

DOI: 10.7511/jslx20230415002

基于 Copula 函数和风速正交分量的风矢量联合概率密度研究

程正兵, 冀晓文*

(北京工业大学 城市与工程安全减灾教育部重点实验室, 北京 100124)

摘要: 准确地表示风矢量联合概率密度对于风能的评估和风机结构设计具有重要意义, 本文将围绕提升风矢量联合概率密度表示结果准确性展开研究。首先, 将风矢量表示为风速正交分量, 采用混合 Normal 分布分别对边缘概率密度进行表示, 并提出了估计混合 Normal 分布参数的方法。然后, 基于 Copula 函数考虑风速正交分量间的相关性, 从而得到二者的联合概率密度, 其中 Copula 参数通过最小二乘法估计, 紧接着通过雅可比变换得到风矢量联合概率密度。最后, 采用印度风能研究所的测点 S3 和 S7 的每小时平均风速和风向数据, 与以往基于 Copula 函数和风速、风向变量得到的风矢量联合概率密度结果进行对比。结果表明, 与以往方法相比, 所提方法得到的风矢量联合概率密度结果准确性有很大的提升, 本文可为风能的开采和利用提供理论依据。

关键词: Copula 函数; 正交分解; 风矢量; 联合概率密度; 混合 Normal

中图分类号: TK81; O35

文献标志码: A

文章编号: 1007-4708(2024)05-0942-06

1 引言

风能是一种清洁的和可再生的能源, 充分获取风速的概率信息对于评估风能^[1]和风机结构设计^[2]十分重要。早在 20 世纪 40 年代, 风速概率模型就曾用于风机的选址^[3]。为了对风速进行准确描述, 历经几十年发展, 众多学者围绕模型的选取、参数估计方法和拟合优度评价^[4]进行了深入研究。近些年, 随着低碳发展成为国际潮流, 各国对清洁能源尤其是风能的需求与日俱增^[5]。

许多常见的分布如 Normal、Rayleigh 和 Weibull 分布等^[6]常用于风速的表示。风速的概率密度亦会表现为两个及以上成分组合而成的特征^[7], 故而组合模型如 GEV-Weibull^[6]和混合 Weibull 等^[7]用于风速的表示。另一方面, 风向也是风的另一项重要特征, 为了对风向进行表示, Coles 等^[8]提出了简谐函数, 而表现更为突出的是混合 von Mises 分布^[9]。

由于风速和风向是不可分割的两个量, 故而应当建立风矢量联合概率密度, 为了考虑两者的相关

性, Carta 等^[10]采用 Weibull-Normal 分布表示风速, 混合 von Mises 分布^[9]表示风向, 最后采用角度-线性模型(AL 模型)^[11]得到了风矢量联合概率密度。为了便于应用风矢量联合分布估计极值风速, Li 等^[12]提出基于 Copula 函数对风速和风向之间的相关性进行考虑, 然而, 该方法得到的风矢量联合概率密度结果与 AL 模型相比并未有过多提高。

此外, 基于风矢量数据, 另一种表示风矢量联合概率密度的做法是将风速进行正交分解^[13], 从而得到两个同类型的风速正交分量。基于风速正交分量, McWilliam 等^[14]通过假设风速正交分量均服从尺度参数相等的 Normal 分布, 且两者相互独立, 从而提出了 Isotropic Normal 模型, 在此基础上, Weber^[15]认为风速正交分量应服从尺度参数不等的 Normal 分布, 从而得到了 Anisotropic Normal 模型, 而 Erdem 等^[16]则假设风速正交分量服从 Lognormal 分布, 进而得到了 Anisotropic Lognormal 模型。

通过分析 Isotropic Normal 模型、Anisotropic

收稿日期: 2023-04-15; 修改稿收到日期: 2023-05-25.

基金项目: 国家自然科学基金(51908014; 52278135)资助项目.

作者简介: 冀晓文* (1990-), 男, 博士, 副教授(E-mail: jixiaowen900308@gmail.com).

