

ORIGINAL INNOVATION

Open Access



Study on the distribution of average wind speeds at a mountainous bridge site for structural durability design

Cheng Cheng¹, Beilong Luo², Yan Jiang^{2,3*} and Jian Liu¹

*Correspondence:
xnjtjiangyan@163.com

¹ Chongqing Wukang
Technology Co., LTD,
Chongqing 404100, China

² College of Engineering
and Technology, Southwest
University, Chongqing 400715,
China

³ State Key Laboratory
of Mountain Bridge and Tunnel
Engineering, Chongqing
Jiaotong University,
Chongqing 400074, China

Abstract

Conventional wind speed distribution methods (e.g., Rayleigh distribution and Weibull distribution) may not adequately capture complex characteristics of wind fields in mountainous areas. To address this problem, this study proposes a semi-parametric mix method for modeling the distribution of average wind speeds based on the combination of nonparametric Kernel Density Estimation (KDE) and Generalized Pareto Distribution (GPD). In the proposed method, KDE focuses on capturing the distribution in the main part of average wind speeds, while GPD aims at performing the distribution in terms of those in the extreme part. The segment point (i.e., the threshold) between KDE and GPD distributions is determined based on the combination of conditional mean excesses criterion and empirical rule. Meanwhile, the selection of modeling parameters should ensure that the mix distribution model is continuous and differentiable at the identified threshold point. Then, the commonly-used conditional probability model is further introduced to describe the wind direction distribution. Finally, a case study based on the measured 10-min average wind speeds at a mountainous bridge site is employed to demonstrate the effectiveness of the proposed method. The results indicate that: (1) the distribution of omnidirectional average wind speeds in the mountainous bridge site exhibits an obviously single-peak characteristic, while those considering wind directionality present a certain bimodal characteristic; (2) the proposed method can effectively describe wind speed distributions with different statistical characteristics, and the fitting accuracy outperforms the frequently-employed Weibull distribution model.

Keywords: Wind speed distribution, Kernel density estimation, Generalized Pareto distribution, Structural durability design, Mountainous wind fields

1 Introduction

1.1 Literature review

With the evolution of mountainous bridges towards the direction of long-span and flexibility, the wind sensitivity of these structures becomes increasingly significant (Tao and Wang 2023). In this context, wind loads have gradually become the dominant factor for the design of long span bridges in mountainous areas (Hu et al. 2020; Tang et al. 2020). Generally, an accurate description of wind field characteristics at the bridge site

is critical for ensuring the operation safety of bridges. However, due to the influence of complex terrains and climatic conditions, it is difficult to explain wind field characteristics in mountainous areas based on the current specification which is formulated to describe the conventional stationary wind field (Ren et al. 2021; Yu et al. 2018; Zhang et al. 2020; Liao et al. 2020). Therefore, it is urgently needed to perform the investigation of characteristics of near-ground wind fields in mountainous areas, thereby improving the wind resistance design level of long-span bridges in complex terrains.

Generally, wind speeds are mainly composed of two components, i.e., average wind speeds and wind fluctuations. Researches regarding these two components primarily focus on the distribution of average wind speeds, power spectrum of wind fluctuations and turbulence intensity (Yang et al. 2002; Zheng et al. 2019a, b; Wu et al. 2017; Wang et al. 2013). Among them, the distribution of average wind speeds has attracted increasing attention, and commonly services for the structural strength design. Therefore, most of relevant explorations take extreme wind speeds (e.g., monthly maximum wind speeds or yearly maximum wind speeds) as the target sample to establish the corresponding probabilistic density function (PDF), thereby inferring the desired design wind speed (i.e., the basic wind speed). The commonly-used models regarding the wind speed distribution include Gumbel distribution (Simiu et al. 2001), Weibull distribution (Wais 2017), Rayleigh distribution (Chiodo and Noia 2020), Generalized Pareto Distribution (GPD) (Ding and Chen 2014), Gamma distribution (Özkan et al. 2020) and lognormal distribution (Kenfack et al. 2021), etc. For example, Zheng et al. (2019a, b) respectively employed three different distribution models (i.e., Gumbel distribution, Weibull distribution and Frechet distribution) to fit the distribution of average wind speeds at the site of Aizhai bridge, and the result showed that the Gumbel distribution had the best fitting performance. Nage (2016) compared the fitting performance of Weibull distribution and Rayleigh distribution, concluding that the former exhibited superior fitting capability than that of the latter. Generally speaking, these above-mentioned distribution models involve in some critical parameters to be determined, and thus can be regarded as parametric models. Their fitting accuracy is closely related to the involved parameters. However, the corresponding parameter determination process requires a prior assumption regarding the distribution type which is difficult to be acquired beforehand in practice.

Compared with a large number of investigations on the distribution of extreme wind speeds, relatively few studies have been conducted on the distribution of parent wind speeds (i.e., samples of consecutive 10-min average wind speeds in a time period of interest), which is critical for the structural durability assessment (Repetto and Solari 2010; Ding et al. 2016). Because of the effect of the inner dependence of parent wind speeds, some of extreme wind speed distribution models cannot be directly applied in the fitting of parent wind speeds. Although Weibull distribution can be used to describe the distribution of parent wind speeds (Wais 2017), those before-mentioned problems in the parametric model still exist. As an effective nonparametric model, Kernel Density Estimation (KDE) can dispense with any parameter or distribution assumption in the modeling process, and has been successfully used in capturing the distribution of wind speeds (Nguyen et al. 2021; Chen and Duan 2018). However, this model can only cover a wide range of probability distributions from the data, and usually fails to reflect the probability distribution of wind speeds in two distribution tails (Takle and Brown 1978).

According to Ding and Chen (2014), the GPD model can provide an approximation description on the PDF of data above a selected high threshold not just the independent peaks (i.e., the distribution in the extreme part of parent wind speeds). Considering the advantage of GPD, the utilization of this model to describe the distribution of wind speeds over a pre-specified threshold may be promising.

It is worth noting that most of these above-mentioned distribution models can only describe the distribution of average wind speeds with the unimodal (i.e., single-peak) distribution characteristic. Affected by climatic factors such as temperature, humidity, atmospheric pressure and topography, the actual distribution of average wind speeds may present more complicated patterns, e.g., bimodal (i.e., double-peak), trimodal (i.e., threefold-peak) or even multimodal (i.e., multi-peak) distribution scenarios (Zhou et al. 2021). In this context, the accuracy of these distribution models is limited to some extent, and it is necessary to develop a mix form PDF to describe the complex average wind speed distribution in practice. Meanwhile, the information of wind direction is also critical for accurately describing the characteristic of wind fields, which is conducive to the refined analysis of structural wind-induced responses (Holmes 2020; Isyumov et al. 2014; Carta et al. 2008; McWilliams et al. 1979). As is well-known, wind speeds at different directions in the same location exhibit a significant difference, and the difference of the structural scale at different directions is also remarkable, especially for that of long-span structures (e.g., bridges, high-rising buildings, transmission tower line system, etc.) (Zhu and Xu 2005; Xu and Zhu 2005). For example, the structural parameters (e.g., stiffness, damping ratio) and vibration performance of long-span bridges along-span direction and vertical-span direction are obviously different from each other. Hence, a reasonable description on the joint distribution of parent wind speeds and wind direction is the essential prerequisite for accurately describing the characteristic of wind loads, thereby realizing a reliable assessment of wind-induced fatigue damages.

Currently, there are mainly three methods for modeling the joint distribution of wind speed and direction including the stationary random process method, maximum wind direction coefficient method, and joint probability distribution method (Cook 2021). Among them, the joint probability distribution method is widely-used because of its simplicity and convenience (Wang and Gu 2009), and the corresponding wind direction models mainly involve conditional probability and von Mises distribution (Erdem and Shi 2011). For example, Xu et al. (2009) estimated the buffeting-induced fatigue damage of the Tsing Ma Bridge based on a continuous damage mechanics model where a combination of Weibull distribution and conditional probability model was established to capture the joint distribution of wind speed and wind direction. Zheng et al. (2019a, b) adopted the multiplication theorem to reveal the joint distribution of wind speed and wind direction in Dali area based on a large number of measured data, where von Mises distribution was introduced to describe the distribution of wind direction.

