IEEE Access
Multidisciplinary | Rapid Review | Open Access Journal

Received July 10, 2020, accepted July 17, 2020, date of publication July 23, 2020, date of current version August 5, 2020.

Digital Object Identifier 10.1109/ACCESS.2020.3011511

Learning-Based Probabilistic Power Flow Calculation Considering the Correlation Among Multiple Wind Farms

XI ZHU, CHUNMING LIU, CHENBO SU, AND JIAOMIN LIU

State Key Laboratory of Alternate Electrical Power System with Renewable Energy Sources, North China Electric Power University, Beijing 102206, China

Corresponding author: Chunming Liu (cm_liu@163.com)

ABSTRACT Wind power generation provides a new route for the sustainable development of energy. However, with the large scale integration of wind farms, the volatile wind energy and the correlation among various wind speeds bring a considerable quantity of uncertainty for the power system operation. In this paper, a probabilistic power flow (PPF) analysis model for the power system incorporating the correlation among various wind speeds is proposed. As distinct from existing studies, in this paper, we introduce the relevance vector machine (RVM) into correlation modeling of wind speeds via historical learning samples to construct bivariate joint distribution. Compared with the conventional parameter estimation methods, the proposed method has higher flexibility and computational efficiency. On this basis, the regular vine copula approach is adopted further to build the multivariate joint distribution model of wind speeds. To calculate the PPF of power system with the integration of wind power, we employ the three-point estimation method (3PEM) while the Rosenblatt Transformation technique is proposed to transform the input variables into independent variables. The effectiveness of the proposed calculation framework is examined through simulation studies, and the obtained results illustrate the advantages of the proposed method.

INDEX TERMS Wind power, probabilistic power flow, correlation, relevance vector machine, Rosenblatt transformation technique.

I. INTRODUCTION

With increasing concerns about sustainable utilization of energy resources, wind energy becomes one of the most popular substitutes for power generation. Compared with traditional generation, wind resources have a tremendous potential to facilitate environmental protection without air pollution, as well as mitigate energy crisis relying on its abundant reserve in nature. However, with the higher penetration of wind energy, the variability and correlation of wind speeds bring many uncertainties to the operation of the power system. Therefore, the conventional deterministic power flow calculation may be no longer appropriate for analyzing the effects of wind generation on power system operation.

To deal with the above issue, extensive research has been conducted to study the effect of uncertainty (e.g., the output of renewable distributed generations, power demand) on power system operation. For example, reference [1] adopted probability distribution to construct the model of uncertain real time line capacities caused by the integration of wind turbines into the power system. To deal with the wind generation uncertainty, reference [2] used a scenario-based approach and made an economic operation scheme based on stochastic optimization. Similarly, the stochastic nature of solar power was described via a scenario-based approach in reference [3]. Unlike the above probability methods, reference [4] adopted uncertainty sets to model the uncertainties in solar and wind output powers. Reference [5] introduced the robust optimization to research the effect of uncertainty produced by high-penetration renewable energy on the transaction cost of a microgrid. Likely, to deal with the issue of renewable energy generation (REG) uncertainty, reference [6] built a multiple-stage robust optimization model and solved the energy management problem caused by unbalanced generation and demand. The above literature [1]–[6] mainly studied the influence of uncertainties on the economy or security of power system operation.

The associate editor coordinating the review of this manuscript and approving it for publication was Canbing Li.

To study the influence of load uncertainty on state variables in the power system, B. Borkowska proposed probabilistic load flow calculation [7] to make up for the deficiency of deterministic power flow calculation. At present, there have been extensive efforts in exploring calculation methods of probabilistic power flow (PPF) with uncertain input variables, including simulation [8]–[11], analytical [12]–[15], approximate methods [16]–[19]. For example, reference [8], [9] developed the Monte Carlo (MC) based PPF analysis approach to evaluate the effects of renewable energy generation (REG) on the operation state of the power system, while considering the randomness and uncertainty of REG. MC simulation is the most accurate method but inapplicable to a practical program with a high computational burden. To decrease the computational time, reference [12] adopted statistical moments and Cornish-Fisher expansion based on linearized power flow equations to conduct PPF calculation incorporating the wind generation. Most analytical methods for PPF are based on the linear power system models and do not adequately consider the correlation of input variables. Unlike MC simulation and analytical methods, approximate methods are computationally effective while ensuring enough accuracy of estimation results such as the point estimate method (PEM), which does not need to evaluate the derivatives of nonlinear power flow equations. Hong's point estimate methods were applied into reference [16] to handle the uncertainty of generation unit outages or load demand for PPF calculation. In reference [17], the MC simulation method was compared with both two-point estimate and Latin hypercube sampling methods. The point estimate method was used to analyze the impact of increasing penetration of intermittent generation resources on load flow of power system in [18]. In this research, different wind generation and loads scenarios were assumed to verify the point estimate method's effectiveness.

The above literature [8]–[19] either did not consider wind power or assumed that there was no correlation among the wind speeds in various wind farms for PPF calculation. However, with the growing demand for wind power, there are usually more than one wind farm in an area. Moreover, the wind farms in the same area are generally influenced by the same wind speed zone. As a result, the wind speeds of multiple wind farms within a particular region have spatiotemporal correlations with each other. The results of PPF without considering the correlation will be too rough to provide reliable information for further planning and operational evaluation of the power system. Facing these problems, the Copula function is proposed to model the correlation among wind speeds via constructing joint distribution function [20]–[28]. For example, reference [20] studied the bivariate correlation of wind speeds using the time-adaptive quantile-copula approach. To simulate the correlation among multiple wind farms, reference [21] proposed the Archimedean copula approach to simulate the correlation of random variables, including photovoltaic and wind power, in a distribution network. In the same vein, the authors of [22] adopted the Normal copula technique to model wind power dependence for the study of the large-scale integration of wind generation in the power system. A fuzzy copula technique was proposed in [23] to express the uncertain wind speed correlation caused by inadequate historical wind speeds data. To construct the dependence structure model of different wind speeds, reference [24] employed a Laplace copula function to study the PPF problem.

Through the above literature review, it can be seen that the spatiotemporal correlation among wind speeds is described by using a single Copula function. However, in reality, the wind speeds series have asymmetrical tail characteristics. It is difficult to capture this feature by just relying on a single Copula function. Specially, the Copula functions are constructed via parameter estimation methods (such as maximum likelihood estimation (MLE), least square (LSQ) and Expectation-Maximization algorithm (EM)). In these parameter estimation methods, since the effects of different kernel functions on the estimation results differ a lot, these methods have a higher requirement for the choice of kernel functions. Besides, the computational process of parameter estimation methods could become so complicated with a large number of kernel functions and samples that the estimation results will be over-fitting.

In practice, to utilize the PEM, the interdependent wind power of different wind farms needs to be transformed into independent variables for further point estimation. Orthogonal transformation and Nataf transformation are commonly used to transform the dependent variables into independent variables [26], [27]. However, orthogonal transformation and Nataf transformation cannot accurately capture the nonlinear relationship between variables due to using a linear correlation coefficient of variables as input [29].

In this paper, a methodological framework for PPF analysis of the power system with incorporating the correlation among various wind speeds is proposed. To accurately model the wind speeds correlation of multiple wind farms, a novel bivariate joint distribution estimate approach based on the relevance vector machine (RVM) has been adopted. Specifically, RVM is applied to convert the nonlinear joint distribution problem into the high dimensional linear problem by thoroughly learning historical samples instead of parameter estimation. As a result, the bivariate joint distribution could be obtained in a shorter time by solving a linear model. On this basis, the Regular vine (R-vine) Copula approach is adopted further to build the multivariate joint distribution model of wind speeds. To calculate the PPF of power system with the integration of wind power, a three-point estimation method (3PEM) is utilized while input variables are transformed into independent variables using the Rosenblatt Transformation technique. Comparing with existing works in the relevant research field, the major contributions of this paper can be summarized as follows:

1) We propose a calculation framework to determine the PPF of the power system incorporating the correlation among multiple wind farms.

2) A learning-based technique integrating the RVM and R-vine copula is developed and used to construct the multivariate joint distribution function of wind speeds in our study. Compared with traditional parameter estimation based correlation analysis, the RVM can effectively avoid over-fitting and dimension disaster with faster computational speed. Hence, it could provide a more accurate estimation for the correlation among various wind speeds, enabling the PPF analysis of the power system to be more realistic.
3) We adopt the Rosenblatt Transformation approach to generate independent input variables for a 3PEM based PPF calculation.

