Applied Energy 212 (2018) 771–785

Contents lists available at ScienceDirect

Applied Energy

journal homepage: www.elsevier.com/locate/apenergy

A conditional model of wind power forecast errors and its application in scenario generation

Zhiwen Wang, Chen Shen*, Feng Liu

Department of Electrical Engineering, Tsinghua University, 100084 Beijing, China

HIGHLIGHTS

- Conditional models of wind power forecast errors under different forecast values.
- The appealing properties of Gaussian mixture model are adequately utilized.
- The non-Gaussianity and temporal-spatial correlation are considered.
- A fast method for generating non-Gaussian interdependent wind power scenarios.

ARTICLE INFO

Keywords:
Conditional distribution
Gaussian mixture model
Scenario generation
Wind power

ABSTRACT

In power system operation, characterizing the stochastic nature of wind power is an important albeit challenging issue. It is well known that distributions of wind power forecast errors often exhibit significant variability with respect to different forecast values. Therefore, appropriate probabilistic models that can provide accurate information for conditional forecast error distributions are of great need. On the basis of Gaussian mixture model, this paper constructs analytical conditional distributions of forecast errors for multiple wind farms with respect to different forecast values. The accuracy of the proposed probabilistic models is verified by using historical data. Thereafter, a sampling method is proposed to generate scenarios from the conditional distributions which are non-Gaussian and interdependent. The efficiency of the proposed sampling method is verified.

1. Introduction

Nowadays, a large amount of wind power has been integrated into power systems. In power system operation, a wind power forecasting tool plays an important role. Since the forecast values always deviate from the true ones more or less, the resulting forecast errors should be taken into account in generation scheduling [1]. In industrial practice, systems operators usually allocate reserves to compensate forecast errors [2]. On one hand, if the forecast errors are overestimated, reserves will be overcommitted, increasing operation costs; on the other hand, if the forecast errors are underestimated, reserves will be undercommitted, causing wind spillage and load shedding. Therefore, modeling the wind power forecast errors is a crucial issue for unit commitment (UC) and economical dispatch (ED).

Given an effective point forecasting tool, distributions of wind power forecast errors are conditioned on forecast values. In the literature [3–7], various probabilistic distributions have been adopted to model conditional distributions of wind power forecast errors. In [3], the authors point out that forecast errors of a single wind farm are far from Gaussian distributions, as the kurtosis could be over 10 (3 for the Gaussian). Beta distribution is suggested to model forecast error uncertainties. In a relevant study [4], the authors combine the Beta distribution and Dirac delta function, and obtain a mixed beta distribution, improving the model accuracy. Further, Bruninx et al. find that Beta distribution is not able to fully characterize the skewed and heavy-tailed forecast errors [5]. To solve this problem, the Levy-stable distribution is adopted. The test results in [5] show that the Levy-stable distribution outperforms Beta distribution. As the distributions of forecast errors are quite various under different forecast levels, the versatile distribution with three adjustable parameters is proposed in [6], achieving higher flexibility. Because the versatile distribution has more adjustable parameters than Beta/Gaussian distributions, it can better represent forecast error uncertainties. Following a similar idea, Menemenlis et al. use the time-varying Gamma-like distribution, whose parameters are adjusted as functions of forecast levels, to model the forecast errors [7]. These detailed conditional models [3–7] help the generation scheduling dynamically adjust reserves to different forecast levels. However, they are applicable to the single wind farm case only.

* Corresponding author.
E-mail address: shenchen@mails.tsinghua.edu.cn (C. Shen).

https://doi.org/10.1016/j.apenergy.2017.12.039
Received 14 July 2017; Received in revised form 21 November 2017; Accepted 6 December 2017

They cannot handle multivariate random variables.

To model a joint distribution for adjacent wind farms, the Copula technique has drawn much attention lately. In [8], the Gaussian Copula is used to model spatial interdependence structure in forecast uncertainties for multiple wind farms across the region. In a similar study [9], applying the Gaussian Copula theory, the authors conduct a multivariate probabilistic analysis for spatial correlated wind generation in the European grid. A remarkable advantage of using the Copula theory to model forecast error uncertainties is made in [10]. Different types of Copula functions, e.g., Gaussian, t, Clayton, Frank, Gumbel, are adopted to model the stochastic dependence of uncertainties. The Copula-based conditional distributions of forecast errors for multiple wind farms are obtained. For applications, authors in [11] propose a Copula-based chance-constrained optimization model for power system planning. Further, in order to deal with different dependency structures between pairs of random variables, e.g., wind and solar, the vine-Copula methods are investigated in [12,13], improving the accuracy in high-dimension cases. Although constructing the Copula-based conditional distributions for multiple wind farms has been investigated in the literature [8–13], it is hard to ensure that the constructed distributions have some desirable attributes. For instance, in terms of the scenario generation,[1] the Gaussian Copula method generates original scenarios from a Gaussian distribution, transforms the original scenarios into the Copula domain, and obtains final scenarios by using inverse transformations of marginal cumulative distribution functions (CDF). The procedure is time-consuming relative to sampling directly from a joint distribution. The scenario generation procedures of other types of Copulas are more complicated.