Inspired by the above literature reviews, the motivation of this study is to develop an innovative method for accurately describing the distribution of average wind speeds in mountainous areas, thereby providing a basis for the wind-induced structural durability design. This method is the mix of nonparametric KDE and GPD, where KDE focuses on capturing the distribution in the main part of average wind speeds, while GPD aims to performing the distribution fitting of those in the extreme part. This mix method

belongs to the semi-parametric model and can well cope with the actual wind speed distribution with the possible unimodal, bimodal or even multimodal distribution situations. Then, the commonly-used conditional probability model is introduced to capture the distribution of wind direction. Finally, a case study based on the measured 10-min average wind speeds at a mountainous bridge site demonstrates the superior performance of the proposed method. The novelty of this study mainly embodies in the following several aspects.

- (1) A semi-parametric mix distribution model (the KDE-GPD distribution model) is proposed to describe the PDF of average wind speeds, which can simultaneously ensure that the function is continuous and derivable at the threshold point. Compared with parametric models, this model involves fewer undetermined parameters.
- (2) The proposed method can well address the complex distribution of average wind speeds in practice including single-peak, double-peak, and even multi-peak scenarios. Case study based on the measured wind speed data in mountainous areas verifies its superiority.

1.2 Organization of the paper

The organization of the study is presented as follows. In Section 2, details of the proposed method are firstly presented. Then, the case study and the subsequent performance analysis are displayed in Section 3. Finally, some main conclusions are drawn in Section 4.

2 Materials and methods

2.1 Study area

Figure 1 shows the overall layout of a certain mountain suspension bridge and its surrounding topography is listed in Fig. 2. The structure of this bridge adopts a double-tower single-span ground-anchored configuration. The total length of the bridge is 1040 m, the span is 166 m + 628 m + 166 m and the rise-to-span ratio is 1/10 (Zhao et al. 2020). The heights of tower on the left and right sides are 153 m and 138 m, respectively. In addition, the bridge has an altitude of 1,860 m and crosses a canyon with a depth of about 400 m. As can be seen from Fig. 2, the bridge runs from southwest to northeast, which is located in a mountainous area with complex terrain and climatic conditions, and significant topographic relief (e.g., including three deep canyons). The terrain on both sides of one of the deep canyons is very steep and V-shaped, running through the entire bridge site area. Therefore, the bridge is not only subject to the wind in the direction of this canyon, but also may be subject to the wind in the direction of the other two canyons (Huang et al. 2019). Obviously, these complex terrain and climatic conditions mentioned above will inevitably cause the bridge to suffer from complex wind field effects during the service. In order to accurately grasp the wind vibration characteristics of the bridge, it is urgent to conduct a more in-depth study of the wind field characteristics at the bridge site.



Fig. 1 Diagram of the bridge layout

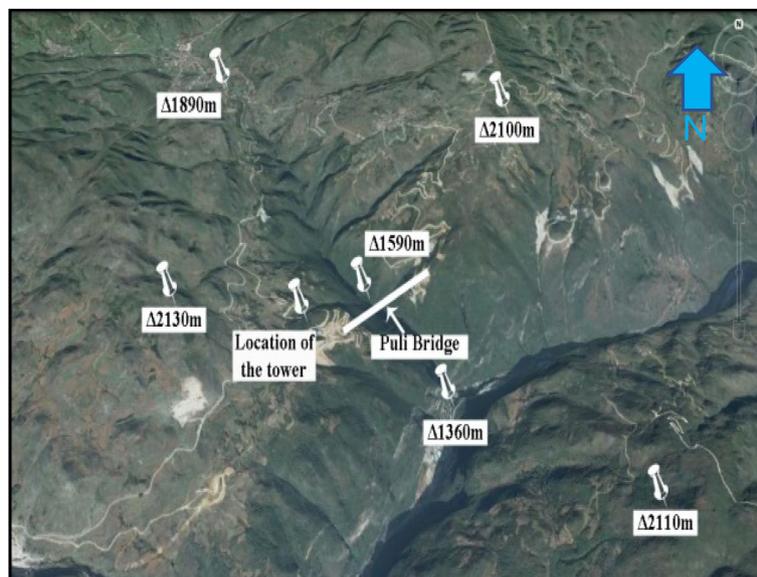


Fig. 2 Surrounding landforms of the bridge

According to Huang et al. (2019), it can be found that these topographic and climatic conditions at the bridge site are liable to produce typical mountain winds. In order to study the characteristics of the wind field, it is necessary to conduct wind field measurement in mountainous areas. Based on the wind field measurement project, a wind speed observation tower with the height of 50 m is established on the slope near the left side in February 2013. The bottom elevation of the tower is 1890 m, and the height of all observation points is therefore above the height of the main beam (1860 m). Figure 3 shows the actual and schematic diagram of the anemometer layout at the observation point.

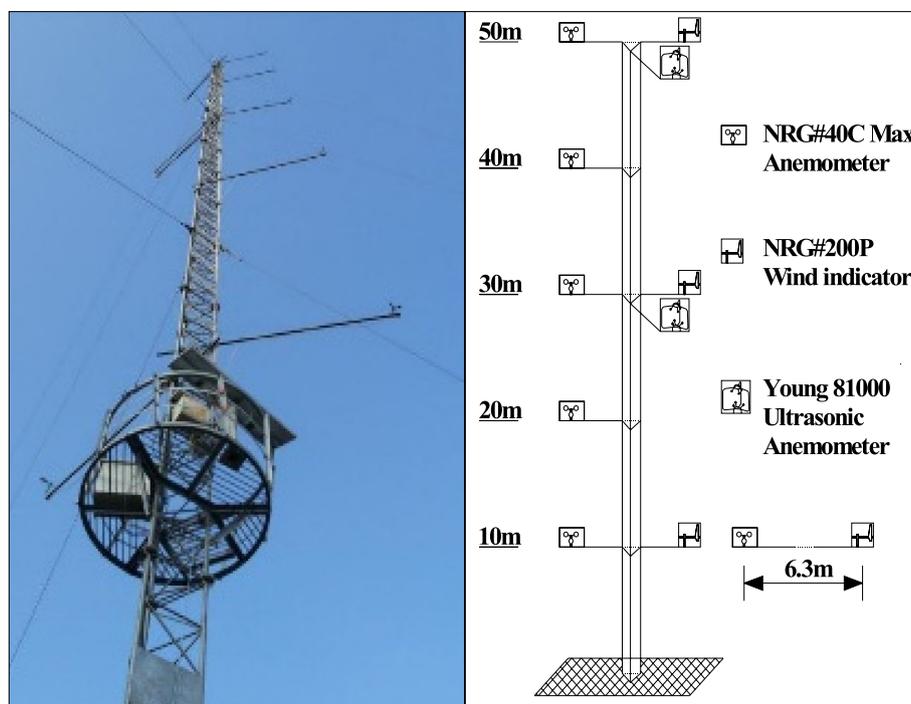


Fig. 3 Layout of anemometer

As shown in Fig. 3, the cup anemometer (NRG) and three dimensions (3D) ultrasonic anemometer (Young 81000) are used in the field measurement. Specifically, five NRG anemometers (marked by the red textbox in Fig. 3) are installed at the heights of 10 m, 20 m, 30 m, 40 m, and 50 m, respectively. Three NRG wind indicators are installed at the heights of 10 m, 30 m, and 50 m, respectively. Meanwhile, two Young anemometers (marked by the blue textbox in Fig. 3) are respectively installed at the heights of 30 m and 50 m, respectively. In addition, the NRG hygrometer, barometer and thermometer are installed at the height of 8 m to obtain humidity, barometric pressure and temperature, respectively. Figure 4 shows the wireless wind data acquisition system. More detailed description on the information of the field measurement can be found in Huang et al. 2015.

Due to the good data acquisition and transmission performance of the NRG anemometer during the field measurement, the measured data recorded by the NRG anemometer are used in this paper. Therefore, 10-min average wind speed and wind direction data for approximately 33 months from February 9, 2013 to October 16, 2015 are used as the experimental data. In addition, it is necessary to exclude the data with zero wind speed recorded in the field measurement during the analysis. On this basis, 139,979 sets of 10-min average wind speed and wind direction data are finally acquired to perform the subsequent investigation on the distribution of average wind speed and direction.