The rest of the paper is organized as follows. Section 2 introduces the method to construct a multivariate joint distribution function of wind speeds. Then section 3 elaborates on the PPF calculation model. Based on the proposed formulation, numerical studies are carried out in section 4. Finally, Section 5 gives the conclusion of the research.

II. CONSTRUCTION OF CORRELATION AMONG WIND FARMS

In this section, the RVM technique for constructing bivariate joint distribution is introduced first, and then the multivariate joint distribution is further modeled via the R-vine Copula model.

A. RVM MODEL

As for a special linear programming model derived from Bayesian theory, the RVM model can convert nonlinear problems in low dimensional space into linear problems in high dimensional space by appropriate kernel function mapping. In the process of iterative calculation of RVM, most of the hyperparameters tend to a large number (could be regarded as infinity). As the hyperparameter approaches infinity, the reciprocal of the hyperparameter tends to zero, and the corresponding kernel function will not affect the regression results. Then a sparse model can be obtained. This could avoid dimension disaster. Therefore, RVM is an ideal method to obtain the joint distribution function of random variables through the marginal distribution of random variables by learning from historical information.

Assumed that $\boldsymbol{x} = [x_1, x_2, \cdots, x_n]^T$ is input vector with N observations values and $Y = y(\boldsymbol{x};\boldsymbol{\omega})$ is corresponding outputs, the model of RVM can be constructed as follows [30]:

$$y(x;\omega) = \sum_{i=1}^{N} \omega_i K(x, x_i) + \omega_0 = \omega^T \phi(x) \tag{1}$$

where $K(\boldsymbol{x}, x_i)$ is a vector composed of multiple kernel functions; $\boldsymbol{\omega} = [\omega_0, \omega_1, \ldots \omega_n]$ represents the weight vector.

According to the corresponding target values $\boldsymbol{t} = (t_1, \cdots, t_n)^{\boldsymbol{T}}$ of given input vector $\boldsymbol{x}$, the likelihood function of the training sample can be obtained via

$$p(t|x_i, \beta, \alpha) = \prod_{i=1}^{N} p(t_i|x_i, \omega, \beta^{-1} I) \tag{2}$$

where α represents the vector of all hyperparameter; Each weight parameter ω_i corresponds to a hyperparameter α_i in the RVM model, which is different in general linear models.

Based on the likelihood function of the training sample, the weight-prior could be described as follow:

$$p(\omega|\alpha) = \prod_{i=1}^{M} N(\omega_i|0, \alpha_i^{-1}) \tag{3}$$

During the above calculation, most elements of vector α tend to infinity, and the posterior distributions of corresponding weight parameters are concentrated at zero.

According to Bayesian theory, the posterior probability distributions of weight parameters satisfy $p(\boldsymbol{\omega} \mid \boldsymbol{t}, x_i, \beta, \boldsymbol{\alpha}) \propto p(\boldsymbol{t} \mid \boldsymbol{\omega}, x_i, \beta, \boldsymbol{\alpha})\, p(\boldsymbol{\omega} \mid \boldsymbol{\alpha})$, and could be obtained via

$$p(\omega|t, x_i, \beta, \alpha) = N(\omega|m, \Sigma) \tag{4}$$

where $\boldsymbol{m}$ and Σ are the mean and covariance of weight parameters respectively, given as follows:

$$m = \beta \Sigma K^T t \tag{5}$$

$$\Sigma = (A + \beta K^T K)^{-1} \tag{6}$$

where A= diag (α_i).

In this paper, we adopt T-copula, Clayton copula, and Gumbel copula function to construct the joint distribution to capture the nonlinear symmetric correlation, the asymmetric correlation, and the tail correlation of joint distribution of two wind farms. T-copula has a symmetric tail correlation structure to capture the nonlinear symmetric correlation between input variables. The distribution of Gumbel copula and Clayton Copula are asymmetric, which can reflect the asymmetric correlation of input variables. And Gumbel copula has a characteristic of strong upper tail correlation, while Clayton copula distribution has a strong lower tail correlation. Suppose that the variables $\boldsymbol{v}_1 = (v_{11}, v_{12}, \ldots, v_{1n})$ and $\boldsymbol{v}_2 = (v_{21}, v_{22}, \ldots, v_{2n})$ stand for wind speed sampled from two wind farms respectively. According to the historical wind speed information of the two wind farms, the marginal distribution functions $F_1(\boldsymbol{v}_1)$ and $F_2(\boldsymbol{v}_2)$ are obtained using kernel density estimation. Based on this, the T-copula, Clayton-copula, and Gumbel-copula can be calculated, respectively denoted as $\zeta\zeta C_T(F_\vartheta(x_\vartheta), F(x)) C_T(F_1(\boldsymbol{v}_1), F_2(\boldsymbol{v}_2))$, $C_C(F_1(\boldsymbol{v}_1), F_2(\boldsymbol{v}_2))$ and $C_G(F_1(\boldsymbol{v}_1), F_2(\boldsymbol{v}_2))$. $C_T(F_1(\boldsymbol{v}_1), F_2(\boldsymbol{v}_2))$, $C_C(F_1(\boldsymbol{v}_1), F_2(\boldsymbol{v}_2))$ and $C_G(F_1(\boldsymbol{v}_1), F_2(\boldsymbol{v}_2))$ take n sampling values respectively to form a sample space denoted as $\mathbf{U}\{U_1, U_2 U_3\}$, in which $U_1 = [u_{11}, u_{12}, \ldots, u_{1n}]^T$, $U_2 = [u_{21}, u_{22}, \ldots, u_{2n}]^T$ and $U_3 = [u_{31}, u_{32}, \ldots, u_{3n}]^T$, are all normalized to control the data range within [0,1]. Then the bivariate joint distribution function of two wind farms could

be described as follows:

$$\hat{C}(v_1, v_2) = \left[\hat{C}_1, \hat{C}_2, \ldots, \hat{C}_n\right]^T = g_1 U_1 + g_2 U_2 + g_3 U_3 \tag{7}$$

where g_1, g_2, and g_3 are the weight parameters of the T-Copula, Clayton-Copula, and Gumbel-Copula, respectively. The element of the bivariate joint distribution function $\zeta\hat{C}_\xi(\boldsymbol{x}_\vartheta, \boldsymbol{x})$ is calculated via

$$\hat{C}_\xi(v_1, v_2) = g_1 u_{1\xi} + g_2 u_{2\xi} + g_3 u_{3\xi}, \xi = 1, 2, \ldots, n \tag{8}$$

According to (1), (7), (8), the RVM model of two wind farms could be obtained as follows:

$$\begin{aligned} & y\left(\hat{C}_\xi(v_1, v_2)\right) \\ &= \sum_{i=1}^{n} \omega_i K\left(\hat{C}_\xi(v_1, v_2), \hat{C}_i(v_1, v_2)\right) + \omega_0 \\ &= \sum_{i=1}^{n} \omega_i \exp\left(-\frac{\left\|\hat{C}_\xi(v_1, v_2) - \hat{C}_i(v_1, v_2)\right\|}{2\sigma_{\hat{C}}^2}\right) + \omega_0, \\ & \xi = 1, 2, \ldots, n \end{aligned} \tag{9}$$

where $\left\|\hat{C}_\xi(\boldsymbol{v}_1, \boldsymbol{v}_2) - \hat{C}_i(\boldsymbol{v}_1, \boldsymbol{v}_2)\right\|$ is the norm of the difference between two elements in bivariate joint distribution function; $\sigma_{\hat{C}}$ is the standard deviation of the bivariate joint distribution.