引用本文: 程正兵, 冀晓文. 基于 Copula 函数和风速正交分量的风矢量联合概率密度研究[J]. 计算力学学报, 2024, 41(5): 942-947.
CHENG Zheng-bing, JI Xiao-wen. Research on joint probability density of wind vector based on Copula function and wind speed orthogonal components[J]. Chinese Journal of Computational Mechanics, 2024, 41(5): 942-947.

Normal 模型和 Anisotropic Lognormal 模型的共同点可知,三者均忽略了风速正交分量间的相关性。针对该问题,本文将基于 Copula 函数,开展对风矢量联合概率密度的研究。首先,采用混合 Normal 分布分别对风速正交分量的边缘概率密度进行表示,并介绍估计参数的方法;然后借助 Copula 函数建立两者的联合概率密度,进而通过雅可比变换得到风矢量联合概率密度;之后,以印度风能研究所的每小时平均风速数据为例,展示所提方法的效果,并与以往基于 Copula 函数和风速、风向变量得到的风矢量联合概率密度结果进行对比。

2 Sklar 理论

Sklar 理论是 Copula 函数^[17]核心,其基本思想为通过 Copula 函数建立二元变量 X_1 和 X_2 的联合概率密度函数(PDF)为

$$f(X_1, X_2) = c(u_i, u_j; \alpha) f(X_1) f(X_2) \quad (1)$$

式中 $c(\cdot, \cdot; \cdot)$ 表示 Copula 函数的 PDF, $u_i = F(X_1)$, $u_j = F(X_2)$, $F(X_1)$ ($F(X_2)$) 表示 X_1 (X_2) 的边缘分布函数(CDF), $f(X_1)$ ($f(X_2)$) 为 X_1 (X_2) 对应的边缘 PDF, α 为 Copula 参数。四种常见的 Copula 函数的 PDF 及参数取值范围列入表 1^[18]。

表 1 Copula 函数的 PDF 及参数范围

Tab. 1 PDF of Copula function and range of parameter

Copulas	$c(\cdot, \cdot; \cdot)$	α
Normal	$\varphi_2(\hat{u}_i, \hat{u}_j; \alpha) / (\varphi_1(\hat{u}_i) \varphi_1(\hat{u}_j))$	$(-1, 1)$
AMH	$[1 - 3\alpha + \alpha(1+u_i)(1+u_j) + \alpha^2(1-u_i) \times (1-u_j)] / [1 - \alpha(1-u_i)(1-u_j)]^3$	$[-1, 1)$
Frank	$[\alpha(1-e^{-\alpha})e^{-\alpha(u_i+u_j)}] / [(1-e^{-\alpha u_i}) \times (1-e^{-\alpha u_j}) - (1-e^{-\alpha})^2]$	$(-\infty, 0) \cup (0, +\infty)$
FGM	$1 + \alpha(1-2u_i)(1-2u_j)$	$[-1, 1]$

注: $\hat{u}_i = \Phi^{-1}(u_i)$, $\hat{u}_j = \Phi^{-1}(u_j)$, $\Phi(\varphi)$ 为标准 Normal 的 CDF(PDF)。

3 风矢量联合概率密度

3.1 以往的方法

假设用 V 表示风速, θ 表示风向,则考虑风速和风向间相关性的风矢量联合 PDF 为

$$f(V, \theta) = c(F(V), F(\theta); \alpha) f(V) f(\theta) \quad (2)$$

式中 $f(V)$ 和 $f(\theta)$ 分别为风速和风向的边缘 PDF, $F(V)$ 和 $F(\theta)$ 分别为相应的边缘 CDF。

在风能领域,风速数据通常取每小时平均风速,此时可采用 Weibull 分布^[6]表示风速边缘 PDF 为

$$f(V) = \frac{h}{\sigma} \left(\frac{V}{\sigma}\right)^{h-1} \exp\left[-\left(\frac{V}{\sigma}\right)^h\right] \quad (3)$$

