

RESEARCH ARTICLE

On the tails of the wind ramp distributions

Adam DeMarco¹ and Sukanta Basu²

¹Department of Marine, Earth, and Atmospheric Sciences, North Carolina State University, Raleigh, NC 27695, USA

²Faculty of Civil Engineering and Geosciences, Delft University of Technology, Delft, the Netherlands

ABSTRACT

We analyzed several multi-year wind speed datasets from four different geographical locations. The probability density functions (pdfs) of wind ramps from all these sites revealed remarkably similar shapes. The tails of the pdfs are much heavier than a Gaussian distribution, and they also systematically depend on time increments. Quite interestingly, from a purely statistical standpoint, the characteristics of the extreme ramp-up and ramp-down events are found to be almost identical. With the aid of extreme value theory, we describe several other inherent features of extreme wind ramps in this paper. Copyright © 2018 John Wiley & Sons, Ltd.

KEYWORDS

Extreme value theory; Hill plot; Pareto distribution; Tail-index

Correspondence

Sukanta Basu (s.basu@tudelft.nl), Faculty of Civil Engineering and Geosciences, Delft University of Technology, Delft, the Netherlands.

Received . . .

1. INTRODUCTION

One of the major challenges facing the wind energy industry is the accurate prediction of sudden and sharp fluctuations in the wind field (a.k.a. wind ramps) near the lower part of the atmospheric boundary layer (ABL; [1, 2, 3, 4]). These not-so-rare and inauspicious events can drastically modulate deficiencies (ramp down) and/or surpluses (ramp-up) in wind power production causing disruptions in operations and energy supply balance. Therefore, as the demand for more reliable wind power increases there is an ever-present need for further advancement in the understanding of how to properly characterize and quantify the ramp events [5]. It is well-known in the literature that various meteorological factors can contribute to ramp events including (but not limited to): thunderstorm outflows, low-level jets, dry lines, cold fronts [6]. Since most of these phenomena are location-specific (for example, dry lines occur predominantly over the southern Great Plains of the US), one would naïvely expect the statistical properties of the ramp events to also be site-dependent. In this paper, we confront this expectation with reality. By making use of long-term observational data from several field sites, we demonstrate that a key trait of wind ramp statistics behaves in a quasi-universal manner.

Before delving into the detailed quantitative aspects, we provide qualitative support for our claim via Fig. 1. Here, we have plotted the probability density functions (pdfs) of wind speed increments (δu) normalized by standard deviation ($\sigma_{\delta u}$). For a specific time increment (τ), the wind speed increments (or ramps) are defined as:

$$\delta u(t) = u(t + \tau) - u(t) \quad (1)$$

where $u(t)$ is the wind speed at time t . Positive (negative) values of δu signify ramp-up (down) events. From a wind energy perspective, τ values on the order of a few minutes to a few hours (the so-called mesoscale regime) are of utmost importance [5, 7, 8, 9]. As illustrative examples, in Fig. 1, we show the observed wind ramp pdfs corresponding to $\tau = 10$ min (left panel) and $\tau = 60$ min (right panel). A number of inferences can be drawn by visual inspection of this figure. First and foremost, all the pdfs, representing four diverse geographical and meteorological conditions (ranging from coastal environment to complex terrain), reveal remarkably similar shapes. They portray strong peakedness near the mode of the distribution, and more importantly, they all possess tails which are much heavier than a Gaussian pdf. These tails seem to

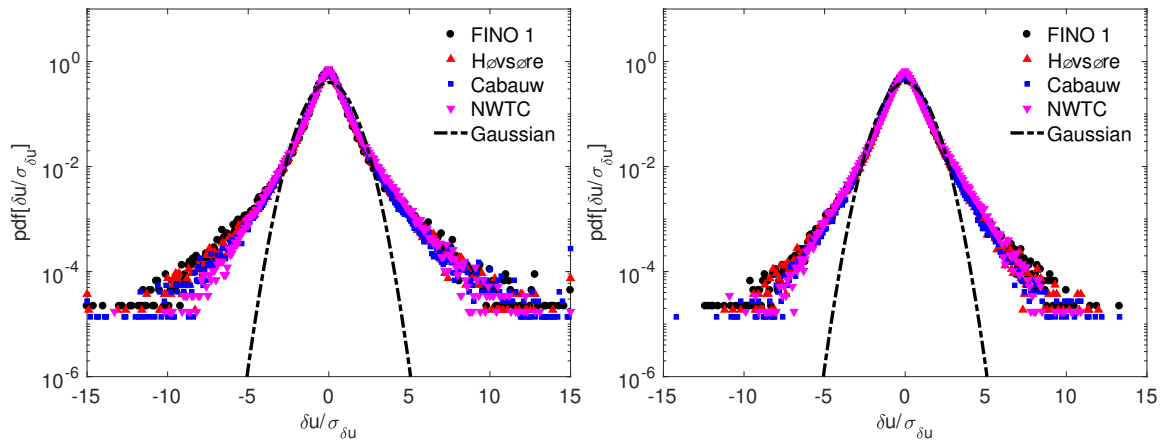


Figure 1. Probability density functions of wind ramps (δu) from four tall-tower sites (FINO1, Høvsøre, Cabauw, and NWTC). Multi-year, 10-min averaged wind data measured by the topmost sensors on these towers are utilized here. Further details are provided in Section 2. The wind increment values are normalized by the corresponding standard deviations ($\sigma_{\delta u}$). The left and right panels represent time increments (τ) of 10 min and 60 min, respectively. A Gaussian pdf is overlaid (dashed line) as a reference.

depend on τ in a subtle yet systematic manner. These unexpected findings inspired us to probe further into this problem by addressing a suite of science questions in this paper:

- Do the pdf tails corresponding to the ramp-up and ramp down events behave differently?
- How do the tails depend on the height (above ground level)?
- What is the impact of aggregation (filtering) on the tails?
- Can the dependence of tail properties on the sample size be quantified?

