



AKADÉMIAI KIADÓ

The impact of parameter estimation uncertainty in extreme wind speed models

Ágnes Kenéz*  and Attila László Joó

Department of Structural Engineering, Faculty of Civil Engineering, Budapest University of Technology and Economics, Budapest, Hungary

Received: April 2, 2022 • Revised manuscript received: June 25, 2022 • Accepted: July 15, 2022

Published online: December 21, 2022

Pollack Periodica •
An International Journal
for Engineering and
Information Sciences

18 (2023) 1, 78–83

DOI:

[10.1556/606.2022.00645](https://doi.org/10.1556/606.2022.00645)

© 2022 The Author(s)

ORIGINAL RESEARCH
PAPER



ABSTRACT

All the available information and uncertainties should be taken into account in a model to give correct answer to a stated problem and evaluate the performance of a structure. This study deals with the impact of parameter estimation uncertainty in extreme wind speeds on the assessed reliability index using frequentist approach. The peak-over-threshold approach with an automated threshold selection method is applied and bootstrapping is used to determine the 95% confidence interval of the estimated reliability index. Based on the results practical recommendations, i.e., a framework of this procedure are derived on how to handle peak-over-threshold in extreme wind speed models for the application of the proposed performance-based wind engineering design.

KEYWORDS

extreme wind speed, parameter estimation uncertainty, peaks-over-threshold method, generalized pareto distribution

1. INTRODUCTION

1.1. Motivation

Uncertainties are unavoidable in the design of engineering structures; therefore the analysis should include the proper treatment and effects of uncertainties. Accordingly, different level of modeling can be distinguished: deterministic, semi-probabilistic (level I), First Order Reliability Method (FORM), (level II) and full probabilistic method (level III). The Eurocode has primarily been based on the deterministic method and the semi-probabilistic (level I) method has been used for further development [1]. Hence, the uncertainties of both load effects and structural resistance model are taken into account in the structural design process nowadays. Although, generally, various assumptions and simplifications are made in the probabilistic model and these uncertainties are included implicitly, e.g., represented by partial factors and characteristic values of actions, material-, geometrical properties, in non-linear models, the precise probabilistic models (level III) provide the only unbiased estimate of the probability failure. In the partial factor method (level I), a verification is made to ensure that no relevant limit state will be exceeded. Nevertheless, if one would like to evaluate the performance of a structure, all the available information should be taken into account and full probabilistic method should be applied [2]. This is the goal of the Performance-Based Design (PBD). While the Performance-Based Earthquake Engineering (PBEE) is well accepted by now, the Performance-Based Wind Engineering (PBWE) is also becoming increasingly available and desirable [3–8] for those structures where the wind is the governing load.

1.2. Problem statement

Extreme wind speed is the dominant component in the reliability of many structures, i.e., the uncertainty in the probabilistic model of the extreme wind speed contributes the most to the failure probability [6, 9]. Therefore, it is of utmost importance to examine the related

*Corresponding author.

E-mail: kenez.agnes@emk.bme.hu

modeling assumptions and to make sure that the model aptly describes reality. Despite its importance, Parameter Estimation Uncertainty (PEU) in extreme wind speed models is rarely considered. PEU comes from the model parameters whose exact values are unknown and whose values cannot be exactly inferred by statistical methods. Yet, some studies have shown that PEU in the modeling of extreme environmental actions can play a considerable role [10–14], e.g., Rózsás and Sýkora, [12] show that neglecting of PEU can lead to a 10% underestimation of the 1000-year return period wind speed value. Other study of Rózsás and Sýkora, [13] shows that the 1000-year return period value might be underestimated by 20% due to the neglect of PEU for ground snow load.

However, Meinen and Steenbergen [14] have focused on the influence of PEU in extreme wind speeds in the context of PBWE. Hence this paper aims to fill this gap and answer the following research question: *What is the impact of PEU in extreme wind speed models on structural reliability?* The question is broken into smaller sub-questions and answers, which are answered accordingly.