式中 h 为形状参数, σ 为尺度参数。

对于风向,混合 von Mises 分布^[9]常用于表示风向边缘 PDF 为

$$f(\theta) = \sum_{n=1}^N \omega_n \frac{\exp[k_n \cos(\theta - \mu_n)]}{2\pi I_0(k_n)} \quad (4)$$

式中 N 为混合数, k_n 为尺度参数,满足 $k_n \geq 0$, μ_n 为位置参数,取值范围为 $0 \leq \mu_n \leq 2\pi$, ω_n 为权重系数,取值范围 $0 < \omega_n < 1$ 且 $\sum_{n=1}^N \omega_n = 1$, $I_0(k_n)$ 为第一类 0 阶的修正 Bessel 函数。

对于 Copula 参数 α , 可通过 Kendall 相关系数^[12] τ 进行估计

$$\tau = 4 \int_0^1 \int_0^1 C(u_i, u_j) dC(u_i, u_j, \alpha) - 1 \quad (5)$$

3.2 提出的方法

风速 V 和风向 θ 是不同类型的变量,单独建立各自的边缘 PDF 十分不易,建立两者的联合 PDF 更加困难。为解决上述问题,Brooks 等^[13]提出将风速做正交分解,得到两个同类型的风速正交分量 V_w 和 V_s , 过程如图 1 所示,相应的转换公式为

$$V_w = V \cdot \sin\theta, \quad V_s = V \cdot \cos\theta \quad (6)$$

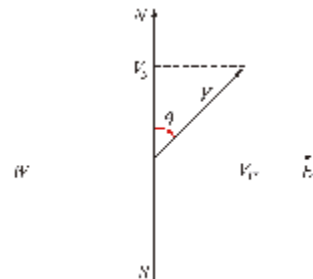


图 1 风矢量和风速正交分量间的转换关系

Fig. 1 Transformation relationship between wind vector and wind speed orthogonal components

此时,采用 Copula 函数考虑风速正交分量 V_w 和 V_s 的相关性,相应的联合 PDF 为

$$g(V_w, V_s) = c(G(V_w), G(V_s); \alpha) g(V_w) g(V_s) \quad (7)$$

式中 $G(V_w)$ 和 $G(V_s)$ 分别为 V_w 和 V_s 的边缘 CDF。 $g(V_w)$ 和 $g(V_s)$ 分别为相应的边缘 PDF。

由于 V_w 和 V_s 的概率密度常表现为多模态的特征,且变量的取值范围为 $-\infty \leq V_w, V_s \leq +\infty$, 可采用混合 Normal 分布^[7]对其进行表示,相应的 PDF 为

$$g(V_r) = \sum_{n=1}^N \omega_n \frac{1}{\sigma_n \sqrt{2\pi}} \exp\left(-\frac{(V_r - \mu_n)^2}{2\sigma_n^2}\right) \quad (8)$$

式中 V_r 为风速正交分量,即 V_w 或 V_s , N 为混合数, ω_n 为权重系数, μ_n 为位置参数, σ_n 为尺度参数。

式(8)的参数可采用最小二乘法估计,在实际应用中,混合数 N 及位置参数 μ_n 的初值选取对于

最终拟合精度有很大影响,而其他参数初值影响较小。为尽可能地提高拟合精度,本文给出了估计混合数及位置参数初值的方法。假设图 2(a)表示的是风速正交分量的概率密度观测值。对图 2(a)的离散概率密度计算一阶导数和二阶导数,结果如图 2(b,c)所示。

估计混合数和位置参数的初值包含以下三个步骤。

(1) 统计图 2(a)中极大值点的个数 N_1 , 并记录对应的横坐标的值 μ_{N_1} 。如图 2(b)所示,图 2(a)的极大值点有 3 个,即 $N_1 = 3$, 对应的横坐标值为 $\mu_{N_1} = \{X_1, X_2, X_3\}$ 。

