COMPLEX SYSTEM MODELING AND SIMULATION
ISSN 2096-9929 06/06 pp 157−168
Volume 3, Number 2, June 2023
DOI: 10.23919/CSMS.2023.0007

Simulating Temporally and Spatially Correlated Wind Speed Time Series by Spectral Representation Method

Qing Xiao, Lianghong Wu*, Xiaowen Wu*, and Matthias Rätsch

Abstract: In this paper, it aims to model wind speed time series at multiple sites. The five-parameter Johnson distribution is deployed to relate the wind speed at each site to a Gaussian time series, and the resultant m-dimensional Gaussian stochastic vector process $\boldsymbol{Z}(t)$ is employed to model the temporal-spatial correlation of wind speeds at m different sites. In general, it is computationally tedious to obtain the autocorrelation functions (ACFs) and cross-correlation functions (CCFs) of $\boldsymbol{Z}(t)$, which are different to those of wind speed times series. In order to circumvent this correlation distortion problem, the rank ACF and rank CCF are introduced to characterize the temporal-spatial correlation of wind speeds, whereby the ACFs and CCFs of $\boldsymbol{Z}(t)$ can be analytically obtained. Then, Fourier transformation is implemented to establish the cross-spectral density matrix of $\boldsymbol{Z}(t)$, and an analytical approach is proposed to generate samples of wind speeds at m different sites. Finally, simulation experiments are performed to check the proposed methods, and the results verify that the five-parameter Johnson distribution can accurately match distribution functions of wind speeds, and the spectral representation method can well reproduce the temporal-spatial correlation of wind speeds.

Key words: multivariate wind speed time series; rank autocorrelation function; rank cross-correlation function; cross-spectral density matrix; five-parameter Johnson distribution

1 Introduction

Due to the increasing integration of wind resources into the power grid, reliable statistical models are invoked to capture the stochastic nature of wind speed[1, 2]. Hitherto, several approaches have been attempted to handle this issue, such as the autoregressive moving average (ARMA) model, Markov chain method, and stochastic differential equation method.

Mathematically, the wind speed can be modeled as a stochastic process $\boldsymbol{X}(t)$, and the wind velocity at an arbitrary time instant t_1 is regarded as a random variable $\boldsymbol{X}(t_1)$[3]. One notable feature of $\boldsymbol{X}(t)$ is that the wind velocity $\boldsymbol{X}(t_1)$ exhibits a certain level of correlation with $\boldsymbol{X}(t_2)$ ($t_1 \neq t_2$), and this temporal correlation of wind speed is generally characterized by the autocorrelation function (ACF).

The ARMA model deploys a set of weighted Gaussian stochastic processes to simulate $\boldsymbol{X}(t)$[4]. By selecting appropriate parameters for the ARMA model, it can well reproduce the ACF of $\boldsymbol{X}(t)$. However, the ARMA model cannot directly generate wind speed samples with a non-Gaussian distribution and fails to well reproduce the actual probability distribution of $\boldsymbol{X}(t)$. In comparison to the ARMA model, Markov chain method can accurately capture the non-Gaussian distribution of wind speed[5, 6], but it does not perform very well for characterizing the time correlation of $\boldsymbol{X}(t)$; more specifically, Markov chain method is

• Qing Xiao, Lianghong Wu, and Xiaowen Wu are with the School of Information and Electrical Engineering, Hunan University of Science and Technology, Xiangtan 411201, China. E-mail: lhwu@hnust.edu.cn; xwu@hnust.edu.cn.
• Matthias Rätsch is with the Image Understanding and Interactive Robotics Group, Reutlingen University, Reutlingen D-72762, Germany.
* To whom correspondence should be addressed.
※ This article was recommended by Associate Editor Wenyin Gong.
Manuscript received: 2023-02-22; revised: 2023-04-06; accepted: 2023-04-10

difficult to accurately reproduce the ACF of $\boldsymbol{X}(t)$[7]. Although this problem can be alleviated by using a Markov chain model with a larger number of states, it would lead to a larger transition matrix and present additional difficulties in computation[8].

In Ref. [9], it suggested a stochastic differential equation based method to simulate $\boldsymbol{X}(t)$. This method employs a sum of exponential functions and sinusoidal functions to represent the ACF of $\boldsymbol{X}(t)$; then, it deploys a weighted sum of stochastic differential equation processes to obtain a Gaussian stochastic process $\boldsymbol{Z}(t)$, which is used to prescribe the ACF of $\boldsymbol{X}(t)$; finally, it uses the marginal transformation to transform $\boldsymbol{Z}(t)$ to $\boldsymbol{X}(t)$ and to match the distribution function of $\boldsymbol{X}(t)$. One point worth noting is that the marginal transformation strategy can also be conjuncted with the ARMA model, whereby the task of modeling wind speed time series $\boldsymbol{X}(t)$ can be decoupled into two independent parts: fitting the distribution function of $\boldsymbol{X}(t)$ and matching the ACF of $\boldsymbol{X}(t)$[10].

In practical settings, wind speed time series at adjacent sites also show a correlation with each other due to similar weather conditions[11]. This spatial correlation can be characterized by the cross-correlation function (CCF)[12]. Among the aforementioned three methods, only the ARMA model has been extended to a vector form, and to capture the CCFs of multivariate wind speed time series. In Refs. [10, 13], it suggested to match the CCFs of multivariate wind speed time series by regulating the correlation among Gaussian stochastic processes of ARMA model. However, because the wind speed generally follows a non-Gaussian probability distribution, it leads to the correlation distortion problem[14], and tedious numerical algorithms should be employed to determine the ACFs and CCFs of wind speeds in the standard normal space[10, 15].

This paper sets out to develop a convenient algorithm to simulate multivariate wind speed time series. The main content and contribution are as follows.

(1) A five-parameter Johnson distribution is employed to reconstruct probability distributions of wind speed samples, which can strictly match the lower and upper bounds of wind velocity and directly map wind speed samples to the standard normal space;

(2) The rank ACF and rank CCF are proposed to measure the temporal-spatial correlation of wind samples at multiple sites. They are invariant under the marginal transformation and allow for analytical procedures to determine the ACF and CCF of wind speeds in the standard normal space;

(3) The spectral representation method is customized for generating samples of wind speed with a prescribed rank temporal-spatial correlation.

The structure of the following parts is organized as follows. Section 2 defines the rank ACF and rank CCF. Section 3 introduces the spectral representation method, the five-parameter Johnson distribution, and the procedures of the proposed method. In Section 4, the case study is performed. Section 5 gives the conclusion.

2 Statistical Features of Multivariate Wind Speed Time Series

Consider wind speeds at m different sites, which can be denoted as an m-dimensional vector $\boldsymbol{X}(t) = (X_1(t), X_2(t), \ldots, X_i(t), \ldots, X_m(t))^{\mathrm{T}}$. The statistical features of $\boldsymbol{X}(t)$ can be characterized by

(1) the distribution functions of wind speed at each site;

(2) the ACF of $X_i(t)$ $(i = 1, 2, \ldots, m)$, which characterizes the temporal correlation;

(3) the CCF of wind speed time series at different sites, which measures the spatial correlation.