1.3. Approach

Briggs et al. [15] make recommendation on the reporting of uncertainty, in terms of both probabilistic methods and deterministic sensitivity analysis techniques. In this study, the Probabilistic Sensitivity Analysis (PSA) is used to quantify the level of confidence in the structural reliability. Input parameter values are the parameters of the extreme value distribution, which are fitted to realizations. Fragility curves are developed and further integrated with the hazard functions to estimate structural failure probabilities. The output is the empirical distribution histogram of the reliability index and it is represented by its point estimate and the 95% confidence interval.

Based on the following results practical recommendations are derived on how to handle PEU in extreme wind speed models.

2. WIND SPEED DATA SET

A 3.5-year record of wind speed data with a sampling interval of 0.9 s, measured at 50 m height above ground level in Sződliget, Hungary is used for the analysis. The data was provided by Hungarian Telekom Telecommunications Plc. Several approaches exist to determine the extreme value distribution of wind speed [16, 17]. In this study Peak-Over-Threshold (POT) method is used, which assumes mutually independent and identically distributed (i.i.d.) random variables. To ensure independence of the observations, an autocorrelation analysis is adopted. The analysis reveals a slight periodicity in case of maxima for 1 and 3 days, though 7-day maxima can be considered as statistically independent events. Therefore, maxima for 1, 3 and 7 days are analyzed in this study to check the impact of weakly dependent observations. These selections comprise 1,326, 441 and 189 samples, respectively.

3. PROPOSED METHODS AND TOOLS

3.1. Hazard and fragility curves

The hazard function is defined as the probability of exceedance within the reference time interval. In the present hazard functions, only the uncertainties associated with the wind speed values are taken into account and the uncertainties of e.g., roughness factor, gust factor, aerodynamic shape factor, etc., are neglected.

Fragility curves represents the probability of exceeding the given limit state, i.e., damage state as a function of the chosen intensity measure parameter, i.e., wind speed. The Log-Normal Distribution (LND) is chosen for the representation of the resistance as proposed by EN 1990:2002 C.6 [18]. Adopting a log-normal distribution for this variable has the advantage that no negative values can occur.

3.2. Parameter estimation

All models have parameters that need to be estimated. Therefore, the estimation of point estimates and uncertainty in parameters is part of the modeling process and the relationship between the parameter uncertainty and the uncertainty of the decision variable should be determined [15]. In this study, maximum likelihood method is used to evaluate the parameters of the different distributions. Three sources of uncertainty can be taken into account during modeling:

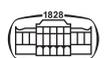
- i. aleatory uncertainty, which cannot be eliminated (arising from the unpredictable nature of a variable);
- ii. epistemic uncertainty, which can be reduced by gathering more data or by refining the model (resulted by the incompleteness and errors of measurements);
- iii. model uncertainties (due to the limitations of our knowledge).

In this study, the impact of the epistemic and model uncertainty in the proposed extreme wind speed model on the reliability index is assessed.

As it was mentioned in Section 2, the applied POT method is based on a conditional distribution (Eq. 1), i.e., exceedances over a specified threshold. The Generalized Pareto Distribution (GPD) is used to model the behavior of the wind speed exceedances over the specified threshold [19–23]. The Cumulative Distribution Function (CDF) of GPD:

$$F_{\xi, \sigma, \eta}(x) = P\{X - u \leq x | X > u\} = \begin{cases} 1 - \left(1 + \frac{\xi(x - \eta)}{\sigma}\right)^{-\frac{1}{\xi}}, & \text{for } \xi \neq 0, \\ 1 - \exp\left(-\frac{x - \mu}{\sigma}\right), & \text{for } \xi = 0, \end{cases} \quad (1)$$

where u is the selected threshold; ξ , σ and η are the shape, scale and location parameters, respectively. The shape parameter ξ of the GPD is the same as for the Generalized Extreme Value (GEV) distribution [19]. According to the recommendation of Thompson et al. [24], for the automated threshold selection, suitable values of equally spaced



candidate thresholds should be chosen between the median and the 98% quantile of the dataset, unless fewer than 100 values exceed this value, in which case the upper bound should be set to the 100th data value in descending order.

