Interval prediction of photovoltaic power generation based on cloud theory

Haibo Zhao¹, Xiaoyang Dong¹, Ai Wang¹, Zheng Wang¹, Yanhui Zhang², Zhi Zhang¹*, Xia Dong³, Xiaojun Song¹, Yaju Wang¹, Yahong Xing¹

¹Economic and Electrical Research Institute of Shanxi Electrical Power Company of SGCC
²Shanxi Electrical Power Company of SGCC
³State grid Tianjin power economic and technology research institute
*Corresponding author’s e-mail: zapletree@163.com

Abstract. With the increasing integration of photovoltaic power into the power system, the reliable photovoltaic power generation prediction is of significant to the security and economics of the power system operation. However, the prediction deviation is unavoidable, therefore this paper presents an interval prediction model for the photovoltaic power generation based on cloud theory. Based on the error analysis of the photovoltaic power prediction, the training samples can be selected to establish the cloud model for each small power generation bin. In this way, the predictive cloud distributions for different predictive power generation can be obtained to generate the prediction intervals at each time slot. To take the photovoltaic power plant in German as example, the proposed model is validated. The results show that the proposed model outperforms the other benchmark method. And the calculation process of the proposed model is simple and short in computation time. The analysis results are expected to be used in the field of power grid dispatching and decision making.

1. Introduction
Photovoltaic power is the fastest growing renewable energy worldwide. A large amount of solar power generation in the power system is challenging to the economics and security of the system operation. Therefore, the prediction of solar power generation is of significance during the system dispatching and requires obtain sufficient reliability to support the decision-making of the system operators [1].

The power prediction methods can be divided into deterministic prediction and uncertainty prediction. The deterministic prediction is the mathematical expectation of power generation in the future, which is the most intuitive expression of the future power generation. However, the prediction error is unavoidable. This greatly limits the application in the operation and management of new energy power systems [2]. Therefore, there are many researchers starting to focus on the photovoltaic power interval prediction. The statistical method uses a large amount of historical data to carry out model training or error fitting, establishes a "function" relationship between input variables and output variables, and predicts future photovoltaic power intervals under given confidence level based on the trained model. In the case of sufficient historical samples, most of them can obtain higher prediction accuracy and generalization ability [3].

A linear programming-based prediction interval model for photovoltaic power generation was established based on extreme learning machine and quantile regression [4]. A novel ultra-short-term
and short-term interval prediction method was presented for solar global horizontal irradiance based on k-means and quantiles [5]. A reliability index was proposed based on prediction interval of solar irradiance and using just-in-time modeling [6]. Multi-Model Ensemble Interval prediction method was proposed based on several data-driven models and different Numerical Weather Prediction input data [7]. Stochastic differential equation for simulating solar radiation uncertainty was proposed to describe the predicted distribution of short-term solar radiance prediction error [8]. Method for predicting the probability density function of the photovoltaic power generation was proposed by using the Bayesian autoregressive time series model and Monte Carlo algorithm to simulate the predicted probability distribution [9]. Independent component analysis of photovoltaic power generation impact factor was used to establish a conditional probability prediction model to predict the interval probability of photovoltaic power generation [10]. The Copula function was used to model the joint probability distribution of PV actual power output and point prediction, and in this way the “conditional probability” of the actual PV power output can be estimated [11]. Fuzzy-random theory was used to predict the photovoltaic power generation in diverse ways to represent its uncertainties [12].

Therefore, the statistical method is the most widely used method and they can be divided into two categories. However, there are still spaces to improve in this field. (1) The reliability and accuracy need to be further improved, especially for the cloudy weather; (2) The adaptability of the model needs to be improved regardless of the deterministic prediction method used, and this requires refined modelling approach to be presented; (3) Most modelling methods need to assume that the prediction error distribution follows a known distribution (i.e. Gaussian distribution), which does not meet the actual situation; (4) The model requires a large number of training samples and long training time.