To the best of our knowledge, none of these questions have been answered in the literature in a comprehensive manner. We do point out that a handful of studies [10, 11, 12, 13, 14, 15, 16] provided important building blocks for our research. Unfortunately, several of these papers focused on wind gusts, and thus, their findings cannot be very relevant for mesoscale wind ramps. More critically, most of these studies utilized very limited amount of observational data (often with durations of a few hours to merely a few days) and came up with conflicting results. For example, by analyzing only a few days worth of data, Liu et al. [15] concluded that wind ramps follow *truncated stable distributions*. In a follow-up study, however, Liu and Hu [16] arrived at an opposite conclusion when they made use of a slightly larger dataset. We argue that, in lieu of converged statistics, the results from these past studies cannot be faithfully generalized.

The present study differs from the others in two areas. First, it primarily relies on rigorous statistical analyses for pdf characterization instead of qualitative visual inspection. Second, it utilizes multi-year wind datasets from four tall-tower sites: FINO 1 (North Sea), Høvsøre (Denmark), Cabauw (the Netherlands), and NWTC (USA). Since these sites are quite diverse in nature, we have more confidence in generalizing the outcomes. In the following section, we briefly describe these datasets.

2. DESCRIPTION OF WIND DATASETS

Over the past decades, the wind energy and boundary layer meteorology communities have invested significant resources in installing and operating a few tall-towers around the world. A multitude of research-grade sensors (e.g., high-fidelity cup-anemometers) are mounted on these towers at various heights. Owing to their periodic calibration and regular maintenance, the meteorological datasets (including wind speed time-series) collected by these sensors are deemed to be of the highest quality. Thus, it is not surprising that these datasets have been heavily utilized to advance our understanding of the lower part of the ABL. For example, very recently, Kiliyanpilakkil et al. [17, 18] conducted rigorous scaling analyses of wind datasets from three of these prominent tall-towers: FINO 1, Cabauw, and NWTC. We also leverage on the same datasets, supplemented by measurements from Høvsøre, to demonstrate the statistical characterization of wind ramps. Table I, along with the following subsections, provide more details into these locations.

2.1. FINO 1

It is an offshore platform in the North Sea [19, 20, 21]. It consists of a 100-m tall meteorological tower equipped with wind speed measurement sensors (cup anemometers) at heights of 33 m, 40 m, 50 m, 60 m, 70 m, 80 m, 90 m, and 100 m. A

Table I. Description of measurement sites

Site	Elevation (m; mean sea level)	Location	No. of Months
FINO 1	0	54.01° N, 6.59° E	91
Høvsøre	0	58.44° N, 8.15° E	132
Cabauw	- 0.7	51.97° N, 4.93° E	170
NWTC	1855	39.91° N, 105.23° W	132

total of 91 months of wind speed data collected over a period of nine years (2004–2012) are utilized in the present study. Each time-series (output rate: 10-min) contains ~ 478 thousand samples.

2.2. Høvsøre

This meteorological tower is situated in a rural area close to the west coast of Jutland, Denmark and played a pivotal role in numerous wind energy studies [22, 23]. We analyze 10-min-average wind data from six levels: 10 m, 40 m, 60 m, 80 m, 100 m, and 116 m collected during the years 2005–2015. In this case, each time-series consists of ~ 567 thousand samples.

2.3. Cabauw

The Cabauw Experimental Site for Atmospheric Research (CESAR) tower is located in the western part of the Netherlands [24, 25, 26]. We use 170 months of 10-min-average wind speed data from the years 2001–2015 (~ 736 thousand samples) measured by propeller wind vanes at heights of 10 m, 20 m, 40 m, 80 m, 140 m, and 200 m.

2.4. NWTC

We analyze multi-year (2004–2014) wind data from a 80-m tall tower (called M2) located at the foothills of the Colorado Rocky near Boulder, Colorado and maintained by the National Renewable Energy Laboratory (NREL) National Wind Technology Center (NWTC). This location represents complex terrain and is prone to various wind flows and disturbances [27]. The NWTC dataset includes 1-min averaged, cup anemometer-based, wind speed time series from four heights: 10 m, 20 m, 50 m, and 80 m. Each time-series is made up of ~ 5.78 million points with virtually no data gaps.

3. METHODOLOGY

In order to investigate the tail features of the wind ramp events, we have borrowed a well-established methodology, called the Hill plot [28], from the extreme value (EV) theory [29, 30, 31]. In this section, we explain this approach in detail by using synthetically generated random variates from two heavy-tailed distributions.

By definition, a heavy-tailed distribution (F) satisfies [32]:

$$\overline{F}(x) = 1 - F(x) \sim \frac{\lambda}{x^\gamma}; \quad x \rightarrow \infty, \gamma > 0, \quad (2)$$

where $\overline{F}$ is the so-called complementary cumulative distribution function (ccdf). λ is a positive constant and γ is known as the tail-index (or shape parameter). In principle, γ can be estimated from the slope, $-\frac{d \log \overline{F}(x)}{d \log x}$ [32]. However, in practice, the most common approach is to invoke the concept of order statistics [33].

The rank-ordered values (in decreasing order) of x can be written as: $\Phi_k = x_k$, where $k = 1, \dots, N$. If the variates x follow Eq. (2), it is expected to exhibit the following power-law behavior (a.k.a. Zipf law; [34]):

$$\Phi_k \propto \left(\frac{k}{N} \right)^{-\frac{1}{\gamma}}. \quad (3)$$

Over the years, several estimators for γ have been proposed in the literature, including (but not limited to) Pickand's estimator [35], Hill estimator [36], and the Dekkers-Einmahl-de Haan estimator [37]. In this work, we use the popular Hill estimator (γ_H):

$$\gamma_H = \left[\frac{1}{k} \sum_{i=1}^k \log \left(\frac{\Phi_i}{\Phi_{k+1}} \right) \right]^{-1}, \quad (4)$$

where $k = 1, \dots, N - 1$. When γ_H is plotted against k , it is known as the Hill plot [28]. For EV distributions (e.g., Pareto), estimated γ_H is supposed to stabilize with increasing values of k . In Fig. 2, we show an illustrative example utilizing generalized Pareto (GP) distributed variates.

