

Article

A New Wind Speed Scenario Generation Method Based on Principal Component and R-Vine Copula Theories

Hui Hwang Goh ^{1,*}, Gumeng Peng ¹, Dongdong Zhang ¹, Wei Dai ¹, Tonni Agustiono Kurniawan ², Kai Chen Goh ³ and Chin Leei Cham ⁴

- ¹ School of Electrical Engineering, Guangxi University, Nanning 530004, China; 1912301042@st.gxu.edu.cn (G.P.); dongdongzhang@gxu.edu.cn (D.Z.); weidai@gxu.edu.cn (W.D.)
² College of Environment and Ecology, Xiamen University, Xiamen 361102, China; tonni@xmu.edu.cn
³ Department of Technology Management, Faculty of Construction Management and Business, University Tun Hussein Onn Malaysia, Parit Raja 86400, Malaysia; kaichen@uthm.edu.my
⁴ Faculty of Engineering (FOE), BR4081, Multimedia University, Persiaran Multimedia, Cyberjaya 63100, Malaysia; clcham@mmu.edu.my
* Correspondence: hhgoh@gxu.edu.cn; Tel.: +86-132-9941-8218

Abstract: The intermittent and uncertain properties of wind power have presented enormous obstacles to the planning and steady operation of power systems. In this context, as an effective technique to study wind power uncertainty, the development of an accurate wind speed scenario generation method is of great significance for evaluating the impact of wind power in the power system. In the case of several wind farms, accurate scenario generation involves precise acquisition of the correlation between wind speeds and the greatest retention of statistical properties of wind speed data. Under this goal, this research provided a new method for scenario development based on principle component (PC) and R-vine copula theories that incorporates the spatiotemporal correlation of wind speeds. By integrating with PC theory, this strategy avoids the dimension disaster induced by employing R-vine copula alone while taking benefit of its flexibility. The simulation results utilizing the historical wind speeds of three adjacent wind farms as samples showed that the method described in this article could effectively preserve the statistical properties of wind speed data. Eight evaluation indicators covering three facets of the scenario generation method were used to compare the proposed method holistically to two other commonly used scenario generation methods. The results indicated that this method's accuracy was increased further. Additionally, the validity and necessity of applying R-vine copula in this model was demonstrated through comparisons to C-vine and D-vine copulas.

Keywords: principal component theory; R-vine copula theory; several wind farms; scenario generation; spatiotemporal correlation



Citation: Goh, H.H.; Peng, G.; Zhang, D.; Dai, W.; Kurniawan, T.A.; Goh, K.C.; Cham, C.L. A New Wind Speed Scenario Generation Method Based on Principal Component and R-Vine Copula Theories. *Energies* **2022**, *15*, 2698. <https://doi.org/10.3390/en15072698>

Academic Editor: Andrés Elías Feijóo Lorenzo

Received: 27 February 2022

Accepted: 4 April 2022

Published: 6 April 2022

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Copyright: © 2022 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

To address issues such as the energy crisis and greenhouse gas emissions, the development and utilization of renewable energy have accelerated in recent decades, and renewable energy sources, including wind power, are becoming more integrated into the power grid. However, wind energy as a form of renewable energy exhibits a high degree of randomness and intermittency, posing certain challenges for the safe and stable operation of the power grid as well as planning [1].

In the existing literature, wind power modeling methods have been widely used. In the study of power generation expansion planning, Sahragard et al. [2] considered the wind power penetration in power generation expansion planning by using the conversion model of wind speed and wind power. Band et al. [3] took the wind energy in the Gulf of Oman as the research object and assessed the change of its power generation potential according to the relevant wind power model. One of the current challenges is determining how to account for wind energy generation uncertainty in the optimization and planning

of grid dispatching with wind energy integration. Additionally, when a grid contains several wind farms, it is critical to accurately capture the temporal and spatial correlation between wind farms for wind power modeling. As a result, in-depth research on wind energy uncertainty modeling in the context of multiple wind farms remains necessary. Scenario generation has been extensively studied as a method to deal with the uncertainty associated with wind energy. Three types of scenario generation techniques exist. The first is a method for simulating wind power time series using the Markov chain model and the autoregressive moving average (ARMA) model. D'Amico et al. [4] developed first and second order semi-Markov chains to generate synthetic wind speed time series, which is a more accurate approach than using a simple Markov chain in reproducing the statistical properties of wind speed data. Chen et al. [5] used Fourier series and an ARMA model to determine the seasonal trend and temporal autocorrelation of wind speed, respectively. Sim et al. [6] compared and analyzed the ability of autoregressive integrated moving averages (ARIMA) to predict wind speed and generate wind speed time series data from historical values. Sun et al. [7] proposed a method for generating multi-wind farm scenarios based on truncated multivariate Gaussian mixture models and Markov chain quasi-Monte-Carlo sampling. Morales et al. [8] decoupled the multivariate ARMA model, simplifying parameter estimation while preserving the statistical properties of wind farm wind speed data sets. Abedi et al. [9] combined the ARMA and fuzzy models to preserve the spatiotemporal correlation of wind farms while studying the impact of wind power correlation on the joint energy and reserve market. Duong et al. [10] demonstrated the efficacy of a hybrid method based on principal component analysis (PCA) and the ARMA model for generating data sets that preserve wind farm temporal and spatial correlation.

The second category of method for studying wind energy scenario generation is the machine learning method represented by generative adversarial networks (GAN). Recent research has applied the Wasserstein GAN (WGAN) [11], a conditionally improved WGAN combined with an unsupervised labeling model [12], a faster and more stable improved GAN [13], a controllable GAN with new evaluation indexes [14], and a GAN combined with reinforcement learning without manual labeling [15] to the generation of wind power scenarios. These findings demonstrate that these types of methods are capable of effectively preserving wind energy output's spatiotemporal correlation.

The third category of wind power scenario generation methods is based on copula theory-based wind power spatiotemporal modeling. Copula theory has attracted the attention of numerous researchers due to its ability to be applied to both linear and nonlinear correlations. By adopting the truncation method for the D-vine copula, Haghi et al. [16] were able to reduce the computational burden while maintaining the flexibility of selecting the appropriate copula function for the varied correlation characteristics of wind farms. Similarly, Becker [17] examined the time autocorrelation of wind energy forecast errors using a D-vine copula. Lin et al. [18] used the t copula to generate the joint distribution of multiple wind farms and combined it with the imprecise Dirichlet model to increase the rationality of the generated scenarios. Borujeni et al. [19] also used t copula to derive the joint probability distribution for hours when examining the tail-dependent structure of wind speed data. Eryilmaz et al. [20] investigated the system reliability of a wind power system with two wind farms by using Gumbel copula to capture the correlation between wind speeds between the two wind farms and evaluating the system's total capacity in relation to the turbines' reliability. To fully account for the impact of wind power uncertainty on economic dispatch, Li et al. [21] sampled the joint distribution of multiple wind farms established via D-vine copula and obtained wind speed data sets that took spatial correlation into account. Deng et al. [22] developed a wind speed scenario generation method that took into account the tail dependency structure of wind speed data. They used t copula and C-vine copula to establish joint distributions of spatial and temporal correlation of the tail structure, respectively, in order to minimize parameter calculation. Qiu et al. [23] used C-vine and D-vine copulas to generate wind speed scenarios from clustered multivariate wind speed data sets. Eventually, scenarios that were more consistent

with historical data were retained. To model the correlation between multiple wind farms, Xu et al. [24] proposed a simplified C-vine copula construct similar to truncation. Probabilistic small signal stability analysis is used to demonstrate the method's effectiveness in wind power probability modeling with high correlation. Henderson et al. [25] estimated copula parameters using a Bayesian approach and demonstrated the effect of the size of the wind speed data sets on the uncertainty associated with parameter estimation via multiple copula functions.

Wang et al. [26] used Gaussian/t copula to capture the temporal correlation between multi-period wind power forecast errors in order to accurately assess the capacity of the energy storage system that must be configured in the system containing the wind farm. Li et al. [27] developed a dynamic copula to model the correlation between wind power forecast errors and wind power fluctuations as another application for studying the correlation of wind power forecast errors. The combination of the ARIMA model and the generalized autoregressive conditional heteroscedasticity (GARCH) model improved the accuracy of the wind power prediction error range obtained using this method. Philippe et al. [28] used step-by-step Gaussian copula and Archimedean copula modeling to estimate temporal and spatial correlations, significantly reducing the number of estimated parameters while maintaining scenario accuracy. Given the copula functions' ability to model both the dependency structure and the marginal distribution separately, Wang et al. [29] combined the highly flexible R-vine copula and probabilistic forecasting to improve forecast quality in the case of multiple wind farms with incomplete sample data. Wang et al. [30] proposed a distance-weighted kernel density estimation method to improve the accuracy of the marginal distribution and combined it with the R-vine copula to accurately model the spatiotemporal correlation of wind farms in another study involving the application of the R-vine copula. Additionally, in the study of probabilistic power flow, copula theory was frequently used to investigate the correlation between wind speeds in order to fully account for the uncertainty associated with wind energy [31–33].

While it is clear that extensive research has been conducted on the scenario generation method for dealing with wind energy uncertainty, there are still some gaps in the existing research. In the case of multiple wind farms, the scenario generation method based on Markov and ARMA model [4–7] lacks the flexibility to account for non-linear correlation and spatial correlation. Additionally, in order to satisfy the ARMA model's requirement for stationary data, it is necessary to perform corresponding transformations or assumptions on the input data, limiting its applicability. While the GAN-based scenario generation method [11–15] is effective at obtaining correlation between data, it is time consuming to train and its efficiency varies significantly depending on the experimental equipment.

As a result, the copula theory has been widely applied, as it is capable of effectively resolving the aforementioned issues. However, existing studies have identified drawbacks such as the loss of statistical characteristics in the data [16,24], the lack of flexibility in applying a single copula [18–20,28,31–33], and the consideration of only temporal or spatial correlation [20,21,23,24]. Furthermore, in the application of vine copula theory, most of them have studied C-vine and D-vine copula with special structure [16,17,21–24], and there is little research on R-vine copula with more general structure. Due to the fact that the few scenario generation methods involving the R-vine copula applied it directly [29,30], the curse of dimensionality problem also occurs when dealing with high-dimensional data. While the independence test mentioned in [30] could assist in reducing the number of variables, when applied to data with correlation, this method contributed little to reducing the computational burden. In light of these limitations and research gaps, we proposed a novel method combining R-vine copula and PC theory for alleviating the R-vine copula's curse of dimensionality while retaining the proposed model's accuracy. R-vine can obtain correlations with a variety of different characteristics and separating spatiotemporal correlation modeling alleviates the problem of dimensional disaster, whereas using PC theory establishes the conditions for separation modeling. The simulation first verified the effectiveness of the scenario generation process proposed in this article, then compared the

proposed method and other methods comprehensively using three aspects of the scenario generation evaluation indicators to verify the accuracy of the proposed method. Finally, by comparing the results obtained with C-vine and D-vine, the advantages of using R-vine in this model were evaluated, as well as the necessity and effectiveness of using R-vine copula to improve the accuracy.