To solve above problems, this paper presents an interval prediction model of solar power generation based on the cloud theory to generate a predictive power interval under given confidence level instead of a single value at each time slot. The reverse cloud generator and forward cloud generator are used to quantify the prediction error distribution by three statistical cloud feature index (expectation, entropy, hyper entropy) and to generate cloud drop map based on the three indexes. By using the quantile theory and the generated cloud drops, the predicted intervals under given confidence level can be obtained. The photovoltaic power plant in German is used to validate the proposed model and indicates better performance comparing to the benchmark method in terms of the reliability and interval range. In this way, the reliability and confidence of the prediction results can be improved.

The rest of this paper is organized as follow. Section 2 introduces the theory of cloud model and the generating process of different cloud models for each small power bin. Section 3.

2. INTERVAL PREDICATION BASED ON CLOUD THEORY

2.1. Basic principle of the cloud model

The concept of the cloud model was proposed on the basis of membership function [11]. The forward and reverse membership cloud generation models are used to realize the uncertain transformation between qualitative concepts and quantitative descriptions by three mathematical features of the membership clouds, which are the expectation, entropy and hyper entropy. This cloud model reflects two types of uncertainty concepts that exist in the objective world, namely, randomness (probability of event occurrence) and ambiguity [12]. Assumed that U is a quantitative domain and is represented by a certain value; C is a qualitative concept on U. If there is a quantitative value \( x \in U \), and \( x \) is a random implementation of qualitative concept C, the certainty degree of \( x \) to C is \( \mu(x) \in [0,1] \), which is a stable and random number. Then, the distribution of \( x \) on the domain \( U \) is called a cloud, and each \( x \) is called a cloud drop.

\[
\text{If } \mu: U \rightarrow [0, 1] \forall x \in U x \rightarrow \mu(x)
\]  
(1)
2.2. Statistical characteristics of the cloud model
The statistical characteristics of the cloud reflect the overall characteristics of the concept. In this paper, the concept is the prediction uncertainty, and the cloud drop is the prediction error by three feature indexes: expectation, entropy, and hyper entropy.

1) Expectation (Ex) is the central value in the space of the universe characterizing the mathematical expectation of the spatial distribution of cloud drops in the domain. It is the most representative point to the qualitative concept reflected in the shape of the cloud drop map - the highest point.

2) Entropy (En) measures the variance of the numerical range of qualitative concepts. In the cloud model, the larger the entropy value, the scope of the qualitative concept measured is wide. It is reflected in the span of the cloud. The larger the entropy, the larger the span.

3) Hyper entropy (He) is the entropy of entropy for characterizing the uncertainty of entropy, which represents the degree of dispersion of cloud drops in a cloud map. It is reflected by the "thickness" of the cloud map. The larger the hyper-entropy value, the larger the dispersion of the cloud drops and the larger the "thickness" of the cloud.

2.3. Cloud drop generation machine
There are two kinds of cloud generators. The converter from qualitative concept to quantitative description is called forward cloud generator. The converter from quantitative description to qualitative concept is called reverse cloud generator [12].

1) Forward Cloud Generator (CG) [13] is a converter from qualitative concept to a quantitative description. The three statistical features of the cloud (expectation, entropy, hyper entropy) and the number of cloud drops would be imported into the forward cloud generator. And, the corresponding number of the cloud drops xi as well as the degree of certainty represented by cloud drops μi can be generated according to the characteristic of the concept.

2) Reverse Cloud Generator (CG-1) [14] is the inverse of the forward cloud generator, which is a converter from quantitative description to qualitative concept. By inputting cloud drops conforming to a certain distribution at the front of the model, three mathematical features (expectation, entropy, hyper entropy) can be exported at the end of the model.

In the reverse cloud model, the expectation, entropy and hyper entropy are calculated according to a certain amount of power prediction error data samples (the difference between the predicted power value and the actual value). Based on the three-feature index, cloud drops are generated through the forward cloud model. Discretize the finite error points into a cloud distribution and simulate the possible values of the error under given predicted power bin. Through such qualitative to quantitative conversion, the degree of abstract prediction uncertainty is quantified into specific cloud drop values to form an error distribution curve, which provides a very reliable analysis basis for the following interval power prediction. The conversion between the two converter generators are shown in Figure 1.

3. PROCESS OF MODEL ESTABLISHMENT
Different predictive model inputs, prediction algorithms and training samples will bring prediction uncertainty to the deterministic power predictions. The establishment of the power interval prediction model quantify these uncertainties from the deterministic prediction. The modeling steps to quantify prediction uncertainty are as shown in Figure 2.