To obtain the weight parameter ω and $\boldsymbol{g}(g_1, g_2, g_3)$, equation (4) could be transformed by using natural logarithm, and then calculate the partial derivative of (10), respectively shown in (10) and (11):

$$\begin{aligned} \ln p(\omega|t, u_i, \beta, \alpha) &= \ln N(\omega|t, 0, C) \\ &= -\frac{1}{2}\left\{N\ln(2\pi) + \ln|C| + tC^{-1}t\right\} \end{aligned} \tag{10}$$

$$\begin{cases} \dfrac{\partial\left(\ln p(\omega|t, u_i, \beta, \alpha, g)\right)}{\partial\omega} = 0 \\ \dfrac{\partial\left(\ln p(\omega|t, u_i, \beta, \alpha, g)\right)}{\partial g} = 0 \end{cases} \tag{11}$$

where $\boldsymbol{t} = (t_1, \cdots, t_N)^{\boldsymbol{T}}$ is sampled from the empirical copula $\hat{C}_n$ of the two wind farms. And $\hat{C}_n$ is defined as follows:

$$\begin{aligned} & \hat{C}_n(F_1(v_{1k}), F_2(v_{2k})) \\ &= \frac{1}{n}\sum_{i=1}^{n} I_{[F_1(v_{1i}) \le F_1(v_{1k})]} I_{[F_2(v_{2i}) \le F_2(v_{2k})]} \end{aligned} \tag{12}$$

where $F_1(v_{1k})$ and $F_2(v_{2k})$ are the marginal distribution; I is an indicator function, if $F_1(v_{1i}) \le F_1(v_{1k})$, or $F_2(v_{2i}) \le F_2(v_{2k})$, $I = 1$; otherwise, $I = 0$.

Besides, the matrix C in (10) is calculated via

$$C = \beta^{-1} I + KA^{-1}K^T \tag{13}$$

where

$$A = diag\left(\alpha_i^{new}\right) = diag\left(\frac{\gamma_i}{m_i^2}\right) \tag{14}$$

$$(\beta^{new})^{-1} = \frac{||t - Km||^2}{N - \sum_i \gamma_i} \tag{15}$$

where γ_i reflects the prediction accuracy of the RVM model and is defined via

$$\gamma_i = 1 - \alpha_i \Sigma_i \tag{16}$$

B. R-VINE MODEL AND MULTIVARIATE DISTRIBUTIONS

Use one space after periods and colons. Hyphenate complex Nonlinear joint distribution of two wind farms is estimated by RVM based method in Section II.A. However, in practice, due to the high penetration of wind power, there are usually more than two correlated wind farms in the real power system. Because of this, this section constructs a high dimensional joint distribution of multiple correlated wind farms based on the above bivariate distribution (9) by introducing the R-vine copula approach.

According to the copula theory [31], the multivariate distributions function of various correlated wind farms can be formed via the marginal cumulative distribution function (CDF) $F_i(\boldsymbol{v}_i)$, described as follows:

$$F(v_1, v_2, \cdots, v_n) = C_j(F_1(v_1), F_2(v_2), \cdots, F_n(v_n)) \tag{17}$$

However, in practice, it is challenging to construct multiple dimensional distribution functions by adopting the conventional Copular model. To solve this problem, *Bedford* proposed R-vine Copular based on a graphical model [32]. A regular vine model with d variables consists of a sequence of nested trees $T_1, \ldots, T_{d-1}$ with nodes $N_1, \ldots, N_{d-1}$ and edges $E_1, \ldots, E_{d-1}$ [32], while satisfies the following conditions:

(1) The tree T_1 has N_1 nodes and E_1 edges.

(2) The tree T_i has N_i nodes equal to the E_i-1 edges in tree T_{i-1}, that is, $N_i = E_{i-1}$, for $i = 1,\ldots, d$-1.

(3) If one edge in tree Ti is connected with another edge in tree T_{i+1}, these two edges share a common node in tree T_i.

Based on the above definition of R-Vine model, the joint probability density function (PDF) $f(\boldsymbol{v}_1, \boldsymbol{v}_2, \ldots, \boldsymbol{v}_d)$ of a d-dimensional wind speed sequence $\boldsymbol{v} = (\boldsymbol{v}_1, \boldsymbol{v}_2, \ldots, \boldsymbol{v}_d)$ can be decomposed into the marginal density function of each variabl e $(f_l(\boldsymbol{v}_l))$ as follows:

$$\begin{aligned} f(v_1, v_2, \ldots, v_d) = & \prod_{l=1}^{d} f_l(v_l) \times \prod_{i=1}^{d-1}\prod_{e\in E_i} \hat{c}_{j(e),k(e)|D(e)} \\ & \times(F(v_{j(e)}|v_{D(e)}), F(v_{k(e)}|v_{D(e)})) \end{aligned} \tag{18}$$

where $j(e)$, $k(e)$ are two conditional nodes of edge e; $D(e)$ is conditional set; The edge $e = j(e),k(e)|D(e)$ belongs to E_i; $\hat{c}_{j(e),k(e)|D(e)}$ is a PDF of conditional bivariate copula which can be used to associate each edge e. Besides, the conditional CDF $F(\boldsymbol{v}_{j(e)}|\boldsymbol{v}_{D(e)})$ can be expressed via

$$F(v_{j(e)}|v_{D(e)}) = \frac{\partial\hat{C}(F(v_{j(e)}), F(v_{D(e)}))}{\partial F(v_{D(e)})} \tag{19}$$

Above all, we can find that precisely estimating the copula function is the key to improve the accuracy of the calculation.

III. PROBABILISTIC POWER FLOW CALCULATION

The uncertain wind power caused by uncertain wind speeds has a significant effect on the power system's operation state. Traditional deterministic power flow cannot provide valid information for further planning and risk evaluation of the power system. To solve this problem, the probabilistic power flow (PPF) of power system needs to be calculated as follows:

$$S = f(H) \tag{20}$$

where S is the output of PPF; H is the input vector of PPF, including correlated wind power $\{h_1, h_2, \ldots, h_d\}$ and other stochastic variables in conventional power system $\{h_{d+1}, h_{d+2}, \ldots, h_n\}$, such as load, the output of traditional generation units. Besides, the correlated wind power can be calculated according to correlated wind speeds obtained by (18). The relationship between wind power and wind speeds is shown as follows [27]:

$$P_w = \begin{cases} 0 & v < v_{wi}, v > v_{wo} \\ \dfrac{v - v_{wi}}{v_r - v_{wi}} P_{\text{rate}} & v_{wi} < v < v_r \\ P_{\text{rate}} & v_r \le v < v_{wo} \end{cases} \tag{21}$$

where v_{wi} is the cut-in wind speed, equal to 3 m/s; v_{wo} is the cut-out wind speed, equal to 24 m/s; v_r is the rated wind speed, equal to 13 m/s.

To obtain more effective power flow analysis, this paper considers spatial and temporal correlations among different wind farms output to calculate the probability power flow by combining the 3PEM with Rosenblatt Transformation technique, described in Section III.A and III.B.

A. THREE-POINT ESTIMATION METHOD

In 3PEM [33], each input random variable $h_{i,k}(k = 1, 2, 3)$ takes three sampling values based on the statistical characteristics of input variables, shown as follows:

$$h_{i,k} = \bar{h}_i + \xi_{h_{i,k}} s_i, \quad k = 1, 2, 3 \tag{22}$$

where $\bar{h}_i$ and s_i are the expectation and variance of each group of input variables respectively; $\xi_{h_{i,k}}$ is the standard sample value of $h_{i,k}$, which could be calculated via

$$\begin{cases} \xi_{h_{i,k}} = \dfrac{\lambda_{h_i}}{2} + (-1)^{3-k} \sqrt{\upsilon_{h_i} - \dfrac{3}{4}\lambda_{h_i}^2} & k = 1, 2 \\ \xi_{h_{i,k}} = 0 & k = 3 \end{cases} \tag{23}$$

where, λ_{h_i} and υ_{h_i} are the skewness coefficient and kurtosis coefficient of random variables respectively, described as follows:

$$\lambda_{h_i} = \frac{1}{n s_i^3} \sum_{i=1}^{n} (h_i - \bar{h}_i)^3 \tag{24}$$

$$\upsilon_{h_i} = \frac{1}{n s_i^4} \sum_{i=1}^{n} (h_i - \bar{h}_i)^4 \tag{25}$$

According to (23)-(25), each sampling result's weight coefficients could be calculated via

$$\begin{cases} w_{h_{i,k}} = \dfrac{(-1)^{3-k}}{\xi_{h_{i,k}}(\xi_{h_{i,1}} - \xi_{h_{i,2}})} & k = 1, 2 \\ w_{h_{i,k}} = \dfrac{1}{n} - \dfrac{1}{\upsilon_{h_i} - \lambda_{h_i}^2} & k = 3 \end{cases} \tag{26}$$

Through the above sampling, the corresponding output random variable Y (i, k) is shown in (27)

$$\begin{aligned} Y(i,k) &= G(\mu_{h1}, \cdots, \mu_{hi-1}, x_{i,k}, \mu_{hi+1}, \cdots, \mu_{hn}), \\ & i = 1, 2, \cdots n, k = 1, 2, 3 \end{aligned} \tag{27}$$

Based on (23) - (27), the above output variables could be evaluated using the deterministic method, described as follows:

$$E(Y^l(i,k)) \approx \sum_{i=1}^{n} \sum_{k=1}^{3} w_{i,k} (Y(i,k))^l \tag{28}$$

where $E\left(Y^l(i,k)\right)$ is the l-order origin moment of the corresponding output variable.