The pdf of the GP distribution can be written as:

$$f(x) = \left(\frac{1}{a}\right) \left(1 + \frac{c(x-b)}{a}\right)^{-1-1/c}; \quad x > b, c > 0, \quad (5)$$

where a, b, c are the parameters of GP. By integrating this equation, one can derive the ccdf of GP as:

$$\bar{F}(x) = \left(1 + \frac{c(x-b)}{a}\right)^{-1/c}. \quad (6)$$

Thus, the ccdf of GP is expected to decline as a power-law with tail index $\gamma = 1/c$.

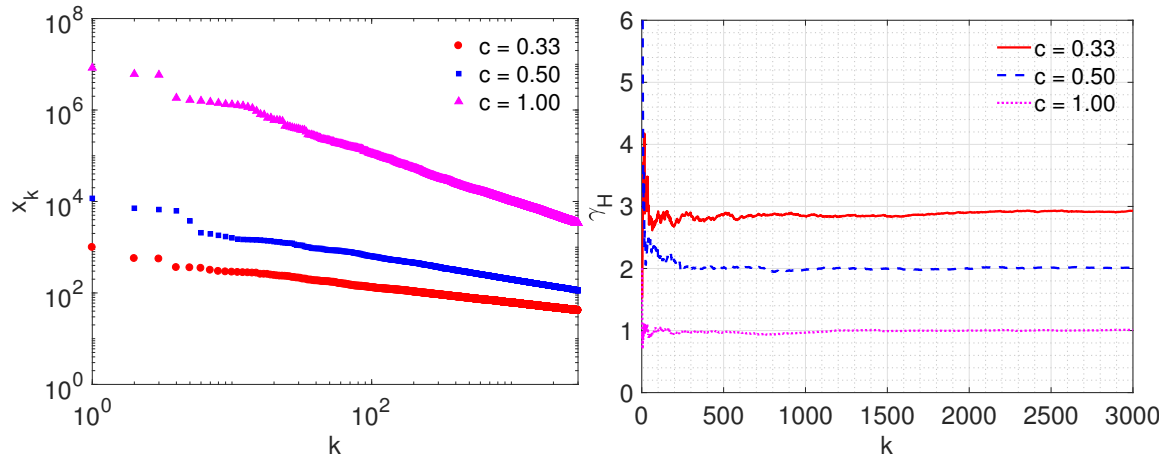


Figure 2. Left panel: rank-order (a.k.a. Zipf) plots for generalized Pareto distributed variates with three different c values. The parameters a and b are assumed to be equal to 1 and 0, respectively. Right panel: the estimated γ_H values for these cases. It is clear that $\gamma_H \approx 1/c$ for $k > 1000$.

In Fig. 2 (left panel), the rank-order plots for the GP distribution are shown for three values of c . For each case, the sample size is 10^7 and the parameters a and b are assumed to be equal to 1 and 0, respectively. By construction, only positive random variates are generated in this case. The tail indices are determined via the Hill plot in the right panel of Fig. 2. Clearly, the γ_H values rapidly stabilize towards $\frac{1}{c}$ for all the three cases, as would be desired. This example attests to the prowess of the Hill plot in estimating the tail indices from a rather simple EV distribution. Next, we investigate the usefulness of the Hill plot using a far more complicated distribution with two distinct tail behaviors.

The generalized hyperbolic skew student's t (GHSST) distribution is often used in financial modeling and risk management [38, 39]. It has the innate ability to fit pdfs with heavy tails and significant asymmetry. A brief overview of this distribution is provided in Appendix A. A realization of the GHSST variates is shown in the top-panel of Fig. 3. Large positive values, signifying a heavy right tail, can be readily observed in this plot.

The pdf of the generated GHSST variates is shown in the middle-left panel of Fig. 3. For comparison, a Gaussian pdf is overlaid on this plot. Clearly, both the left and right tails of the GHSST variates are much heavier than the Gaussian pdf. They also show different decaying behaviors as evident in the middle-right panel. For large values of x , the right tail portrays a quasi-linear appearance in this log-log representation. In other words, the right tail is characterized by a power-law distribution which is in-line with the asymptotic limit discussed in Appendix A. In contrast, the left tail strongly departs from linearity highlighting its mixed-exponential-power-law behavior. The rank-order plot, shown in the bottom-left panel of Fig. 3, provides further supporting evidence.

At this point, we would like to point out that Eq. (2) has limited applicability in real-world scenarios. For such cases, this equation should be generalized as:

$$\bar{F}(x) = 1 - F(x) \sim \frac{L(x)}{x^\gamma}; \quad x \rightarrow \infty, \gamma > 0, \quad (7)$$

where, L is a slowly varying function (e.g., exponential). For large x , $L(x)$ may be approximated as a constant λ .