This article is structured as follows. Section 2 discusses the theoretical foundations of PC theory and R-vine copula. Section 3 presents a method for generating wind speed scenarios. Section 4 introduces the wind speed data used in this study and evaluates the effectiveness of the method. Section 5 draws a conclusion.

2. Principal Component and Vine Copula Theories

2.1. Principal Component Generation Process

PC theory can extract the primary characteristics of historical data by transforming the correlated historical data for each dimension into an uncorrelated set of PC values. For a given matrix, \mathbf{X} containing n -dimensional sample data:

$$\mathbf{X} = [x_1; x_2; \dots; x_n] \quad (1)$$

$$x_m = [x_m^1, x_m^2, \dots, x_m^k], m = 1, 2, \dots, n \quad (2)$$

where n is the dimension of input data, x_m is the m th dimensional vector of input data, and x_m^k is the k th observation value of x_m .

The basic generation process of PC is as follows [10]:

1. Centralized input data:

$$\mathbf{X}' = \mathbf{X} - \bar{x} \quad (3)$$

$$\bar{x} = \begin{pmatrix} \bar{x}_1 & \bar{x}_1 & \cdots & \bar{x}_1 \\ \bar{x}_2 & \bar{x}_2 & \cdots & \bar{x}_2 \\ \vdots & \vdots & \ddots & \vdots \\ \bar{x}_n & \bar{x}_n & \cdots & \bar{x}_n \end{pmatrix} \quad (4)$$

$$\bar{x}_m = \frac{1}{k} \sum_{d=1}^k x_m^d, m = 1, 2, \dots, n \quad (5)$$

2. Calculation of the covariance matrix of the input data:

$$\mathbf{C} = \begin{pmatrix} \text{cov}(x_1, x_1) & \text{cov}(x_1, x_2) & \cdots & \text{cov}(x_1, x_n) \\ \text{cov}(x_2, x_1) & \text{cov}(x_2, x_2) & \cdots & \text{cov}(x_2, x_n) \\ \vdots & \vdots & \ddots & \vdots \\ \text{cov}(x_n, x_1) & \text{cov}(x_n, x_2) & \cdots & \text{cov}(x_n, x_n) \end{pmatrix} \quad (6)$$

$$\text{cov}(x_p, x_q) = \frac{1}{k} \sum_{d=1}^k (x_p^d - \bar{x}_p)(x_q^d - \bar{x}_q), p, q = 1, 2, \dots, n \quad (7)$$

3. Calculation of eigenvalue vector λ and eigenvector matrix \mathbf{U} of matrix \mathbf{C} :

$$\lambda = [\lambda_1, \lambda_2, \dots, \lambda_n], \lambda_1 > \lambda_2 > \dots > \lambda_n \quad (8)$$

$$\mathbf{U} = [u_1, u_2, \dots, u_n] \quad (9)$$

4. Calculation of principal components:

$$\mathbf{Z} = \mathbf{U}^T \mathbf{X}' \quad (10)$$

Each row in \mathbf{Z} is a PC, and the order of PC is the number of the row it is in.

2.2. Vine Copula Theory

2.2.1. Copula Theory

Copula is a type of function that employs the marginal distribution function to create a joint distribution capable of capturing all correlated information between variables. Along with linear correlation, copula functions have the distinct advantage of being able to describe nonlinear correlation. Sklar's theorem, which serves as the theoretical foundation for copula functions, defines the joint distribution of variables. According to Sklar's theorem [23], for n -ary random variables $\mathbf{X} = [X_1, X_2, \dots, X_n]$ with continuous cumulative distribution functions (CDF) $F_1(x_1), F_2(x_2), \dots, F_n(x_n)$, there is a unique copula function to obtain the joint distribution among variables which can be formulated as:

$$F(x_1, x_2, \dots, x_n) = C(F_1(x_1), F_2(x_2), \dots, F_n(x_n)) \quad (11)$$

The joint probability density can be formulated as:

$$f(x_1, x_2, \dots, x_n) = c(f_1(x_1), f_2(x_2), \dots, f_n(x_n)) \cdot \prod_{m=1}^n f_m(x_m) \quad (12)$$

where C is the copula function, c is the density function of C , and $f_1(x_1), f_2(x_2), \dots, f_n(x_n)$ is the probability density functions of \mathbf{X} .

As illustrated in Figure 1, copula functions can be classified into two families: elliptical copula family and Archimedean copula family. The Archimedean copula family contains a number of different copula functions, the most frequently used of which are the Gumbel copula, the Clayton copula, and the Frank copula. The Gaussian copula and the t copula are the most frequently encountered in the family of elliptical copulas.

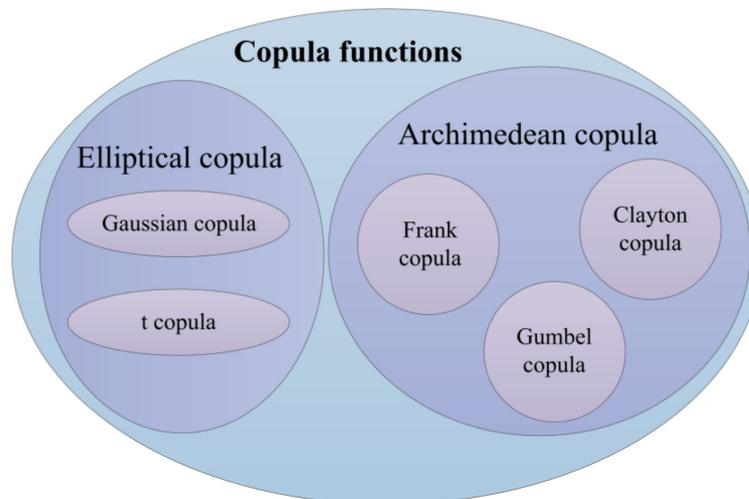


Figure 1. The main types of copula functions.

2.2.2. R-Vine Copula

Copula functions, as discussed in Section 2.2.1, exhibit a variety of characteristics [22]. While Gaussian, t , and Frank copulas all have a symmetric structure, only the t copula is suitable for capturing upper and lower tail dependency structures. Both the Clayton copula and the Gumbel copula are asymmetric structures that can be used to represent lower and upper tail dependency structures, respectively. However, only the Gaussian and t copulas are suitable for describing the joint distribution of multivariate variables, whereas the other functions are restricted to establishing correlation between binary variables. As a result, a single copula function has limitations in capturing the correlation of multivariate variables with varying tail dependency structures.

The emergence of vine copula enables the solution of the problem of high-dimensional copula. The fundamental concept of vine copula is to decompose the multivariate copula probability density expression into a binary copula joint probability density product, as defined in Equation (13) [22]. As can be seen from Equation (13), the decomposition is not unique when it comes to n -ary ($n \geq 3$) copula functions.

$$f(x|v) = c_{x,v_j|v_{-j}}(F(x|v_{-j}), F(v_j|v_{-j})) \cdot f(x|v_{-j}) \tag{13}$$

where v is vector composed of a set of variables, v_j is a variable in v , and v_{-j} is the vector formed by removing v_j from v .

Different decomposition methods produce trees with a variety of distinct structures, each with one less layer than the amount of variables. And the structure of the tree is referred to as R-vine structure. The $n-1$ layer trees $T = (T_1, T_2, \dots, T_{n-1})$ are used to represent the R-vine copula model of n -ary random variables, and the nodes set and edges set of the tree $T_i (i = 1, 2, \dots, n-1)$ are N_i and E_i , respectively. The nodes and edges of each layer tree are formulated as:

$$\begin{cases} N_1, E_1 & i = 1, \\ N_i = E_{i-1} & i = 2, 3, \dots, n-1. \end{cases} \tag{14}$$

It should be stressed that when two edges on the tree T_i are used as two connecting nodes on the tree T_{i+1} , it must be ensured that there is a common node between the two edges [34]. Taking $n = 4$ as an example, Figure 2 shows a certain R-vine structure in the case of four-dimensional variables. Therefore, it can be deduced that the R-vine copula joint probability density of Equation (12) is [34]:

$$f(x_1, x_2, \dots, x_n) = \prod_{k=1}^n f_k(x_k) \prod_{l=1}^{n-1} \prod_{e \in E_l} c_{a(e), b(e)|D_e}(F_{a(e)|D_e}(x_{a(e)}|x_{D_e}), F_{b(e)|D_e}(x_{b(e)}|x_{D_e})) \tag{15}$$

where D_e is the set corresponding to the conditioning variables, $a(e)$ and $b(e)$ are any other variables excluding the variables in D_e , x_{D_e} is the set of variables corresponding to D_e , and $c_{a(e), b(e)|D_e}$ is the copula probability density of variable a and b .

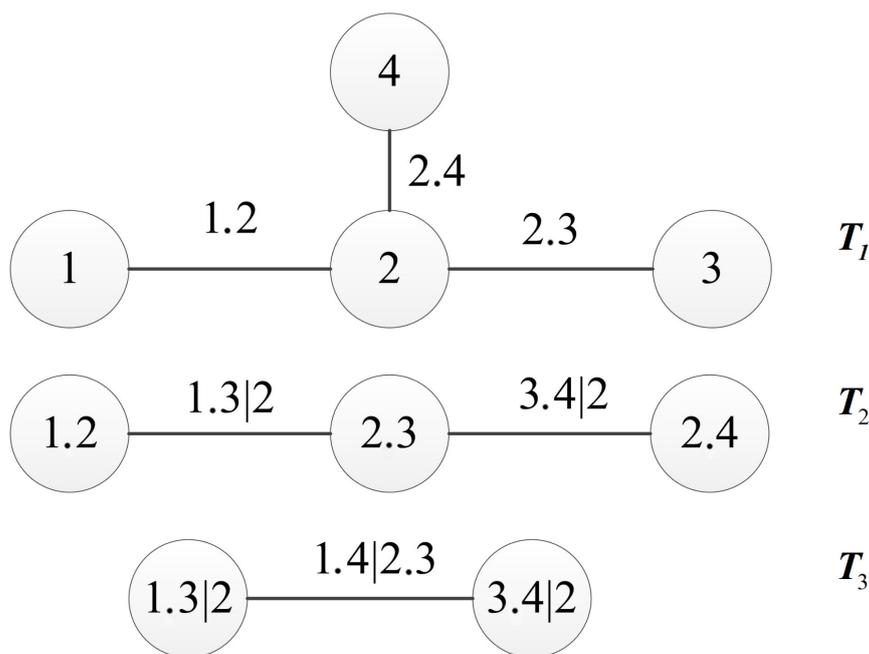


Figure 2. A four-dimensional R-vine copula structure.