Based on the first-order moment and second-order moment of output variable obtained by (28), the expectation $(\mu_{y,i})$ and variance $(\sigma_{y,i})$ of output variable can be expressed as follow:

$$\mu_{y,i} = E(Y(i,k)) \quad i = 1, 2, \cdots, n \; k = 1, 2, 3 \tag{29}$$

$$\sigma_{y,i} = \sqrt{E\left(Y^2(i,k)\right) - \mu_{y,i}^2} \quad i = 1, 2, \cdots, n \; k = 1, 2, 3 \tag{30}$$

B. ROSENBLATT TRANSFORMATION BASED 3PEM

The above 3PEM model in Section 3.1 requires that the input variables must be independent with each other. Therefore, correlated wind speeds need to be transformed into independent variables. In this paper, we used the Rosenblatt transformation technique to obtain independent input variables for the 3PEM model while maintaining the correlation among wind power reflected by the joint distribution function. The Rosenblatt transformation method is not limited by a specific type of distribution. And no matter whether the correlation of variables is linear or not [30], the Rosenblatt transformation has better performance.

Referring to (2), the joint probability density function (PDF) $f = (x_1, x_2, \ldots, x_d)$ of the d-dimensional random vector $\boldsymbol{x} = (x_1, x_2, \ldots, x_d)$ can be reformulated as follows:

$$\begin{aligned} & f(x_1, x_2, \cdots, x_d) \\ &= \prod_{k=1}^{n} f_k(x_k) \cdot \prod_{j=1}^{n-1} \prod_{i=1}^{n-j} \hat{c}_{i,i+j|i+1,\cdots,i+j-1} \\ &\quad \cdot (F_{i|i+1,\cdots,i+j-1}(x_i | x_{i+1}, \cdots, x_{i+j-1}), \\ &\quad \times F_{i+j|i+1,\cdots,i+j-1}(x_{i+j} | x_{i+1}, \cdots, x_{i+j-1})) \end{aligned} \tag{31}$$

To solve the above PDF, taking n=2 as an example, the conditional probability density function needs to be calculated as follows:

$$f_{2|1}(x_2|x_1) = \frac{\hat{c}_{1,2} f_1(x_1) \cdot f_2(x_2)}{f_1(x_1)} = \hat{c}_{1,2} f_2(x_2) \tag{32}$$

The conditional probability cumulative function ($F_{1|2}(x_1|x_2)$) can be expressed as follows:

$$F_{1|2}(x_1|x_2) = \int_{-\infty}^{x_2} f_{2|1}(x_2|x_1)dx_2 = \int_{-\infty}^{x_2} \hat{c}_{1,2} f_2(x_2)dx_2 \tag{33}$$

When n=3, conditional probability density function can be obtained as follows:

$$\begin{aligned} f_{3|1,2}(x_3|x_1, x_2) &= \frac{f_{123}(x_1, x_2, x_3)}{f_1(x_1) \cdot f_{2|1}(x_2|x_1)} \\ &= \frac{\hat{c}_{1,2} \cdot \hat{c}_{1,3} \cdot \hat{c}_{2,3|1} f_1(x_1) \cdot f_2(x_2) \cdot f_3(x_3)}{f_1(x_1) \cdot \hat{c}_{1,2} \cdot f_2(x_2)} \\ &= \hat{c}_{1,3} \cdot \hat{c}_{2,3|1} \cdot f_3(x_3) \end{aligned} \tag{34}$$

And the conditional probability cumulative function $F_{i|1,2,\ldots,i}(x_i|x_1, x_2, \ldots, x_i)$ can be deduced as follows:

$$\begin{aligned} &F_{i|1,2,\ldots,i}(x_i|x_1, x_2, \ldots, x_i) \\ &\quad = \int_{-\infty}^{x_i} \hat{c}_{1,i}\hat{c}_{2,i|1} \cdots \hat{c}_{i-1|1,2,\ldots,i-2} f_i(x_i)dx_i \end{aligned} \tag{35}$$

Based on the conditional probability cumulative function and the Rosenblatt Transformation technique, the correlated input variables could be transformed into independent variables $\Gamma = (\tau_1, \tau_2, \ldots, \tau_d)^T$.

$$\begin{cases} \tau_1 = \Phi^{-1}[F_1(x_1)] \\ \tau_2 = \Phi^{-1}\left[F_{2|1}(x_2|x_1)\right] \\ \vdots \\ \tau_d = \Phi^{-1}\left[F_{d|1,2,\cdots,d-1}(x_d|x_1, x_2, \cdots, x_{d-1})\right] \end{cases} \tag{36}$$

where Φ is the CDF of the standard normal distribution.

The procedures for calculating the PPF of power system with considering the correlation of wind speeds are presented as follow:

Step 1) Based on historical sample data, obtain marginal CDF and PDF of wind speeds in each wind farm using the kernel estimation method.

Step 2) Calculate the bivariate joint distribution function and PDF of every two wind farms via the RVM technique according to (7)-(16), denoted as $\hat{C}(F_1(x_1), F_2(x_2))$ and $\hat{c}(F_1(x_1), F_2(x_2))$ respectively.

Step 3) According to R-Vine model and bivariate joint distribution function obtained in *Step 2)*, construct high dimensional joint distribution function by (18) and (19).

Step 4) Transform the multiple dimensional joint distributions into various independent distribution based on the Rosenblatt Transformation technique via (36), and obtain independent wind speeds sequence by inversely converting the Weibull distribution of wind speeds.

Step 5) According to the relationship between wind speeds and wind power, calculate the corresponding wind power. Based on this, analyze the PPF of power system with considering the correlated wind speeds via adopting 3PEM, described in (22)-(30).

IV. CASE STUDY

A. TEST DATA

The effectiveness of the proposed PPF analysis approach is tested on an IEEE 118-bus test system with four wind farms connected to buses 5, 8, 18, and 23, respectively. The specific data of the IEEE 118-bus test system can refer to [34]. In this paper, we assume that load fluctuation follows the normal distribution, and the rated power of one wind farm is 80 MW with a constant power factor of 0.98. Besides, we adopt wind farms data of HeiBei province in China to construct the joint distribution functions of wind speeds. The obtained data covers 365 days with a time interval of one hour. Therefore, there is a total of 8760 wind speed points on each wind farm.

Based on the historical data of each wind farm, the marginal distribution of wind speeds could be obtained by the kernel estimation method, which is compared with corresponding empirical distribution, shown in Fig.1.

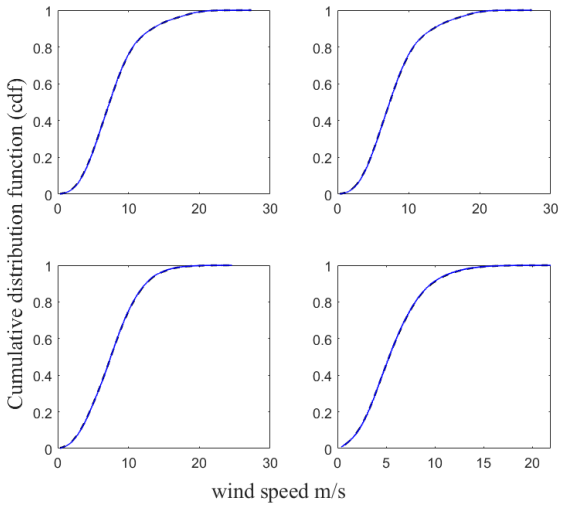

FIGURE 1. Kernel estimation of the marginal distribution of wind speeds.

Furthermore, Weibull distribution is adopted to inversely convert the independent distributions of wind speeds to independent wind speeds v_w for the further PPF analysis, shown as follows:

$$v_w = W^{-1}(\tau) = c_w^{k_w}\sqrt{\ln\frac{1}{1-\tau}} \tag{37}$$

where the shape parameter c_w and scale parameter k_w of Weibull distribution for each wind farm can be estimated via the real data of the studied wind farms, shown in Table.1.