The procedure of the applied analysis is discussed in detail in Section 3.3. As the 95% confidence interval of the estimated reliability index is not necessarily symmetric, it is assessed using bootstrapping.

3.3. Framework of the procedure

Since this study focuses on the influence of parameter estimation uncertainty on the estimated reliability index, the failure probability is quantified in terms of functions characterizing the wind effect and a general structure corresponding to a level of safety for reliability class RC2 with reliability index β of 3.8 [18] (Eq. 2). Figure 1 shows the framework of the analysis, which can be summarized in the following steps:

1. For the representation of the action side, GEV distribution or GPD is applied for synthetic data or for real observation, respectively;
2. For the representation of the resistance side, the LND is applied. The Coefficient of Variation (CoV) of the LND is chosen and the mean is calculated with the intention to reach the recommended minimum value of the reliability index 3.8 for ultimate limit state verification;
3. In order to investigate the impact of a large number of wind speed realizations on the assessed uncertainty, synthetic data is generated using a pseudorandom number generator. However, for other calculations, real observations are used;
4. A bootstrapping technique with 1,000 resamples is applied for the estimation of the distribution of the reliability index. A particular sample is randomly chosen with replacement from the data;
5. The POT approach with an automatic threshold selection method [24] is used to determine parameters of the Probability Density Function (PDF) of the wind speed.

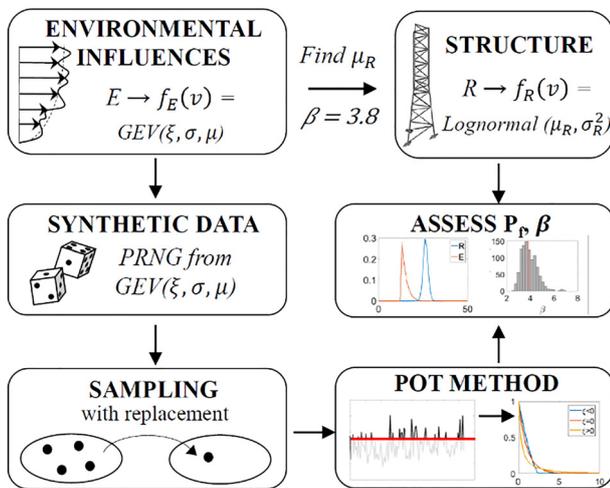


Fig. 1. Framework of the procedure

This automated technique is computationally inexpensive and simple;

6. The probability of failure and the corresponding reliability index are calculated according to the following equations (Eq. (2) for GEV, Eq. (3) for GPD):

$$P_f = \int_0^\infty f_E(v)F_R(v)dv, \tag{2}$$

$$P_f = \int_0^\infty f_E(v|v>u)P(v>u)F_R(v)dv, \tag{3}$$

$$\beta = -\Phi^{-1}(P_f), \tag{4}$$

where $f_E(v)$ is the PDF of the GEV distribution, $F_R(v)$ is the CDF of LND, $f_E(v|v > u)$ is the probability density function (PDF) of GPD, $Pr(v > u)$ is approximated by the ratio of number of exceedances to the number of observations and Φ is the standard normal distribution function. After that, the bootstrapped 95% confidence interval of the reliability index is produced.

4. ANALYSIS RESULTS

PSA is carried out and statistical uncertainty due to the use of limited wind speed samples during the application of POT method is determined and presented below.

4.1. The effect of the number of realizations

First, the effect of number of realizations on the estimated reliability index is investigated on synthetic data. For the representation of the action side, GEV distribution with shape parameter -0.1706 , scale parameter 2.7 and location parameter 11.69 is chosen (based on real observation). For the representation of the resistance side, LND with CoV of 0.1 is chosen and the mean of this log-normal distribution is 28.16 m s^{-1} .

The expected value of the estimate converges to the true value as the number of realizations increases and the bandwidth shrinks accordingly (Fig. 2). The horizontal line represents the target reliability index β of 3.8. The associated 95% confidence interval is reduced by about 70% when the number of realizations is increased from 100 to 1,000. Further reduction ($\sim 50\%$) can be achieved when the number of realizations is increased by an order of magnitude.