2.2 KDE-GPD distribution

The PDF distributions of random samples typically include parametric and non-parametric models. Unlike parametric models which require the assumption of distribution types or

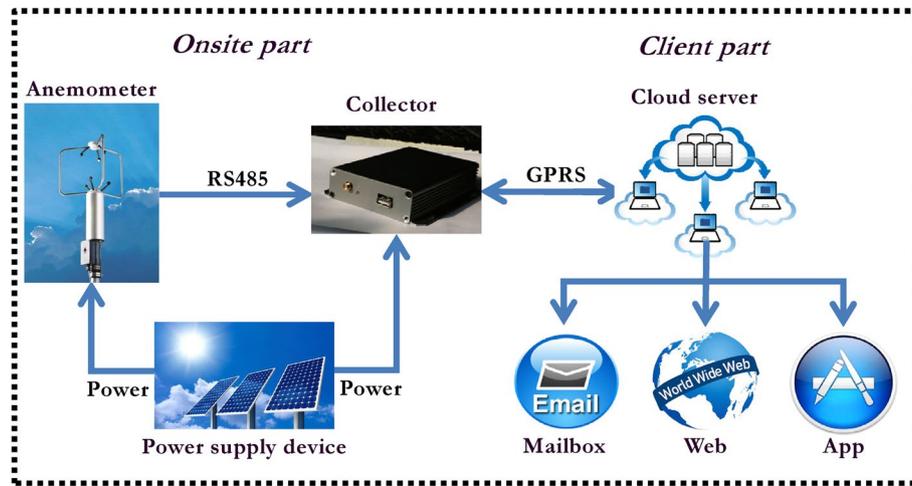


Fig. 4 The wireless transmission system

some key parameters in advance, nonparametric models do not require any assumptions and have obvious data-driven attribute. With the continuous enrichment of data, the fitting distribution of nonparametric models becomes closer to the true distribution situation (Jiang et al. 2020).

KDE is a good nonparametric model for estimating PDF. This model can not only better describe the probability distribution of a single variable (one-dimensional estimation), but also can establish the conditional probability among multiple variables (i.e., multidimensional estimation) (Hyndman et al. 1996). In this paper, one-dimensional kernel density estimation is used to reflect the distribution of the main part of average wind speeds, as follows.

Let $\{u\} = (u_1, u_2, \dots, u_M)$ be an independent and identically distributed random variables, and then the probability density estimation function of KDE is given by (Jiang and Huang 2017):

$$g_{KDE}(u) = \frac{1}{Mh'} \sum_{i=1}^M K\left(\frac{u - u_i}{h'}\right) \tag{1}$$

where, $g_{KDE}(u)$ is the PDF estimated via the KDE model; M is the number of samples; $h'(h' > 0)$ is the bandwidth parameter; and $K(\cdot)$ is a symmetric kernel function with the integration equal to one and has many possible choices (Epanechnikov 1969). Therefore, the estimation result of KDE is closely related to the total number of samples M , the bandwidth parameter h' , and the kernel function $K(\cdot)$. Typically, two types of kernel functions are used, i.e.,

Gaussian kernel function: $K(u) = \frac{1}{\sqrt{2\pi}} e^{-\frac{u^2}{2}}$ (2)

Epanechnikov kernel function : $K(u) = \begin{cases} 3 \times (1 - u^2)/4, & |u| \leq 1 \\ 0 & |u| > 1 \end{cases}$ (3)

Among them, Gaussian kernel function has advantages of smoothness and strict positive definiteness, and thus is adopted in this paper (Liu et al. 2022). Theoretically speaking, the selection of bandwidth parameter and kernel function has some effect on the fitting performance (Hyndman et al. 1996). By contrast, the effect of bandwidth parameter is much more significant on the estimation result. For example, a larger bandwidth will reduce the proportion of actual data on the fitting curve, while a smaller one will lead to the increase of the corresponding proportion, thereby bringing out a steep fitting curve. This paper determines the bandwidth parameter based on the standard parameter determination criterion (Jiang et al. 2019; Zhang et al. 2014):

$$h' = \left(\frac{4\tilde{\sigma}^5}{3M} \right)^{\frac{1}{5}} \tag{4}$$

where $\tilde{\sigma}$ denotes the standard deviation of the target sample data.

Although the KDE model has the advantages such as high fitting accuracy and strong applicability, it can only describe the probability distribution of the main part of the data (Ding and Chen 2014). In comparison, the GPD model can better address the extreme value problem (Zhang and Chen 2015). Therefore, this study describes the probability distribution of the main part of average wind speeds based on the KDE model in the first place, and then employs the GPD model to fit the probability distribution of the extreme part of average wind speeds, and the specific modeling process is listed as follows:

Let $(u_1, u_2, \dots, u_M, \dots, u_n)$ be independent and identically distributed random variables, where the PDF distribution of (u_1, u_2, \dots, u_M) is described by the KDE model, and the probability density distribution of $(u_{M+1}, u_{M+2}, \dots, u_n)$ is captured by the GPD model. Among them, the cumulative distribution function $F_{GPD}(\cdot)$ and the probability density function $f_{GPD}(\cdot)$ of GPD are respectively defined as (Ding and Chen 2014):

$$\begin{cases} F_{GPD}(u) = P(U \leq u | U \geq x_T) = 1 - [1 + (c(u - x_T)/d)]^{-\frac{1}{c}} \\ f_{GPD}(u) = \frac{1}{d} [1 + (c(u - x_T)/d)]^{-\frac{1}{c}-1}, c \neq 0 \end{cases} \tag{5}$$

in which x_T is the threshold value; c and d are the model shape parameter and scale parameter, respectively, which can be estimated using the maximum likelihood estimation (Holmes and Moriarty 1999). Then, the mix cumulative distribution function and probability density distribution function of KDE-GPD $f_U(u)$ is presented as follows:

$$P_U(u) = \begin{cases} G_{KDE}(u) & u \leq x_T \\ F_{GPD}(u) \cdot [1 - G_{KDE}(x_T)] + G_{KDE}(x_T) & u > x_T \end{cases} \tag{6}$$

$$f_U(u) = \begin{cases} g_{KDE}(u) & u \leq x_T \\ f_{GPD}(u) \cdot [1 - G_{KDE}(x_T)] & u > x_T \end{cases} \tag{7}$$

where $G_{KDE}(u)$ is the cumulative distribution function estimated by the KDE model at $u \leq x_T$; $G_{KDE}(x_T)$ is the cumulative probability value of $G_{KDE}(u)$ at $u = x_T$, which can be expressed as $G_{KDE}(x_T) = M/n$, M is the number of data when the wind speed u is less than or equal to the threshold x_T and n is the number of all data.

From the above equation, it can be seen that the PDF being continuous at this threshold point x_T is the key to accurately obtaining the mix cumulative distribution function

(Ragan and Manuel 2008). This paper determines the threshold value x_T based on two methods, namely the threshold value x_{T1} based on the extreme value analysis of the wind turbine response (Luo et al. 2021):

$$x_{T1} = \bar{u} + 1.4\tilde{\sigma} \tag{8}$$

where \bar{u} is the mean value of the sample. The other threshold value x_{T2} is determined according to the Conditional Mean Excesses (CME) criterion (Harris 2005). If the data exceeding the threshold x_{T2} follows the GPD distribution, then for any data exceeding this threshold, there is a variable $Y_i = U - u_i | U > u_i$ that also follows the GPD distribution, and the mean value \bar{y}_i of Y_i and $u_i - x_{T2}$ has the following linear relationship when x_{T2} is the initial point, i.e.,

$$\bar{y}_i = E(U - u_i | U > u_i) = -\frac{c(u_i - x_{T2} + d)/(1 + c)}{(1 + c)} \tag{9}$$