Let $\{x_{i,k}\}$ $(k = 1, 2, \ldots, n)$ denote samples of wind speed at the i-th site, and the empirical CDF of $X_i(t)$ is defined by

$$\begin{aligned} &\hat{F}_i(X_i) = \frac{1}{n}\sum_{k=1}^{n} I(x_{i,k} \leqslant X_i), \\ &I(x_{i,k} \leqslant X_i) = \begin{cases} 1, & \text{if } x_{i,k} \leqslant X_i; \\ 0, & \text{else} \end{cases} \end{aligned} \tag{1}$$

In this paper, instead of the commonly used linear ACF and CCF, the rank ACF and rank CCF are employed to measure the time-spatial correlation of $\boldsymbol{X}(t)$. With the empirical CDF in Eq. (1), it has

$$u_{i,k} = \widehat{F}_i(x_{i,k}), \quad k = 1, 2, \ldots, n \tag{2}$$

Then

$$\rho_{i,i}(\tau)\,|_{\tau = s\cdot\Delta\tau} = \frac{\sum_{k=M}^{N} (u_{i,k} - \mu_i)(u_{i,k+s} - \mu_i)}{(N - M + 1)\cdot\sigma_i^2} \tag{3}$$

where $\rho_{i,i}(\tau)$ denotes the rank ACF of $X_i(t)$ $(i = 1, 2, \ldots, m)$, and τ and $\Delta\tau$ are the time lag and the recording interval, respectively. N is the number of the used wind samples, and $(N + |s|) \leqslant n$ and $(M + s) > 0$. μ_i is the mean of $\{u_{i,k}\}$ $(k = 1, 2, \ldots, n)$, and σ_i is the standard deviation. M and s are integers.

The proposed rank CCF is given by

$$\rho_{i,j}(\tau)\,|_{\tau=s\cdot\Delta\tau}=\frac{\sum_{k=M}^{N}(u_{i,k}-\mu_i)(u_{j,k+s}-\mu_j)}{(N-M+1)\cdot\sigma_i\sigma_j} \tag{4}$$

where $\rho_{i,j}(\tau)$ is the rank CCF between $X_i(t)$ and $X_j(t)$ $(i,j=1,2,\ldots,m;\ i\neq j)$.

With the rank ACF in Eq. (3) and the rank CCF in Eq. (4), the rank cross-correlation matrix can be established to characterize the temporal-spatial correlation of $\boldsymbol{X}(t)$.

$$\boldsymbol{\rho}(\tau)=\begin{pmatrix}\rho_{1,1}(\tau) & \cdots & \rho_{1,j}(\tau) & \cdots & \rho_{1,m}(\tau)\\ \vdots & & \vdots & & \vdots\\ \rho_{i,1}(\tau) & \cdots & \rho_{i,j}(\tau) & \cdots & \rho_{i,m}(\tau)\\ \vdots & & \vdots & & \vdots\\ \rho_{m,1}(\tau) & \cdots & \rho_{m,j}(\tau) & \cdots & \rho_{m,m}(\tau)\end{pmatrix} \tag{5}$$

3 Spectral Representation Method Based on Translation Model

In this section, a spectral representation method is proposed to match the empirical CDFs and rank cross-correlation matrix of $\boldsymbol{X}(t)$.

3.1 Translation model

Let $\boldsymbol{Z}(t)=(Z_1(t),Z_2(t),\ldots,Z_i(t),\ldots,Z_m(t))^{\mathrm{T}}$ be an m-variate Gaussian stochastic vector process, and $Z_i(t)$ $(i=1,2,\ldots,m)$ follow the standard normal distribution. Based on marginal transformation, the multivariate wind speed time series $\boldsymbol{X}(t)=(X_1(t),X_2(t),\ldots,X_i(t),\ldots,X_m(t))^{\mathrm{T}}$ can be simulated by

$$\begin{pmatrix}X_1(t)\\ \vdots\\ X_i(t)\\ \vdots\\ X_m(t)\end{pmatrix}\xleftarrow{X_i(t)=F_i^{-1}\{\Phi[Z_i(t)]\}}\begin{pmatrix}Z_1(t)\\ \vdots\\ Z_i(t)\\ \vdots\\ Z_m(t)\end{pmatrix} \tag{6}$$

where $F_i^{-1}(\cdot)$ is the inverse CDF of $X_i(t)$ $(i=1,2,\ldots,m)$. $\Phi(\cdot)$ is the CDF of $Z_i(t)$ $(i=1,2,\ldots,m)$.

Denote the linear cross-correlation matrix of $\boldsymbol{Z}(t)$ as

$$\boldsymbol{\rho}(\tau)=\begin{pmatrix}\rho_{z;1,1}(\tau) & \cdots & \rho_{z;1,j}(\tau) & \cdots & \rho_{z;1,m}(\tau)\\ \vdots & & \vdots & & \vdots\\ \rho_{z;i,1}(\tau) & \cdots & \rho_{z;i,j}(\tau) & \cdots & \rho_{z;i,m}(\tau)\\ \vdots & & \vdots & & \vdots\\ \rho_{z;m,1}(\tau) & \cdots & \rho_{z;m,j}(\tau) & \cdots & \rho_{z;m,m}(\tau)\end{pmatrix} \tag{7}$$

where $\rho_{z;i,i}(\tau)$ is the linear ACF of $Z_i(t)$ $(i=1,2,\ldots,m)$, and $\rho_{z;i,j}(\tau)$ is the linear CCF between $Z_i(t)$ and $Z_j(t)$ $(i,j=1,2,\ldots,m;\ i\neq j)$.

Because the rank correlation coefficient is invariant under the transformation $X_i(t)=F_i^{-1}\{\Phi[Z_i(t)]\}$ in Formula (6), the rank cross-correlation matrix of $\boldsymbol{Z}(t)$ would also be $\boldsymbol{\rho}(\tau)$ in Eq. (5). Then, the linear cross-correlation matrix of $\boldsymbol{Z}(t)$ can be calculated by

$$\boldsymbol{\rho}_z(\tau)=2\sin\left(\frac{\pi}{6}\boldsymbol{\rho}(\tau)\right) \tag{8}$$

3.2 Spectral representation method

The rank ACF $\rho_{i,i}(\tau)$ in Eq. (3) is an even function

$$\rho_{i,i}(\tau)=\rho_{i,i}(-\tau),\quad i=1,2,\ldots,m \tag{9}$$

The rank CCF in Eq. (4) has the following property:

$$\rho_{i,j}(\tau)=\rho_{j,i}(-\tau),\quad i,j=1,2,\ldots,m;\ i\neq j \tag{10}$$

According to Eqs. (9) and (10), it has

$$\begin{cases}\rho_{z;i,i}(\tau)=\rho_{z;i,i}(-\tau),\quad i=1,2,\ldots,m;\\ \rho_{z;i,j}(\tau)=\rho_{z;j,i}(-\tau),\quad i,j=1,2,\ldots,m;\ i\neq j\end{cases} \tag{11}$$

Denote $S_{Z;i,i}(\omega)$ as the power spectral density function (PSDF) of $Z_i(t)$, it has

$$S_{Z;i,i}(\omega)=\frac{1}{2\pi}\int_{-\infty}^{\infty}\rho_{z;i,i}(\tau)\cos(\omega\tau)\mathrm{d}\tau\approx\frac{\Delta\tau}{\pi}\sum_{s=0}^{\infty}\rho_{z;i,i}(s\cdot\Delta\tau)\cos(s\cdot\Delta\tau\omega) \tag{12}$$

Denote $S_{Z;i,j}(\omega)$ as the cross PSDF between $Z_i(t)$ and $Z_j(t)$, it has

$$S_{Z;i,j}(\omega)=\frac{1}{2\pi}\int_{-\infty}^{\infty}\rho_{z;i,j}(\tau)\mathrm{e}^{-I\cdot\omega\tau}\mathrm{d}\tau=\mathrm{Re}[S_{Z;i,j}(\omega)]-I\cdot\mathrm{Im}[S_{Z;i,j}(\omega)] \tag{13}$$

where I is the imaginary unit. $\mathrm{Re}[S_{Z;i,j}(\omega)]$ is the real part of $S_{Z;i,j}(\omega)$, and $\mathrm{Im}[S_{Z;i,j}(\omega)]$ is the imaginary part of $S_{Z;i,j}(\omega)$.