The tail indices from the GHSST variates are estimated via the Hill plot and shown in the bottom-right panel of Fig. 3. These results should be interpreted carefully by taking into consideration Eq. (7). In the case of right tail, the γ_H values stabilize rapidly as in the case of GP distributed variates. For the left tail, the γ_H values are significantly higher; also,

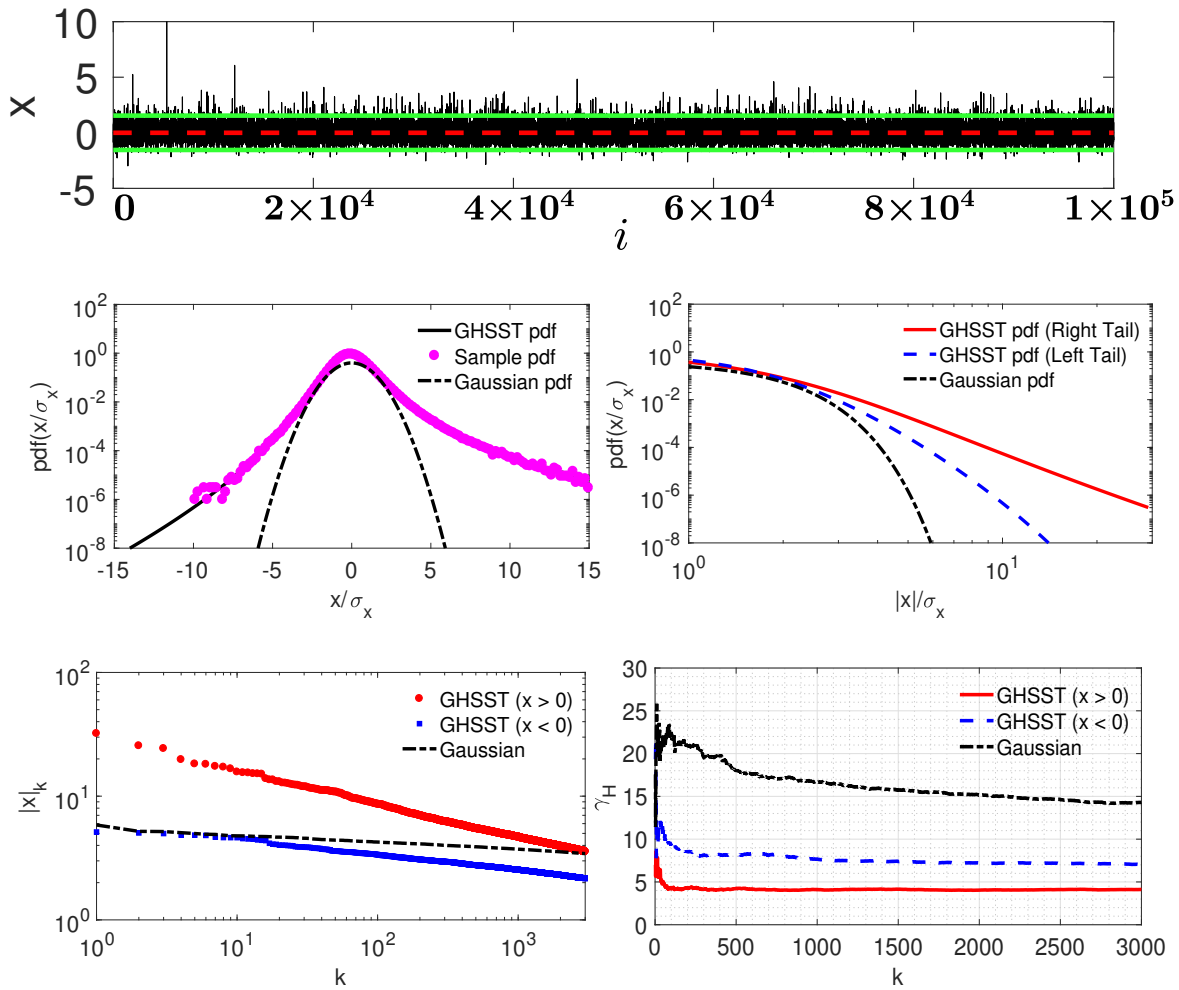


Figure 3. A realization of the GHSST variates (sample size = 10^7) is generated using the following parameters: $\nu = 6$, $\beta = 0.5$, $\mu = -0.125$, and $\delta = 1$. A subset of these variates is shown in the top panel as an illustrative example. The mean of the variates is depicted by the dashed red line. The green lines denote three times the standard deviation around the mean. The analytical (black line; [39]) and sample (magenta circles) pdfs are shown in the middle-left panel. For comparison, a Gaussian pdf (dot-dashed line) with zero mean and unit variance is overlaid on this panel. The tails of the GHSST and Gaussian pdfs are shown in the middle-right panel. Since the right and the left tails of the GHSST pdf behave differently, they are shown separately. Clearly, the right tail exhibits a linear behavior in this log-log representation. Rank-order (a.k.a. Zipf) plots for the GHSST and Gaussian distributed variates are shown in the bottom-left panel. Estimated tail indices (γ_H) utilizing the Hill plot are documented in the bottom-right panel.

the stabilization is slightly slower (difficult to detect in this figure without zooming in). In this case, the exponential term modulates the power-law tail. Please note that the estimated γ_H values are very high for the random Gaussian variates and they never stabilize as the tails simply follow exponential behavior. In Fig. 4, we document the influence of sample size on γ_H estimation using the GHSST variates. In the case of the right tail (exhibiting power-law behavior), γ_H values are more-or-less insensitive to sample size (N). In contrast, for the left tail, γ_H keeps increasing as N increases. This non-convergence essentially corroborates the fact that the left tail of the GHSST pdf does not exhibit a purely power-law behavior; rather, it follows a mixed-exponential-power-law.*

In summary, with the aid of randomly generated variates, we have demonstrated that the Hill plots can be very effective in characterizing different types of tail behaviors. It is also computationally very efficient. For these reasons, in the following section, we will invoke this methodology to address the science questions posed in the Introduction. The idealized examples shown in Figs. 2, 3, and 4 will provide guidance for interpreting the wind ramp characteristics observed within various observational datasets.

*In a recent paper [40], we reported a similar trend for the normal inverse Gaussian (NIG) distribution, which also possesses mixed-exponential-power-law tails.