Then the initial point of the linear relationship (i.e., x_{T2}) can be considered as the threshold value. Therefore, the final threshold x_T takes the maximum of the above two results, i.e.,

$$x_T = \max(x_{T1}, x_{T2}) \tag{10}$$

To ensure that the cumulative distribution function and the PDF are continuous at the threshold point x_T , the function values of the PDF at the threshold point not only must be equal, but also their first derivative is equal at the threshold point, namely:

$$\begin{aligned} g_{KDE}(x_T) &= [1 - G_U(x_T)] \cdot f_{GPD}(x_T); \\ g'_{KDE}(x_T) &= [1 - G_U(x_T)] \cdot f'_{GPD}(x_T) \end{aligned} \tag{11}$$

where $g'_{KDE}(x_T)$ and $f'_{KDE}(x_T)$ are the derivative values of $g_{KDE}(x_T)$ and $f_{KDE}(x_T)$ at the threshold point x_T , respectively. Due to the use of Gaussian kernel function, the above equation can be changed to:

$$\begin{cases} \frac{1}{n} \sum_{i=1}^n \left(\frac{A}{\sqrt{2\pi}h'} \right) - \frac{1-G_{KDE}(x_T)}{d} = 0 \\ \frac{1}{n} \sum_{i=1}^n \left(A \frac{u_i - x_T}{\sqrt{2\pi}h'^3} \right) + \frac{1-G_{KDE}(x_T)}{d^2} (c + 1) = 0 \\ A = \exp \left[-\frac{(x_T - u_i)^2}{2h'^2} \right] \end{cases} \tag{12}$$

Therefore, in addition to satisfying the requirement of maximum likelihood estimation, the parameters should also meet the requirement of Eq. (12), i.e., the determined parameters c and d should minimize the value of the variable Q :

$$\begin{aligned}
 Q &= S + m_1 \left(\frac{1}{n} \sum_{i=1}^n \left(\frac{A}{\sqrt{2\pi}h'} \right) - \frac{1 - G_{KDE}(x_T)}{d} \right)^2 + \\
 & m_2 \left(\frac{1}{n} \sum_{i=1}^n \left(A \frac{u_i - x_T}{\sqrt{2\pi}h'^3} \right) + \frac{1 - G_{KDE}(x_T)}{d^2} (c + 1) \right)^2 \tag{13} \\
 S &= \log(d)^{(n-M)} + \frac{c + 1}{c} \sum_{i=M+1}^n \log \left(1 + \frac{c(u_i - x_T)}{d} \right)
 \end{aligned}$$

in which $m_1 > 0; m_2 > 0$. Finally, $P_U(u)$ and $f_U(u)$ can be obtained based on the determined parameters c, d , and x_T . In summary, the KDE-GPD model can be considered as a semi-parametric mix model. Compared to parametric models, it typically requires fewer parameters. Therefore, this mix model may have a high practicality.

2.3 Joint wind speed and direction distribution

Considering the directionality of the wind is significant for accurately calculating the buffeting response and fatigue damage of bridges (Wang and Gu 2009). This paper adopts the conditional probability model to describe the corresponding joint distribution of average wind speed and wind direction. In the conditional probability model, wind speed and wind direction are assumed to be mutually independent and the wind speed distribution is a conditional distribution with a fixed wind direction, while satisfying the following assumptions (Xu et al. 2009): (1) The distribution of wind speeds for any given wind direction obeys the KDE-GPD distribution; (2) The interdependence of wind distribution in different wind directions can be reflected by the relative frequency of occurrence of wind:

$$P_{U,\Theta}(u, \theta) = P_{\Theta}(\theta) \int f_{U|\Theta=\theta}(u, \vartheta[\theta]) du = \iint f_{U,\Theta}(u, \theta) dud\theta \tag{14}$$

$$\begin{aligned}
 f_{U,\Theta}(u, \theta) &= f_{U|\Theta=\theta}(u, \vartheta[\theta])f_{\Theta}(\theta); \\
 P_{\Theta}(\theta) &= \int_0^{\theta} f_{\Theta}(\theta) d\theta
 \end{aligned} \tag{15}$$

where $P_{U,\Theta}(u, \theta)$ is the joint cumulative distribution function of wind speed and direction; $f_{U,\Theta}(u, \theta)$ is the corresponding joint probability density function; $f_{U|\Theta=\theta}(u, \vartheta[\theta])$ is the probability density function of wind speeds in the wind direction $\theta (0 \leq \theta \leq 2\pi)$; $P_{\Theta}(\theta)$ is the relative frequency of occurrence of wind in the wind direction θ ; $f_{\Theta}(\theta)$ is the wind direction probability density function; $\vartheta[\theta]$ is the parameter vector in the wind direction θ . In terms of KDE-GPD, $\vartheta[\theta] = [x_T(\theta), c(\theta), d(\theta), h'(\theta)]$. These parameters and the relative frequency of wind direction are fitted by the harmonic function (Xu et al. 2009).

3 Results

3.1 Result in full wind direction

To begin with, without considering the effect of wind direction i.e., the direction of all wind speeds is assumed to be completely consistent, statistical characteristics of all wind speed data are shown in Table 1. From the table, it can be seen that the average value, variance, skewness, and kurtosis of the data are 2.05 m/s, 1.69 m/s, 1.22,

Table 1 Statistical characteristics of 10-min average wind speed data (all wind directions)

Methods	Mean	Standard deviation	Maximum value	Minimum value	Skewness	Kurtosis
Measured	2.05	1.69	13.10	0.40	1.22	4.65
KDE-GPD	2.01	1.67	13.02	0.38	1.20	4.66
Weibull	1.96	1.76	12.89	0.36	1.16	4.58

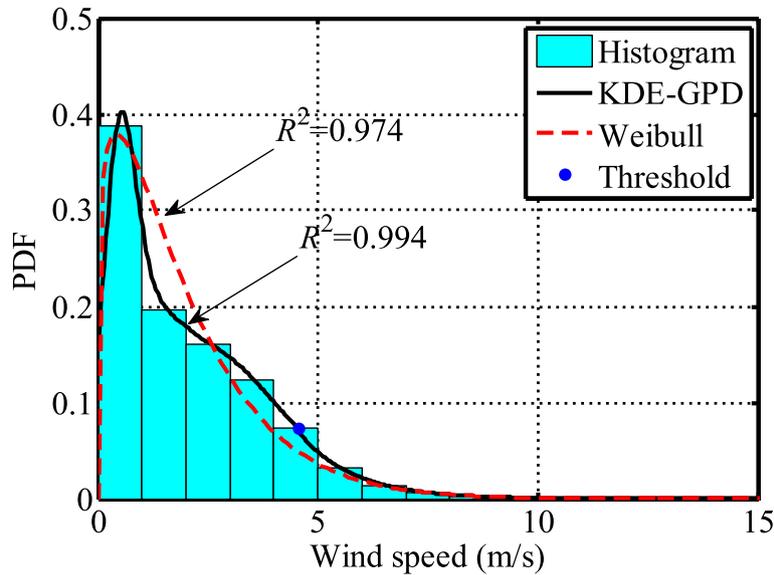


Fig. 5 Comparison of Weibull and KDE-GPD with the actual one (all wind directions)

and 4.65, respectively. Obviously, these data feature a significant non-Gaussian characteristic (Note that: the skewness and kurtosis of Gaussian distribution are equal to 0 and 3, respectively). In order to illustrate the superiority of the KDE-GPD distribution model, the commonly-used Weibull distribution model is employed in this paper for the comparative study, and the corresponding results are shown in Fig. 5. From the figure, it can be seen that the wind speed distribution in all directions belongs to a typical single peak distribution. In Fig. 5, the histogram shows the distribution frequency of actual 10-min average wind speeds, the blue dot denotes the selected threshold value, the red dashed line stands for the fitting result of the Weibull distribution, and the black solid line corresponds to the fitting result of the KDE-GPD distribution. By comparison, the KDE-GPD distribution model can better describe the distribution of the actual wind speeds than that of the Weibull distribution model, and the distribution function is smooth at the threshold point. The corresponding model parameters of the KDE-GPD distribution are $h' = 0.17$; $c = 0.033$; $d = 1.24$; $x_T = 6.90$. Figure 6 shows the threshold of x_{T2} determined by the CME criterion where $x_{T1} = 4.42$ according to Eq. (8). Therefore, the final threshold x_T should be set as 6.90 based on Eq. (10).

