambda), \quad (3)$$

where

$$J_i(\lambda) = \left\{ \Phi_{SN}^{-1} \left(\frac{\ell - 0.5}{N} \right) - \left(\frac{z(\ell) - \hat{\mu}_\lambda}{\hat{\sigma}_\lambda} \right) \right\}^2 \quad (4)$$

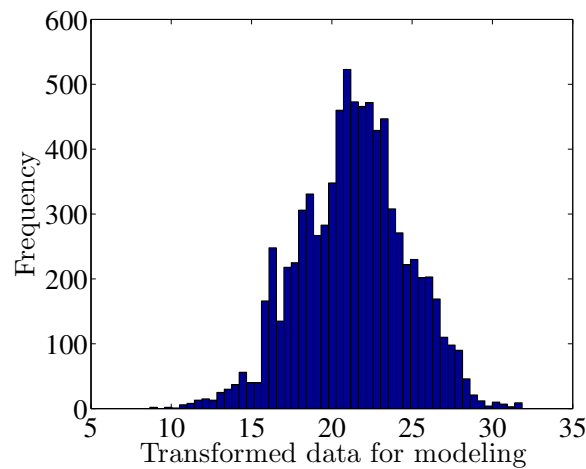


Fig. 4. Histogram of the transformed wind power data at three hours period where sampling time is one second

and $\Phi_{SN}^{-1}(\cdot)$ represents an inverse cumulative density function of the standard normal distribution. Note that the transformed data $z(\ell)$ in (4) is sorted in ascending order that is $z(1) \leq z(2) \leq \dots \leq z(N)$. The cost function (4) means the error between standardized time series data and a quantile corresponding to the data.

Next, applying the Box-Cox transformation to the time series data $\{y(k)\}$, we obtain the transformed data $\{z(k)\}$. Since the mean value of the transformed data $\{z(k)\}$ does not become zero, it is further transformed into time series data $\{\tilde{z}(k)\}$ whose mean value is zero as

$$\tilde{z}(k) = z(k) - \hat{\mu}_{\lambda^*}.$$

Finally, we construct the AR model (1) from the transformed data $\{\tilde{z}(k)\}$. In a similar way to the conventional modeling methods^[5], we find the model parameter a_i by solving the Yule-Walker equation that is determined by $\{\tilde{z}(k)\}$.

3. Prediction result of wind power output

In this section, we apply the proposed modeling method to wind power output data and show the effectiveness of the method.

3.1. Modeling

We apply the proposed modeling method to wind power output data in Fig. 1. The data at Minami-Awaji city, Japan, was measured for every 1 second interval. The time series data from 12:00 to 15:00 in February 3rd, 2015, is plotted in Fig. 1. The first 75 percent of the data, denoted by $\{y_1(k)\}$, was utilized for modeling and the remaining data, denoted by $\{y_2(k)\}$, was for a prediction experiment.

For the modeling data $\{y_1(k)\}$, we solved the optimization problem, which was given in (3). We performed the grid search to solve the problem under the additional constraint: $\lambda > 0$, and obtained $\lambda^* = 0.355$. The histogram of $\{y_1(k)\}$ and that of the transformed data $\{z_1(k)\}$ achieved by the Box-Cox transformation with λ^* are shown in Figs. 3, 4, respectively. We see that the distribution of $\{z_1(k)\}$ in Fig. 4 is closer to a Gaussian distribution than that of $\{y_1(k)\}$ in Fig. 3.

After we transformed the original modeling data $\{y_1(k)\}$, we modeled an AR model from the transformed modeling data $\{\tilde{z}_1(k)\}$. In the AR model, the model order was set to $n = 4$ on the basis of Akaike's information criterion^[6]. Then, we designed the model parameter as $[a_1, a_2, a_3, a_4] = [-1.232, -0.0523, 0.1152, 0.1735]$.

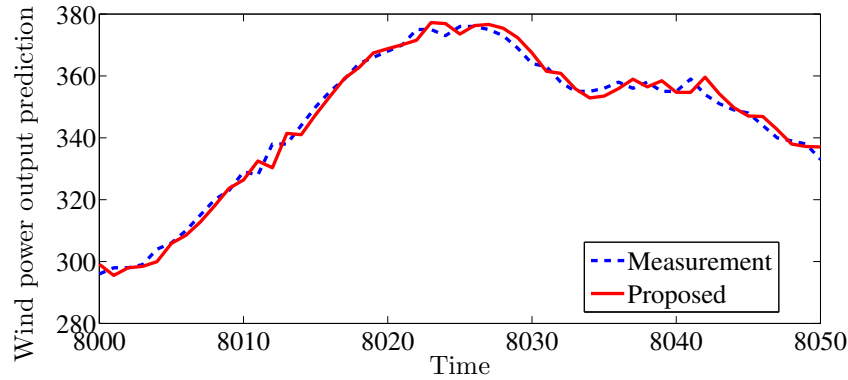


Fig. 5. One-step ahead prediction for 50 seconds period

Table 1. Prediction result in Fig. 5

	AR	ARI	Proposed
RMSE [kW]	6.439	6.459	6.496
ℓ_∞ [kW]	58.158	62.202	54.934

3.2. Result of one-step ahead prediction

We performed one-step ahead prediction of the wind power output using the constructed model. The prediction $\hat{y}_2(k)$ was obtained by applying the inverse Box-Cox transformation to the prediction $\hat{z}_2(k)$ that was obtained as the output of the AR model. Additionally, we utilized a saturation function to avoid an unrealistic prediction value. Details of this method is summarized in the Appendix.A

Fig. 5 shows the result for one-step ahead prediction. We applied the constructed model to the remaining data $\{y_2(k)\}$ and the prediction values were partially plotted in Fig. 5. In this figure, we see that the prediction values approximately capture the measurement data.