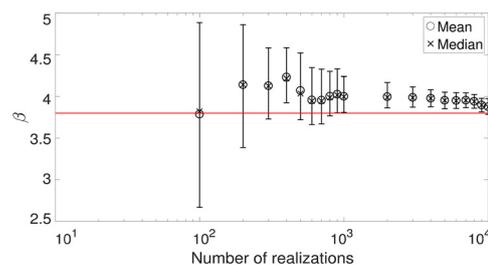


Fig. 2. 7-day maxima, various record length and upper bound of suitable threshold



Empirical distribution histograms of the reliability index β can differ from normal distribution, especially in case of small number of realizations (less than 100 realizations).

4.2. Short time series

With the advancement of technology, the installation of monitoring systems is becoming the norm in the industry. Site-specific probability distribution of wind can be taken into account for the performance-based wind design of structures. Therefore, from a practical point of view, the influence of short time series length available for the analysis on the final results can be interesting to determine a minimum observation-length, i.e., minimum number of realizations or exceedances.

In case of the abovementioned 7-day maxima and shorter dataset, 198 data or fewer are given, which is not sufficient for the recommendation of Thompson et al. [24]. Hence, some modification should be made to prevent the fitting of parameters based on low number of observations. Since reliable results are obtained in the range of 8.5 m s^{-1} and 12.5 m s^{-1} for 3-day maxima (40–80% quantiles), 10.5 m s^{-1} and 13 m s^{-1} for 7-day maxima (40–70% quantiles), lower bound should be still the median, and the chosen upper bound is the 70% quantile for shorter (than 200 data) dataset. Maximum numbers of exceedances of the potential upper bounds are shown in Table 1.

Realizations are resampled from the observation with replacement to obtain the reliability indices and assess the bootstrapped confidence interval. The exact discrepancy between results using different maxima can be seen in Fig. 3.

According to the expectation, the case of 7-day maxima carries a higher level of uncertainty due to fewer data points. The 95% confidence interval of β is reduced by about 30–40% using 3-day maxima and 20–44% using 1-day maxima. Moreover, the associated confidence interval is reduced by ~40, 30 and 40% when the observation length is increased from 1 to 3.5 year for 1-, 3- and 7-day maxima, respectively. The accuracy may not increase with the frequency of observation due to the different threshold.

Table 1. Range of numbers of exceedances

	1 year	2 year	3 year	3.5 year
1-day-maxima	113–189	227–378	341–568	398–663
2-day-maxima	38–63	75–126	113–189	132–220
3-day-maxima	16–27	32–54	48–81	57–94

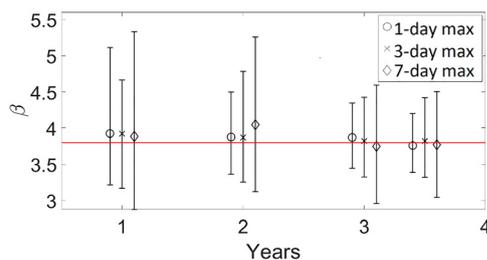


Fig. 3. Assessed β of 1–3 and 7-day maxima, various record length

4.3. Various CoV of the resistance side

Parameters of the GPD are fitted on the real 3.5-year wind velocity record. The probability density function (PDF) of the GPD and various LND are shown in Fig. 4. Realizations are resampled from the observation with replacement.

The confidence interval of the reliability index is decreasing with increasing CoV of the LND due to spread of the distribution function. In other words, uncertainties of the action side have smaller impact on the assessed uncertainty of the reliability index when there is a larger uncertainty associated with the resistance side (Fig. 5). For example, the CoV of the distribution of tensile strength of concrete can be assumed to be 0.3, while the CoV corresponding to material properties of structural steel is about 0.03–0.07 [25].