In this simulation, we achieve the PPF calculation via Matlab on a PC (with 3.20GHz Intel Core i7, 8GB RAM).

B. SIMULATION RESULTS

1) EFFECTIVENESS OF PROPOSED CORRELATION ESTIMATION METHOD

In this study, we propose an RVM approach to estimate the joint distribution of wind speeds for every two wind farms. For simplicity, two groups of bivariate joint distribution

TABLE 1. Estimated value of weibull distribution parameters.

No.	c_W	k_W
1	8.9828	2.1142
2	8.8971	2.1092
3	7.4522	1.2036
4	7.1601	1.7866

functions of wind speeds and frequency histogram of empirical distribution are shown in Fig.2.

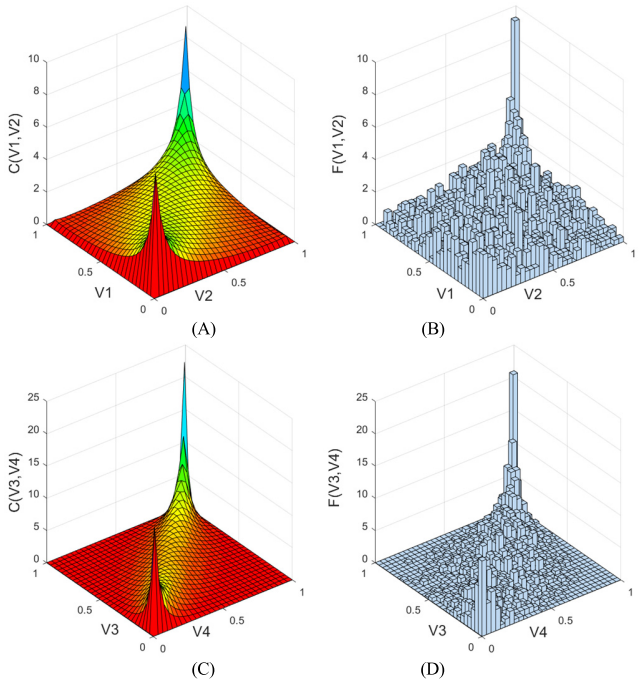

FIGURE 2. Joint distribution and frequency histogram of empirical distribution of two wind farms. (A) NO.1 AND NO.2 WIND FARMS. (B) NO.1 AND NO.2 WIND FARMS. (C) AND NO.3, NO.4 WIND FARMS.

To test the performance of the RVM method, we evaluate the difference between estimated joint distribution function and empirical distribution function by adopting Kolmogorov–Smirnov (K-S) test [35] and Euclidean Distance [36]. With high estimation accuracy, the estimated distribution should reach the corresponding empirical distribution as closely as possible. The formulation of the Kolmogorov–Smirnov (K-S) test can be expressed as follows:

$$K = \sup_{v} |C_n(v) - C(v)| \tag{38}$$

And the Euclidean Distance of two probability distribution is shown as follows:

$$d = |C_n(v) - C(v)| \tag{39}$$

where $C_n(v)$ is the 4-dimensional empirical distribution function of wind speeds; $C(v)$ is the 4-dimensional joint distribution function calculated by different joint distribution estimation methods.

According to the definition of K-S test and Euclidean Distance, a smaller value of K and a smaller value of d demonstrate that the estimated joint distribution is more similar to the corresponding empirical distribution.

Besides, to evaluate the effectiveness of the proposed method, we make a comparison of K-S test results, Euclidean Distance, and P-P diagram of the proposed method with those of EM and OLS methods.

The K and d values of the estimated distribution obtained by RVM, EM, and OLS methods are reported in Table.2. And for simplicity, the P-P diagrams of 4-dimensional joint distribution are illustrated in Fig.3.

TABLE 2. K-S test results and euclidean distance using different distribution estimation methods.

Method	d	K
$C_{\mathrm{RVM}}(v_1,v_2,v_3,v_4)$	1.7752	0.168
$C_{\mathrm{OLS}}(v_1,v_2,v_3,v_4)$	2.1872	0.192
$C_{\mathrm{EM}}(v_1,v_2,v_3,v_4)$	2.0513	0.183

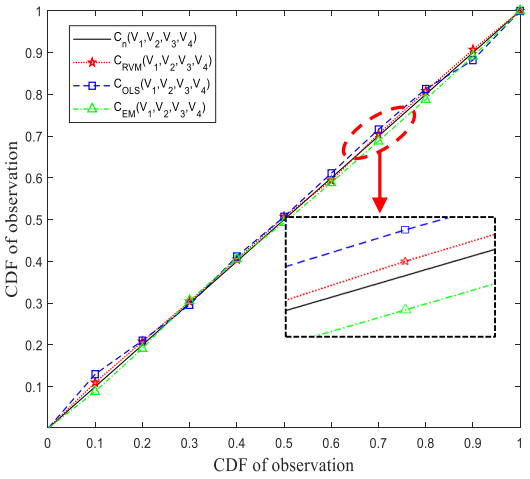

FIGURE 3. The comparison among p-p diagram of joint cdf estimated by different methods and p-p diagram of empirical distribution of two correlated wind farms.(A) NO.1 AND NO.2 WIND FARMS. (B) NO.3 AND NO.4 WIND FARMS.

As depicted in Table.2, the values K and values d of RVM method are all smaller than those of the EM method and OLS method. This indicates that the 4-dimensional joint distributions estimated by RVM are more consistent with the corresponding empirical distribution.

Also, the P-P diagrams of the joint distribution of 4-dimensional wind farms are illustrated in Fig.3. As can be seen, the P-P diagrams of three different estimation methods (including RVM, EM, and OLS) are all approximately linear and very close to the corresponding empirical distribution. However, compared with EM and OLS methods, the RVM method based distribution is more consistent with empirical distribution after magnifying the P-P diagrams locally (shown in Fig.3). This proves that the joint distribution estimated

via RVM can reflect the actual distribution better. Therefore, the proposed method has a more satisfactory performance for the multi-dimensional joint distribution estimation of wind speeds.

Furthermore, to illustrate the effect of the proposed correlation estimation method on the PPF analysis results, we make a comparison of the relative errors of PPF calculation results based on RVM, the non-parameter kernel density estimation (KDE) and OLS. The relative errors of the expected value and standard deviation of the PDF of voltage amplitude, voltage phase angle, active power, and reactive power are described as follow:

$$\varepsilon_{1\mu}^{x} = \left| \frac{\mu_{C_n}^{x} - \mu_{CEM}^{x}}{\mu_{C_n}^{x}} \right| \times 100\% \tag{40}$$

$$\varepsilon_{1\sigma}^{x} = \left| \frac{\sigma_{C_n}^{x} - \sigma_{CEM}^{x}}{\sigma_{C_n}^{x}} \right| \times 100\% \tag{41}$$

where, $\mu_{C_n}^{x}$, $\sigma_{C_n}^{x}$ and μ_{CEM}^{x}, σ_{CEM}^{x} represent the expected value and standard deviation respectively obtained from the empirical distribution and the above correlation estimation methods(RVM, KDE, and OLS). Besides, we use statistical characteristics (average value and maximum value) of the relative errors to illustrate the effectiveness of the calculation results obtained from the proposed method.

The average values and maximum values of the relative errors of PDF of the system state variables calculated based on RVM, KDE, and OLS are shown in Table.3.

TABLE 3. Relative error comparison of different correlation estimation methods.

	RVM		KDE		OLS	
	Average	Maximum	Average	Maximum	Average	Maximum
$\varepsilon_{1\mu}^{V}$	0.0015	0.0112	0.0016	0.0157	0.0019	0.0188
$\varepsilon_{1\sigma}^{V}$	1.1120	2.0522	1.1221	2.3055	1.1311	2.7642
$\varepsilon_{1\mu}^{\theta}$	0.5962	1.4425	0.5991	1.4632	0.5986	1.4857
$\varepsilon_{1\sigma}^{\theta}$	1.8633	2.6241	1.9562	2.6793	1.9902	2.7243
$\varepsilon_{1\mu}^{P}$	0.4286	1.9564	0.4398	2.1576	0.4567	2.3312
$\varepsilon_{1\sigma}^{P}$	0.7655	2.3228	0.8012	2.4523	0.8954	2.8864
$\varepsilon_{1\mu}^{Q}$	0.4533	4.2263	0.4823	4.7882	0.5231	5.1322
$\varepsilon_{1\sigma}^{Q}$	0.8924	6.0132	0.9441	6.8784	0.9852	7.0121

As observed in Table.3, the maximum difference of relative errors of all system state variables (voltage amplitude, voltage phase angle, and active power) between RVM and KDE is 0.87%, and that between RVM and OLS is 1%. This indicates that, different correlation estimation methods has a different influence on PPF analysis of power system with wind power integration. Besides, compared with relative errors of PPF analysis results based on the traditional methods (KDE and OLS), all the average and maximum values of relative errors PPF analysis results based on the proposed method (RVM) are smaller. This demonstrates that, the PDF of the system state variables obtained based on RVM is closer to the real distribution. Therefore, RVM has better performance than traditional methods in reflecting the impact of correlation among wind farms on the power system's PPF analysis.