To evaluate the prediction accuracy quantitatively, we define the root mean squares error

$$\text{RMSE} := \left(\sum_{k=1}^N (\tilde{y}_2(k) - \hat{y}_2(k))^2 \right)^{\frac{1}{2}}$$

and the ℓ_∞ norm of the prediction error

$$\ell_\infty := \max_k |\tilde{y}_2(k) - \hat{y}_2(k)|.$$

The accuracy evaluation is summarized in Table 1. In this table, the prediction accuracy achieved by the proposed method is compared with that by the conventional AR model and an autoregressive integrated (ARI) model, which is an AR model designed by using difference data. The model order of the AR and ARI models were 4 and 3, respectively. In Table 1, we see that the proposed modeling method achieves the highest accuracy in the sense of ℓ_∞ norm, while the ARI model achieves the highest accuracy in RMSE.

3.3. Analysis of prediction result for prediction horizon

We analyze the dependency of the prediction accuracy on the prediction horizon. In order to analyze the dependency, we decimate the original time series data with sampling period $T_s = 1$ second at every rT_s second interval, where $r = 1, 2, 5, 10$. Furthermore, we utilized averaged values of RMSEs and ℓ_∞ norms of the prediction error to evaluate the prediction accuracy. We analyzed the dependency of the

Table 2. Analysis of prediction result for every prediction horizon, $r = 1, 2, 5, 10$

r		AR	ARI	Proposed
1	RMSE	7.234	7.200	* 7.157
	ℓ_∞	82.644	* 79.840	82.272
2	RMSE	13.169	13.370	* 13.000
	ℓ_∞	125.92	* 122.66	125.28
5	RMSE	36.675	38.993	* 35.894
	ℓ_∞	238.75	246.06	* 238.19
10	RMSE	67.393	69.785	* 66.474
	ℓ_∞	319.01	326.96	* 313.23

prediction accuracy on the prediction horizon for 9 sets of 3 hours data that were measured at Minami-Awaji city, Japan in October, 2014, January and February, 2015.

In a similar way as mentioned in 3.2, the first 75 percent of the data was utilized for modeling and the remaining data was for prediction experiment. Table 2 shows the prediction accuracy for every prediction horizon $r = 1, 2, 5, 10$. In this table, the averages of RMSEs and ℓ_∞ norms of all 9 data sets are listed and the asterisk symbols indicate the best values in each prediction horizon. The ARI model achieves the highest accuracy in ℓ_∞ norm at $r = 1, 2$ while the proposed modeling method achieves the highest accuracy in both evaluation index at $r = 5, 10$. Therefore, we showed the effectiveness of the proposed method at a longer prediction horizon.

4. Conclusion

In this paper, we proposed a modeling method of wind power output. The proposed modeling method was composed of a nonlinear function part and an AR model part. The nonlinear function was designed such that the distribution of its output matches with that of the wind power output. We applied the proposed method to measured wind power output and illustrated the effectiveness of the proposed method, in particular, for long step-ahead prediction.

In future work, we consider a more general nonlinear function to shape the distribution of the model output more accurately.

References

- [1] Yo M, Ono M, Adachi S, Murayama D and Okita N 2014 in *Proceedings of the 53rd IEEE Conference on Decision and Control* pp 928–934
- [2] Tores J, Garcia A, Blas M and Francisco A 2005 *Solar Energy* **79** 65–77
- [3] Ishikawa T, Kojima T and Namerikawa T 2015 *IEEE Transactions on Electronics, Information and Systems* **135** 81–89
- [4] Pinson P, Madsen H, Nielsen H, Papaefthymiou G and Klockl B 2009 *Wind Energy* **12** 51–62
- [5] Hamilton J 1994 *Time Series Analysis* (Princeton university press Princeton)
- [6] Akaike H 1998 *Information theory and an extension of the maximum likelihood principle* (Springer New York)

Appendix A. Constraint on prediction value

The prediction value obtained by the proposed model takes an unrealistic value. To avoid this problem, we impose two limitations on the prediction value and its difference value. Let $\text{sat}_{(\gamma_1, \gamma_2)}(\cdot)$ denote a saturation function whose lower bound and upper bounds are γ_1 and γ_2 , respectively. Then, define $\text{sat}_{(0, p_{\max})}(\cdot)$ and $\text{sat}_{(-\beta, \beta)}(\cdot)$ to express the saturations on the prediction value and its difference value, respectively. Function $\text{sat}_{(0, p_{\max})}(\cdot)$ means that p_{\max} is the maximum rating of the wind power output, while $\text{sat}_{(-\beta, \beta)}(\cdot)$ means a constraint of the difference value of the output. For example, we choose the parameter β as the maximum value of the difference of the modeling data.