(2) 统计图 2(a)中满足一定条件的拐点个数

N_2 , 并记录对应的横坐标值 μ_{N_2} 。如在图 2(b)中找到纵轴值大于 0 的连续区域,并在该区域中找到图 2(b)的极小值点的位置,该过程可在图 2(c)中完成,如 X_4 ; 在图 3 中定位纵轴值小于 0 的连续区域,并在该区域找到图 2(b)极大值的位置,如 X_5 。通过统计满足上述条件的拐点可知 $N_2 = 2, \mu_{N_2} = \{X_4, X_5\}$ 。

(3) 混合数为 $N = N_1 + N_2 = 5$, 位置参数初值为 $\mu_N = \{\mu_{N_1}, \mu_{N_2}\} = \{X_1, \dots, X_5\}$ 。

式(7)的 Copula 参数,不同于以往方法通过 Kendall 相关系数估计,而是采用最小二乘法估计。

最后,风矢量联合 PDF 通过雅可比变换得到

$$f(V, \theta) = Vg(V_w, V_s) \quad (9)$$

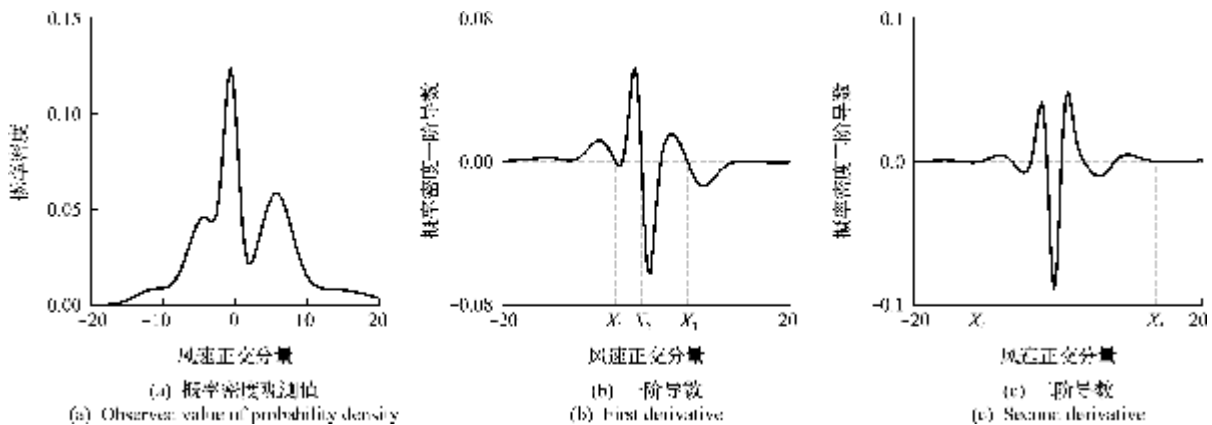


图 2 风速正交分量边缘概率密度观测值及导数

Fig. 2 Observed and derivative value of marginal probability density for wind speed orthogonal components

4 实例分析

4.1 数据来源

数据来源于印度风能研究所^[19]记录的印度 Tamil Nadu 地区附近名为 S1~S12 的 12 个测点的风速和风向数据。原始数据记录了海拔 100 m 高度处,10 min 为间隔的平均风速。取测点 S3 和 S7 位置的每小时平均风速和风向数据,并取其中的 S7 测点进行详细展示,相应的散点如图 3 所示。

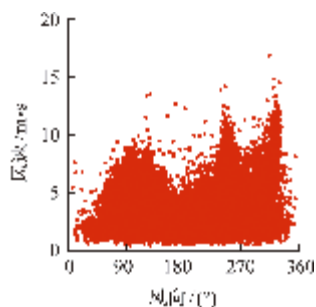


图 3 风速和风向数据散点

Fig. 3 Scatter points of wind speed and wind direction observations

4.2 数据预处理

确定边缘 PDF 和 Copula 函数的参数需基于风矢量数据进行预处理得到观测值,步骤如下。

(1) 构造风速 V 和风向 θ 的网格空间,保证间隔 $\Delta V = V_{i+1} - V_i$ 和 $\Delta \theta = \theta_{j+1} - \theta_j$, 如图 4(a)所示。