2) PERFORMANCE OF THE PROPOSED PPF ANALYSIS METHOD

In this paper, Rosenblatt transformation based 3PEM is utilized to carry out PPF analysis of the power system with integrating multiple wind farms. To examine the effectiveness of the proposed method, we make a comparison of the calculation results obtained via the proposed method with that via the MC simulation, Nataf transformation-based 3PEM, and 3PEM without Rosenblatt transformation (assuming that the distribution of wind speeds is independent).

For this purpose, we calculate the relative errors of the expected value and standard deviation of the probability density distribution of system state variables (voltage amplitude, voltage phase angle, and active power) obtained from above three PPF analysis methods respectively. The relative errors of the expected value and standard deviation are described as follow:

$$\varepsilon_{2\mu}^{x} = \left| \frac{\mu_{C_n}^{x} - \mu_{PPF}^{x}}{\mu_{C_n}^{x}} \right| \times 100\% \tag{42}$$

$$\varepsilon_{2\sigma}^{x} = \left| \frac{\sigma_{C_n}^{x} - \sigma_{PPF}^{x}}{\sigma_{C_n}^{x}} \right| \times 100\% \tag{43}$$

where μ_{PPF}^{x}, σ_{PPF}^{x} represent the expected value and standard deviation respectively obtained from the above PPF analysis methods.

The average values and maximum values of the relative errors of the system state variables calculated via Rosenblatt transformation based 3PEM, the MC simulation, Nataf transformation based 3PEM, and 3PEM without Rosenblatt transformation (3PEM only) are shown in Table.4. Moreover, the PDF of voltage and power belong to bus 8 and 18 are presented from Fig.4 to Fig.7.

TABLE 4. Relative error comparison of different PPF analysis methods.

	Rosenblatt		MC simulation		Nataf		3PEM only	
	Ave	Max	Ave	Max	Ave	Max	Ave	Max
$\varepsilon_{2\mu}^{V}$	0.0015	0.0112	0.0015	0.0107	0.0016	0.1180	0.0024	0.0288
$\varepsilon_{2\sigma}^{V}$	1.1120	2.0522	1.0651	1.9632	1.1281	2.4533	12.241	16.676
$\varepsilon_{2\mu}^{\theta}$	0.5962	1.4425	0.5827	1.3955	0.5966	1.4428	0.7843	1.9324
$\varepsilon_{2\sigma}^{\theta}$	1.8633	2.6241	1.7549	2.6104	1.9011	2.7003	9.0132	12.479
$\varepsilon_{2\mu}^{P}$	0.4286	1.9564	0.4246	1.8756	0.4288	1.9563	0.8667	2.9581
$\varepsilon_{2\sigma}^{P}$	0.7655	2.3228	0.7231	2.2681	0.7822	2.7863	8.5641	12.435
$\varepsilon_{2\mu}^{Q}$	0.4533	4.2263	0.4510	4.1025	0.4583	4.2389	0.7831	6.2342
$\varepsilon_{2\sigma}^{Q}$	0.8924	6.0132	0.8813	5.9886	0.9102	6.1614	10.627	16.083

It can be seen from Table.4 that, both average and maximum values of relative errors caused by the MC simulation are the smallest among these three PPF analysis methods.

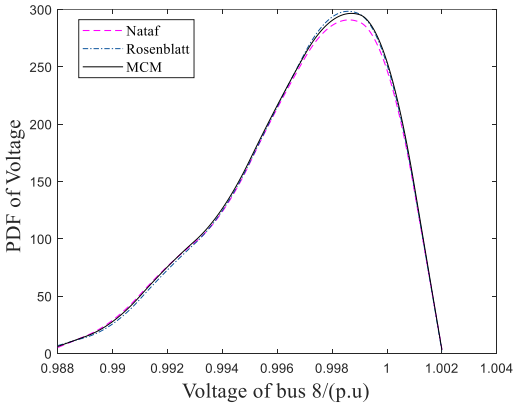

FIGURE 4. Pdf of voltage on bus 8.

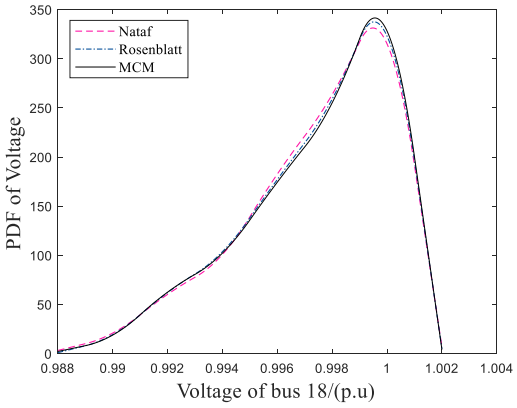

FIGURE 5. Pdf of voltage on bus 18.

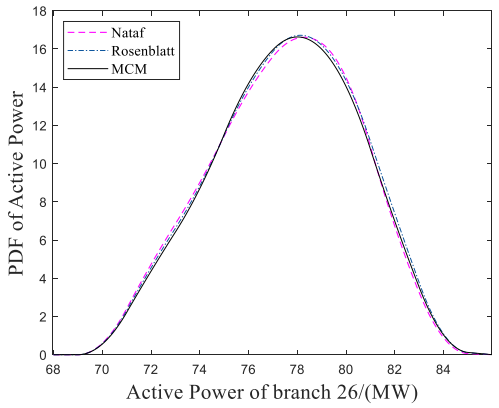

FIGURE 6. Pdf of active power on branch 26.

This shows that, the probability density distribution of system state variables calculated via the MC simulation is most consistent with the real operation. Besides, both the average and maximum values of relative errors of Rosenblatt transformation based 3PEM are smaller than those of Nataf transformation based 3PEM. And the average values of relative errors of all system state variables (voltage amplitude, voltage phase angle and active power) obtained by the proposed method are less than 2%. In comparison, the maximum difference of relative errors of all system state variables between the proposed

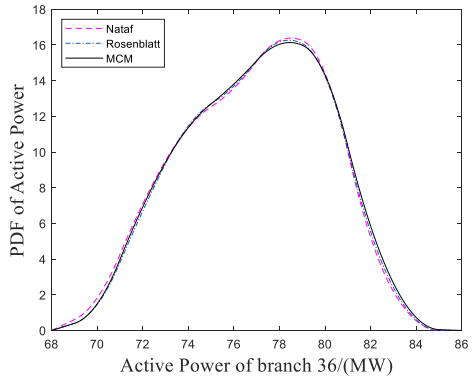

FIGURE 7. Pdf of active power on branch 36.

method with the MC simulation is only 0.12%. Also, both average and maximum values of relative errors of the PPF calculation results without considering the correlation of wind speeds (3PEM only) are much larger than those of the PPF calculation results considering the correlation of wind speeds (Rosenblatt transformation based 3PEM, the MC simulation, Nataf transformation based 3PEM). And the relative errors of standard deviation obtained by 3PEM without considering the correlation of wind speeds are very large. This indicates that, only using 3PEM to analyze PPF has a significant error. The correlation of wind speeds should be considered by using a transformation approach. And the input variables converted via the Rosenblatt transformation method could reflect the correlation of original variables more completely than that via the Nataf transformation method. Compare with the Nataf transformation based 3PEM, the PDF of system state variables obtained from the Rosenblatt transformation based 3PEM is more similar to that from the MC simulation (shown in Fig.4-7).

In addition, the average computational time required for Rosenblatt transformation based 3PEM (35.76 seconds) is shorter than that for the MC simulation (1124.56 seconds). This means that, the PPF analysis using the proposed method could be more than thirty times faster than that using the MC simulation.

Therefore, the proposed method is more accurate with considering the correlation among wind speeds, while having better performance in actual application with higher computational efficiency.