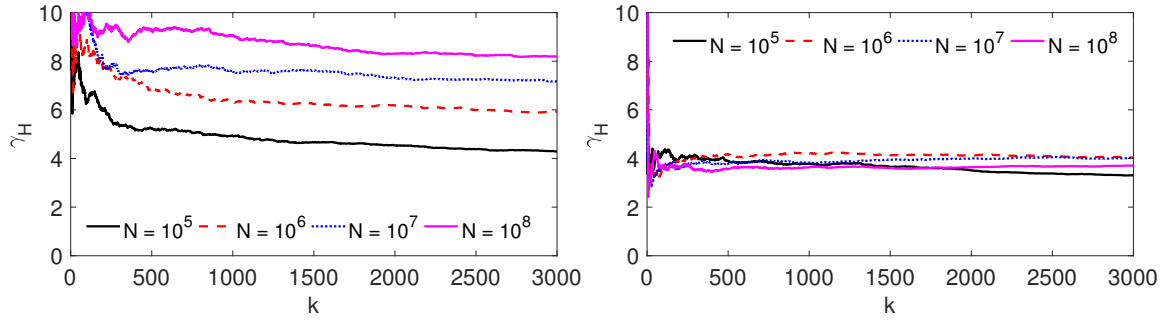


Figure 4. Sensitivity of estimated γ_H values with respect to sample size. All the GHSST variates are generated using the same parameters as in Fig. 3. The left and right panels correspond to the left and right tails, respectively.

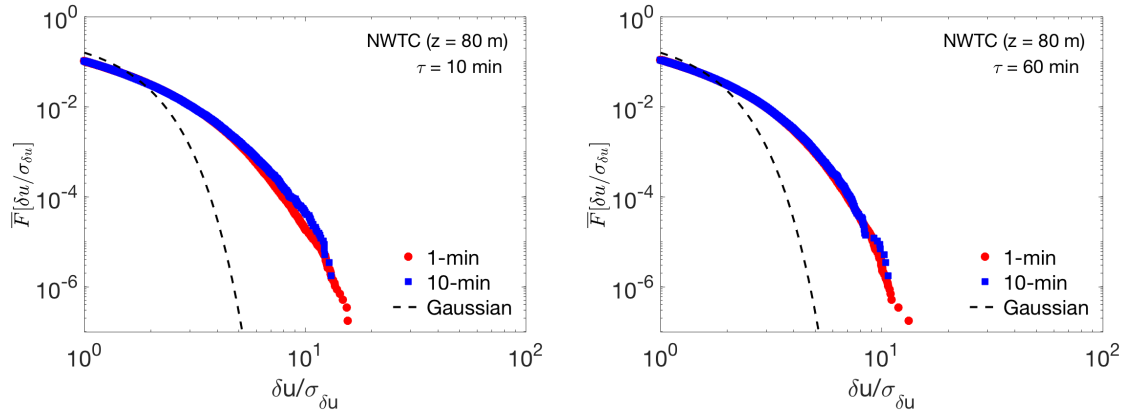


Figure 5. Complementary cumulative distribution function ($\bar{F}$) from the NWTC_{1min} and NWTC_{10min} datasets at $z = 80$ m are shown. The left and right panels represent time increments (τ) of 10 min and 60 min, respectively. The wind increment values are normalized by the corresponding standard deviations ($\sigma_{\delta u}$). A Gaussian pdf is overlaid (dashed line) as a reference.

4. RESULTS

Given that the NWTC dataset offers the largest sample size, we select it first for comprehensive analysis. The original granularity of the wind time-series is 1-min. Henceforth, we refer to this series as NWTC_{1min}. In order to study the effects of aggregation on the tail characteristics, we created a 10-min-average series (sample size: ~ 578 thousand) by simple moving averaging (followed by downsampling) of the NWTC_{1min} series. This new series will be identified as NWTC_{10min}. In addition to the results reported in the current section, these series are also utilized in Appendices B and C to investigate the issues of nonstationarity and correlation.

The ccdf ($\bar{F}$) for both the NWTC_{1min} and NWTC_{10min} time-series are shown in Fig. 5. The left and right panels represent $\tau = 10$ min, and $\tau = 60$ min, respectively. On these plots, $\bar{F}$ for a Gaussian distribution is also shown for comparison. In this log-log representation, we only focus on the right tail (ramp-up) of the pdf. Several remarks can be made from this figure. First of all, both the NWTC_{1min} and NWTC_{10min} cases clearly portray non-Gaussian tails. The implication of this heavy-tail behavior is rather crucial for the wind energy community. For example, the exceedance probability of a strong ramp-up event of magnitude $5\sigma_{\delta u}$ is very small (much less than 10^{-6}) if one assumes Gaussianity. According to observations, however, the exceedance probability is almost 10^{-2} . In other words, the assumption of Gaussianity leads to severe underestimation of extreme wind ramp events.

In Fig. 1, we reported that the tails of the wind ramp pdfs systematically depend on τ . Thus, it is not surprising that the same dependence is also evident from the ccdfs. From Fig. 5, one can discern that, in comparison with $\tau = 10$ min, the right tail decays faster in the case of $\tau = 60$ min. Later on, we will establish that this trend is actually monotonic in the range of $\tau = 10 - 360$ min.

According to Fig. 5, the agreement between NWTC_{1min}- and NWTC_{10min}-based $\bar{F}$ curves are excellent up to $\delta u \approx 5\sigma_{\delta u}$. Beyond that point, the NWTC_{1min}-based $\bar{F}$ curve starts to decline faster. With the aid of the Hill plots, we will further investigate if this discrepancy is due to the disparity in sample sizes or it is an artifact of aggregation.