(2) 计算落在网格区域 $[V_i, V_{i+1}] \times [\theta_j, \theta_{j+1}]$ 中样本的个数,记作 m_{ij} 。此时, $f(V, \theta)$ 的观测值 $\tilde{f}(V_i, \theta_j) = m_{ij} / M \Delta V \Delta \theta$, M 为数据总个数。

(3) 丢弃风速等于 0 的样本,即当风速 $V_i = 0$ 时,令 $\tilde{f}(V_i, \theta_j) = 0$, 并对观测值 $\tilde{f}(V_i, \theta_j)$ 归一化处理。

(4) 构造正交分量 V_w 和 V_s 的网格空间,保证 $\Delta V_w = \Delta V_s$, 如图 4(b)所示。

(5) 将 V_w 和 V_s 的网格交点 $\{V_{w,k}, V_{s,l}\}$ 通过式(6)得到对应在 V 和 θ 空间的交点 $\{V_{kl}, \theta_{kl}\}$ 。交点 $\{V_{kl}, \theta_{kl}\}$ 处的观测值 $\tilde{f}(V_{kl}, \theta_{kl})$ 可通过对 $\tilde{f}(V_i, \theta_j)$ 插值得到。

(6) 得到观测值 $\tilde{f}(V_{kl}, \theta_{kl})$ 后,交点 $\{V_{w,k}, V_{s,l}\}$ 处的观测值 $\tilde{g}(V_{w,k}, V_{s,l})$ 可通过 $\tilde{f}(V_{kl}, \theta_{kl})$ 和式(9)得到。

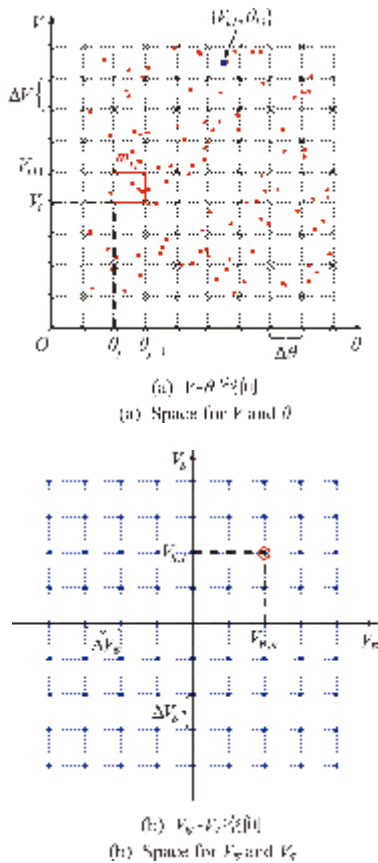


图4 矢量风速和风速正交分量的网格空间
Fig. 4 Grid space for wind vector and wind speed orthogonal components

根据上述步骤,风速间隔取 $\Delta V = 1 \text{ m/s}$,风向间隔 $\Delta\theta = 10^\circ$,风速正交分量 $\Delta V_w = \Delta V_v = 1 \text{ m/s}$,最终,图3的风速和风向数据对应的风矢量联合PDF观测值等高线如图5所示。

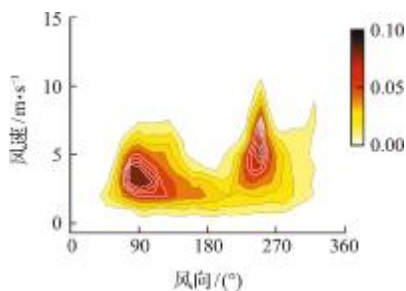


图5 风矢量联合概率密度观测值
Fig. 5 Observed value for joint probability density of wind vector

4.3 结果对比

对于以往的方法,采用 Weibull 分布表示的风速边缘 PDF 如图 6(a)所示,采用混合 von Mises 分布表示风向边缘 PDF 的结果如图 6(b)所示,再通过式(2)得到风矢量联合 PDF,采用 Normal Copula 函数得到的风速矢量联合 PDF 等高线如图 7 所示。