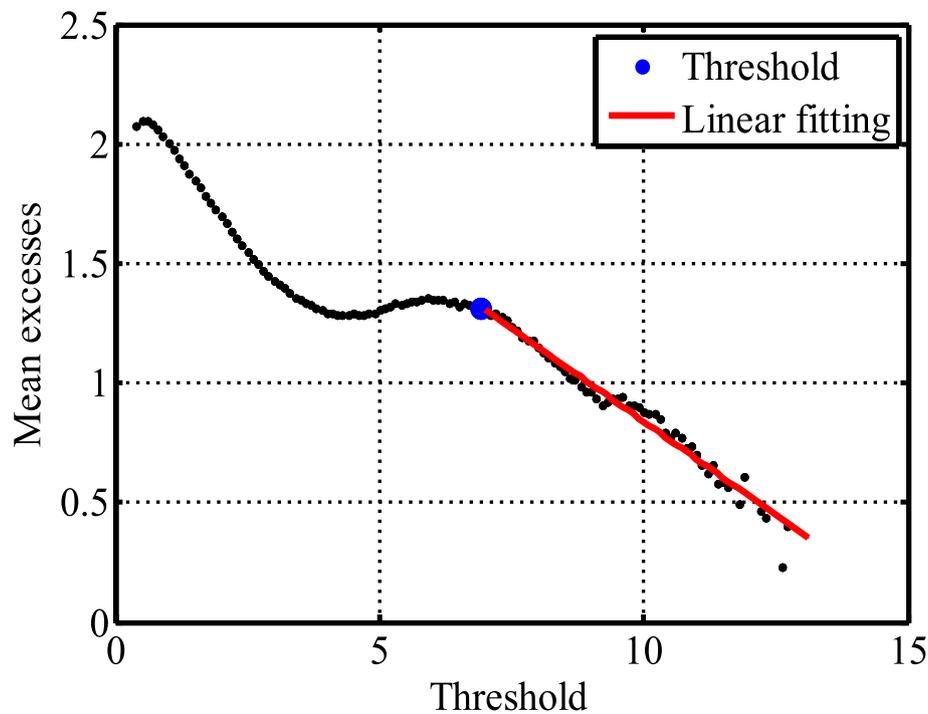


Fig. 6 The threshold selection based the CME criterion

Table 2 Probability distributions of Weibull and KDE-GPD in different wind speed intervals

Interval(m/s)	0–2	2–4	4–6	6–8	8–10	10–12
Weibull	0.65	0.26	0.075	0.019	4.40e-03	9.35e-04
KDE-GPD	0.56	0.30	0.111	0.021	5.50e-03	9.24e-04
Measured	0.55	0.31	0.110	0.021	5.49e-03	9.26e-04

To further illustrate the performance of the KDE-GPD model, the overall wind speed range is divided by the interval $\Delta u = 2 \text{ m/s}$, and then seven different wind speed intervals are generated (Xu et al. 2009). On this basis, the actual wind speed distributions and two probability distribution fittings (i.e., KDE-GPD and Weibull) in each interval are calculated, respectively, and the corresponding results are shown in Table 2. From the table, it can be seen that the main part of the wind speed data obtained by these two distributions is located in the interval $[0 \text{ m/s}, 8 \text{ m/s})$, and the actual wind speed distribution is in good agreement with the fitting results of the KDE-GPD distribution model. Meanwhile, the probability of Weibull distribution is greater than that of KDE-GPD distribution in two extreme value intervals $[0 \text{ m/s}, 2 \text{ m/s})$ and $[10 \text{ m/s}, 14 \text{ m/s})$, and the probability of distribution according to KDE-GPD is larger in the interval $[2 \text{ m/s}, 10 \text{ m/s})$. Therefore, using the Weibull distribution for response analysis will not only overestimate the effect of the extreme part of wind speeds, but also underestimate the effect of the main part of the wind speed. Meanwhile, statistical characteristics estimated by KDE-GPD and Weibull are

also given in Table 1. Obviously, KDE-GPD presents a better statistical approximation to the true results than the Weibull distribution.

3.2 Result of considering directionality

In order to investigate the impact of wind direction, first of all, this section divides the wind speed according to the wind direction, and then fits the wind speed data in each wind direction. Theoretically, the narrower the wind direction partition is, the more accurate the results are obtained. However, too many partitions will not only increase the workload, but also are not conducive to the practical application. At the same time, when there is insufficient wind speed data, it will lead to insufficient data in each wind direction zone, which will make the fitting results significantly deviate from the actual situation. Generally, the reasonable number of wind partitions is set in the range of [8, 16]. Figure 7 divides the wind direction into 16 sectors using $\Delta\theta = 22.5^\circ$. 0° stands for the due north direction and the wind direction angle increases clockwise; "E, W, S, N" denotes due east, due west, due south and due north, respectively; The dashed line represents the wind direction corresponding to each sector, which is "N, NNE, NE, ENE, E, ESE, SE, SSE, S, SSW, SW, WSW, W, WNW, NW, NNW".

Figure 7 also shows the wind rose diagram obtained from the measured data at the bridge site divided into 16 sectors (i.e., the relative frequency of wind direction). Among them, the frequency of winds coming from the due east is significantly higher than that of the other directions. This wind direction can therefore be considered as the dominant wind direction. However, the wind speed data in this direction is mostly below 10 m/s. At the same time, it can also be observed that strong winds mainly come from two southwest directions (i.e., SSW and SW directions) with a

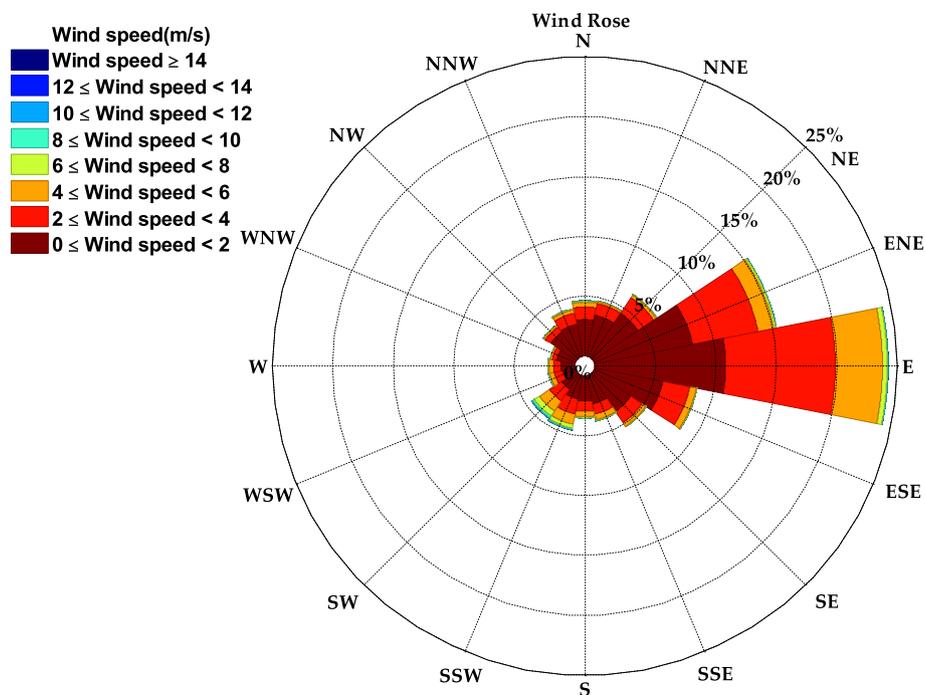


Fig. 7 Wind direction division using the 16-sector

Table 3 Descriptions of wind speed data characteristics in some wind directions

Wind direction	Average value (m/s)	Variance (m/s) ²	Maximum value (m/s)	Minimum value (m/s)	Skewness	Kurtosis
N	1.75	1.71	11.50	0.40	1.60	5.31
NNE	1.26	1.11	7.80	0.40	1.78	6.61
NE	1.35	0.97	7.10	0.40	1.31	4.95
ENE	2.07	1.35	8.60	0.40	0.74	3.23
E	2.47	1.55	10.50	0.40	0.36	2.42
ESE	1.38	1.10	6.40	0.40	1.18	3.78
SE	1.31	1.24	6.70	0.40	1.53	4.56
SSE	1.77	1.67	8.10	0.40	1.14	3.21
S	1.86	1.74	9.40	0.40	1.17	3.48
SSW	3.34	2.45	13.10	0.40	0.68	2.88
SW	4.13	2.56	12.70	0.40	0.26	2.39
WSW	2.42	1.62	8.90	0.40	0.35	2.29
W	2.28	1.82	12.20	0.40	0.77	3.37
WNW	1.19	1.26	10.30	0.40	2.04	7.73
NW	1.12	0.96	9.10	0.40	2.21	10.27
NNW	1.68	1.39	10.00	0.40	1.25	4.38

maximum wind speed of 13.10 m/s (i.e., the SSW direction). In addition, Table 3 provides a description of the wind speed data characteristics within 16 sectors. Comparing Tables 1 and 3, it can be seen that the wind speed characteristics in each direction differ significantly from the all-directional wind speed characteristics, and all of them exhibit certain non-Gaussian characteristics, especially in the NW direction which has the maximum skewness and kurtosis. This comparison contributes to examining the effectiveness and superiority of the KDE-GPD model in a comprehensive way. Table 4 in Appendix presents the annual cumulative frequency of wind speed and wind direction provided by the KDE-GPD distribution.