According to Eq. (11), it has

$$\rho_{z;i,j}(\tau)=\frac{\rho_{z;i,j}(\tau)+\rho_{z;i,j}(-\tau)}{2}+\frac{\rho_{z;i,j}(\tau)-\rho_{z;i,j}(-\tau)}{2}=\frac{\rho_{z;i,j}(\tau)+\rho_{z;j,i}(\tau)}{2}+\frac{\rho_{z;i,j}(\tau)-\rho_{z;j,i}(\tau)}{2} \tag{14}$$

where $\dfrac{\rho_{i,j}(\tau)+\rho_{j,i}(\tau)}{2}$ is an even function, and $\dfrac{\rho_{i,j}(\tau)-\rho_{j,i}(\tau)}{2}$ is an odd function. Then, $\mathrm{Re}[S_{Z;i,j}]$ and $\mathrm{Im}[S_{Z;i,j}]$ can be calculated by

$$\begin{aligned}\mathrm{Re}[S_{Z;i,j}] &= \frac{1}{\pi}\int_0^{\infty}\rho_{z;i,j}(\tau)\cos(\omega\tau)\mathrm{d}\tau \approx \\ &\frac{\Delta\tau}{\pi}\sum_{s=0}^{\infty}[\rho_{z;i,j}(s\cdot\Delta\tau)+\rho_{z;j,i}(s\cdot\Delta\tau)]\cos(s\cdot\Delta\tau\omega), \\ \mathrm{Im}[S_{Z;i,j}] &= \frac{1}{\pi}\int_0^{\infty}\rho_{z;i,j}(\tau)\sin(\omega\tau)\mathrm{d}\tau \approx \\ &\frac{\Delta\tau}{\pi}\sum_{s=0}^{\infty}[\rho_{z;i,j}(s\cdot\Delta\tau)-\rho_{z;j,i}(s\cdot\Delta\tau)]\sin(s\cdot\Delta\tau\omega)\end{aligned} \tag{15}$$

With the PSDF from Eq. (12) and the cross PSDF from Eq. (13), the cross spectral density matrix of $\boldsymbol{Z}(t)$ can be established

$$\boldsymbol{S}_{\boldsymbol{Z}}(\omega)=\begin{pmatrix} S_{Z;1,1}(\omega) & \cdots & S_{Z;1,j}(\omega) & \cdots & S_{Z;1,m}(\omega) \\ \vdots & & \vdots & & \vdots \\ S_{Z;i,1}(\omega) & \cdots & S_{Z;i,j}(\omega) & \cdots & S_{Z;i,m}(\omega) \\ \vdots & & \vdots & & \vdots \\ S_{Z;m,1}(\omega) & \cdots & S_{Z;m,j}(\omega) & \cdots & S_{Z;m,m}(\omega) \end{pmatrix} \tag{16}$$

then, $\boldsymbol{H}(\omega)$ can be obtained by performing Cholesky decomposition on $\boldsymbol{S}_{\boldsymbol{Z}}(\omega)$

$$\boldsymbol{S}_{\boldsymbol{Z}}(\omega)=\boldsymbol{H}(\omega)\cdot\boldsymbol{H}^{\mathrm{T}}(\omega) \tag{17}$$

where $\boldsymbol{H}(\omega)$ is a lower triangular matrix, and $\boldsymbol{H}^{\mathrm{T}}(\omega)$ is the conjugate transpose of $\boldsymbol{H}(\omega)$.

The equation for generating samples of $\boldsymbol{Z}(t)$ is[16]

$$Z_i(t)=2\sum_{j=1}^{i}\sum_{k=1}^{N}|H_{i,j}(\omega_{j,k})|\sqrt{\Delta\omega}\cos\left[\omega_{j,k}t-\varphi_{i,j}(\omega_{j,k})+\theta_{j,k}\right] \tag{18}$$

The symbols in Eq. (18) are explained in Table 1.

With a large value of N, $\boldsymbol{Z}(t)=(Z_1(t),Z_2(t),\ldots,Z_i(t),\ldots,Z_m(t))^{\mathrm{T}}$ from Eq. (18) would be an m-variate Gaussian vector process with a cross spectral density matrix $\boldsymbol{S}_{\boldsymbol{Z}}(\omega)$ in Eq. (16), the linear cross correlation matrix of $\boldsymbol{Z}(t)$ would be $\boldsymbol{\rho}_z(\tau)$ in Eq. (7), and the rank cross-correlation matrix of $\boldsymbol{Z}(t)$ would be $\boldsymbol{\rho}(\tau)$ in Eq. (5). Because $\boldsymbol{X}(t)$ and $\boldsymbol{Z}(t)$ in Formula (6) share the same rank cross-correlation matrix, along with Eq. (16), Formula (6) allows for matching CDFs and the rank cross-correlation matrix of $\boldsymbol{X}(t)$.

Table 1 Meaning of symbols in Eq. (18).

Symbol	Definition
ω_u	An upper cutoff frequency, for $\mid\omega\mid>\omega_u$, it has $S_{Z;i,j}(\omega)\approx 0$, $(i,j=1,2,\ldots,m)$
$\Delta\omega$	$\Delta\omega=\dfrac{\omega_u}{N}$
$\omega_{j,k}$	$(j-1)\cdot\Delta\omega+\dfrac{j}{m}\Delta\omega,\quad k=1,2,\ldots,N$
t	Time instant: $t=s\cdot\Delta\tau$
$\theta_{j,k}$	Uniform random variable over $[0,2\pi]$
$\mid H_{i,j}(\omega_{j,k})\mid$	$\sqrt{\left(\mathrm{Re}[H_{i,j}(\omega_{j,k})]\right)^2+\left(\mathrm{Im}[H_{i,j}(\omega_{j,k})]\right)^2}$
$\varphi_{i,j}(\omega_{j,k})$	$\varphi_{i,j}(\omega_{j,k})=\tan^{-1}\left(\dfrac{\mathrm{Im}[H_{i,j}(\omega_{j,k})]}{\mathrm{Re}[H_{i,j}(\omega_{j,k})]}\right)$

In the case where only wind speed samples are available, the marginal transformation in Formula (6) can be approximated by the following five-parameter Johnson distribution[17]

$$X_i(t)=F_i^{-1}\{\Phi[Z_i(t)]\}\approx\xi+\lambda\cdot\left[1+\exp\left(\frac{\gamma-Z_i(t)}{\delta}\right)\right]^{-\alpha} \tag{19}$$

where ξ, λ, γ, δ, and α are parameters, which can be determined by the algorithm in Ref. [17].

3.3 Procedure

This section presents the detailed procedures of the spectral representation method for simulating temporally and spatially correlated wind speed time series.

(1) With the five-parameter Johnson distribution in Eq. (19), fit distributions to wind speed time series at each site[17].

(2) With Eq. (2), calculate the rank ACFs and rank CCFs by Eqs. (3) and (4), and construct the rank cross-correlation matrix $\boldsymbol{\rho}(\tau)$ in Eq. (5) for wind speed time series at multiple sites.

(3) Determine the linear cross-correlation matrix $\boldsymbol{\rho}_z(\tau)$ by Eq. (8), and construct the cross spectral density matrix $\boldsymbol{S}_{\boldsymbol{Z}}(\omega)$ in Eq. (16).

(4) Generate samples of $\boldsymbol{Z}(t)$ by Eqs. (17) and (18), and transform samples of $\boldsymbol{Z}(t)$ into samples of wind speed time series by Formula (6) and Eq. (19).

These procedures have also been presented in Fig. 1.

4 Case Study

4.1 Recovering distribution functions of wind samples

In this section, a case study is performed based on wind data collected from three different sites[18], and the five-parameter Johnson distribution in Eq. (19) is employed to fit distributions to wind speed at each site.