4.4. Mean and short-term velocity

Ideally, 10 min mean of a continuously measured wind speed is required for the Eurocode standard [26], but in practice this is rarely achieved. Therefore, results of two cases are compared to see its effect; extreme value analysis is carried out on:

- i. 10 min averaged wind data points;
- ii. instantaneous wind speed (sampling interval of 0.9 s).

In Fig. 6, the large influence of extreme events on the final result can be seen. Taking into account these extremes in the analysis may lead to different inference results and this behavior causes the increased statistical uncertainty of

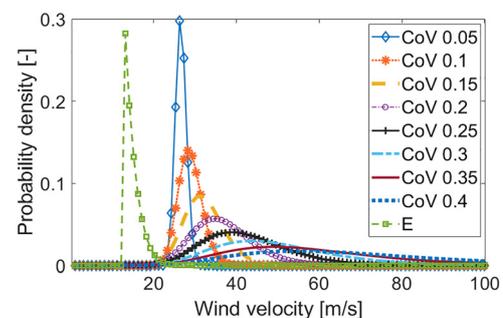


Fig. 4. The PDF of the GPD and various LND representing the resistance

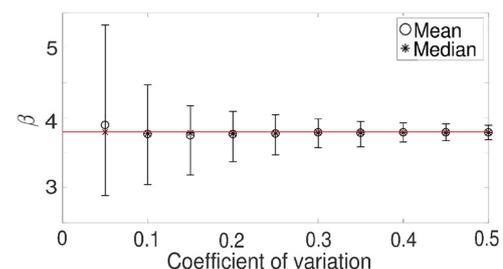


Fig. 5. Assessed β associating with various CoV of the LND function



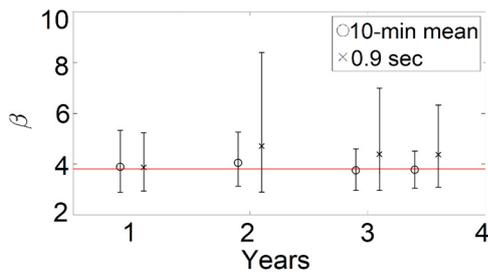


Fig. 6. 7-day maxima, measured and 10-min averaged wind speeds, various record length

the 0.9-s gust dataset. If one would like to examine 10-min mean wind speeds, then this uncertainty should be included in the exposure factor.

5. DISCUSSION

It should be noted that the reality of multiple observations suggests that reliance on a single study probably underestimate the actual uncertainty. This suggests a more extensive uncertainty analysis than based on study data alone [15]. Moreover, adding one or two extreme events to the sample may have significant influence on the final result. The present statements are valid for the dataset under consideration and for similar climatic conditions.

The traditional method to derive extreme wind speeds is the Annual Maximum (AM) approach. The great advantage of this approach is that few decisions are required during the calculation of the distribution parameters. However, the main drawback is the data reduction; therefore, the wind measurement must be long. At least 10–20 extremes should be used to determine reliable results according to [27]. This drawback can be overcome through applying POT method, which is based on a conditional distribution, i.e., the exceedances over a specified threshold. In this way, more events per year can be applied for the estimation. Therefore, the POT approach is applied in this study to assess parameters of the extreme value distribution due to the relatively short length of data series. The complexity of POT method mainly is from the determination of both physical and statistical threshold since variance decreases and the bias increases with lower threshold, higher threshold results that the bias decreases and the variance increases. Hence, this decision may have strong impact on the estimated values.

In this study, the contribution of uncertainty associated with aerodynamic and aero-elastic phenomena to the failure probability is not taken into account, since it is focused on the epistemic uncertainties in the distribution parameters and quantile estimates of extreme wind speeds. Nevertheless, the interaction between the relevant properties of the structure and the wind field are essential and cannot be disregarded. Also, non-environmental actions and the presence of nearby structures can influence the structural

response by modifying the aerodynamic and aero-elastic characteristics of the structure.