Based on the above zoning results, Fig. 8 shows the results of KDE-GPD parameters $\vartheta[\theta]$ for different wind directions. Since strong winds mainly occur in the SSW and SW directions, the thresholds are larger in these two directions. However, the scale parameter and shape parameter are not only related to the threshold, but also ensure that the mix function is continuous and derivable at the threshold point, as shown in Eq. (12). In addition, the value of the bandwidth parameter under each wind direction is related to the total amount of data in the corresponding direction and the standard deviation of the data, as shown in Eq. (3). The fitting regarding the relative frequency of wind direction and the parameter vector is performed via the harmonic function (Xu et al. 2009). Figure 9 displays the histogram of the relative frequency of wind direction (see Fig. 7) and the corresponding harmonic function fitting result.

Due to space constraint, Fig. 10 shows the fitting results based on Weibull distribution and KDE-GPD distribution in some typical directions where the goodness-of-fit results are also provided. As can be seen from Fig. 10, when considering the wind directionality, the distribution of wind speeds can be either unimodal or bimodal, i.e., the PDF with single-peak or double-peak. By comparison, the KDE-GPD distribution can better describe the actual wind speed data, and the goodness of fitting result is

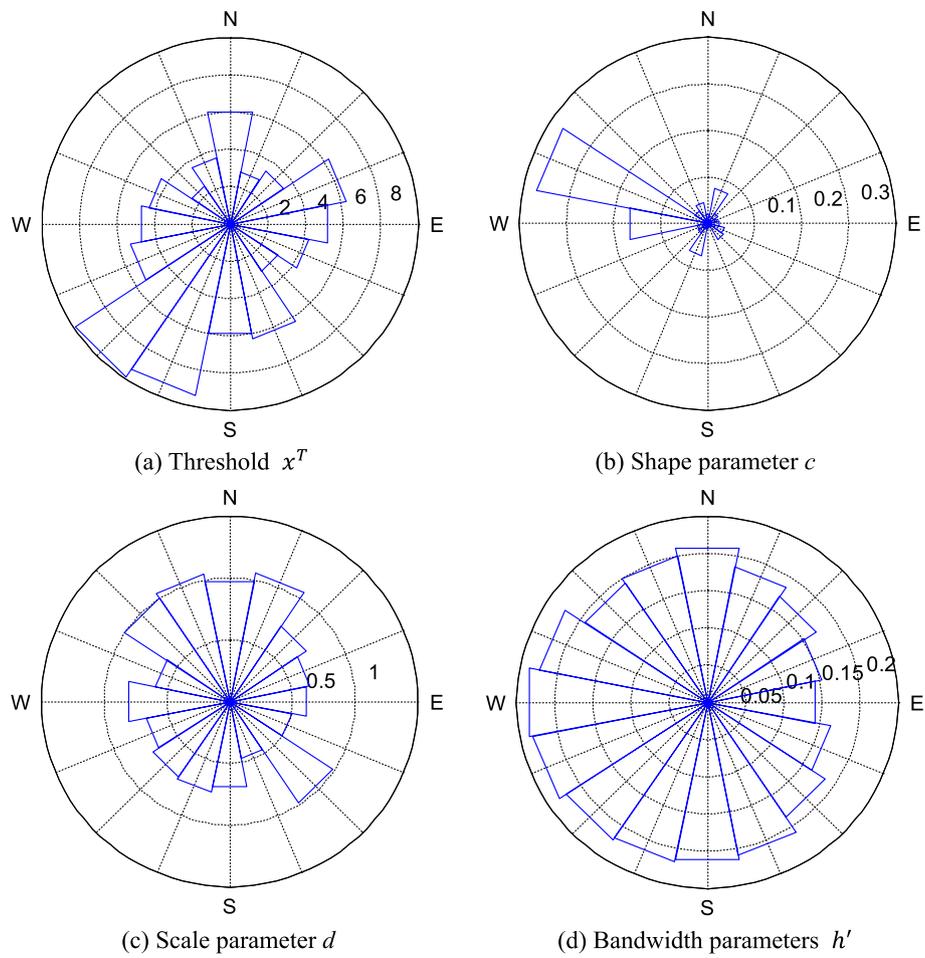


Fig. 8 Parameters of KDE-GPD in different directions

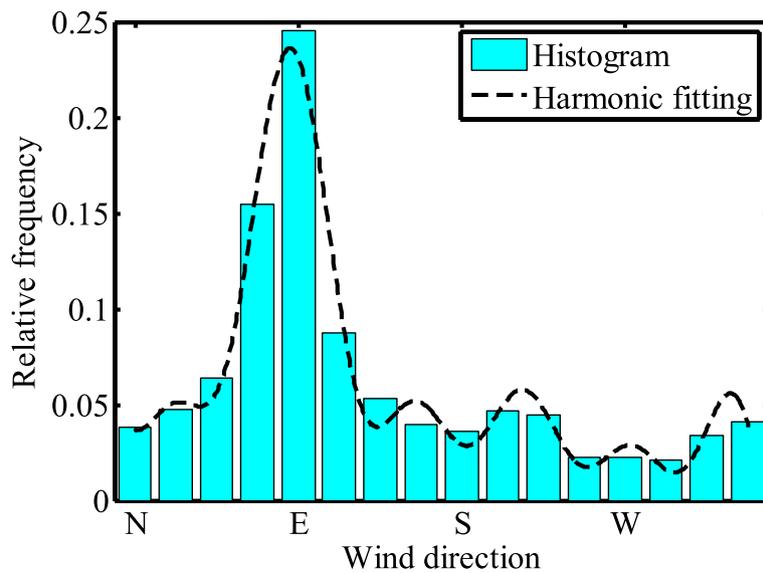


Fig. 9 Harmonic fitting of relative frequency of wind direction

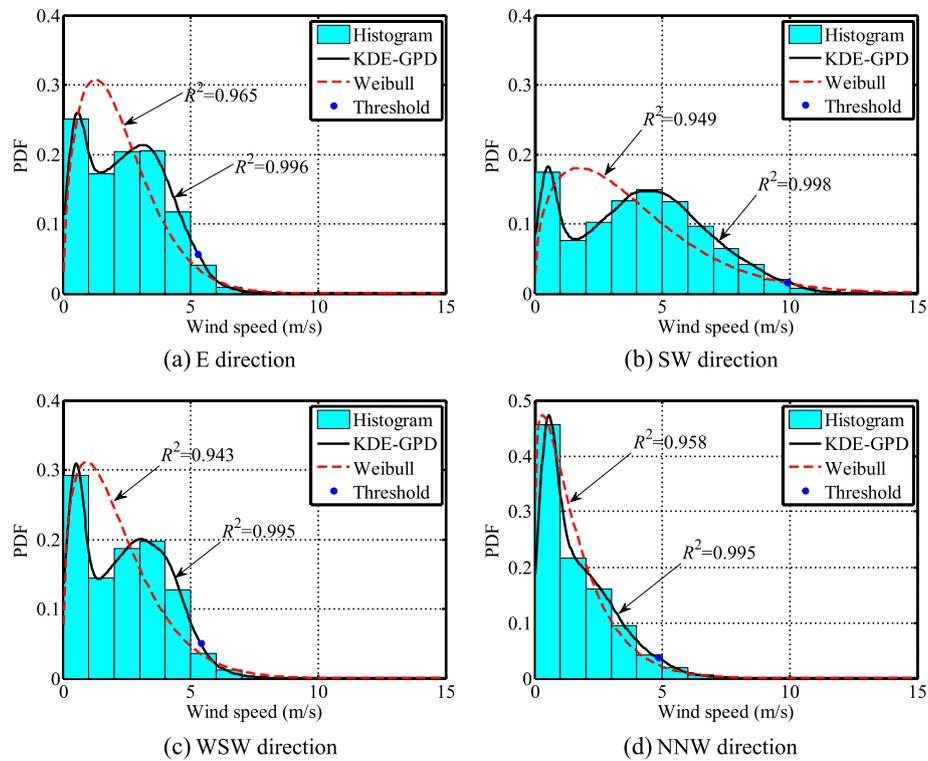


Fig. 10 Comparison of Weibull and KDE-GPD distributions with the actual one (considering wind directions)