1) To determine the range of predicted power bin for each time slot in the testing set and to select the corresponding power data and prediction error data from the training data set. For each power prediction value at the “current” time slot, the cloud model is first required to generate some numerical cloud drops to qualitatively describe the features of the error distribution using mathematical expectation, entropy, and hyper entropy, as shown in Figure 3.

2) Reverse Cloud and Forward Cloud Model Processing. The actual prediction error distribution for all the power output bins does not follow any known error distribution. However, the error distribution for a small power bin is more closed to the normal distribution. Therefore, the cloud drops can be obtained based on the small predicted power bins to model the error distribution in a more accurate
manner. The reverse cloud is used to describe the distribution law of the input cloud drop selected in the above steps. These cloud drops of prediction error can be selected in the training set based on the prediction power on the testing samples to prepare for the cloud modelling.

3) The forward cloud model is used to generate the cloud drop distribution map based on the mathematical features (Ex, En, He) as shown in Figure 3.

4) Using the quantile theory to calculate the upper and lower limits of different prediction intervals for different confidence intervals. Before this, cloud model for predicted power for each testing sample has been established, at a given confidence level. And then, the upper and lower limits of the predicted power error are intercepted on the cloud drop map and the prediction interval can be obtained.

4. CASE STUDY

4.1. Data and evaluation criteria
Taking a solar power plant in Germany as an example, the predicted power and actual power from January 1, 2015 to December 31, 2016 were used. The time resolution was 15 minutes. The support vector machine (SVM) algorithm was used to establish the deterministic power prediction model. Data from January 1 to August 31, 2015 was used to train the deterministic model and data from September 1, 2015 to August 31, 2016 to make the deterministic prediction for training the error distribution of different predictive PV power bins during the following interval prediction. Data from September 1 to December 31, 2015 was used as testing samples for the interval prediction. The quantile regression algorithm is used as the benchmark. Three confidence levels are implemented (97%, 95%, 90%).

The evaluation index – reliability [17] is used to verify the proposed interval prediction model. Reliability index is to measure the difference between the proportion of valid data and the pre-set confidence level. If the actual power output point falls within the prediction interval, the result is valid.
If the ratio of the valid result is greater than or equal to the nominal confidence level, the proposed model is reliable. If it is greater than zero, the reliable value is higher than the theoretical level [18].

4.2. Results and discussion

Figure 4-6 show the results of interval prediction on four randomly selected and consecutive days under 90%, 95%, 97% confidence levels respectively. It can be seen that for each confidence level that in the first three days the weather was sunny and clear while in the last two days the weather was overcast with apparently reduced power output. In both weather types, the predicted intervals cover the actual power and predicted power in every confidence level, which validates the proposed algorithm. Besides, the predicted interval with high confidence level would be relatively large comparing to other low confidence level. In this way, the nominal confidence can be achieved.

Figure 5. Interval prediction under 95% confidence level

Figure 6. Interval prediction under 97% confidence level

Table 1 and Table 2 are the comparisons of two methods – the proposed method and the mainstream algorithm quantile regression in terms of the power interval and the percentage that exceeds the upper or lower limits for each confidence level. Even with better reliability index, the proposed cloud-based method has smaller predicted power interval. This indicates the sharpness of the proposed model.

Table 1. The percentage of two methods exceeding the upper or lower limits for each confidence level.

| Models                      | Ratio that out of the limits |
|-----------------------------|-------------------------------|
|                            | Cloud based method           | Quantile regression         |
| Ratio of out of 97% upper limits | 0%                            | 1.5%                        |
| Ratio of out of 97% lower limits | 0.6%                          | 1.6%                        |
| Ratio of out of 95% upper limits | 0%                            | 2.3%                        |
| Ratio of out of 95% lower limits | 1.8%                          | 2.1%                        |
| Ratio of out of 90% upper limits | 0.2%                          | 3.7%                        |
| Ratio of out of 90% lower limits  | 3.2%                          | 3.6%                        |