6. SUMMARY AND CONCLUSIONS

During the application of the PBWE design, one main source of epistemic uncertainty associated with the extreme wind speed model is the statistical uncertainty due to the use of limited samples. It can be reduced by using more relevant data, i.e., longer measurement or alternative approaches, e.g., POT method instead of the Annual Maximum (AM). In this study the impact of the parameter estimation uncertainty of the wind speed function using POT method with automated threshold selection on the assessed reliability index is evaluated.

The failure probability is quantified in terms of the GPD PDF of wind speeds and the LND of a general structure corresponding to a level of safety for reliability class RC2 with reliability index β of 3.8. In this study the POT approach with an automated threshold selection method is applied. The bootstrapping is used to determine the 95% confidence interval of the estimated reliability index β .

Maxima for 1 and 3 days can be considered as weakly dependent observations, and 7-day maxima can be treated as statistically independent events. Weakly dependent realizations do not cause considerable discrepancy, however the increased amount of data can reduce the statistical uncertainty. The 95% confidence interval can be reduced by about 70 and 85% when the number of realizations is increased from 100 to 1,000 and 10,000, respectively. It was found that at least approx. 500 realizations, i.e., about two years of observation is needed for the analysis to provide sufficiently reliable results when the CoV of the LND of the resistance side is about 0.1. However, regression techniques or simulation modeling should be used to extend the available record for extreme value analysis by comparison with neighboring stations or using synthetic time series. If more than 3,000 realizations are available, then the effect of PEU can be neglected. However, increasing the uncertainty associated with the resistance side, uncertainties of the action side have smaller impact on the assessed uncertainty of the failure probability. Moreover, the available data type can have considerable influence on the final result. The confidence interval of the reliability index calculated using gust wind speed, is wider than the bandwidth corresponding to 10-min mean wind speeds. Therefore, considering 10-min mean wind speed records, this uncertainty should be taken into account in the exposure factor.

In conclusion, the main drawback of the traditional AM approach is the data reduction, which can be overcome through applying other techniques, such as the proposed, but not typically used in civil engineering, POT method presented in this paper. It may lead to more accurate return level estimates, i.e., basic wind velocity with their uncertainty properly qualified due to the increased number of samples available for analysis from a given time series. This epistemic,



statistical uncertainty is quantified, and the framework of this analysis is demonstrated.