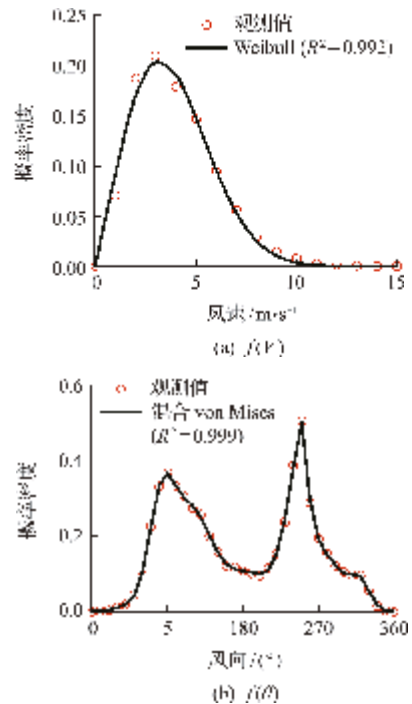


图6 风速和风向边缘概率密度的观测值和模型值
Fig. 6 Observed and model values of marginal probability density for wind speed and direction

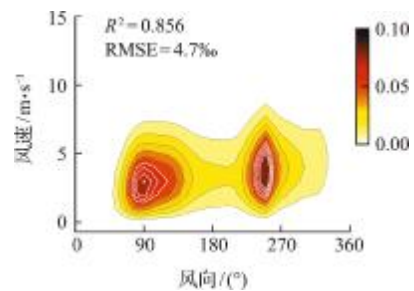


图7 通过以往方法得到的风矢量联合概率密度模型值
Fig. 7 Model value for joint probability density of wind vector by previous method

风速和风向边缘 PDF 表示结果均很理想,如图 6 所示,然而实际参数估计过程十分繁琐。对于风速,当测点发生变化时,Weibull 分布可能不足以充分对风速边缘 PDF 表示,此时需要进行多个分布函数的尝试才能找出最优风速分布。对于风向,混合数和参数值的确定过程也较为复杂。所以,以往方法在确定边缘 PDF 时效率往往较为低下。

对于本文方法,采用混合 Normal 分布表示的风速正交分量的边缘 PDF 如图 8 所示。得到风速正交分量的边缘 PDF 后,通过式(7)得到风速正交分量的联合 PDF,再通过式(9)得到风矢量联合 PDF,采用 Normal Copula 函数得到的风矢量 PDF 等高线如图 9 所示。最后,将 S3 和 S7 测点以及表 1 中其他 Copula 函数的决定系数 (R^2) 结果列入表 2。其中测点 S3 的风速和风向边缘 PDF 的拟

合优度分别为 0.996 和 0.984,而风速正交分量边缘 PDF 的拟合优度均达到了 0.999。

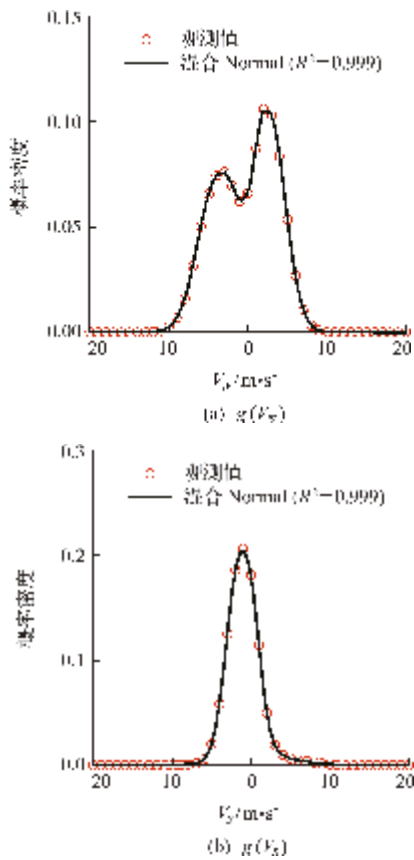


图 8 风速正交分量边缘概率密度的观测值和模型值

Fig. 8 Observed and model values of marginal probability density for wind speed orthogonal components