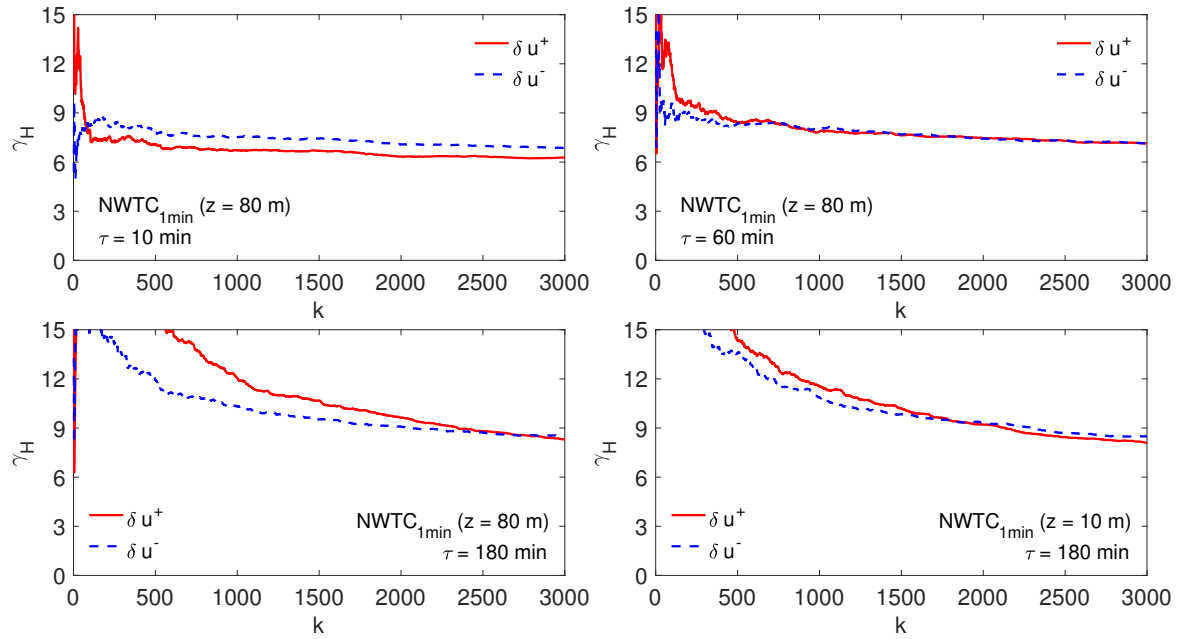


Figure 6. The Hill plots for wind ramp distributions based on the $NWTC_{1min}$ time-series. The red (solid) and blue (dashed) lines represent ramp-up (δu^+) and ramp-down (δu^-) cases, respectively. The top-left, top-right, bottom-left, and bottom-right panels correspond to the following scenarios, respectively: (i) $\tau = 10$ min, $z = 80$ m; (ii) $\tau = 60$ min, $z = 80$ m; (iii) $\tau = 180$ min, $z = 80$ m; and (iv) $\tau = 180$ min, $z = 10$ m.

The Hill plots for $NWTC_{1min}$ are shown in Fig. 6. The values of γ_H are found to be noticeably higher for larger τ values. In other words, the wind ramp pdf tails decay faster for larger τ values, which is in-line with our earlier finding. Both the ramp-up (noted as δu^+) and ramp-down (δu^-) cases follow similar trends. However, the γ_H curves never fully stabilize for either case in the considered range ($1 < k \leq 3000$). Thus, we can deduce that the wind ramp pdf tails do not obey power-laws. Later on, we will explore further if these NWTC data-based results also hold for other datasets.

In order to quantify the effect of sample size on the γ_H values, we adopted a Monte-Carlo-type strategy. From the $NWTC_{1min}$ time-series (sample size 5.78 million), we extract one hundred contiguous subsets from random locations. Each subset is called $NWTC_{sub}^i$ and contains 578 thousand samples. The index i varies from 1 to 100 to demarcate each subset. We then perform Hill plot analysis on each subset separately and compute ensemble statistics. These ensemble Hill plots are shown in Fig. 7. The ramp-up and ramp-down cases are shown separately in the left and right panels, respectively.

The overall trends of the γ_H values reported in Fig. 6 and Fig. 7 are qualitatively very similar. However, the magnitudes of γ_H for the $NWTC_{sub}^i$ cases are significantly lower than the $NWTC_{1min}$ series. We would like to remind the readers that a similar sample size dependency was reported earlier in the case of the left tail (depicting a mixed-exponential-power-law behavior) of the GHSST pdf (refer to bottom-left panel of Fig. 4). Thus, on the basis of the Hill plot analyses and ccdf plots, we can confidently claim that the wind ramp distributions do not exhibit power-law tails.

In order to further bolster our claim, we have computed γ_H for three other locations: FINO 1, Høvsøre, and Cabauw. Wind data from the topmost sensor levels are utilized. Figs. 8–10 show the Hill plots for three different values of τ , 10 min, 60 min and 180 min, respectively. In each of these figures, we also included the Hill plot for the $NWTC_{10min}$ series. Based on these figures, several assertions can be made. First of all, the Hill plots from all the locations look remarkably similar. For (almost) all the cases, the differences between the ramp-up and ramp-down events are marginal. At the same time, for all plots, the values of γ_H do not stabilize and continue to decrease with increasing k . Thus, we can safely rule out power-law being a viable candidate for wind ramp distributions. In-line with our earlier finding, the γ_H values exhibit dependence on τ . Thus, stable distributions [33] should not be used to characterize wind ramp distributions.

Earlier, we have concluded that the γ_H values strongly depend on sample size. Now, in order to probe the impact of aggregation, we compare the γ_H values for the $NWTC_{10min}$ series (bottom right panels of Figs. 8–10) against the corresponding values from the $NWTC_{sub}^i$ series (reported in Fig. 7). Please note that these series have identical sample size; albeit, they have different granularity. From the plots, it is quite evident that the γ_H values from the $NWTC_{10min}$ and $NWTC_{sub}^i$ series are very much comparable. Thus, within the limited filtering range of 1-min to 10-min, the aggregation effect is negligible. However, one should not downplay the effects of sample size.