Table 2. The power interval of two methods for each confidence level

| Models          | Power Interval (Normalized value) |
|-----------------|----------------------------------|
|                 | Cloud based method | Quantile regression |
| 97% upper limits| 0.0467              | 0.0452              |
| 97% lower limits| 0.0355              | 0.0477              |
| 95% upper limits| 0.0304              | 0.0404              |
| 95% lower limits| 0.0296              | 0.0385              |
| 90% upper limits| 0.0316              | 0.0361              |
| 90% lower limits| 0.0249              | 0.0381              |
5. CONCLUSION
In this paper, a cloud-based interval prediction model is proposed for estimating the prediction uncertainty of the photovoltaic power generation. Firstly, the prediction error samples are partitioned according to the predicted power so that the error distribution of each predicted power bin is approximated by a Gaussian distribution. And, the dynamic cloud models with given predicted power are established to obtain the cloud drop map of the prediction error. Finally, at a given confidence level, the upper and lower limits of the predicted value are calculated based on the quantile theory.

Taking the measured data of a photovoltaic power plant in Germany as an example, the proposed model is validated. The results show that the model built in this paper has achieved better reliability performance comparing to the benchmark method.

References
[1] Pelland S, Remund J, Kleissl J, et al. (2013) Photovoltaic and solar forecasting: state of the art. IEA PVPS Task,
[2] Long H, Zhang Z, Su Y. (2014) Analysis of daily solar power prediction with data-driven approaches. Applied Energy, 126: 29-37.
[3] Gong YF, Lu ZX, Qiao Y, et al. (2016) Photovoltaic power prediction technology. Automation of Electric Power Systems, Feb. 25, 40(4): 140-149.
[4] Wan C, Lin J, Song Y, et al. (2017) Probabilistic Forecasting of Photovoltaic Generation: An Efficient Statistical Approach. IEEE Transactions on Power Systems, 32(3):2471-2472.
[5] Scolari E, Sossan F, Paolone M. (2016) Irradiance prediction intervals for PV stochastic generation in microgrid applications. Solar Energy, 139:116-129.
[6] Yamazaki T, Homma H, Wakao S, et al. (2016) Estimation Prediction Interval of Solar Irradiance Based on Just-in-Time Modeling for Photovoltaic Output Prediction. Electrical Engineering in Japan, 195(3):1-10.
[7] Pierro M, Bucci F, De Felice M, et al. (2016) Multi-Model Ensemble for day ahead prediction of photovoltaic power generation. Solar Energy,134:132-146.
[8] Iversen E B, Morales J M, Møller J K, et al. (2014) Probabilistic forecasts of solar irradiance using stochastic differential equations. Environmetrics, 25(3):152–164.
[9] Bracale A, Caramia P, Carpinelli G, et al. (2013) A Bayesian Method for Short-Term Probabilistic Forecasting of Photovoltaic Generation in Smart Grid Operation and Control. Energies, 6(2):733-747.
[10] Zhou TX, Zhou SL. (2017) The interval probability prediction of photovoltaic generation. Journal of Tongling University, (2):108-110.
[11] Zhao WJ, Zhang N, Kang CQ, et al. (2015) A Method of Probabilistic Distribution Estimation of Conditional Forecast Error for Photovoltaic Power Generation. Automation of Electric Power Systems, (16): 8-15.
[12] Wang K. (2013) Study on photovoltaic power prediction considering multiple uncertainties North China Electric Power University.
[13] Bin Fu, Li DG, Wang MK. (2011) Review and prospect on research of cloud model. Application Research of Computers, 28(2): 420-426.
[14] Li DY, Meng HJ. (1995) Membership clouds and membership cloud generators. Computer R&D, (6):15-20.
[15] Li DY, Liu CY. (2004) Study on the Universality of the Normal Cloud Model. ENGINEERING SCIENCE, 6(8):28-34.
[16] Ye Q, Li SY, Zhang YY, et al. (2011) Cloud model and application overview. Computer Engineering and Design, 32(12):4198-4201.
[17] Pinson P, Nielsen HA, Møller JK, et al. (2007) Nonparametric probabilistic forecasts of wind power: required properties and evaluation. Wind Energy, 10: 497-516.
[18] Yan J, Liu YQ, Han S, Wang YM, Feng SL. (2015) Reviews on uncertainty analysis of wind power forecasting. Renewable & Sustainable Energy Reviews, December 52: 1322- 1330.