3. Formulation of Wind Speed Scenarios

3.1. Reification of the Structure of the R-Vine Copula Model

The approach proposed by Dißmann et al. [34] for determining the structure of the R-vine copula model was used in this research. The process of developing an R-vine copula model is divided into three stages: the first stage involves determining the specific structure of each layer of the copula, the second stage involves estimating the parameters of the binary copula corresponding to each edge, and the third stage involves evaluating the goodness of fit of each binary copula. The R-vine copula model's trees are structured using the inverse standard of the minimal spanning tree approaches provided by the Prim's and Kruskal's algorithms. Prim's approach is used to estimate the greatest spanning tree of each layer of the R-vine copula model, which is more suitable for dense networks. Additionally, the weight assigned to each edge in Prim's algorithm is the empirical Kendall's rank correlation coefficient τ for each binary copula function, which may be calculated as:

$$\tau = \frac{2}{n(n-1)} \sum_{1 \leq i, j \leq n}^n \zeta[(x_i - x_j)(y_i - y_j)] \quad (16)$$

$$\zeta[(x_i - x_j)(y_i - y_j)] = \begin{cases} 1 & (x_i - x_j)(y_i - y_j) > 0, \\ 0 & (x_i - x_j)(y_i - y_j) = 0, \\ -1 & (x_i - x_j)(y_i - y_j) < 0. \end{cases} \quad (17)$$

where x_i and y_i ($i = 1, 2, \dots, n$) are the samples corresponding to binary copula, respectively, and n is the total number of samples in each sample set.

The Akaike information criterion (AIC) is used to assist in determining the optimal binary copula for each node of each tree for the R-vine copula model, which may be expressed as follows [28]:

$$AIC = 2k - 2 \sum_{i=1}^n \log c(u_i, v_i) \quad (18)$$

where k is the number of parameters of the corresponding copula pair, c is the copula density function, and u_i and v_i are the CDF values of the i th value of two sample sets, respectively.

Maximum likelihood estimation (MLE) is applied to estimate the parameters of the binary copulas. For the two marginal CDFs $U = F_1(x)$ and $V = F_2(y)$, the maximum likelihood estimate of the copula functions can be expressed as [19]:

$$\gamma = \arg \max \sum_{i=1}^n \log c(u_i, v_i) \quad (19)$$

In [34], a more extensive discussion of the R-vine copula model construction and parameter calculation is provided. Figure 3 depicts the flowchart for building the R-vine copula model structure.

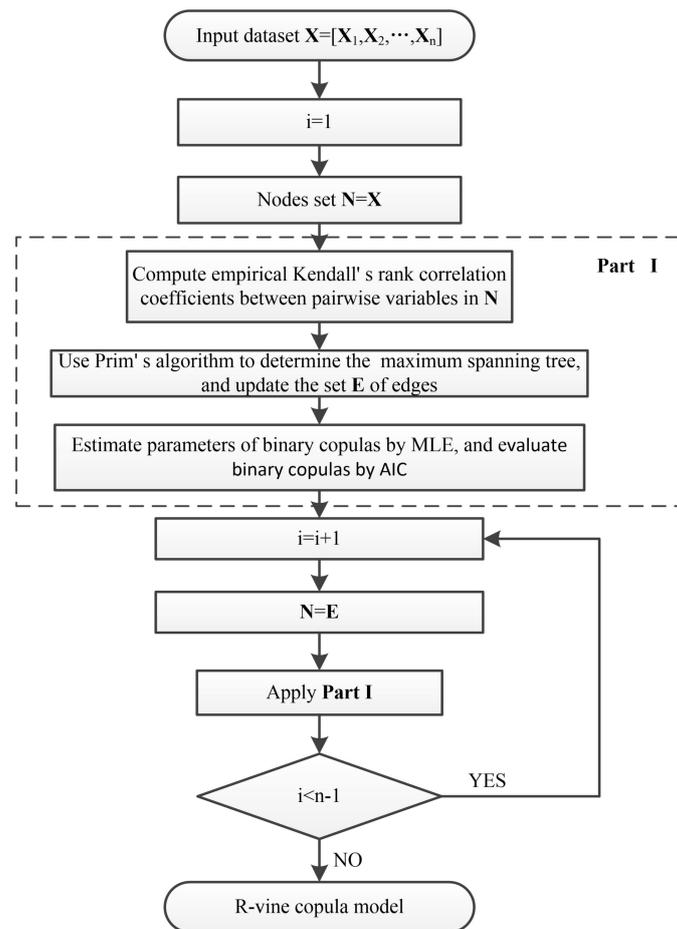


Figure 3. The flowchart depicting the formation of the R-vine copula model.

3.2. Procedure of Scenario Generation Method

According to the introductions in Sections 2 and 3.1, the following are the precise processes for generating the wind speed scenarios based on the historical wind speed of p wind farms in q years:

Step 1: $W_N(N = 1, 2, \dots, p)$ represents dividing the historical data of the N th wind farm by time $T = 1, 2, \dots, 24$ h, where the i th ($i = 1, 2, \dots, 24$) row W_N^i is the historical wind speed set of wind farm N at the i th hour in $q \times 365$ days.

$$W_N = [W_N^1; W_N^2; \dots; W_N^{24}] \quad (N = 1, 2, \dots, p) \quad (20)$$

Step 2: For the wind speed data sets of p wind farms in the same hour $X^i = [W_1^i; W_2^i; \dots; W_p^i]$ ($i = 1, 2, \dots, 24$), the PC generation process introduced in Section 2.1 is utilized to transform the p -dimensional spatially correlated wind speed sets into PC $Z^i = [Z_1^i; Z_2^i; \dots; Z_p^i]$ [10].

Step 3: The kernel density estimation (KDE) method [21] is applied to estimate the marginal cumulative distribution function (CDF) of Z_N^i ($i = 1, 2, \dots, 24, N = 1, 2, \dots, p$), and the values of Z_N^i is transformed into the values uniformly distributed over $[0, 1]$.

Step 4: Combined with Sections 2.2 and 3.1, the corresponding R-vine copula structure is determined for the transformed matrix $Z_N = [Z_N^1; Z_N^2; \dots; Z_N^{24}]$ ($N = 1, 2, \dots, p$). According to this, 24-dimensional temporal correlated scenarios $S_N = [S_N^1; S_N^2; \dots; S_N^{24}]$ are generated.

Step 5: The inverse transformation is applied to the marginal CDF in Step 3 to transform the scenario S_N^i ($i = 1, 2, \dots, 24, N = 1, 2, \dots, p$) into the PC values.

Step 6: Through the data reconstruction as shown in Equation (21) [10], the scenario matrix $S^i = [S_1^i; S_2^i; \dots; S_p^i]$ ($i = 1, 2, \dots, 24$) of each time stamp transformed to PC values are transformed into wind speed values WS^i that restore the spatial correlation among wind farms.

$$WS^i = U^i S^i + \bar{x}^i, i = 1, 2, \dots, 24 \quad (21)$$

where \bar{x}^i and U^i are the mean and the eigenvector matrix when the PCs are generated using the wind farms' data of the hour, which are shown as Equations (4) and (9).

The pseudo code of the scenario generation is as follows (Algorithm 1):

Algorithm 1 Mixed method to generate wind speed scenarios based on PC and R-vine copula theories.

Input: Historical wind speed of p wind farms in q years.

Output: Wind speed scenarios of p wind farms.

1: **for** $i = 1 : p$ **do**

2: Divide the samples of wind farm i into 24 sample sets by time $T = 1, 2, \dots, 24$.

3: **end for**

4: **for** $j = 1 : 24$ **do**

5: Apply PC theory to the matrix consisting of sample sets of p wind farms at the j th hour.

6: Transform the values of each PC into CDF values using KDE.

7: **end for**

8: **for** $k = 1 : p$ **do**

9: Apply the R-vine copula model to generate PC scenarios according to the matrix consisting of the k th PCs of each hour.

10: **end for**

11: **for** $l = 1 : 24$ **do**

12: Restore the p generated PCs of the l th hour to wind speed scenarios by data reconstruction of PC theory and the inverse transformation of the kernel density function.

13: **end for**

4. Results and Analysis

This section begins by introducing wind speed measurements. Along with the five major copula functions illustrated in Figure 1, the R-vine structure also employs two additional significant copula functions, namely the Joe copula and the Ali-Mikhail-Haq copula. The validity of the scenario creation approach used in this article is then established, and the created 1000 wind speed scenarios are compared to the samples using three-dimensional figures. Finally, the other two models are used to generate the identical amount of scenarios, and some metrics are utilized to compare and examine the three models' accuracy.

4.1. Data Sources

The wind speeds of three adjacent wind farms were simulated and studied in this article. Wind speed data were provided by the National Renewable Energy Laboratory (NREL), and the three wind farms can be identified by the site IDs 604171, 605146, and 606121, respectively [35]. Wind speed data for the three wind farms were collected over a six year period from 1 January 2007 to 31 December 2012 using wind turbines with a hub height of 100 m. The sampling interval for the sample values was one hour. Kendall's rank correlation coefficients between two wind speed sample sets for these three wind farms were 0.8992, 0.7872, and 0.8743, respectively, indicating that there were strong spatial correlations between the three wind farms.

4.2. Evaluation of the Process for Generating Wind Speed Scenarios

Wind speed samples were transformed into equivalent PC values using PC theory's orthogonal transformation and then converted to CDF values using KDE. Because the distribution of PC values in different orders is quite different, and the first PC frequently contains the majority of the information in the samples, in order to unify the legend,

taking the first PCs at hour 7, 15, and 23 as examples, Figure 4 illustrates the probability density plots of the KDE and the corresponding frequency histogram of the PC values. To maintain consistency with the examples, the accompanying analysis used the first PCs of the three-time stamps [22]. As illustrated in Figure 4, the proposed estimate method achieved the corresponding continuous probability distribution using discrete samples while retaining the numerical distribution's statistical properties. To demonstrate that the orthogonal transformation of PC theory eliminated correlation between PCs over the same time period, Figure 5 shows scatter plots of wind speed samples and corresponding PCs at the 7th, 15th, and 23rd h for three wind farms that have all been transformed to CDF values using KDE. The goal of translating all data into CDF values before to plotting was to unify the comparison's dimensions. As illustrated in Figure 5, the three PCs created from the sample data of three wind farms with clear spatial correlation were dispersed uniformly in the value space with no correlation. This conclusion established an effective rationale for focusing exclusively on the temporal correlation between PCs.