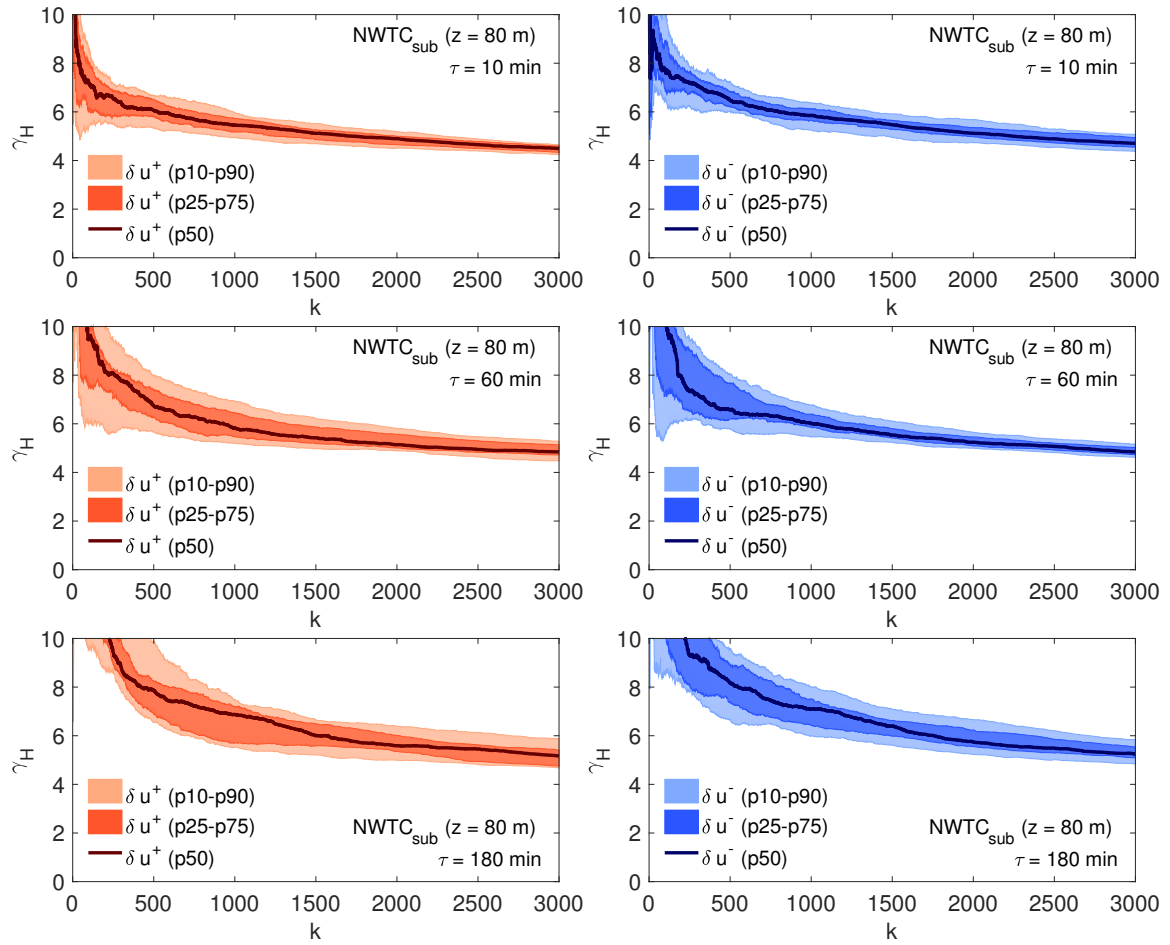


Figure 7. The Hill plots for wind ramp distributions based on the subsets of 1-min-average NWTC wind speed time-series. Each subset contains contiguous 578 thousand samples. A total of one hundred randomly selected subsets are utilized for these plots. The top, middle, and bottom panels represent $\tau = 10$ min, 60 min, and 180 min, respectively. Ramp-up and ramp-down results are shown in left and right panels, respectively. The solid lines, dark shaded areas, and the light shaded areas correspond to the medians, 25th-75th percentile ranges, and 10th-90th percentile ranges, respectively.

Thus far, we have only focused on wind data from the topmost sensors of all the four meteorological towers. It would be interesting to find out if/how the γ_H values depend on sensor height. Instead of plotting several individual Hill plots, we opt for plotting averaged γ_H values so as to report all the results succinctly in Figs. 11 and 12. The values of $\langle \gamma_H \rangle$ are computed for $2000 \leq k \leq 3000$. We intentionally (and incorrectly) assume that over this range the values of γ_H have fully stabilized. Despite this ad-hoc assumption, the results are quite revealing. For both the ramp-up and ramp-down cases, $\langle \gamma_H \rangle$ increase monotonically from ~ 4 (at $\tau = 10$ min) to $\sim 6 - 7$ (at $\tau = 360$ min). The diversity in $\langle \gamma_H \rangle$ values across various heights is rather small (especially for $\tau < 180$ min). Among all the locations, the spread of $\langle \gamma_H \rangle$ is the most significant at Cabauw; the $\langle \gamma_H \rangle$ curves from the top two sensors, located at heights of 140 m and 200 m, seem to branch out from others for $\tau > 180$ min. We speculate that certain meteorological processes (e.g., low-level jets) influence the wind ramp statistics at higher altitudes. However, we need more observational datasets from higher altitudes (possibly collected by lidars and/or sodars) to shed further light on this intriguing finding.