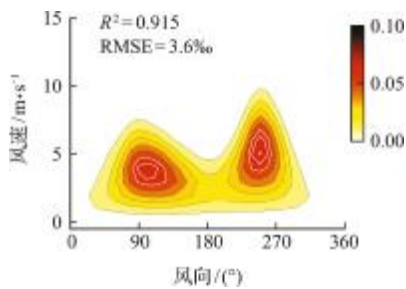


图 9 通过本文方法得到的风矢量联合概率密度模型值

Fig. 9 Model value for joint probability density of wind vector by proposed method

表 2 不同方法得到的 R^2

Tab. 2 Value of R^2 obtained by different methods

测点	方法	Normal	AMH	Frank	FGM
S3	以往	0.759	0.750	0.775	0.767
	本文	0.852	0.828	0.892	0.853
S7	以往	0.856	0.848	0.859	0.863
	本文	0.915	0.927	0.928	0.918

由表 2 中风矢量联合 PDF 可知,无论采用何种 Copula 函数,本文方法得到的结果均优于以往方法。原因是风矢量经正交分解后,考虑同类型的

正交分量间的相关性要比不同类型变量风速和风向间的相关性更加容易。

5 结论

本文通过 Copula 函数考虑了风速正交分量间的相关性,进而开展了风矢量联合 PDF 的研究,主要结论如下。

(1) 对于边缘 PDF,直接采用混合 Normal 分布对风速正交分量边缘 PDF 进行表示的做法要比以往分别对风速和风向边缘 PDF 的做法更加便捷。此外,混合 Normal 分布可对风速正交分量的边缘 PDF 准确表示。

(2) 对于联合 PDF,基于 Copula 函数考虑同类型的风速正交分量间的相关性比考虑风速和风向间相关性的做法更加合理,具体表现为风矢量联合 PDF 结果更加准确。

附录：风矢量随机数的生成

在工程应用中,通常需要生成符合风矢量联合 PDF 特征的随机数以方便分析,步骤如下。

(1) 生成服从标准均匀分布的二元独立随机数 Z_1 和 Z_2 。

(2) 假设 $u_1 = Z_1$,得到其中一个变量的 CDF 样本。

(3) 令 $Z_2 = u_{2|1} = h(u_1, u_2; \alpha)$,则 $u_2 = h^{-1}(u_1, u_{2|1}; \alpha)$ 。

(4) 此时风速正交分量随机数 $\hat{V}_W = G^{-1}(u_1)$, $\hat{V}_S = G^{-1}(u_2)$ 。

步骤(3)中若 Copula 函数的 CDF 为 $C(u_i, u_j; \alpha)$,则偏导数 $u_{j|i} = h(u_i, u_j; \alpha) = \partial C(u_i, u_j; \alpha) / \partial u_i$,而 $h^{-1}(u_i, u_{j|i}; \alpha)$ 表示 $h(u_i, u_j; \alpha)$ 的逆函数^[18]。

参考文献(References):

[1] Dong Z, Wang X J, Zhu R Z, et al. Improving the accuracy of wind speed statistical analysis and wind energy utilization in the Ningxia Autonomous Region, China[J]. *Applied Energy*, 2022, **320**:119256.

[2] Chen P, Han D Z. Effective wind speed estimation study of the wind turbine based on deep learning[J]. *Energy*, 2022, **247**:123491.

[3] Putnam P C. *Power From the Wind*[M]. New York: Van Nostrand Reinhold, 1948.