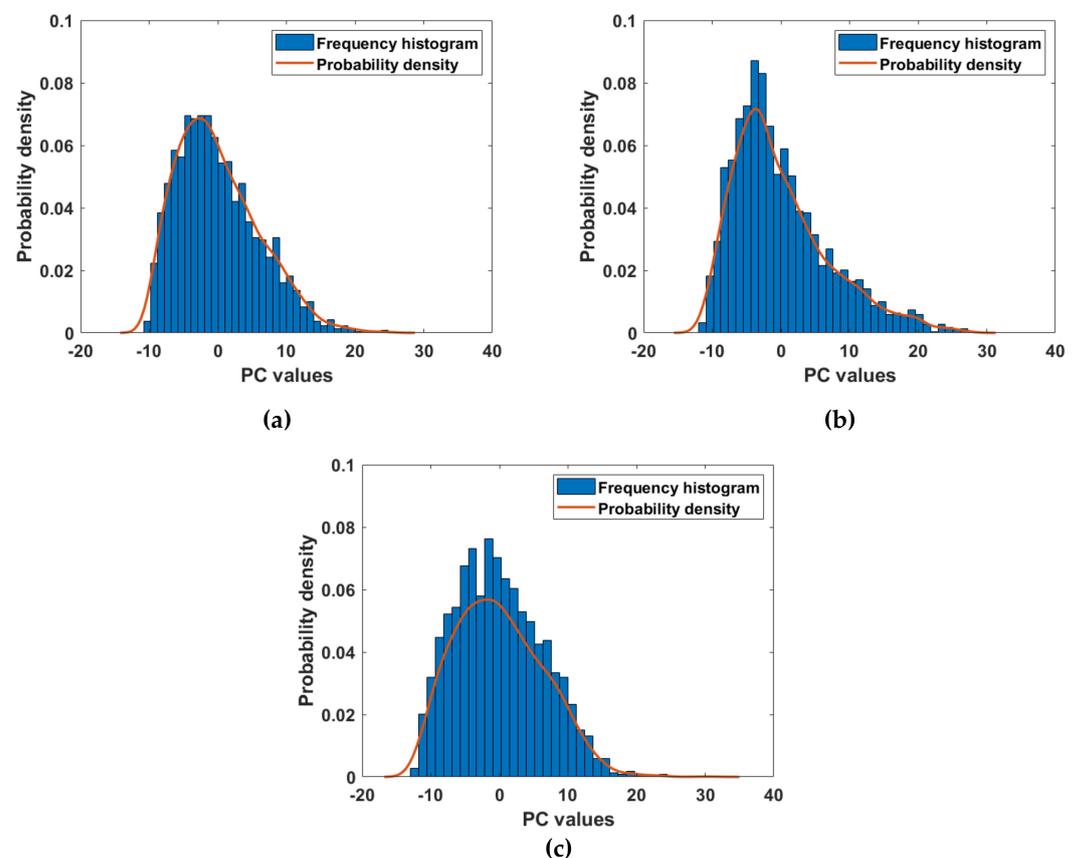


Figure 4. Histogram of frequency and probability plot of the first PC (a) at the 7th hour, (b) at the 15th hour, (c) at the 23rd hour.

To examine the PCs' temporal correlation across time, Figure 6 presents scatter plots of the CDF values of wind speed samples and the corresponding initial PCs for each wind farm (WF) at the 7th, 15th, and 23rd hours. As can be observed, the distribution space and degree of dispersion of the samples and PCs were nearly identical. This demonstrates that the PCs could still reflect the temporal correlation of samples. To summarize, the PC generation process used for the three wind farm samples in each period preserved the temporal correlation between the samples but also eliminated the spatial correlation between wind farms, providing theoretical support for subsequent consideration of only the temporal correlation among the PCs.

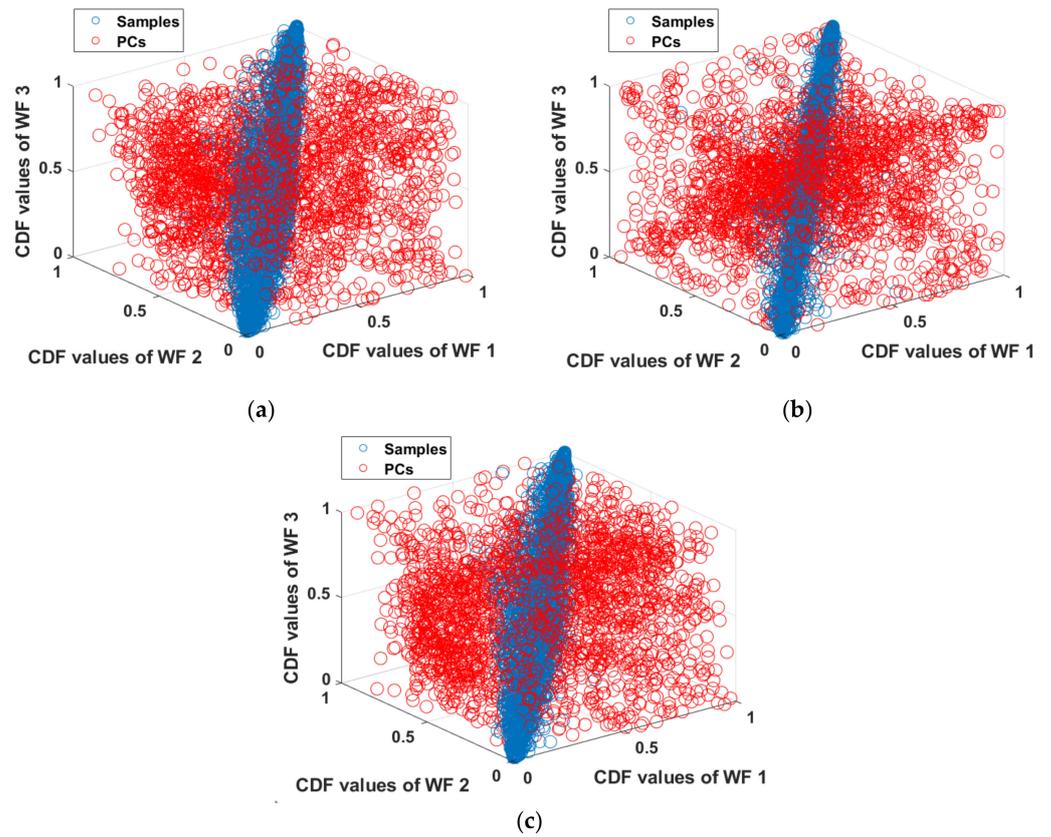


Figure 5. Sample scatter plots with related PCs (a) at the 7th hour, (b) at the 15th hour, (c) at the 23rd hour.

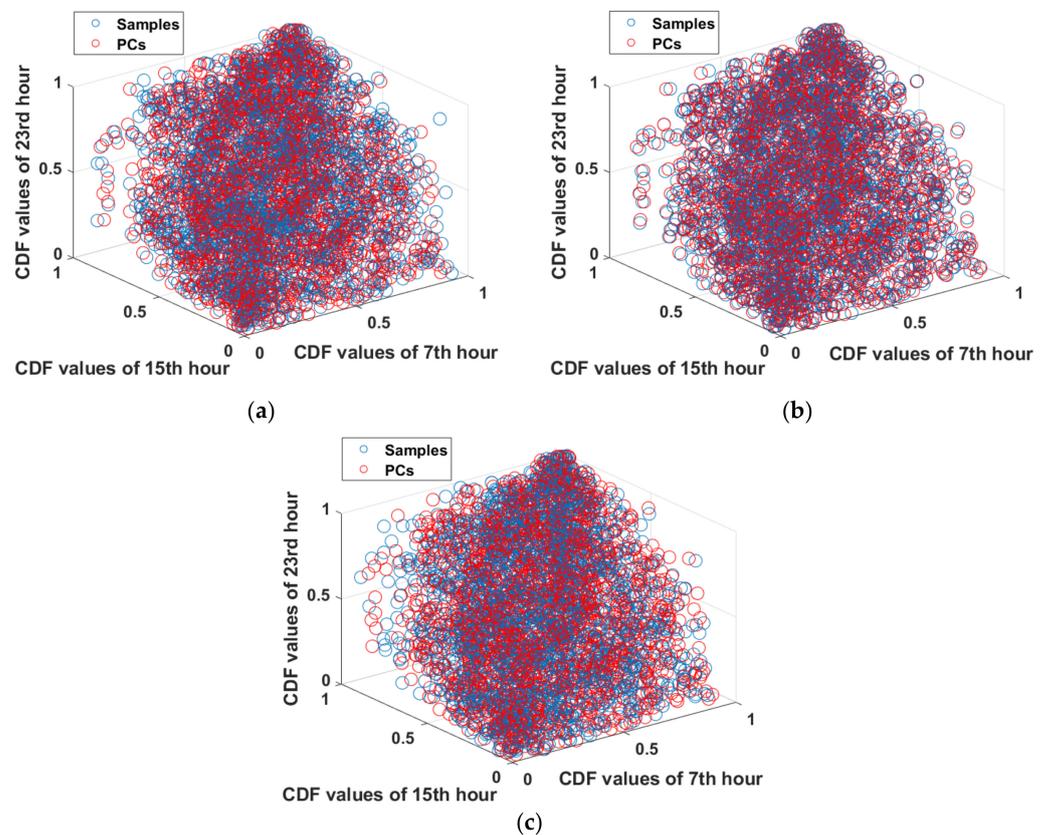


Figure 6. Sample scatter plots with related first PCs (a) for WF 1, (b) for WF 2, (c) for WF 3.

The final wind speed scenarios were derived by performing inverse transformations on the PC scenarios generated using the R-vine copula theory. Figure 7 depicts scatter plots of wind speed samples and scenarios for each wind farm at the 7th, 15th, and 23rd hours. As illustrated in Figure 7, the generated wind speed scenarios were concentrated in the first half of the coordinate axis, and their distribution space and degree of dispersion were essentially identical to the wind speed samples. Figure 8 illustrates the link between wind speed samples and scenarios of the three wind farms for the 7th, 15th, and 23rd hours, respectively, demonstrating the same linear relationship between samples and scenarios. By comparing the wind speed samples and the generated scenarios, it was clear that the wind speed scenarios generated using the method described in this article retained all of the potential correlation characteristics of the wind speed samples and had a distribution interval for the values that was nearly identical to the distribution interval for the samples.