Rewrite Eq. (19) as

$$Z_i=\Omega_i(X_i)=\gamma+\delta\cdot\ln\left[\frac{(X-\xi)^{\frac{1}{\alpha}}}{\lambda^{\frac{1}{\alpha}}-(X-\xi)^{\frac{1}{\alpha}}}\right] \tag{20}$$

Then, the CDF of wind speed at the i-th site is

$$F_i(X_i)=\Phi[\Omega_i(X_i)],\quad \xi\leqslant X_i\leqslant(\xi+\lambda) \tag{21}$$

The PDF of wind speed is

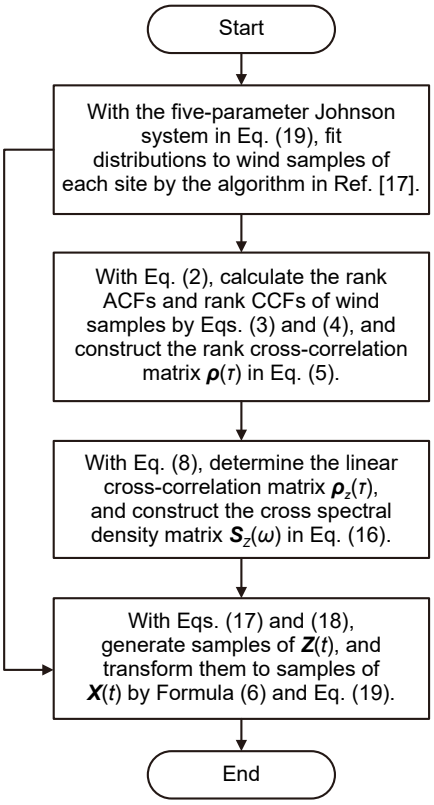


Fig. 1 Procedure of the proposed methods.

$$f_i(X_i)=\frac{\mathrm{d}Z_i}{\mathrm{d}X_i}\cdot\phi[\Omega_i(X_i)],\quad \phi(t)=\frac{1}{\sqrt{2\pi}}\mathrm{e}^{-\frac{t^2}{2}} \tag{22}$$

Here, the PDF and CDF of wind speeds can also be fitted by the four-parameter Johnson distribution in Ref. [19].

$$X_i=\xi+\lambda-\frac{\lambda}{1+\exp\left(\frac{Z_i-\gamma}{\delta}\right)} \tag{23}$$

The Johnson distribution in Eq. (23) would be denoted as "Johnson-IV"; the five-parameter Johnson distribution in Eq. (19) would be denoted as "Johnson-V". Table 2 presents the parameters of Johnson distributions.

According to Johnson distributions in Eqs. (19) and (23), the variation range of wind speed would be $X_i\in[\xi,\ (\xi+\lambda)]$ $(i=1,2,3)$. Figure 2 depicts the PDFs of wind speeds at Site 1, Site 2, and Site 3. Besides, Figs. 3a, 3c, and 3e show the CDFs of wind speeds; Figs. 3b, 3d, and 3f present the absolute errors between the empirical CDF from Eq. (1) and those from Johnson distributions. As can be seen, the Johnson-V yields a better fitting than Johnson-IV and provides a more accurate matching for the variation range of wind speeds.

Table 2 Parameters of Johnson distributions.

Distribution	Site No.	ξ	λ	γ	δ	α
Johnson-IV	1	−3.7557	112.6225	3.7190	1.7278	–
	2	−1.6681	109.3165	4.4072	1.8190	–
	3	−4.6994	176.1336	6.1299	2.5057	–
Johnson-V	1	0	37.0000	1.5332	0.8426	0.8192
	2	0	37.0000	1.9000	1.1073	0.8812
	3	0	40.0000	1.9710	1.1442	0.7914

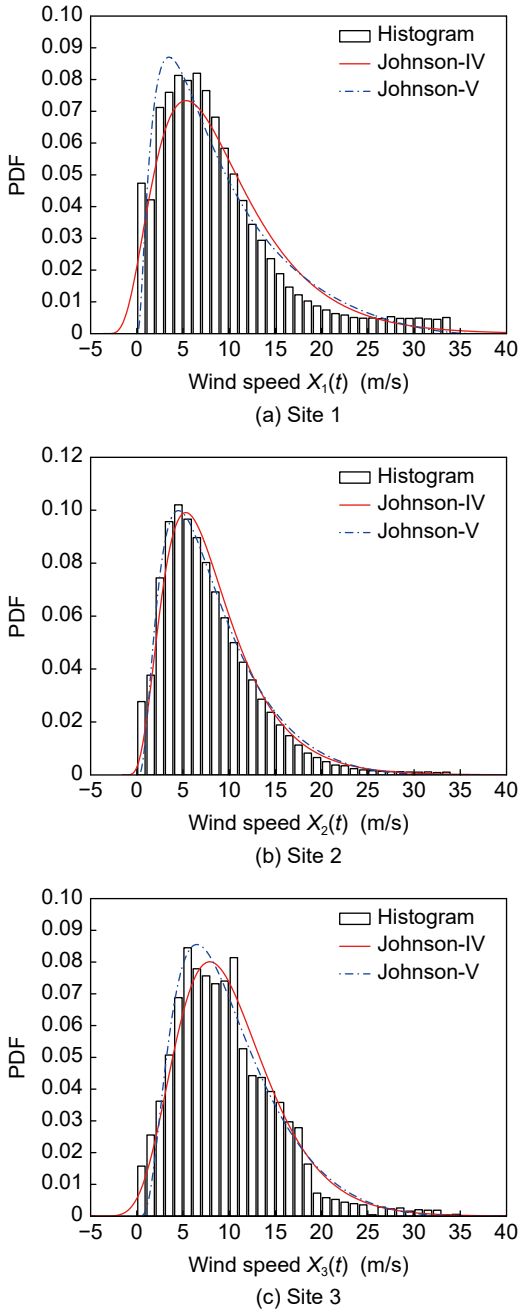


Fig. 2 PDFs of $X_i(t)$ (i = 1, 2, 3).