[4] Carta J A, Ramírez P, Velázquez S. A review of wind speed probability distributions used in wind energy analysis; Case studies in the Canary Islands[J]. *Renewable and Sustainable Energy Reviews*, 2009, **13**

- (5):933-955.
- [5] Doblinger C, Surana K, Li D Y, et al. How do global manufacturing shifts affect long-term clean energy innovation? A study of wind energy suppliers[J]. *Research Policy*, 2022, **51**(7):104558.
- [6] Jung C, Schindler D. Global comparison of the goodness-of-fit of wind speed distributions [J]. *Energy Conversion and Management*, 2017, **133**:216-234.
- [7] Hu Q H, Wang Y, Xie Z X, et al. On estimating uncertainty of wind energy with mixture of distributions [J]. *Energy*, 2016, **112**:935-962.
- [8] Coles S G, Walshaw D. Directional modelling of extreme wind speeds [J]. *Applied Statistics*, 1994, **43**(1):139-157.
- [9] Carta J A, Bueno C, Ramírez P. Statistical modelling of directional wind speeds using mixtures of von mises distributions; Case study [J]. *Energy Conversion and Management*, 2008, **49**(5):897-907.
- [10] Carta J A, Ramírez P, Bueno C. A joint probability density function of wind speed and direction for wind energy analysis [J]. *Energy Conversion and Management*, 2008, **49**(6):1309-1320.
- [11] Johnson R A, Wehrly T E. Some angular-linear distributions and related regression models [J]. *Journal of the American Statistical Association*, 1978, **73**(363):602-606.
- [12] Li H N, Zheng X W, Li C. Copula-based joint distribution analysis of wind speed and direction [J]. *Journal of Engineering Mechanics*, 2019, **145**(5):04019024.
- [13] Brooks C E P, Durst C S, Carruthers N. Upper winds over the world. Part I: The frequency distribution of winds at a point in the free air [J]. *Quarterly Journal of the Royal Meteorological Society*, 1946, **72**(311):55-73.
- [14] McWilliams B, Newmann M M, Sprevak D. The probability distribution of wind velocity and direction [J]. *Wind Engineering*, 1979, **3**(4):269-273.
- [15] Weber R. Estimator for the standard deviation of wind direction based on moments of the cartesian components [J]. *Journal of Applied Meteorology*, 1991, **30**(9):1341-1353.
- [16] Erdem E, Shi J. Comparison of bivariate distribution construction approaches for analysing wind speed and direction data [J]. *Wind Energy*, 2011, **14**(1):27-41.
- [17] Nelsen R B. *An Introduction to Copulas* [M]. Array New York: Springer, 2006.
- [18] Ji X. Multivariate extreme wind loads: Copula-based analysis [J]. *Journal of Engineering Mechanics*, 2023, **149**(1):04022082.
- [19] Bharani R, Sivaprakasam A. A large volume wind data for renewable energy applications [J]. *Data in Brief*, 2019, **25**:104291.

Research on joint probability density of wind vector based on Copula function and wind speed orthogonal components

CHENG Zheng-bing, JI Xiao-wen*

(Key Laboratory of Urban Security and Disaster Engineering of Ministry of Education, Beijing University of Technology, Beijing 100124, China)

Abstract: Accurately representing the joint probability density of wind vector is significant for wind energy assessment, as well as structural design of wind turbine. This study will focus on improving the accuracy of describing joint probability density of wind vector. Firstly, the wind vector is represented as orthogonal components, a mixed Normal distribution is used to represent their marginal probability density, and a method for estimating the initial values of parameters is proposed. Then, based on Copula function, the correlation between the orthogonal components of wind speed is considered to obtain their joint probability density, the Copula parameter is estimated by the Least Squares estimation, and then joint probability density of wind vector is obtained through Jacobian transformation. Finally, the hourly average wind speed and direction data in point S3 and S7 from the Indian Wind Energy Research Institute is used to compare with previous joint probability density results of wind vector obtained based on the Copula function and wind speed and direction variables. The results show that the joint probability density of wind vector obtained by proposed method is more accuracy than previous method. This study can provide a theoretical basis for the exploitation and utilization of wind energy.

Key words: copula function; orthogonal decomposition; wind vector; joint probability density; mixed Normal