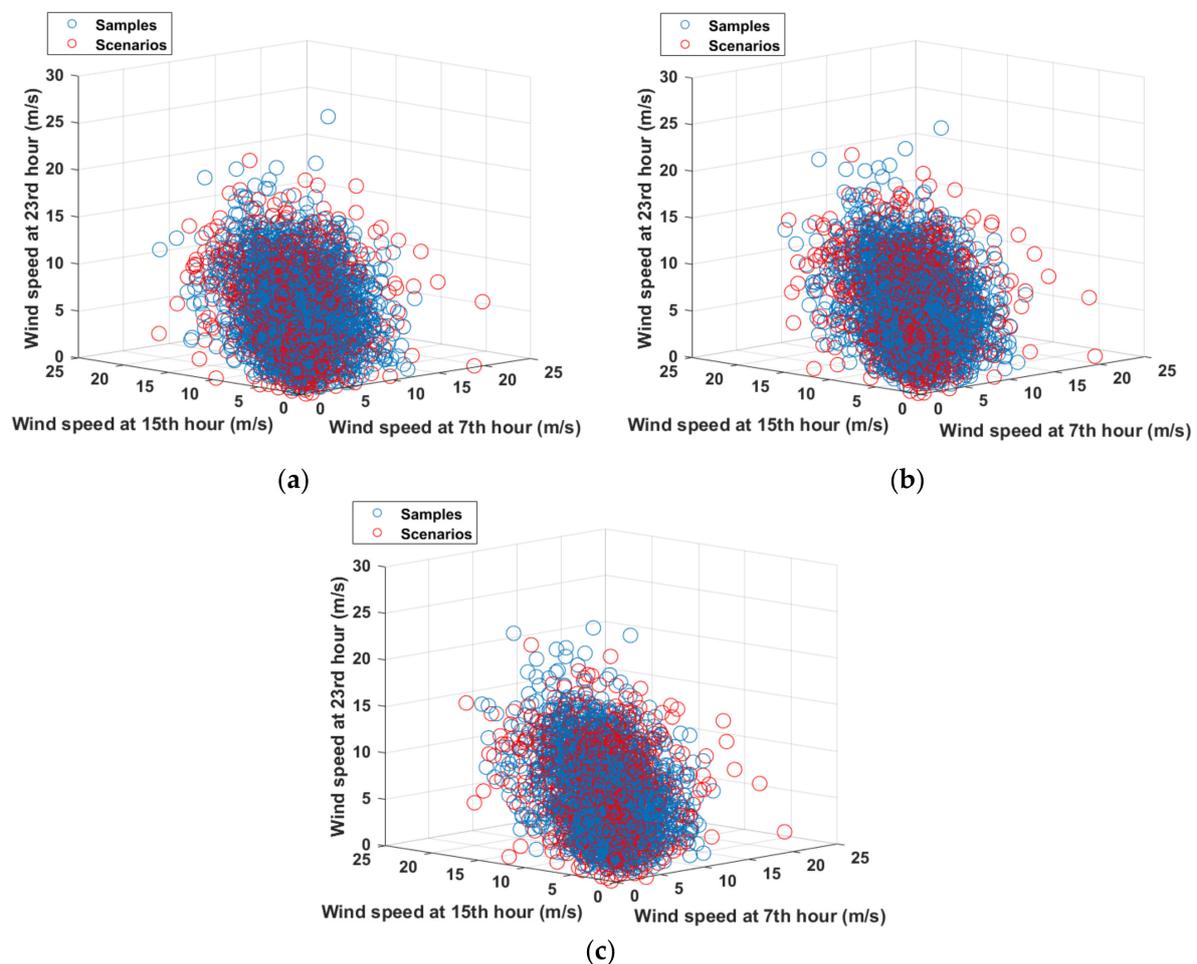


Figure 7. Wind speed samples and scenarios at the 7th, 15th, and 23rd hours (a) for WF 1, (b) for WF 2, (c) for WF 3.

It should be noted that due to the article's length constraint, only the 7th, 15th, and 23rd hours were randomly chosen for graphical example display. The comparison of the final wind speed scenarios to the wind speed samples demonstrated that this method was capable of obtaining the correlation and numerical characteristics of the sample data effectively. As a result, the results obtained using the method described in this article were consistent across all time periods. Their results are identical for time periods shown and not shown.

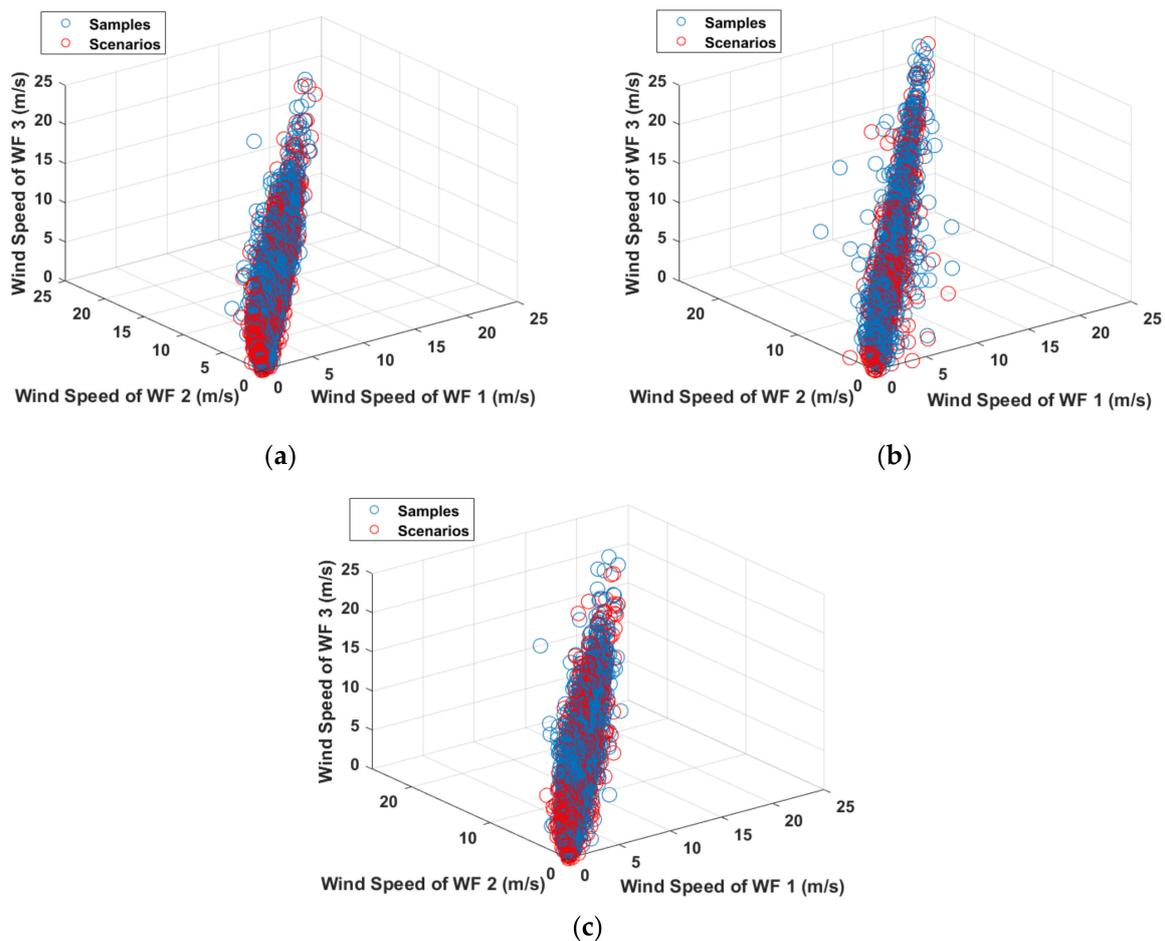


Figure 8. Wind speed samples and scenarios for the three wind farms (a) at the 7th hour, (b) at the 15th hour, (c) at the 23rd hour.

4.3. Comparison of Typical Scenario Generation Methods

Following a comparison of the wind speed samples to the created scenarios, this part compared the proposed model, dubbed the PC-R-vine model, against two other approaches using the same samples, applying multiple assessment criteria to assess each method's performance. As the first study to use PC theory to the generation of wind speed scenarios, this article made use of the PCA-ARMA model [10] as one of the comparison approaches. Additionally, the Hourly Mixed Copula Model (HMCM) [28], which has been proven to be superior to several earlier proposed approaches for generating wind speed scenarios, was used to compare with the proposed method in this research.

Three types of evaluation indicators can be used to categorize scenario generation: output-based evaluation, distribution-based evaluation, and event-based evaluation [36]. To fully compare the advantages and disadvantages of various models, the three types of evaluation indicators mentioned above were used.

5. Output-based evaluation

In contrast to the mean absolute error which only indicates the range of error variation, the mean relative error can be used to indicate the accuracy of the scenario. As a result, this article used the average relative error indicator to compare the mean, standard deviation, skewness, and kurtosis of various numerical characteristics across scenarios and samples.

The mean shows the data's overall numerical level and its relative error E_{mean} in relation to the scenarios and samples can be expressed as follows [22]:

$$E_{mean} = \frac{\sum_{N=1}^p \sum_{i=1}^T \left| \frac{M_{N,i}^{sce} - M_{N,i}^{samp}}{M_{N,i}^{samp}} \right|}{p \times T} \quad (22)$$

where $p = 3$, $T = 24$, and $M_{N,i}^{sce}$ and $M_{N,i}^{samp}$ are the generated scenarios' and samples' average values of wind farm N at the i th hour, respectively.

Similarly, the relative errors between wind speed samples and scenarios with respect to standard deviation, skewness, and kurtosis, E_{std} , E_{ske} , and E_{kur} , respectively, are shown as Equations (23)–(25):

$$E_{std} = \frac{\sum_{N=1}^p \sum_{i=1}^T \left| \frac{St_{N,i}^{sce} - St_{N,i}^{samp}}{St_{N,i}^{samp}} \right|}{p \times T} \quad (23)$$

$$E_{ske} = \frac{\sum_{N=1}^p \sum_{i=1}^T \left| \frac{Sk_{N,i}^{sce} - Sk_{N,i}^{samp}}{Sk_{N,i}^{samp}} \right|}{p \times T} \quad (24)$$

$$E_{kur} = \frac{\sum_{N=1}^p \sum_{i=1}^T \left| \frac{Kur_{N,i}^{sce} - Kur_{N,i}^{samp}}{Kur_{N,i}^{samp}} \right|}{p \times T} \quad (25)$$

where $St_{N,i}^{sce}$, $Sk_{N,i}^{sce}$, and $Kur_{N,i}^{sce}$ are the generated scenarios' standard deviation, skewness, and kurtosis of wind farm N at the i th hour, respectively; $St_{N,i}^{samp}$, $Sk_{N,i}^{samp}$, and $Kur_{N,i}^{samp}$ are the samples' standard deviation, skewness, and kurtosis of wind farm N at the i th hour, respectively.