also larger than that of Weibull distribution. This KDE-GPD distribution model can not only ensure an accurate description of the distribution of the main part of wind speeds, but also better reflect the distribution of the extreme part of wind speeds. It is worth stating that although occurrences of strong winds are infrequent, the fatigue damage under their effect is significant; On the other hand, the fatigue damage under the main part wind speeds is minimal, but its occurrences are more frequent. Therefore, an accurate description of the wind speed distribution is critical for ensuring reliable wind-induced buffeting response calculation as well as its fatigue damage estimation. In summary, the mix wind speed distribution model based on KDE-GPD still has a good fitting accuracy with the consideration of the impact of wind direction.

4 Conclusion

In order to obtain the distribution of wind speed and direction for structural durability design (e.g., wind-induced fatigue analysis) in mountainous areas, this paper proposes a new mix distribution model based on the combination of nonparametric Kernel Density Estimation (KDE) and Generalized Pareto Distribution (GPD) for addressing the complex wind speed distribution. Then, the distribution of wind direction is described via the conditional probability model. Finally, the measurement of 10-min average wind speeds at a mountainous bridge site is used as the experimental data to evaluate the performance of the proposed method. The main conclusions are as follows:

- (1) A semi-parametric mix distribution model (the KDE-GPD distribution model) is proposed to describe the probability density function of parent average wind speeds, and the calculation method of model parameters is given. The mix model can simultaneously ensure that the function is continuous and derivable at the threshold point. In this model, the KDE model is used to fit the main part of wind speeds, while the GPD model is employed to describe the extreme part of wind speeds.
- (2) According to the measured wind field characteristics in this mountainous area, it is found that when the wind directionality is not considered, the wind speed presents a significant single-peak distribution; and when considering the wind directionality, the wind speed distribution can exhibit a bimodal characteristic in some scenarios.
- (3) Based on the analysis of measured wind speed samples with single and double peaks shows that the model proposed in this paper can better describe the distribution of wind speed and direction with different wind speed data characteristics, and its performance is obviously better than the commonly-used Weibull distribution model.
- (4) The research in this paper can provide some references for guiding the analysis of engineering problems such as the probability distribution of wind speed in other regions for wind-induced fatigue of structures. However, the research is based on measured wind speed data in a mountainous area, and may not be entirely applicable to other regions. Meanwhile, the influence of parameters and kernel functions should be estimated. Additionally, the distribution of wind direction also deserves further research using some advanced models (e.g., Von Mises or coupla model) for better showing the joint PDF of wind speed and direction.

Appendix

Table 4 Annual cumulative frequency of wind speed and wind direction (KDE-GPD)

Wind direction	0–2 (m/s)	2–4 (m/s)	4–6 (m/s)	6–8 (m/s)	8–10 (m/s)	10–12 (m/s)	12–14 (m/s)	Total
N	1368	397	187	62	8	1	0	2023
NNE	1956	453	68	11	1	0	0	2489
NE	2582	725	54	4	0	0	0	3365
ENE	4225	3129	718	63	3	0	0	8138
E	5230	5355	2174	138	7	0	0	12905
ESE	3346	1112	139	3	0	0	0	4600
SE	2122	544	120	15	2	0	0	2802
SSE	1337	421	273	37	1	0	0	2068
S	1176	433	234	50	3	0	0	1896
SSW	838	705	542	257	104	17	2	2464
SW	555	559	665	390	150	28	2	2348
WSW	496	463	217	14	1	0	0	1191
W	582	365	213	23	4	1	0	1189

Wind direction	0–2 (m/s)	2–4 (m/s)	4–6 (m/s)	6–8 (m/s)	8–10 (m/s)	10–12 (m/s)	12–14 (m/s)	Total
WNW	891	174	54	4	1	0	0	1125
NW	1514	242	27	5	1	0	0	1788
NNW	1410	608	128	20	3	0	0	2169
Total	29627	15685	5811	1096	288	49	4	52560

Acknowledgements

The supports by “China Postdoctoral Science Foundation (grant number: 2023M730431)”, “Special Funding of Chongqing Postdoctoral Research Project (grant number: 2022CQBSHTB2053)”, “the Science and Technology Research Program of Chongqing Municipal Education Commission (grant number: KJZD-M202300201)”, “State Key Laboratory of Mountain Bridge and Tunnel Engineering (grant number: SKLBT-2202)”, “Chongqing graduate education teaching reform research general project (grant number: YJG233030)”, “Wind Engineering Re-search Center of Sichuan Key Laboratory (grant number: WEKLS202303)” and “the Funda-mental Research Funds for the Central Universities (grant number: SWU-KT22021)” are greatly acknowledged.

Authors’ contributions

Conceptualization, Cheng Cheng, Yan Jiang and Jian Liu; Data curation, Cheng Cheng, Beilong Luo, Yan Jiang and Jian Liu; Formal analysis, Cheng Cheng, Beilong Luo, Yan Jiang and Jian Liu; Funding acquisition, Yan Jiang; Investigation, Cheng Cheng, Beilong Luo and Yan Jiang; Methodology, Cheng Cheng and Yan Jiang; Resources, Cheng Cheng and Ming Li; Software, Cheng Cheng and Yan Jiang; Supervision, Beilong Luo and Yan Jiang; Validation, Yan Jiang and Ming Li; Visualization, Cheng Cheng; Writing – original draft, Cheng Cheng, Beilong Luo and Yan Jiang; Writing – review & editing, Cheng Cheng, Beilong Luo and Yan Jiang. All authors have reviewed and approved the final manuscript. All authors have read and agreed to the published version of the manuscript.

Funding

This research was funded by “China Postdoctoral Science Foundation (grant number: 2023M730431)”, “Special Funding of Chongqing Postdoctoral Research Project (grant number: 2022CQBSHTB2053)”, “the Science and Technology Research Program of Chongqing Municipal Education Commission (grant number: KJZD-M202300201)”, “State Key Laboratory of Mountain Bridge and Tunnel Engineering (grant number: SKLBT-2202)”, “Chongqing graduate education teaching reform research general project (grant number: YJG233030)”, “Wind Engineering Re-search Center of Sichuan Key Laboratory (grant number: WEKLS202303)” and “the Funda-mental Research Funds for the Central Universities (grant number: SWU-KT22021)”.

Availability of data and materials

The data used to support the findings of this study are available from the corresponding author upon request.

Declarations

Competing interests

All authors declare that they have no competing interests.