3) IMPACT OF CORRELATION OF WIND SPEEDS

Compared with existing studies, the correlations among multiple wind farms and their potential influence on the PPF of the power system has been carefully reflected in this paper. To demonstrate this impact, a comparative analysis is conducted in this section, wherein the PPF of the power system is analyzed, with and without incorporating the correlation among various wind speeds.

The influence of the correlation among four wind farms on the PDF of power system state variables is illustrated in Fig.8 and Fig.9.

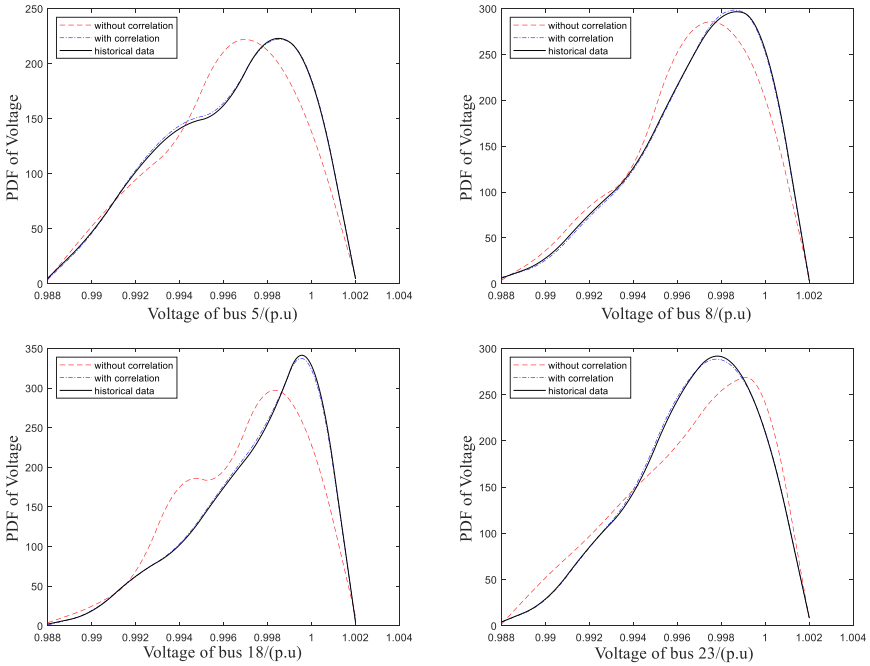

FIGURE 8. The comparison between the pdf of node voltage of four wind farm nodes with and that without considering correlation.

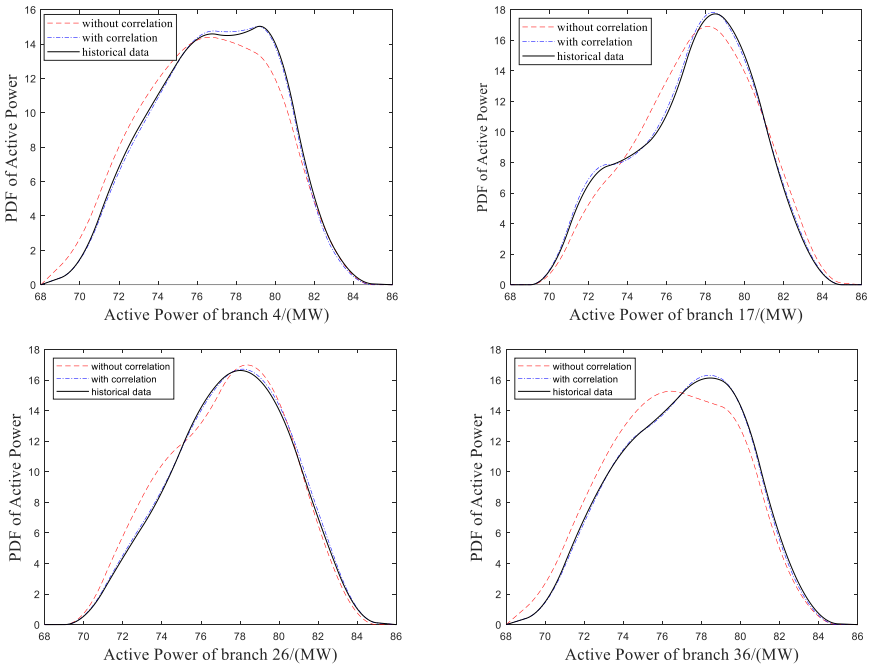

FIGURE 9. The comparison between the pdf of active power of two branches with and that without considering correlation.

As depicted in Fig.8, for nodes 5, 8, 18, and 23, the PDFs of node voltage amplitude without incorporating correlation are considerably different from those with correlation among various wind speeds. Besides, the curve of historical data represents the PDF of the node voltage in the real situation. It can be seen that the PDF of the node voltage with considering correlation is very close to the real PDF of the corresponding node voltage. This indicates that considering the correlation among multiple wind farms is more accordant with the actual condition in PPF calculation.

The PDFs of the active power of branches 4, 17, 26, and 36 are shown in Fig.9. Similar to node voltage amplitude,

there are apparent differences between the PDFs of active power between with and without incorporating correlation of the four wind farms. And the PDFs of active power with considering correlation are more consistent with the PDFs of the real situation (all for branches 4, 17, 26, and 36). This signifies that the correlation of various wind speeds has an essential impact on the PDF of active power.

In summary, the correlation of various wind speeds significantly influences the PPF calculation of the power system. Therefore, it is more realistic to consider the correlation of multiple wind farms for PPF analysis of the power system.

V. CONCLUSION

In this paper, we present a PPF calculation framework for quantifying the impact of the correlation among multi-dimensional wind farms on the power system. Compare with the existing studies, the major contributions of this work is that we introduce a novel learning-based distribution estimation approach to determine the multiple dimensional joint distribution functions to model the correlated wind speeds, which does not require parameter estimation. On this basis, Rosenblatt transformation-based 3PEM is employed to implement PPF calculation. According to the simulation results from, some major findings of this research are summarized as follows:

- Compared with traditional estimation parameter methods, the proposed distribution estimation method more accurately construct the joint distribution functions for modeling the correlation among various wind farms.
- Rosenblatt transformation technique helps improve the computation accuracy of 3PEM and make the obtained PPF analysis results reflect the influence of real wind speeds better.
- Rosenblatt transformation-based 3PEM not only captures the correlation among wind speeds accurately but also has advantages of engineering applications with higher computational efficiency.
- The correlation of multiple wind speeds has a high impact on the PPF calculation results. Therefore, it is indispensable to incorporate the correlation of wind speeds for conducting the PPF calculation.

In this study, due to space restrictions, the uncertainty of photovoltaic power generation is not considered in the proposed PPF calculation model. However, in practice, not only wind but also photovoltaic energy is the main renewable energy resource for power generation. And the output characteristic of photovoltaic is different from that of wind turbines, while there could be coupling between the output power of wind and photovoltaic. Besides, incentive-based and price-based energy loads are also stochastic variables in integrated energy systems, including electricity, natural gas, heat, cool, and other energy carriers. It is desirable to develop the PPF calculation of the power system into probabilistic energy flow calculation of multiple energy systems. As such, these are valuable issues, and we would focus on these for our future research.