The Euclidean distance between the correlation coefficients of the wind speed samples and the scenarios can indicate the degree to which the wind speed data has been restored to their original correlation. This article assessed the Euclidean distance of the correlation between samples and situations in terms of temporal and spatial correlation, as stated in [28]:

$$E_{temp} = \frac{\sqrt{\sum_{N=1}^p \sum_{i=1}^T \sum_{j=i+1}^T \left(\rho_{N,(i,j)}^{samp} - \rho_{N,(i,j)}^{sce} \right)^2}}{p \times T} \quad (26)$$

$$E_{spa} = \frac{\sqrt{\sum_{i=1}^T \sum_{N=1}^p \sum_{M=N+1}^p \left(\rho_{i,(N,M)}^{samp} - \rho_{i,(N,M)}^{sce} \right)^2}}{p \times T} \quad (27)$$

where $\rho_{N,(i,j)}^{samp}$ and $\rho_{N,(i,j)}^{sce}$ are the Spearman's correlation coefficients for the i th hour and j th hour wind speed data for wind farm N in the samples and scenarios, respectively; $\rho_{i,(N,M)}^{samp}$ and $\rho_{i,(N,M)}^{sce}$ are the Spearman's correlation coefficients of wind speed data for wind farm N and wind farm M at the i th hour in the samples and scenarios, respectively. The smaller the above-mentioned six evaluation metrics, the more favorable the outcome.

Table 1 presents these metrics for the three models discussed in this section. Except for the standard deviation, all other evaluation metrics for PC-R-vine were superior to those for the other two models, as shown in Table 1. Specifically, when compared to the PCA-ARMA and HMCM models, the PC-R-vine model produced the lowest relative errors for the mean value, skewness, and kurtosis of the wind speed scenarios, indicating that the numerical range and probability distribution characteristics of the wind speed data generated by the PC-R-vine model were closer to those of the wind speed samples. In terms of standard deviation, the errors of the PC-R-vine and HMCM models relative to wind speed samples were much smaller than those of the PCA-ARMA model, but the PC-R-vine

model had a bigger standard deviation than the HMCM model. This finding indicates that, when compared to the samples, the fluctuation degree of wind speed values generated by the PC-R-vine model was significantly less than that generated by the PCA-ARMA model, but slightly greater than that provided by the HMCM model. While HMCM had a lower relative error for standard deviation, its mean had a higher relative error, which means that while the generated wind speed scenarios had less fluctuation, their numerical level was lower than the sample data. Smaller fluctuations around low numerical values, which have a detrimental effect on the numerical distribution of the scenarios, were generated by this model to some extent. This is also why the relative errors for kurtosis and skewness associated with HMCM were greater. Meanwhile, it demonstrated that an effective model testing process necessitates a thorough evaluation using a variety of indicators.

Table 1. Each model's evaluation metrics.

	PC-R-Vine	PCA-ARMA	HMCM
E_{mean}	0.0143	0.0701	0.0175
E_{std}	0.0287	0.0435	0.0224
E_{ske}	0.0983	0.4272	0.3807
E_{kur}	0.0411	0.1981	0.2009
E_{temp}	0.0421	0.0844	0.3945
E_{spa}	0.0019	0.0088	0.0033

Additionally, the Euclidean distances in terms of temporal and spatial correlations in the PC-R-vine model was significantly smaller than that in the PCA-ARMA and HMCM models, indicating that the wind speed scenarios generated by the PC-R-vine model better retained the correlation characteristics of wind speed samples.

In summary, while the PC-R-vine model did not outperform the HMCM models in terms of standard deviation, it outperformed the two models in terms of other assessment criteria. As a result, it can be stated that the wind speed scenarios generated by the PC-R-vine model retained a higher degree of statistical fidelity to the wind speed samples.

6. Distribution-based evaluation

Along with numerical characteristics, it is necessary to investigate the distribution differences between scenarios and samples. To demonstrate the consistency of wind speed samples and scenarios of each wind farm, Quantile-Quantile plots (QQ plots) between the wind speed values derived by each model and the corresponding sample data are shown in Figures 9–11. As illustrated in Figure 9, for wind farm 1, the variation in probability distribution between the scenarios generated by HMCM model and the wind speed samples was the most noticeable. In comparison to the PC-R-vine model, the PCA-ARMA model's difference in probability distributions varied more prominently in the right corner of Figure 9b. Through comparison of Figure 10a–c, the discrepancy between the probability distributions of wind speed scenarios provided by the PC-R-vine model and wind speed samples was minimal for wind farm 2. This result was also supported by the wind speed data for wind farm 3, as illustrated in Figure 11a–c. As a result of the preceding, it can be inferred that the PC-R-vine model's probability distribution for wind speed situations is the most similar to the wind speed samples.

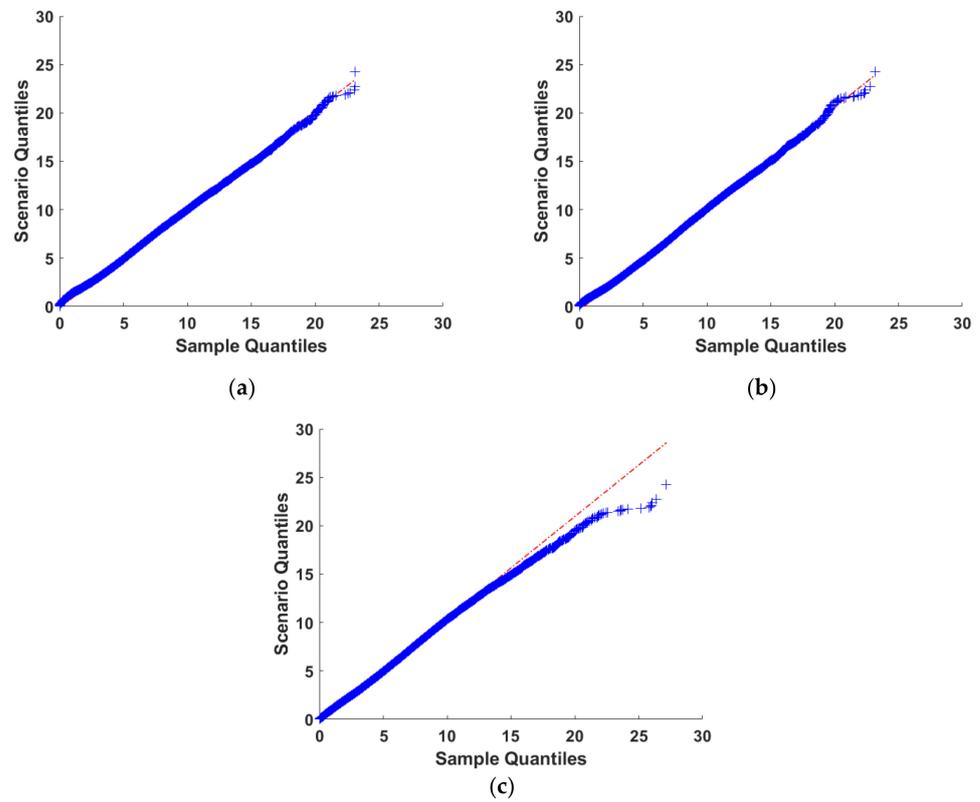


Figure 9. QQ plot for WF 1 of samples and scenarios generated using (a) the PC-R-vine model, (b) the PCA-ARMA model, (c) the HMCM model.

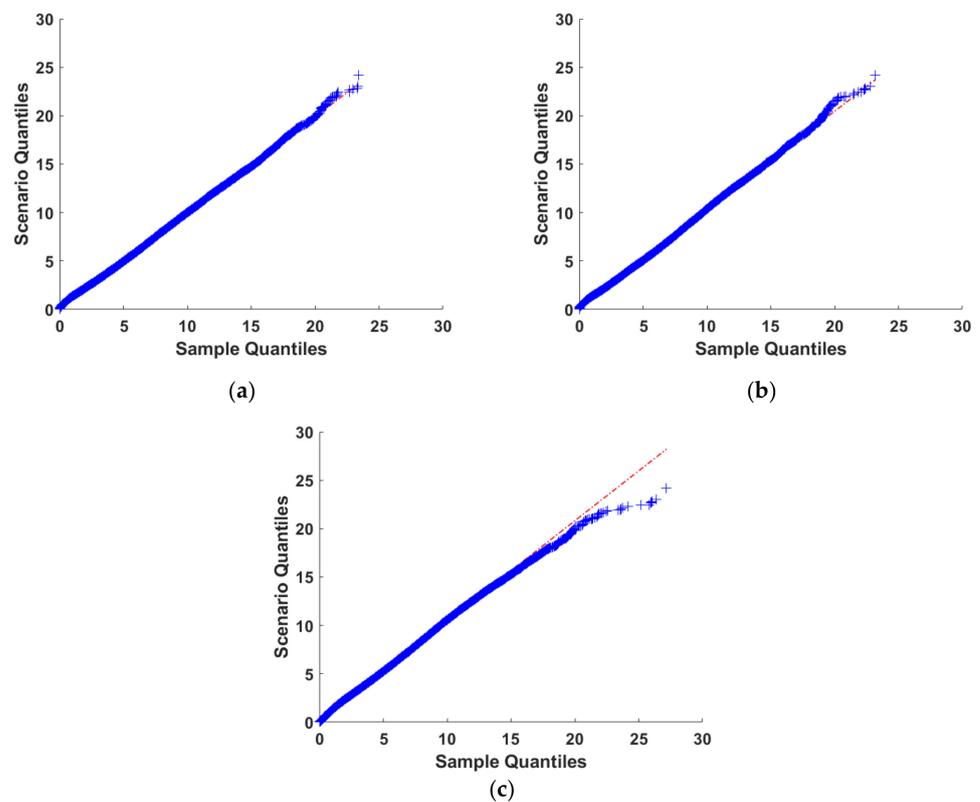


Figure 10. QQ plot for WF 2 of samples and scenarios generated using (a) the PC-R-vine model, (b) the PCA-ARMA model, (c) the HMCM model.