Received: 14 March 2024 Accepted: 15 April 2024

Published online: 15 May 2024

References

- Carta JA, Ramirez P, Bueno C (2008) A joint probability density function of wind speed and direction for wind energy analysis. *Energy Convers Manage* 49(6):1309–1320
- Chen Y, Duan Z (2018) A statistical dynamics track model of tropical cyclones for assessing typhoon wind hazard in the coast of southeast China. *J Wind Eng Ind Aerodyn* 172:325–340
- Chiodo E, Noia LPD (2020) Stochastic extreme wind speed modeling and bayes estimation under the inverse Rayleigh distribution. *Appl Sci* 10(16):5643
- Cook NJ (2021) Implications of the OEN mixture model of the mean wind vector for the generation of synthetic time-series and for the assessment of extremes. *J Wind Eng Ind Aerodyn* 208:104424
- Ding J, Chen X (2014) Assessment of methods for extreme value analysis of non-Gaussian wind effects with short-term time history samples. *Eng Struct* 80:75–88
- Ding J, Chen X, Zuo D, Hua J (2016) Fatigue life assessment of traffic-signal support structures from an analytical approach and long-term vibration monitoring data. *J Struct Eng* 142(6):04016017
- Epanechnikov VA (1969) Non-parametric estimation of a multivariate probability density. *Theory Probab Appl* 14(1):153–158
- Erdem E, Shi J (2011) Comparison of bivariate distribution construction approaches for analysing wind speed and direction data. *Wind Energy* 14(1):27–41
- Harris I (2005) Generalised Pareto methods for wind extremes. Useful tool or mathematical mirage. *J Wind Eng Ind Aerodyn* 93(5):341–360
- Holmes JD (2020) Comparison of probabilistic methods for the effects of wind direction on structural response. *Struct Saf* 87:101983

- Holmes JD, Moriarty WW (1999) Application of the generalized Pareto distribution to extreme value analysis in wind engineering. *J Wind Eng Ind Aerodyn* 83(1–3):122–128
- Hu P, Han Y, Xu G, Cai CS, Cheng W (2020) Effects of inhomogeneous wind fields on the aerostatic stability of a long-span cable-stayed bridge located in a mountain-gorge terrain. *J Aerosp Eng* 33(3):04020006
- Huang G, Peng L, Su Y, Liao H, Li M (2015) A wireless high-frequency anemometer instrumentation system for field measurements. *Wind Struct* 20(6):739–749
- Huang G, Jiang Y, Peng L, Solari G, Liao H, Li M (2019) Characteristics of intense winds in mountain area based on field measurement: focusing on thunderstorm winds. *J Wind Eng Ind Aerodyn* 190:166–182
- Hyndman RJ, Bashtannyk DM, Grunwald GK (1996) Estimating and visualizing conditional densities. *J Comput Graph Stat* 5(4):315–336
- Isumov N, Ho E, Case P (2014) Influence of wind directionality on wind loads and responses. *J Wind Eng Ind Aerodyn* 133:169–180
- Jiang Y, Huang G (2017) Short-term wind speed prediction: hybrid of ensemble empirical mode decomposition, feature selection and error correction. *Energy Convers Manage* 144:340–350
- Jiang Y, Zhao N, Peng L, Liu S (2019) A new hybrid framework for probabilistic wind speed prediction using deep feature selection and multi-error modification. *Energy Convers Manage* 199:111981
- Jiang Y, Liu S, Zhao N, Xin J, Wu B (2020) Short-term wind speed prediction using time varying filter-based empirical mode decomposition and group method of data handling-based hybrid model. *Energy Convers Manage* 220:113076
- Kenfack-Sadem C, Tagne R, Pelap FB, Nfor Bawe G (2021) Potential of wind energy in Cameroon based on Weibull, normal, and lognormal distribution. *Int J Energy Environ Eng* 12(4):761–786
- Liao H, Jing H, Ma C, Tao Q, Li Z (2020) Field measurement study on turbulence field by wind tower and windcube lidar in mountain valley. *J Wind Eng Ind Aerodyn* 197:104090
- Liu S, Luo Y, Peng L, Jiang Y, Meng E, Li B (2022) Wind pressure field reconstruction based on unbiased conditional kernel density estimation. *J Wind Eng Ind Aerodyn* 223:104947
- Luo Y, Huang G, Han Y, Cai CS (2021) Evaluation of full-order method for extreme wind effect estimation considering directionality. *Wind Struct* 32(3):193
- McWilliams B, Newmann MM, Sprevak D (1979) The probability distribution of wind velocity and direction. *Wind Eng* 269–273
- Nage GD (2016) a Comparative study of Weibull to Rayleigh probability density function; a case of two sites in Ethiopia. *Am J Modern Energy* 2(3):10–16
- Nguyen CH, Owen JS, Franke J, Neves LC, Hargreaves DM (2021) Typhoon track simulations in the North West Pacific: informing a new wind map for Vietnam. *J Wind Eng Ind Aerodyn* 208:104441
- Özkan R, Sen F, Balli S (2020) Evaluation of wind loads and the potential of Turkey's south west region by using log-normal and gamma distributions. *Wind Struct* 31(4):299–309
- Ragan P, Manuel L (2008) Statistical extrapolation methods for estimating wind turbine extreme loads. *J Sol Energy Eng* 130(3):031011
- Ren W, Pei C, Ma C, Li Z, Wang Q, Chen F (2021) Field measurement study of wind characteristics at different measuring positions along a bridge in a mountain valley. *J Wind Eng Ind Aerodyn* 216:104705
- Repetto MP, Solari G (2010) Wind-induced fatigue collapse of real slender structures. *Eng Struct* 32(12):3888–3898
- Simiu E, Heckert NA, Filliben JJ, Johnson SK (2001) Extreme wind load estimates based on the Gumbel distribution of dynamic pressures: an assessment. *Struct Saf* 23(3):221–229
- Takle ES, Brown JM (1978). Note on the use of Weibull statistics to characterize wind-speed data. *J Appl Meteorol* (1962-1982):556-559
- Tang H, Li Y, Shum KM, Xu X, Tao Q (2020) Non-uniform wind characteristics in mountainous areas and effects on flutter performance of a long-span suspension bridge. *J Wind Eng Ind Aerodyn* 201:104177
- Tao T, Wang H (2023) Efficient buffeting analysis of long-span bridges under non-stationary winds: a 2D interpolation enhanced approach. *J Sound Vib* 559:117754
- Wais P (2017) A review of Weibull functions in wind sector. *Renew Sustain Energy Rev* 70:1099–1107
- Wang Q, Gu M (2009) Study on the joint distribution function of wind speed and wind direction for structural wind fatigue analysis. *Struct Eng* 25(06):98–103
- Wang H, Wang L, Fan X, Tao T, Zong Z (2013) Study of joint distribution of wind speed and direction of sutong bridge based on SHMS. *Bridge Construct* 43(05):55–61
- Wu J, Zheng Q, Fu J (2017) Reliability analysis of wind-induced response of high-rise buildings based on the action of joint probability distribution of wind speed and wind direction. *J Build Struct* 38(10):88–94
- Xu YL, Zhu LD (2005) Buffeting response of long-span cable-supported bridges under skew winds. Part 2: case study. *J Sound Vib* 281(3–5):675–697
- Xu Y, Liu T, Zhang W (2009) Buffeting-induced fatigue damage assessment of a long suspension bridge. *Int J Fatigue* 31(3):575–586
- Yang Y, Ge Y, Xiang H (2002) Statistical analysis of mean wind for joint wind speed and direction distribution. *Struct Eng* 46(03):29–36
- Yu J, Li M, Li S, Li H (2018) Determination of design wind speed for mountain bridges considering joint wind speed and wind direction distribution and terrain effect. *Chin J Highways* 31(08):122–128
- Zhang X, Chen X (2015) Assessing probabilistic wind load effects via a multivariate extreme wind speed model: a unified framework to consider directionality and uncertainty. *J Wind Eng Ind Aerodyn* 147:30–42
- Zhang F, Liu Y, Chen C et al (2014) Fault diagnosis of rotating machinery based on Kernel density estimation and Kullback-Leibler divergence. *J Mech Sci Technol* 28(11):4441–4454
- Zhang J, Zhang M, Li Y, Qin J, Wei K, Song L (2020) Analysis of wind characteristics and wind energy potential in complex mountainous region in southwest China. *J Clean Prod* 274:123036
- Zhao N, Huang G, Liu R, Peng L (2020) A remote long-term and high-frequency wind measurement system: design, comparison and field testing. *Wind Struct* 31(1):21

- Zheng X, Li H, Li C, Liu Y, Zhang H (2019a) Joint probability distribution and application of wind speed and direction based on multiplication rule and AL model. *Eng Mech* 36(10):50–57+85
- Zheng G, Han Y, Cai C (2019b) Study on the joint distribution of wind speed and wind direction of Qiaozhai Bridge. *China Foreign Highway* 39(05):80–85
- Zhou Q, Wang H, Wang W (2021) Modeling of wind speed probability distribution based on peak pattern identification. *J Solar Energy* 42(08):355–360
- Zhu L, Xu Y (2005) Buffeting response of long-span cable-supported bridges under skew winds. Part 1: theory. *J Sound Vib* 281(3–5):647–673

Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.