REFERENCES

[1] B. Banerjee, D. Jayaweera, and S. M. Islam, "Probabilistic optimisation of generation scheduling considering wind power output and stochastic line capacity," in *Proc. 22nd Australas. Univ. Power Eng. Conf. (AUPEC)*, Bali, Indonesia, 2012, pp. 1–6.
[2] W. Bai and K. Y. Lee, "Modified optimal power flow on storage devices and wind power integrated system," in *Proc. IEEE Power Energy Soc. Gen. Meeting (PESGM)*, Boston, MA, USA, Jul. 2016, pp. 1–5.
[3] M. Bazrafshan and N. Gatsis, "Decentralized stochastic optimal power flow in radial networks with distributed generation," *IEEE Trans. Smart Grid*, vol. 8, no. 2, pp. 787–801, Mar. 2017.
[4] M. Hosseinzadeh and F. R. Salmasi, "Robust optimal power management system for a hybrid AC/DC micro-grid," *IEEE Trans. Sustain. Energy*, vol. 6, no. 3, pp. 675–687, Jul. 2015.
[5] Y. Zhang, N. Gatsis, and G. B. Giannakis, "Robust energy management for microgrids with high-penetration renewables," *IEEE Trans. Sustain. Energy*, vol. 4, no. 4, pp. 944–953, Oct. 2013.
[6] B. Zhao, X. Wang, D. Lin, M. M. Calvin, J. C. Morgan, R. Qin, and C. Wang, "Energy management of multiple microgrids based on a system of systems architecture," *IEEE Trans. Power Syst.*, vol. 33, no. 6, pp. 6410–6421, Nov. 2018.
[7] B. Borkowska, "Probabilistic load flow," *IEEE Trans. Power App. Syst.*, vol. PAS-93, no. 3, pp. 752–759, May 1974.
[8] Z. Ren, K. Wang, W. Li, L. Jin, and Y. Dai, "Probabilistic power flow analysis of power systems incorporating tidal current generation," *IEEE Trans. Sustain. Energy*, vol. 8, no. 3, pp. 1195–1203, Jul. 2017.
[9] G. E. Constante-Flores and M. S. Illindala, "Data-driven probabilistic power flow analysis for a distribution system with renewable energy sources using Monte Carlo simulation," *IEEE Trans. Ind. Appl.*, vol. 55, no. 1, pp. 174–181, Jan. 2019.
[10] R. Y. Rubinstein, D. P. Kroese, *Simulation and the Monte Carlo Method*. New York, NY, USA: Wiley, 2011, pp. 5–20.
[11] A. B. Rodrigues and M. G. Da Silva, "Probabilistic assessment of available transfer capability based on Monte Carlo method with sequential simulation," *IEEE Trans. Power Syst.*, vol. 22, no. 1, pp. 484–492, Feb. 2007.
[12] J. Usaola, "Probabilistic load flow with correlated wind power injections," *Electr. Power Syst. Res.*, vol. 80, no. 5, pp. 528–536, May 2010.
[13] W. Wu, X. Jiang, Z. Wang, G. Li, and K. Wang, "Probabilistic load flow calculation using cumulants and multiple integrals," *IET Gener., Transmiss. Distrib.*, vol. 10, no. 7, pp. 1703–1709, May 2016.
[14] P. Zhang and S. T. Lee, "Probabilistic load flow computation using the method of combined cumulants and Gram-Charlier expansion," *IEEE Trans. Power Syst.*, vol. 19, no. 1, pp. 676–682, Feb. 2004.
[15] R. N. Allan, A. M. Da Silva, and R. C. Burchett, "Evaluation methods and accuracy in probabilistic load flow solutions," *IEEE Trans. Power App. Syst.*, vol. PAS-100, no. 5, pp. 2539–2546, May 1981.
[16] J. M. Morales and J. Perez-Ruiz, "Point estimate schemes to solve the probabilistic power flow," *IEEE Trans. Power Syst.*, vol. 22, no. 4, pp. 1594–1601, Nov. 2007.
[17] M. J. Vahid-Pakdel and B. Mohammadi-Ivatloo, "Probabilistic assessment of wind turbine impact on distribution networks using linearized power flow formulation," *Electr. Power Syst. Res.*, vol. 162, pp. 109–117, Sep. 2018.
[18] N. Gupta, "Probabilistic load flow with detailed wind generator models considering correlated wind generation and correlated loads," *Renew. Energy*, vol. 94, pp. 96–105, Aug. 2016.
[19] C.-L. Su, "Probabilistic load-flow computation using point estimate method," *IEEE Trans. Power Syst.*, vol. 20, no. 4, pp. 1843–1851, Nov. 2005.
[20] R. J. Bessa, V. Miranda, A. Botterud, Z. Zhou, and J. Wang, "Time-adaptive quantile-copula for wind power probabilistic forecasting," *Renew. Energy*, vol. 40, no. 1, pp. 29–39, Apr. 2012.
[21] H. Valizadeh Haghi, M. Tavakoli Bina, M. A. Golkar, and S. M. Moghaddas-Tafreshi, "Using copulas for analysis of large datasets in renewable distributed generation: PV and wind power integration in Iran," *Renew. Energy*, vol. 35, no. 9, pp. 1991–2000, Sep. 2010.
[22] G. Papaefthymiou and D. Kurowicka, "Using copulas for modeling stochastic dependence in power system uncertainty analysis," *IEEE Trans. Power Syst.*, vol. 24, no. 1, pp. 40–49, Feb. 2009.
[23] C. Sun, Z. Bie, M. Xie, and J. Jiang, "Fuzzy copula model for wind speed correlation and its application in wind curtailment evaluation," *Renew. Energy*, vol. 93, pp. 68–76, Aug. 2016.
[24] S. Zhou, Q. Xiao, and L. Wu, "Probabilistic power flow analysis with correlated wind speeds," *Renew. Energy*, vol. 145, pp. 2169–2177, Jan. 2020.

[25] Q. Xiao and S. Zhou, "Matching a correlation coefficient by a Gaussian copula," *Commun. Statist.-Theory Methods*, vol. 48, no. 7, pp. 1728–1747, Apr. 2019.
[26] Y. Chen, J. Wen, and S. Cheng, "Probabilistic load flow method based on Nataf transformation and Latin hypercube sampling," *IEEE Trans. Sustain. Energy*, vol. 4, no. 2, pp. 294–301, Apr. 2013.
[27] C. Chen, W. Wu, B. Zhang, and H. Sun, "Correlated probabilistic load flow using a point estimate method with nataf transformation," *Int. J. Electr. Power Energy Syst.*, vol. 65, pp. 325–333, Feb. 2015.
[28] Q. Xiao, "Evaluating correlation coefficient for Nataf transformation," *Probabilistic Eng. Mech.*, vol. 37, pp. 1–6, Jul. 2014.
[29] M. E. Tipping, "Sparse Bayesian learning and the relevance vector machine," *J. Mach. Learn. Res.*, vol. 1, no. 3, pp. 211–244, Sep. 2001.
[30] P.-L. Liu and A. Der Kiureghian, "Multivariate distribution models with prescribed marginals and covariances," *Probabilistic Eng. Mech.*, vol. 1, no. 2, pp. 105–112, Jun. 1986.
[31] A. Sklar, "Fonctions de repartition an dimensions et leurs marges," *Publication l'Institut Statistique l'Universite Paris*, vol. 8, pp. 229–231, 1959.
[32] T. Bedford and R. M. Cooke, "Vines—A new graphical model for dependent random variables," *Ann. Statist.*, vol. 30, no. 4, pp. 1031–1068, Aug. 2002.
[33] H. Yang and B. Zou, "A three-point estimate method for solving probabilistic power flow problems with correlated random variables," *Autom. Electr. Power Syst.*, vol. 36, no. 15, pp. 51–56, 2012.
[34] R. Christie. *Power System Test Case Archive*. Accessed: Jan. 16, 2020. [Online]. Available: http://www.ee.washington.edu/research/pstca/
[35] *Kolmogorov-Smirnov Test*. Accessed: Jan. 16, 2020. [Online]. Available: https://en.wikipedia.org/wiki/Kolmogorov%E2%80%93Smirnov_test
[36] S. Lele, "Euclidean distance matrix analysis (EDMA): Estimation of mean form and mean form difference," *Math. Geol.*, vol. 25, no. 5, pp. 573–602, Jul. 1993.

XI ZHU received the B.S. and M.S. degrees in electrical engineering from North China Electric Power University, Beijing, China, in 2014 and 2017, respectively, where she is currently pursuing the Ph.D. degree.

Her research interests include power system operation and integrated energy system planning.

CHUNMING LIU was born in Hebei, China, in 1972. He received the B.S. degree in mechanical engineering from the Hebei University of Technology, China, in 1993, and the M.S. and Ph.D. degrees in electrical engineering from North China Electric Power University, in 2000 and 2009, respectively.

He was a Professor with North China Electric Power University. His current research interests include geomagnetically induced currents in power grids, including monitoring, modeling, and assessing the influence on the security of power systems.

CHENBO SU received the B.S. and M.S. degrees in electrical engineering from North China Electric Power University, Beijing, China, in 2014 and 2017, respectively, where he is currently pursuing the Ph.D. degree.

His research interests include wind farm planning and stability of grid connection of MMC.

JIAOMIN LIU received the B.S. degree in automation and the Ph.D. degree from the Hebei University of Technology, Tianjin, China, in 1982 and 1998, respectively.

He is currently a Professor with the Department of Electrical and Electronic Engineering, North China Electric Power University, Baoding, China. His research interests include power system stability analysis and control, including the renewable energy and power system reliability.