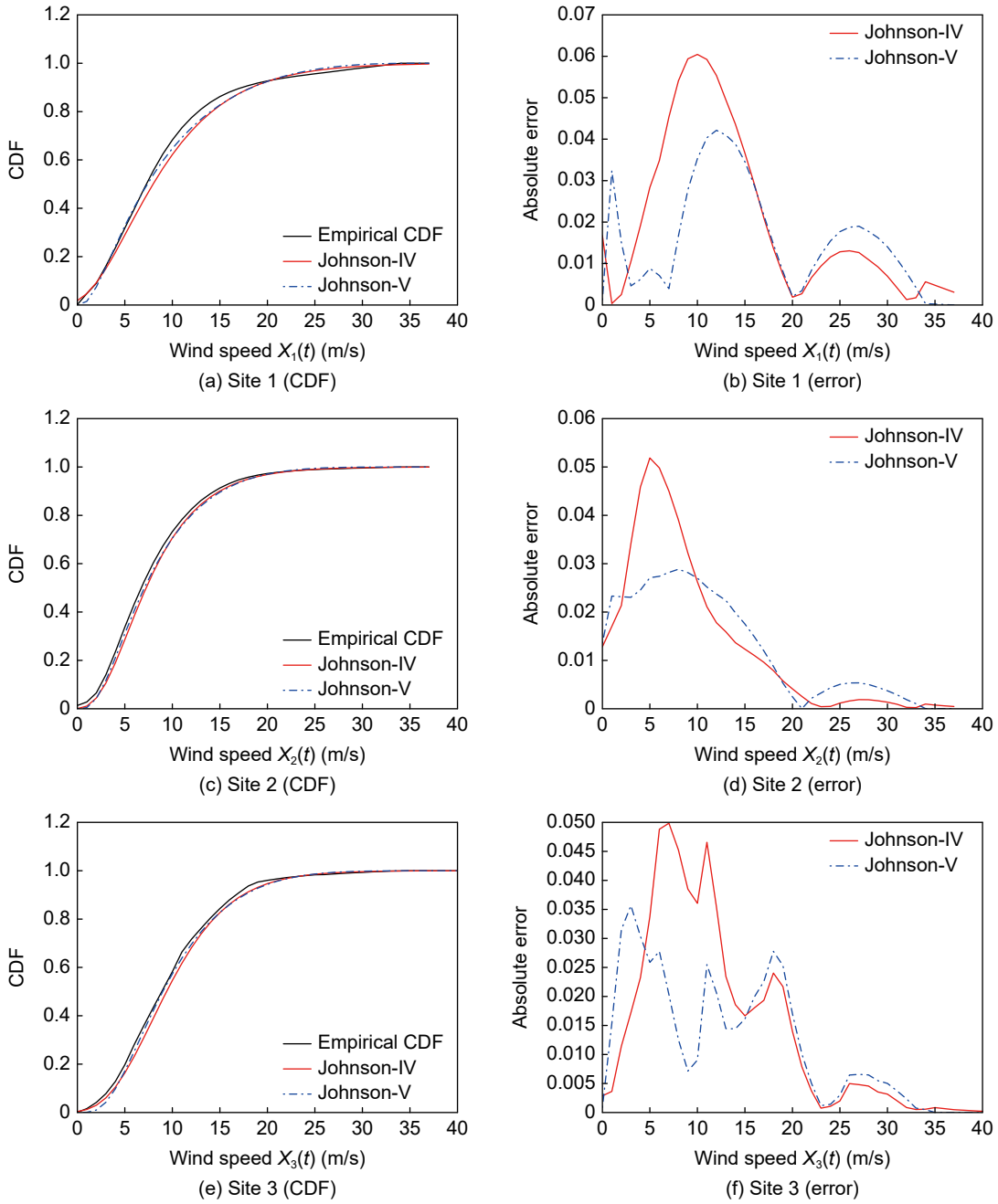


Fig. 3 CDFs and absolute error of $X_i(t)$ (i = 1, 2, 3).

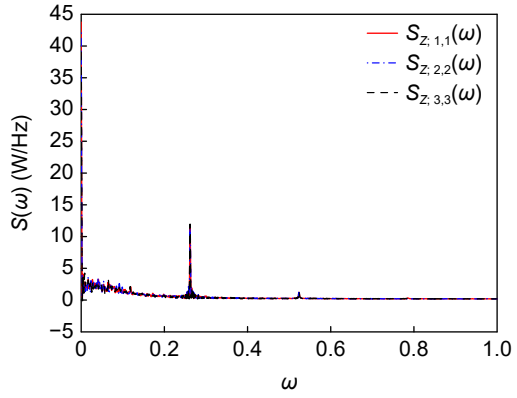
4.2 Cross spectral density matrix of wind speed samples

According to Formula (6), $\boldsymbol{X}(t)$ can be modeled by a Gaussian stochastic process $\boldsymbol{Z}(t)$, and the temporal-spatial correlation of $\boldsymbol{X}(t)$ is represented by the cross spectral density matrix of $\boldsymbol{Z}(t)$.

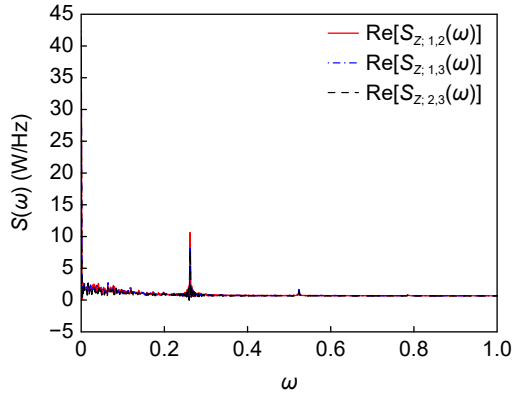
Following Eqs. (3)–(5), the rank cross-correlation matrix of wind speeds is established. Then, the PSDFs and cross PSDFs of $\boldsymbol{Z}(t)$ can be obtained by Eqs. (8), (12), and (13). With Eqs. (11) and (15), it can be seen that

$$\begin{aligned} &S_{Z;i,i}(\omega) = S_{Z;i,i}(-\omega),\\ &\mathrm{Re}[S_{Z;i,j}(\omega)] = \mathrm{Re}[S_{Z;i,j}(-\omega)],\\ &\mathrm{Im}[S_{Z;i,j}(\omega)] = -\mathrm{Im}[S_{Z;i,j}(-\omega)],\\ &\mathrm{Re}[S_{Z;i,j}(\omega)] = \mathrm{Re}[S_{Z;j,i}(\omega)],\\ &\mathrm{Im}[S_{Z;i,j}(\omega)] = -\mathrm{Im}[S_{Z;j,i}(\omega)] \end{aligned} \tag{24}$$

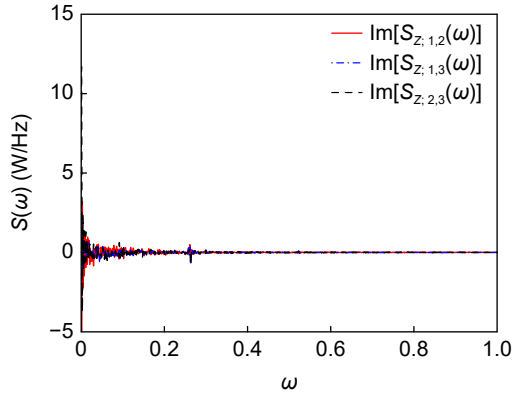
In Fig. 4, it shows the PSDFs, the real parts of cross PSDFs, and the imaginary parts of cross PSDFs. As can be seen, values of $S_{Z;i,j}(\omega)$ $(i, j = 1, 2, 3)$ approach 0 for $\omega \geqslant 1$. Thus, ω_u in Table 1 is set to be 1, and



(a) PSDFs of wind speeds



(b) Real parts of cross PSDFs



(c) Imaginary parts of cross PSDFs

Fig. 4 PSDFs and cross PSDFs of $\boldsymbol{Z}(t)$.

$N = 1000$. According to Eq. (17), Cholesky decomposition of $\boldsymbol{S}_Z(\omega)$ should be performed. It is possible that $\boldsymbol{S}_Z(\omega)$ may not be positive semi-definite, which can be handled by the method in the Appendix.

4.3 Sample generation

The parameters in Eq. (18) are as follows:

$$\omega_u = 1, \quad N = 2000.$$

Because samples of $\boldsymbol{Z}(t)$ from Eq. (18) are stationary and ergodic, samples of $X_i(t)$ $(i = 1,2,3)$ from Eq. (19) are also expected to be stationary; that is to say, the CDF and rank ACF of $X_i(t)$ would not change over time $(i = 1,2,3)$.

Here, it generates 1.5×10^5 samples, the following two sets of wind speed samples are considered:

Set-I: It comprises of the first 10^5 samples;

Set-II: It comprises of the $(5\times10^4+1)$-th to (1.5×10^5)-th samples.

Figures 5a, 5c, and 5e depict the empirical CDFs of the original wind speed samples from Ref. [18], Set-I and Set-II; Figs. 5b, 5d, and 5f present the absolute errors between these empirical CDFs. Figures 6a, 6c, and 6e show the rank ACFs of these three sets of wind speed samples; Figs. 6b, 6d, and 6f present the absolute errors between these rank ACFs. An inspection of Figs. 5 and 6 indicates that:

(1) The CDFs and rank ACFs of samples from Eqs. (18) and (19) are in accordance with those of the original wind speed samples;

(2) The CDFs and rank ACFs of Set-I and Set-II are in close agreement, and samples from Eqs. (18) and (19) are stationary and ergodic.