In order to incorporate forecast error uncertainties into UC and ED, scenarios generated from the conditional distributions of multiple wind farms are needed [14,15]. Generally, when random variables are non-Gaussian and interdependent, generating scenarios, i.e., sampling, from their joint distribution is difficult [16]. Many existing techniques are either not efficient enough or less accurate [17,18]. For example, the acceptance-rejection method and conditional sampling method need many steps and multiple transformations, which are time-consuming. The affine transformation method does not ensure that the generated scenarios strictly follow the predefined joint distribution, which may lead to inaccurate results. The Nataf technique is used in [19] to produce wind power scenarios. A time-varying correlation matrix is used in [20] for generating short-term wind uncertainty scenarios. Neither of them proves that the generated scenarios follow a predefined joint distribution. Using historical time series data of wind power and the kernel density estimator, Xydas et al. propose a generation method for forecast scenarios [21]. Alternatively, Morales et al. adopt the autoregressive model to generate time series data for wind power scenarios [22]. However, these techniques do not retain the original distributions of uncertainties [12]. To the best of the authors knowledge, when the conditional forecast error distribution of multiple wind farms is available, there is not an accurate and efficient sampling method that can generate scenarios from the non-Gaussian and interdependent joint distribution.

To address these important issues, this paper aims at a systematic methodology that can accurately model conditional distributions and generate scenarios. The original contributions are twofold:

(1) On the basis of Gaussian mixture model (GMM), this paper constructs conditional distributions of wind power forecast errors for multiple wind farms under different forecast values. With the proposed distributions, non-Gaussianity and correlations of forecast error uncertainties can be handled. Whats more, operator can conveniently obtain the conditional distribution of the aggregated forecast errors across a region.

(2) Based on the proposed probabilistic model, a method is developed to generate scenarios for wind power forecast errors with high accuracy and efficiency. The method is proved to be an exactly accurate sampling method for interdependent random variables. With the proposed sampling method, tens of thousands of scenarios can be generated within milliseconds.

The rest of this paper is organized as follows. In Section 2, a framework for the proposed methodology is provided. In Section 3, the GMM is used to represent a joint distribution of actual wind power outputs and forecast values for multiple wind farms. In Section 4, analytical formulae are derived to construct a conditional distribution of wind power forecast errors. Advantages of the proposed probabilistic model are discussed. In Section 5, a method that generates scenarios from the constructed conditional distribution is proposed. Case study results are presented in Section 6. Conclusions and limitations are shown in Section 7.

2. Framework

The proposed methodology consists of three phases. A flow chart is shown in Fig. 1.

Phase 1: Modeling a probability density functions (PDF) of the actual wind power outputs and forecast values by a GMM. Without loss of generality, let a random vector $\boldsymbol{X}$ denote actual wind power outputs of multiple wind farms, $\boldsymbol{Y}$ denote the corresponding forecast values, and $\boldsymbol{Z}$ denote the forecast errors. A GMM is used to represent the joint PDF of an aggregated random vector $[\boldsymbol{X}^{\mathrm{T}}\ \boldsymbol{Y}^{\mathrm{T}}]^{\mathrm{T}}$.

Phase 2: Constructing conditional distributions of wind power with respect to forecast values $[\boldsymbol{X}|\boldsymbol{Y}]$, as well as the conditional distributions of wind power forecast errors $[\boldsymbol{X}|\boldsymbol{Y}]$.

Phase 3: Generating scenarios from the constructed conditional

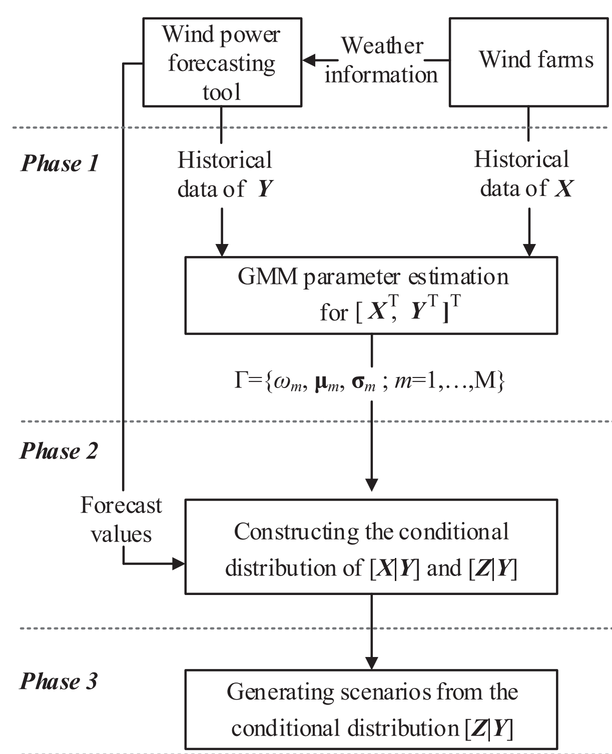

Fig. 1. Implementation procedure of the proposed methodology.

[1] A scenario can be understood as a plausible realization of uncertainty [14]. The uncertainty could be formulated as random variables or a stochastic process. Wind power scenario generation means producing a set of possible realization of wind power uncertainty. From the prospective of the probability theory, the scenario generation indeed means generating samples from a given probabilistic distribution. In this sense, we abuse the terminology "sampling" to stand for "scenario generation" in this paper. A rigorous definition and several illustrative examples can be found in [14].