REFERENCES

- [1] EN 1990:2002, Eurocode – Basis of structural design, standard owner, CEN, Brussels, 2002.
- [2] O. Abusalem, N. Bertalan, and B. Kocsi, “Implementing quantitative techniques to improve decision making in construction projects: A case study,” *Pollack Period.*, vol. 14, no. 3, pp. 223–234, 2019.
- [3] C. Paulotto, M. Ciampoli, and G. Augusti, “Some proposals for a first step towards a Performance Based Wind Engineering,” in *Proceedings of the International Forum in Engineering Decision Making*, Stoos, Switzerland, Dec. 5–9, 2004, pp. 50–60.
- [4] F. Petrini and M. Ciampoli, “Performance-based wind design of tall buildings,” *Struct. Infrastruct. Eng.* vol. 8, pp. 954–966, 2012.
- [5] M. F. Huang, C. M. Chan, and W. J. Lou, “Optimal performance-based design of wind sensitive tall buildings considering uncertainties,” *Comput. Struct.* vols. 98–99, pp. 7–16, 2012.
- [6] R. K. Tessari, H. M. Kroetz, and A. T. Beck, “Performance-based design of steel towers subject to wind action,” *Eng. Struct.* vol. 143, pp. 549–557, 2017.
- [7] W. C. Chuang and S. M. J. Spence, “A performance-based design framework for the integrated collapse and non-collapse assessment of wind excited buildings,” *Eng. Struct.* vol. 150, pp. 746–758, 2017.
- [8] H. M. Kroetz, R. K. Tessari, and A. T. Beck, “Performance of global metamodeling techniques in solution of structural reliability problems,” *Adv. Eng. Softw.* vol. 114, pp. 394–404, 2017.
- [9] S. Gerasimidis, E. Efthymiou, and C. C. Baniotopoulos, “On the application of robustness criteria to steel lattice masts,” *Pollack Period.*, vol. 4, no. 1, pp. 17–28, 2009.
- [10] M. Baravalle and J. Köhler, “On the probabilistic representation of the wind climate for calibration of structural design standards,” *Struct. Saf.*, vol. 70, pp. 115–127, 2018.
- [11] F. P. Bakker, N. De Koker, and C. Viljoen, “Preconditioning wind speeds for standardised structural design,” *Eng. Structures*, vol. 238, 2021, Paper no. 111856.
- [12] Á. Rózsás and M. Sýkora, “Effect of statistical uncertainties on predicted extreme wind speeds,” in *Proceedings on 7th Int. Work. Reliab. Eng. Comput.*, Bohum, Germany, June 15–17, 2016, pp. 489–503.
- [13] Á. Rózsás and M. Sýkora, “Effect of statistical uncertainties in ground snow load on structural reliability,” in *IABSE Conf. on Structural Engineering: Providing Solutions too Global Chalanges*, Geneva, Switzerland, Sept. 23–25, 2015, pp. 220–227.
- [14] N. Meinen and R. D. J. M. Steenbergen, “A reliability based assessment of the partial factor for wind loads in the Eurocode framework,” *Heron*, vol. 63, no. 3, pp. 243–302, 2015.
- [15] A. H. Briggs, M. C. Weinstein, E. A. L. Fenwick, J. Karnon, M. J. Sculpher, and A. D. Paltiel, “Model parameter estimation and uncertainty analysis: A report of the ISPOR-SMDM modeling good research practices task force working group-6,” *Value in Health*, vol. 15, pp. 835–842, 2012.
- [16] J. P. Palutikof, B. Brabson, D. Lister, and S. T. Adcock, “A review of methods to calculate extreme wind speeds,” *Meteorol. Appl.* vol. 6, pp. 119–132, 1999.
- [17] D. K. Dey and J. Yan, Eds, *Extreme value Modeling and Risk Analysis: Methods and Applications*. Taylor and Francis Group, 2016.
- [18] EN 1990:2002, Eurocode: Basis of Structural Design, CEN, Brussels, 2002.
- [19] J. Pickands, III, “Statistical inference using extreme order statistics,” *Ann. Stat.* vol. 3, pp. 119–131, 1975.
- [20] J. D. Holmes and W. W. Moriarty, “Application of the generalized Pareto distribution to extreme value analysis in wind engineering,” *J. Wind Eng. Ind. Aerodyn.* vol. 83, nos. 1–3, pp. 1–10, 1999.
- [21] N. J. Cook and R. I. Harris, “Discussion on application of the generalized Pareto distribution to extreme value analysis in wind engineering by J.D. Holmes, W.W. Moriarty,” *J. Wind Eng. Ind. Aerodyn.* vol. 89, no. 2, pp. 215–224, 2001.
- [22] J. D. Holmes and W. W. Moriarty, “Response to discussion by N. J. Cook and R. I. Harris,” *J. Wind Eng. Ind. Aerodyn.* vol. 89, no. 2, pp. 225–227, 2001.
- [23] S. Kang and J. Song, “Parameter and quantile estimation for the generalized Pareto distribution in peaks over threshold framework,” *J. Korean Stat. Soc.* vol. 46, no. 4, pp. 487–501, 2017.
- [24] P. Thompson, Y. Cai, D. Reeve, and J. Stander, “Automated threshold selection methods for extreme wave analysis,” *Coast. Eng.* vol. 56, no. 10, pp. 1013–1021, 2009.
- [25] JCSS Probabilistic Model Code, Joint Committee on Structural Safety, 2000. [Online]. Available: <https://www.jcss-lc.org/jcss-probabilistic-model-code/>. Accessed: April 12, 2022.
- [26] EN 1991-1-4:2005, Eurocode 1: Actions on structures – Part 1-4: General actions – Wind Actions, CEN, Brussels, 2005.
- [27] N. J. Cook, *Designer’s Guide to Wind Loading of Building Structures, Part 1: Background, Damage Survey, Wind Data and Structural Classification*. Butterworth-Heinemann, 1986.