In this paper, the rank CCF in Eq. (4) is introduced to measure the spatial correlation of $\boldsymbol{X}(t)$. With Eq. (10), it can be seen that

$$\rho_{1,2}(\tau) = \rho_{2,1}(-\tau),$$
$$\rho_{1,3}(\tau) = \rho_{3,1}(-\tau),$$
$$\rho_{2,3}(\tau) = \rho_{3,2}(-\tau).$$

Thus, it just requires to check how well $\rho_{1,2}(\tau)$, $\rho_{1,3}(\tau)$ and $\rho_{2,3}(\tau)$ can match the expected CCFs. Figure 7 shows the rank CCFs of generated samples and their absolute errors with respect to those of original wind samples. The results illustrate the effectiveness of the proposed method for reproducing the spatial correlation of multivariate wind speed time series.

4.4 Discussion

Although this paper aims to match the rank ACFs and CCFs of wind samples, the spectral representation method also allows for reproducing the linear ACFs and CCFs of target wind speed samples. The procedures are as follows.

(1) Fit distributions to wind speed samples at each site by the five-parameter Johnson distribution;

(2) Generate samples of $\boldsymbol{X}(t)$ by Eq. (20), and calculate the linear ACFs and CCFs of $\boldsymbol{Z}(t)$;

(3) Follow Steps (3) and (4) in Section 3.3 to produce sample realizations of wind speed time series.

But this approach cannot be used to generate wind

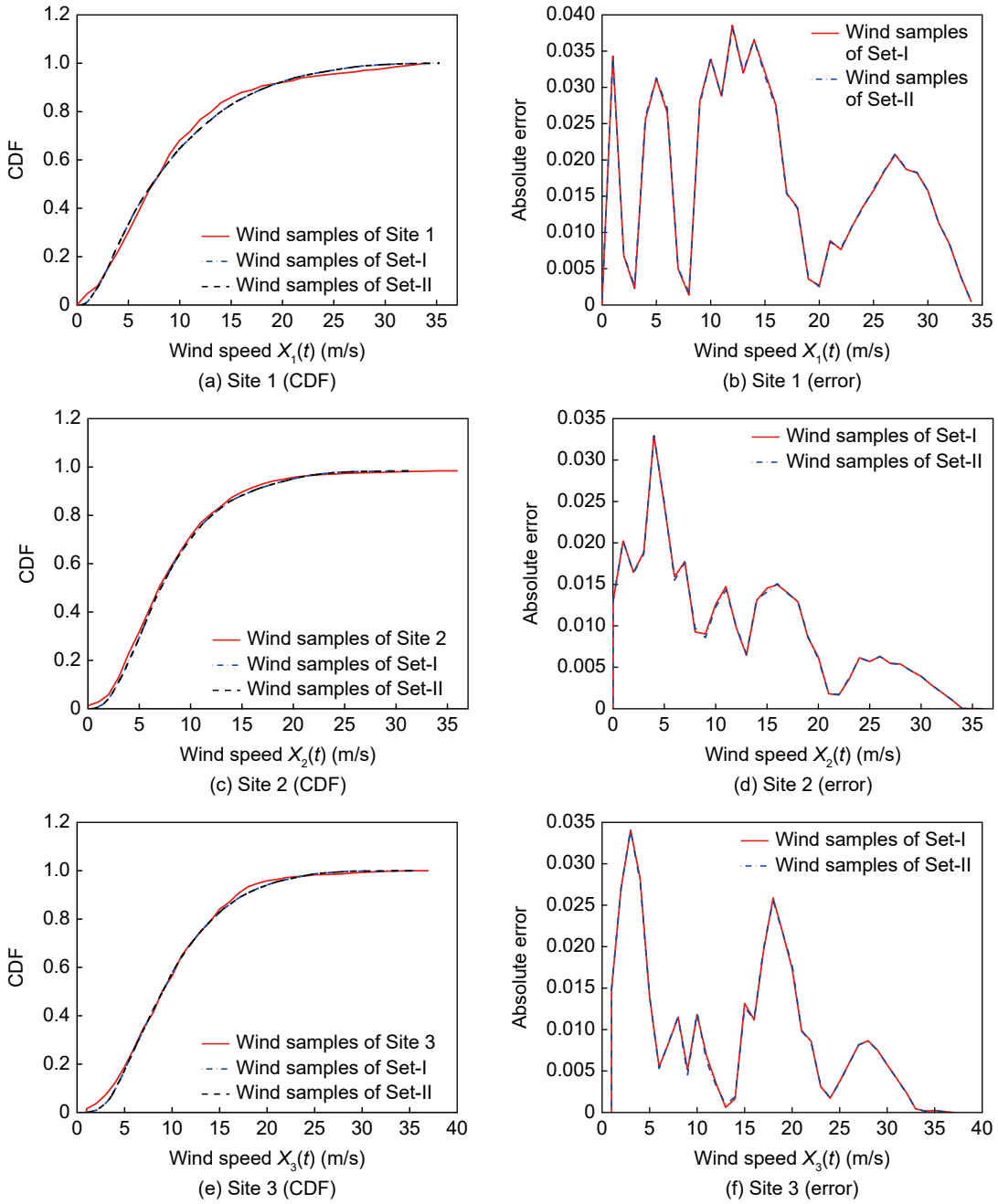


Fig. 5 Comparison for empirical CDFs and absolute error of samples of $X_i(t)$ (i = 1, 2, 3).

speed samples with a prescribed linear cross-correlation matrix.

If wind speed time series are not stationary, statistical properties of wind speed would vary in time. In this case, the marginal transformation in Formula (6) should take the time instant t into consideration

$$X_i(t) = F_i^{-1}\{\Phi[Z_i(t)]; t\} \tag{25}$$

As for the temporal-spatial correlation of wind speeds, they can be represented by the evolutionary spectrum[20], and samples of non-stationary wind speed time series can be generated by the evolutionary spectrum representation method[21].

5 Conclusion

In conjunction with the five-parameter Johnson system, this paper presents a spectral representation method to capture statistical features of multivariate wind speed time series. Through theoretical discussions and numerical experiments, the following conclusions can be drawn.