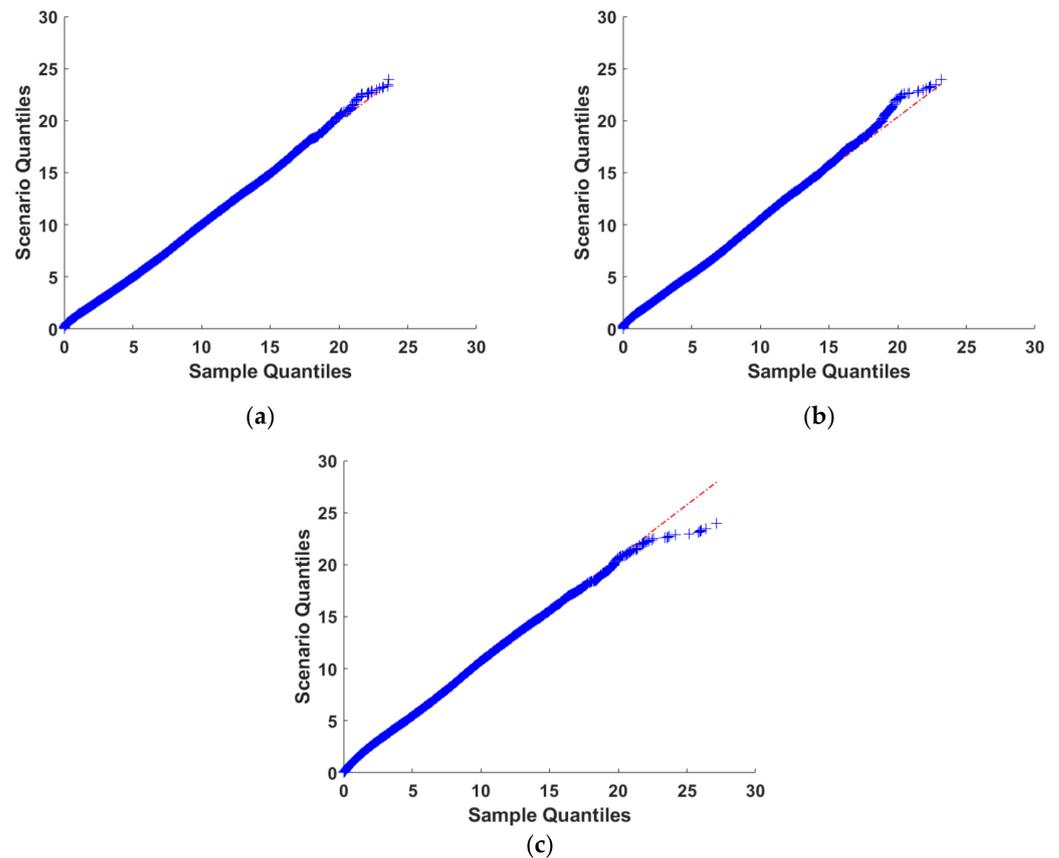


Figure 11. QQ plot for WF 3 of samples and scenarios generated using (a) the PC-R-vine model, (b) the PCA-ARMA model, (c) the HMCM model.

7. Event-based evaluation

The purpose of the event-based evaluation index is to conduct comparative analysis on the entire dataset. The coverage rate concept was used in this article to determine the degree to which the sample data was included in the generated scene. Due to the fact that the majority of sample data fell within the numerical range of the generated scenarios, this article appropriately modified the concept of coverage rate by converting it to an uncovered percentage metric (UPM). It can be stated in the following manner [36]:

$$c = \frac{\sum_{N=1}^p \sum_{t=1}^T \mathbf{I}(p_{\min,t}^N \leq P_t^N \leq p_{\max,t}^N)}{T \times NS_{total}} \quad (28)$$

$$p_{\min,t}^N = \min(p_{s,t}^N), N = 1, 2, \dots, p, s = 1, 2, \dots, SN \quad (29)$$

$$p_{\max,t}^N = \max(p_{s,t}^N), N = 1, 2, \dots, p, s = 1, 2, \dots, SN \quad (30)$$

where p is the number of wind farms; T is the total number of time stamps; $p_{\min,t}^N$ and $p_{\max,t}^N$ are the maximum and minimum wind speeds of wind farm N in all scenarios at time t ; NS_{total} is the total number of comparisons in each time period; $\mathbf{I}(\cdot)$ is a binary variable that is zero if the conditions in the parentheses are satisfied, and one otherwise; SN is the number of scenarios.

Table 2 summarizes the UPM results for the three models. As shown in Table 2, the PC-R-vine method produced the lowest result with a UPM value of 0.22%, indicating that the wind speed scenarios generated by the PC-R-vine model contained more wind speed samples than the other two methods. As a result, the wind speed scenarios generated by

the method proposed in this article were more representative of the data as a whole, and their overall accuracy and reliability were higher.

Table 2. Each model's uncovered percentage results.

	PC-R-Vine	PCA-ARMA	HMCM
UPM (%)	0.22	0.46	0.49

The proposed method was compared to two other models using three different types of evaluation metrics in this article. The results indicated that the model proposed in this article outperformed the other two models in numerous aspects of these three indicator categories. By and large, the PC-R-vine was superior to the PCA-ARMA model in most dimensions. This is mostly owing to the flexibility of the copula theory utilized in the first model, which allows for the capture of nonlinear correlations, whereas the latter is limited to linear correlations. Additionally, the PC-R-vine model is preferable than the HMCM model because the R-vine copula is adaptable and flexible in terms of copula function selection, and PC theory and its inverse transformation assure the effective retention of correlations.

It should be noted that both the PC and R-vine copula theories used in this article are applicable to high-dimensional data, which means that as the number of wind farms increases, the results will remain stable and the accuracy will remain constant. As a result, this model is still applicable to a greater number of wind farms. Additionally, the primary goal of scenario generation methods is to ensure model accuracy, followed by runtime acceptability. Therefore, the duration of various scenario generation processes varies depending on the specific methods used, as long as they remain within an acceptable range. While scenario generation methods that do not use the vine structure have some advantages in runtime, their accuracy is frequently inferior to the method combined with the vine copula. The purpose of this article was to provide a method for scenario generation that is computationally efficient while maintaining accuracy, and thus there was no comparison of model running time. Although the running time of the scenario generation method proposed in this article was slightly inferior to that of the HMCM and PCA-ARMA models, it was completely within an acceptable range, and the method's accuracy was significantly higher than that of the other two models, which is consistent with the article's expectations and objectives.

4.4. Comparison of Several Vine Structure Models

To demonstrate the importance and effectiveness of R-vine in this article, the most frequently used C-vine and D-vine structures were applied to the scenario generation process introduced in Section 3.2. The R-vine copula was replaced with C-vine and D-vine copulas, and the accuracy of the wind speed scenarios generated by each was compared and analyzed. Due to the maturity of C-vine and D-vine copula techniques and the article's limited space, this article did not delve into detail about the structure of C-vine and D-vine copula. The accuracy of C-vine and D-vine was evaluated using the evaluation indicators introduced in Section 4.3, and the results are shown in Table 3.

Table 3. Each vine structure's evaluation metrics.

	R-Vine	C-Vine	D-Vine
E_{mean}	0.0143	0.0308	0.0315
E_{std}	0.0287	0.0568	0.0559
E_{ske}	0.0983	0.0944	0.0905
E_{kur}	0.0411	0.0670	0.0521
E_{temp}	0.0421	0.0367	0.0371
E_{spa}	0.0019	0.0023	0.0021
UPM (%)	0.22	0.29	0.29

As shown in Table 3, there was little difference between the results obtained with C-vine and D-vine, which is due to the fixed structure of the two, limiting the difference between the two. In terms of skewness and temporal correlation, D-vine and C-vine had the smallest mean relative errors in Euclidean distance, respectively. From a single perspective, this demonstrated that D-vine deviated the least from the sample distribution and C-vine was the most capable of capturing temporal correlations. Nonetheless, as demonstrated in Section 4.3, their poor performance in other areas means that their advantages in a single area did not translate to the accuracy of the final results.

In comparison to the C-vine and D-vine copulas, the R-vine copula outperformed them in terms of mean, standard deviation, kurtosis, Euclidean distance with respect to spatial correlation, and UPM, indicating that the R-vine results were superior to those obtained by the other two structures in terms of numerical level, degree of fluctuation, ability to obtain spatial correlation, and degree of sample data inclusion. While R-vine copula performed slightly worse than C-vine and D-vine copula in terms of temporal correlation acquisition, R-vine copula generated correlation scenarios that were closer to the numerical level of the sample data. While C-vine and D-vine retained the greatest degree of temporal correlation, the sample data error was greater, resulting in a greater difference between the distribution and numerical characteristics of the final result and the sample data. As a result, the results obtained by R-vine were more precise than those obtained by C-vine and D-vine.

In addition, for the wind speed data of p wind farms with time resolution T , the parameter estimation amount of the method proposed in this article was $p \times (T \times (T - 1)/2)^2 + C$. When only R-vine is used, the required parameter estimation amount is $(p \times T \times (p \times T - 1)/2)^2 + D$, where C and D are constants. It can be seen from this that the method can greatly reduce the computational burden of parameters to be estimated. Therefore, the method proposed in this article was an effective scenario generation method that took into account both accuracy and computational efficiency.

5. Conclusions

By utilizing an accurate generation approach for wind speed scenarios, it is possible to significantly increase the reliability of incorporating wind power output while developing a power system that includes wind power. This article provided a new method for creating wind speed scenarios in the case of many wind farms that takes into account the spatiotemporal correlation of wind speeds. To begin, PC theory was used to turn each wind farm's 24 h wind speed data into PC values, thereby temporarily eliminating spatial correlation between wind farms. The simulation results indicated that the temporal correlation of the wind speed data between the PCs was maintained during this stage. Then, the R-vine copula was utilized to capture the temporal correlation between the collected PCs, and PC scenarios were constructed that account for the temporal correlation. Finally, using the inverse process of PC generation, the collected findings were translated into the final wind speed scenarios, restoring the spatial correlation between wind speed data. The simulation examination of wind speed data from three wind farms provided by NREL demonstrated that the proposed method was more accurate at capturing the correlation relationship than the HMCM and PCA-ARMA models. Additionally, the generated scenarios' statistical properties were more similar to the original data, and the probability distribution gap between the generated scenarios and the original data was lower. All of these data suggest that the strategy proposed in this article could produce more detailed wind speed scenarios. This method can be applied to scenario-based power system stochastic programming problems, providing support for the development of power system with wind power.