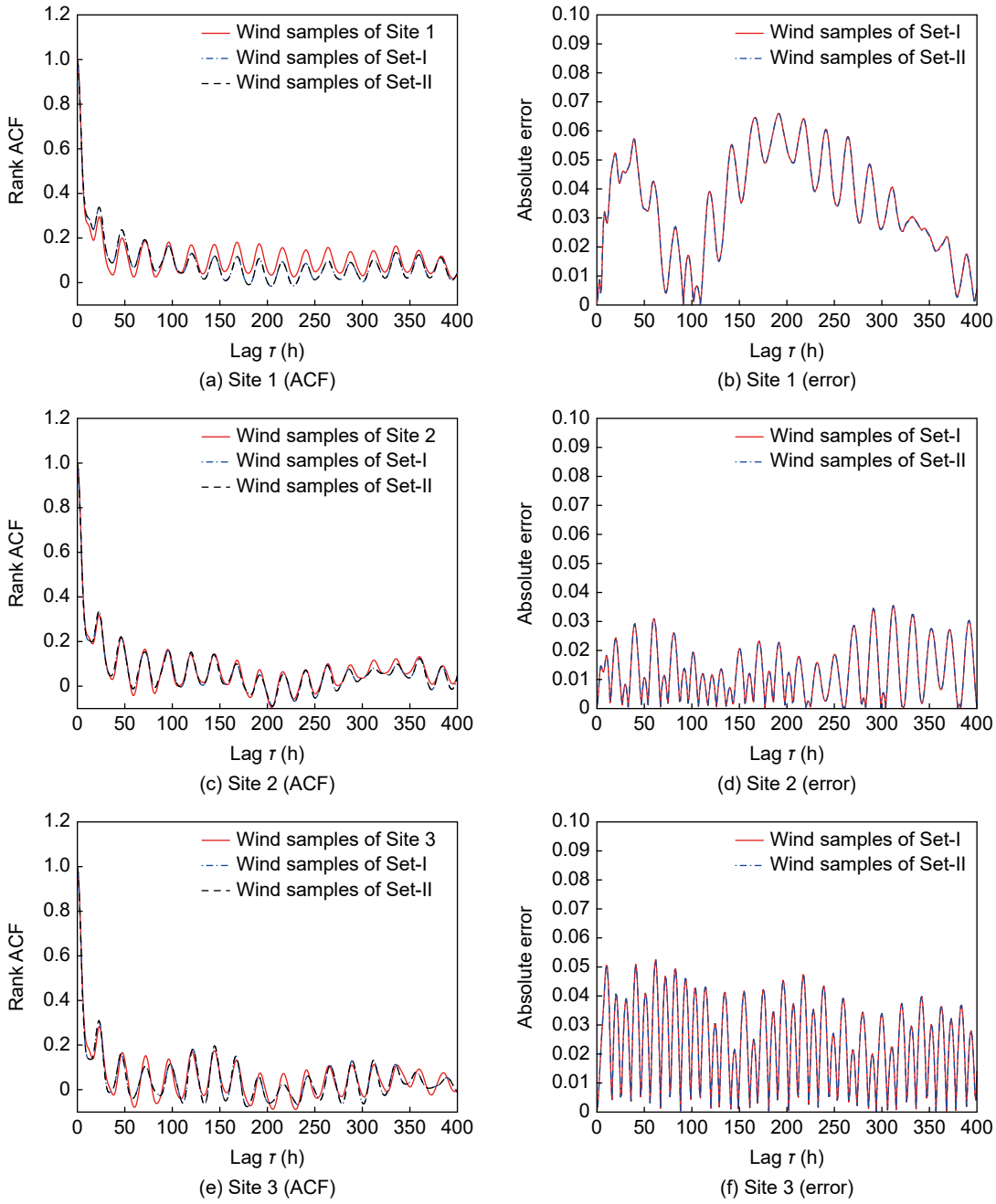


Fig. 6 Comparison for rank ACFs and absolute error of $X_i(t)$ (i = 1, 2, 3).

(1) The distribution functions of wind speeds can be well fitted by the five-parameter Johnson distribution, and the lower and upper bounds of wind velocity are strictly matched by the five-parameter Johnson distribution.

(2) If the translation model is employed to simulate multivariate wind speed time series, the rank ACF and rank CCF serve useful tools to measure the temporal-spatial correlation of wind speeds at multiple sites, as well as to facilitate the simulation process

(3) The case study verifies the effectiveness of spectral representation method, and the PSDF and cross PSDF can be used to reproduce the rank temporal-spatial correlation of wind speeds at multiple sites.

Appendix

A Remedying cross spectral density matrix

Denote the eigenvalues of $\boldsymbol{S}_{\boldsymbol{Z}}(\omega)$ in Eq. (16) as $\{\lambda_1,\lambda_2,\ldots,\lambda_i,\ldots,\lambda_m\}$. Because $\boldsymbol{S}_Z(\omega)$ is an Hermite matrix, if $\boldsymbol{S}_Z(\omega)$ is not semi-definite positive, at least one eigenvalue of $\boldsymbol{S}_Z(\omega)$ would be negative. Denote

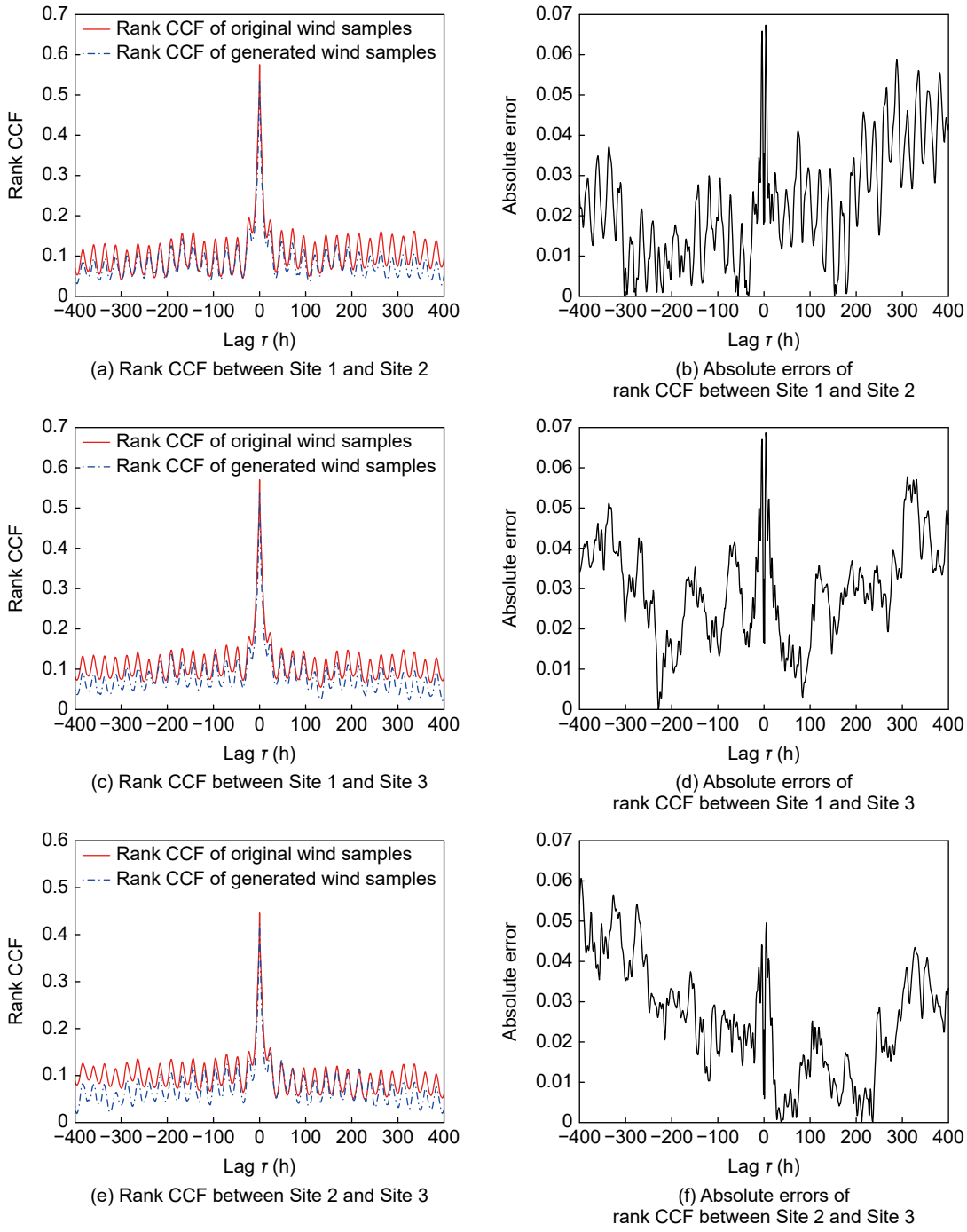


Fig. 7 Rank CCFs and their absolute errors of wind speed samples.

$$\widehat{\lambda} = \min\{\lambda_1, \lambda_2, \ldots, \lambda_i, \ldots, \lambda_m\} \tag{A1}$$

Note that an Hermite matrix is positive semi-definite if and only if all eigenvalues are not negative, then $\boldsymbol{S}_Z(\omega)$ can be remedies by

$$\widehat{\boldsymbol{S}}_Z(\omega) = \frac{\boldsymbol{S}_Z(\omega) - \mathrm{diag}(\widehat{\lambda}, \ldots, \widehat{\lambda}, \ldots, \widehat{\lambda})}{1 - \widehat{\lambda}} \tag{A2}$$

where $\mathrm{diag}(\widehat{\lambda}, \ldots, \widehat{\lambda}, \ldots, \widehat{\lambda})$ is a diagonal matrix of size $m \times m$.

B Nomenclature

CDF: Cumulative distribution function;
PDF: Probability distribution function;
ACF: Autocorrelation function;
CCF: Cross-correlation function;
PSDF: Power spectral density function;
$\boldsymbol{X}(t)$: Wind speed time series;
$X(t_1/t_2)$:Wind velocity at the time instant t_1/t_2;
$\boldsymbol{X}(t)$: Wind speed time series at multiple sites;

$X_i(t)$: Wind speed time series at the i-th site;
$Z_i(t)$: Gaussian stochastic process;
$\boldsymbol{Z}(t)$: Gaussian stochastic vector process;
$\widehat{F}_i(\cdot)$: Empirical CDF of $X_i(t)$;
$F_i^{-1}(\cdot)$: Inverse CDF of $X_i(t)$;
$\Phi(\cdot)$: CDF of standard normal variable;
$\rho_{i,i}(\tau)$: Rank ACF of $X_i(t)$;
$\rho_{i,j}(\tau)$: Rank CCF between $X_i(t)$ and $X_j(t)$;
$\boldsymbol{\rho}(\tau)$: Rank cross-correlation matrix of $\boldsymbol{X}(t)$;
$\rho_{z;i,i}(\tau)$: Linear ACF of $Z_i(t)$;
$\rho_{z;i,j}(\tau)$: Linear CCF between $Z_i(t)$ and $Z_j(t)$;
$\boldsymbol{\rho}_{\boldsymbol{z}}(\tau)$: Linear cross-correlation matrix of $\boldsymbol{Z}(t)$;
$S_{Z;i,i}(\omega)$: PSDF of $Z_i(t)$;
$S_{Z;i,j}(\omega)$: Cross PSDF between $Z_i(t)$ and $Z_j(t)$;
$\boldsymbol{S}_{\boldsymbol{Z}}(\omega)$: Cross spectral density matrix of $\boldsymbol{Z}(t)$;
$\boldsymbol{H}(\omega)$: The lower triangular matrix.

Acknowledgment

This work was supported by the National Natural Science Foundation of China (No. 12271155), Doctoral Research Start-Up Fund of Hunan University of Science and Technology (No. E52170), and Hunan Science and Technology Talent Promotion Project (No. 2020TJ-N08).

References

[1] T. Long and Q. -S. Jia, Matching uncertain renewable supply with electric vehicle charging demand—A bi-level event-based optimization method, *Complex System Modeling and Simulation*, vol. 1, no. 1, pp. 33–44, 2021.

[2] Y. Shen, L. Yu, and J. Li, Robust electric vehicle routing problem with time windows under demand uncertainty and weight-related energy consumption, *Complex System Modeling and Simulation*, vol. 2, no. 1, pp. 18–34, 2022.

[3] C. Wang, Z. Liu, H. Wei, L. Chen, and H. Zhang, Hybrid deep learning model for short-term wind speed forecasting based on time series decomposition and gated recurrent unit, *Complex System Modeling and Simulation*, vol. 1, no. 4, pp. 308–321, 2021.

[4] D. S. Callaway, Sequential reliability forecasting for wind energy: Temperature dependence and probability distributions, *IEEE Transactions on Energy Conversion*, vol. 25, no. 2, pp. 577–585, 2010.

[5] K. Xie, Q. Liao, H. Tai, and B. Hu, Non-homogeneous Markov wind speed time series model considering daily and seasonal variation characteristics, *IEEE Transactions on Sustainable Energy*, vol. 8, no. 3, pp. 1281–1290, 2017.

[6] S. Miao, K. Xie, H. Yang, H. -M. Tai, and B. Hu, A Markovian wind farm generation model and its application to adequacy assessment, *Renewable Energy*, vol. 113, pp. 1447–1461, 2017.

[7] R. Carapellucci and L. Giordano, A new approach for synthetically generating wind speeds: A comparison with the Markov chains method, *Energy*, vol. 49, pp. 298–305, 2013.

[8] J. Tang, A. Brouste, and K. L. Tsui, Some improvements of wind speed Markov chain modeling, *Renewable Energy*, vol. 81, pp. 52–56, 2015.

[9] G. M. Jónsdóttir and F. Milano, Data-based continuous wind speed models with arbitrary probability distribution and autocorrelation, *Renewable Energy*, vol. 143, pp. 368–376, 2019.

[10] J. M. Morales, R. Minguez, and A. J. Conejo, A methodology to generate statistically dependent wind speed scenarios, *Applied Energy*, vol. 87, no. 3, pp. 843–855, 2010.

[11] P. Chen, P. Siano, B. Bak-Jensen, and Z. Chen, Stochastic optimization of wind turbine power factor using stochastic model of wind power, *IEEE Transactions on Sustainable Energy*, vol. 1, no. 1, pp. 19–29, 2010.

[12] D. C. Hill, D. McMillan, K. R. Bell, and D. Infield, Application of auto-regressive models to UK wind speed data for power system impact studies, *IEEE Transactions on Sustainable Energy*, vol. 3, no. 1, pp. 134–141, 2011.

[13] Y. Li, K. Xie, and B. Hu, Copula-ARMA model for multivariate wind speed and its applications in reliability assessment of generating systems, *Journal of Electrical Engineering and Technology*, vol. 8, no. 3, pp. 421–427, 2013.

[14] K. Yunus, P. Chen, and T. Thiringer, Modelling spatially and temporally correlated wind speed time series over a large geographical area using VARMA, *IET Renewable Power Generation*, vol. 11, no. 1, pp. 132–142, 2016.

[15] D. D. Le, G. Gross, and A. Berizzi, Probabilistic modeling of multisite wind farm production for scenario-based applications, *IEEE Transactions on Sustainable Energy*, vol. 6, no. 3, pp. 748–758, 2015.

[16] G. Deodatis, Simulation of ergodic multivariate stochastic processes, *Journal of Engineering Mechanics*, vol. 122, no. 8, pp. 778–787, 1996.

[17] J. Bacon-Shone, Fitting five parameter Johnson SB curves by moments, *Journal of the Royal Statistical Society Series C*: *Applied Statistics*, vol. 34, no. 1, pp. 95–100, 1985.

[18] Iowa State University, http://mesonet.agron.iastate.edu/request/awos/1min.php, 2021.

[19] I. D. Hill, R. Hill, and R. L. Holder, Fitting Johnson curves by moments, *Journal of the Royal Statistical Society Series C*: *Applied Statistics*, vol. 25, no. 2, pp. 180–189, 1976.

[20] B. A. Benowitz, M. Shields, and G. Deodatis, Determining evolutionary spectra from non-stationary autocorrelation functions, *Probabilistic Engineering Mechanics*, vol. 41, pp. 73–88, 2015.

[21] M. D. Shields and G. Deodatis, Estimation of evolutionary spectra for simulation of non-stationary and non-Gaussian stochastic processes, *Computers & Structures* , vol. 126, pp. 149–163, 2013.

Qing Xiao received the PhD degree from Hunan University of Science and Technolgoy, Xiangtan, China in 2022. He is currently a lecturer at the School of Information and Electrical Engineering, Hunan University of Science and Technology, China. His research interests include statistical modeling in power systems, probabilistic power flow computation, and probabilistic optimal power flow computation.

Lianghong Wu is currently a professor with the School of Information and Electrical Engineering, Hunan University of Science and Technology. His research interests include artificial intelligence, evolutionary computation, and applications.

Xiaowen Wu received the PhD degree from Wuhan University, China in 2013. He is currently an associate professor at the School of Information and Electrical Engineering, Hunan University of Science and Technology, China. His research interests include state monitoring and fault diagnosis of power apparatus and noise and vibration control of electric power facilities.

Matthias Rätsch is currently a professor of computer vision and intelligent robots at the Image Understanding and Interactive Robotics Group, Reutlingen University, Reutlingen, Germany. His research interests are in the fields of artificial intelligence and machine learning.