Although the combination of PC and R-vine copula theories increased calculation efficiency while maintaining the highest possible accuracy, it has some drawbacks. Restricted by the fact that PC theory is limited to the transformation of linear correlations, there are limitations in applying PC theory to model complex nonlinear correlations of wind farms, influencing its applicability in this case. The method proposed in this article reduced the accuracy of capturing complex nonlinear correlations when seeking a balance

with computational efficiency. Future research can enhance PC theory's ability to deal with nonlinear correlation data sets by combining it with other transformations, allowing it to convert correlation data into completely independent data, as well as expand its applicability to data sets with different correlations, thereby increasing the accuracy of the method proposed in this article.

Author Contributions: Conceptualization, H.H.G.; methodology, H.H.G.; software, H.H.G., D.Z. and W.D.; validation, H.H.G., D.Z. and W.D.; formal analysis, K.C.G.; data curation, G.P.; writing—original draft preparation, G.P.; writing—review and editing, H.H.G. and K.C.G.; visualization, H.H.G., C.L.C. and T.A.K.; supervision, H.H.G. and D.Z.; funding acquisition, H.H.G. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by Jun Wu Scholar Grant Scheme, Vot: A3020051008. The APC was funded by Guangxi University.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: Not applicable.

Conflicts of Interest: The authors declare no conflict of interest.

References

- Goh, H.H.; He, R.; Zhang, D.; Liu, H.; Dai, W.; Lim, C.S.; Kurniawan, T.A.; Teo, K.T.K.; Goh, K.C. A multimodal approach to chaotic renewable energy prediction using meteorological and historical information. *Appl. Soft Comput.* **2022**, *118*, 108487. [\[CrossRef\]](#)
- Sahragard, A.; Falaghi, H.; Farhadi, M.; Mosavi, A.; Estebarsari, A. Generation Expansion Planning in the Presence of Wind Power Plants Using a Genetic Algorithm Model. *Electronics* **2020**, *9*, 1143. [\[CrossRef\]](#)
- Band, S.S.; Bateni, S.M.; Almazroui, M.; Sajjadi, S.; Chau, K.-w.; Mosavi, A. Evaluating the potential of offshore wind energy in the Gulf of Oman using the MENA-CORDEX wind speed data simulations. *Eng. Appl. Comput. Fluid Mech.* **2021**, *15*, 613–626. [\[CrossRef\]](#)
- D'Amico, G.; Petroni, F.; Praticco, F. First and second order semi-Markov chains for wind speed modeling. *Phys. A Stat. Mech. Its Appl.* **2013**, *392*, 1194–1201. [\[CrossRef\]](#)
- Chen, J.; Rabiti, C. Synthetic wind speed scenarios generation for probabilistic analysis of hybrid energy systems. *Energy* **2017**, *120*, 507–517. [\[CrossRef\]](#)
- Sim, S.K.; Maass, P.; Lind, P.J.E. Wind Speed Modeling by Nested ARIMA Processes. *Energies* **2018**, *12*, 69. [\[CrossRef\]](#)
- Sun, W.; Zamani, M.; Zhang, H.T.; Li, Y. Probabilistic Optimal Power Flow With Correlated Wind Power Uncertainty via Markov Chain Quasi-Monte-Carlo Sampling. *IEEE Trans. Ind. Inform.* **2019**, *15*, 6058–6069. [\[CrossRef\]](#)
- Morales, J.M.; Mínguez, R.; Conejo, A.J. A methodology to generate statistically dependent wind speed scenarios. *Appl. Energy* **2010**, *87*, 843–855. [\[CrossRef\]](#)
- Abedi, A.; Rahimiyan, M. Day-ahead energy and reserve scheduling under correlated wind power production. *Int. J. Electr. Power Energy Syst.* **2020**, *120*, 105931. [\[CrossRef\]](#)
- Le, D.D.; Gross, G.; Berizzi, A. Probabilistic Modeling of Multisite Wind Farm Production for Scenario-Based Applications. *IEEE Trans. Sustain. Energy* **2015**, *6*, 748–758. [\[CrossRef\]](#)
- Chen, Y.; Wang, Y.; Kirschen, D.; Zhang, B. Model-Free Renewable Scenario Generation Using Generative Adversarial Networks. *IEEE Trans. Power Syst.* **2018**, *33*, 3265–3275. [\[CrossRef\]](#)
- Zhang, Y.; Ai, Q.; Xiao, F.; Hao, R.; Lu, T. Typical wind power scenario generation for multiple wind farms using conditional improved Wasserstein generative adversarial network. *Int. J. Electr. Power Energy Syst.* **2020**, *114*, 105388. [\[CrossRef\]](#)
- Jiang, C.; Mao, Y.; Chai, Y.; Yu, M.; Tao, S. Scenario Generation for Wind Power Using Improved Generative Adversarial Networks. *IEEE Access* **2018**, *6*, 62193–62203. [\[CrossRef\]](#)
- Qiao, J.; Pu, T.; Wang, X. Renewable scenario generation using controllable generative adversarial networks with transparent latent space. *CSEE J. Power Energy Syst.* **2021**, *7*, 66–77.
- Liang, J.; Tang, W. Sequence Generative Adversarial Networks for Wind Power Scenario Generation. *IEEE J. Sel. Areas Commun.* **2020**, *38*, 110–118. [\[CrossRef\]](#)
- Haghi, H.V.; Lotfifard, S. Spatiotemporal Modeling of Wind Generation for Optimal Energy Storage Sizing. *IEEE Trans. Sustain. Energy* **2015**, *6*, 113–121. [\[CrossRef\]](#)
- Becker, R. Generation of Time-Coupled Wind Power Infeed Scenarios Using Pair-Copula Construction. *IEEE Trans. Sustain. Energy* **2018**, *9*, 1298–1306. [\[CrossRef\]](#)
- Lin, Z.; Chen, H.; Wu, Q.; Huang, J.; Li, M.; Ji, T. A data-adaptive robust unit commitment model considering high penetration of wind power generation and its enhanced uncertainty set. *Int. J. Electr. Power Energy Syst.* **2021**, *129*, 106797. [\[CrossRef\]](#)

19. Borujeni, M.S.; Foroud, A.A.; Dideban, A. Wind speed scenario generation based on dependency structure analysis. *J. Wind Eng. Ind. Aerodyn.* **2018**, *172*, 453–465. [[CrossRef](#)]
20. Eryilmaz, S.; Kan, C. Reliability based modeling and analysis for a wind power system integrated by two wind farms considering wind speed dependence. *Reliab. Eng. Syst. Saf.* **2020**, *203*, 107077. [[CrossRef](#)]
21. Li, M.S.; Lin, Z.J.; Ji, T.Y.; Wu, Q.H. Risk constrained stochastic economic dispatch considering dependence of multiple wind farms using pair-copula. *Appl. Energy* **2018**, *226*, 967–978. [[CrossRef](#)]
22. Deng, J.; Li, H.; Hu, J.; Liu, Z. A new wind speed scenario generation method based on spatiotemporal dependency structure. *Renew. Energy* **2021**, *163*, 1951–1962. [[CrossRef](#)]
23. Qiu, Y.; Li, Q.; Pan, Y.; Yang, H.; Chen, W. A scenario generation method based on the mixture vine copula and its application in the power system with wind/hydrogen production. *Int. J. Hydrogen Energy* **2019**, *44*, 5162–5170. [[CrossRef](#)]
24. Xu, J.; Wu, W.; Wang, K.; Li, G. C-Vine pair copula based wind power correlation modelling in probabilistic small signal stability analysis. *IEEE CAA J. Autom. Sin.* **2020**, *7*, 1154–1160. [[CrossRef](#)]
25. Henderson, S.B.; Shahirinia, A.H.; Tavakoli Bina, M. Bayesian estimation of copula parameters for wind speed models of dependence. *IET Renew. Power Gener.* **2021**, *15*, 3823–3831. [[CrossRef](#)]
26. Wang, C.; Liang, Z.; Liang, J.; Teng, Q.; Dong, X.; Wang, Z. Modeling the temporal correlation of hourly day-ahead short-term wind power forecast error for optimal sizing energy storage system. *Int. J. Electr. Power Energy Syst.* **2018**, *98*, 373–381. [[CrossRef](#)]
27. Li, L.; Miao, S.; Tu, Q.; Duan, S.; Li, Y.; Han, J. Dynamic dependence modelling of wind power uncertainty considering heteroscedastic effect. *Int. J. Electr. Power Energy Syst.* **2020**, *116*, 105556. [[CrossRef](#)]
28. Philippe, W.P.J.; Zhang, S.; Eftekharnajad, S.; Ghosh, P.K.; Varshney, P.K. Mixed Copula-Based Uncertainty Modeling of Hourly Wind Farm Production for Power System Operational Planning Studies. *IEEE Access* **2020**, *8*, 138569–138583. [[CrossRef](#)]
29. Wang, Z.; Wang, W.; Liu, C.; Wang, Z.; Hou, Y. Probabilistic Forecast for Multiple Wind Farms Based on Regular Vine Copulas. *IEEE Trans. Power Syst.* **2018**, *33*, 578–589. [[CrossRef](#)]
30. Wang, Z.; Wang, W.; Liu, C.; Wang, B. Forecasted Scenarios of Regional Wind Farms Based on Regular Vine Copulas. *J. Mod. Power Syst. Clean Energy* **2020**, *8*, 77–85. [[CrossRef](#)]
31. Xiao, Q.; Zhou, S. Probabilistic power flow computation considering correlated wind speeds. *Appl. Energy* **2018**, *231*, 677–685. [[CrossRef](#)]
32. Zhou, S.; Xiao, Q.; Wu, L. Probabilistic power flow analysis with correlated wind speeds. *Renew. Energy* **2020**, *145*, 2169–2177. [[CrossRef](#)]
33. Xiao, Q.; Zhou, S.; Xiao, H. Probabilistic optimal power flow analysis incorporating correlated wind sources. *Int. Trans. Electr. Energy Syst.* **2020**, *30*, e12441. [[CrossRef](#)]
34. Dißmann, J.; Brechmann, E.C.; Czado, C.; Kurowicka, D. Selecting and estimating regular vine copulae and application to financial returns. *Comput. Stat. Data Anal.* **2013**, *59*, 52–69. [[CrossRef](#)]
35. NREL. Wind Speed Dataset. Available online: <https://maps.nrel.gov/wind-prospector/> (accessed on 14 January 2022).
36. Li, J.; Zhou, J.; Chen, B. Review of wind power scenario generation methods for optimal operation of renewable energy systems. *Appl. Energy* **2020**, *280*, 115992. [[CrossRef](#)]