5. CONCLUSIONS

In this study, we analyzed several long-term wind speed datasets comprised of four different geographical locations, from offshore to complex terrain. We showed that the wind ramp pdfs from all the sites reveal amazingly similar shape characteristics. Most interestingly, the tails of the wind ramp pdfs are much heavier than Gaussian and decay faster as time increments increase. With the aid of the Hill plots, we showed that the extreme ramp-up and ramp-down events behave

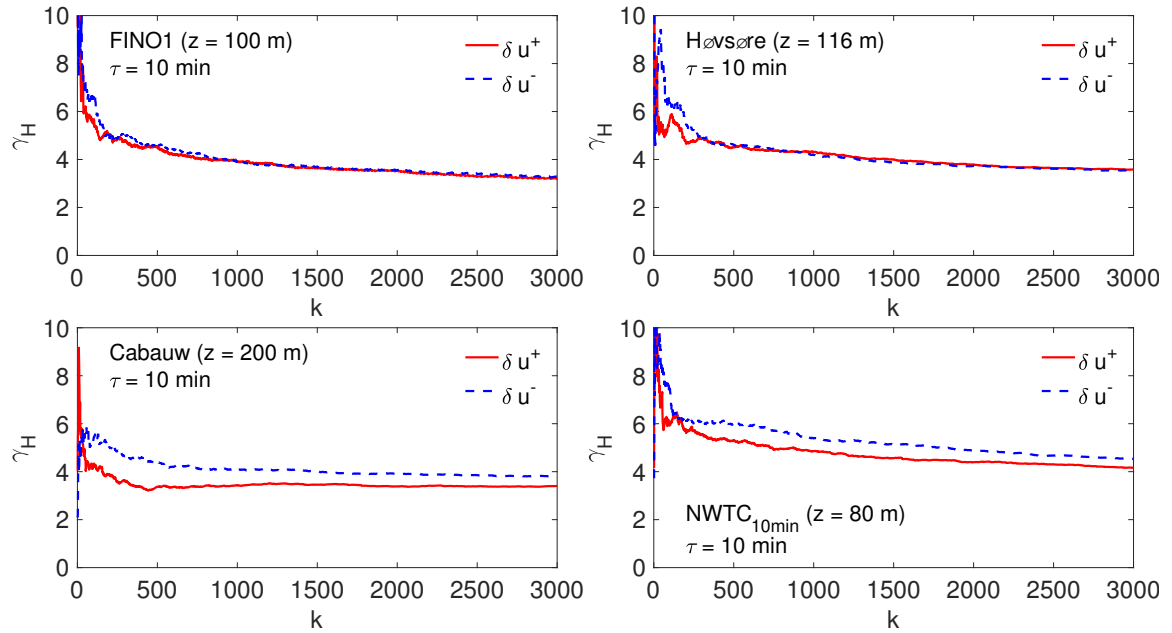


Figure 8. The Hill plots of wind ramp distributions from four field sites: FINO1 (top-left panel; sensor height: 100 m), Høvsøre (top-right panel; sensor height: 116 m), Cabauw (bottom-left panel; sensor height: 200 m), NWTC (bottom-right panel; sensor height: 80 m). For all the cases, 10-min-averaged wind speed are utilized. The time increment (τ) is 10 min. The ramp-up (δu^+) and ramp-down (δu^-) statistics are denoted by red (solid) and blue (dashed) lines, respectively.

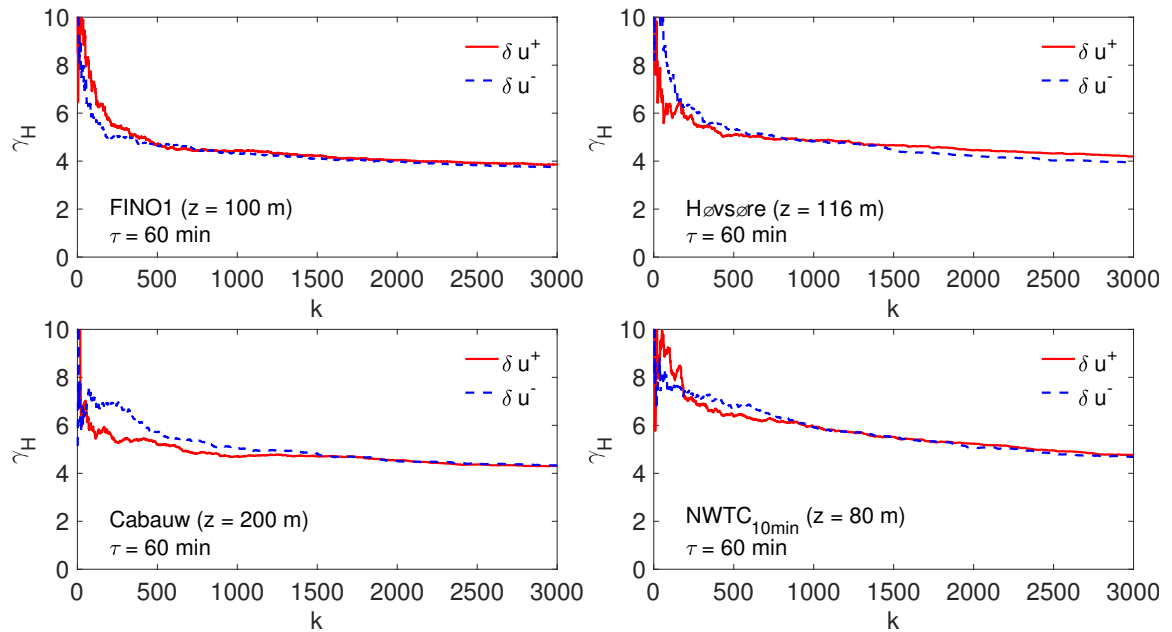


Figure 9. Same as Fig. 8, except for time increment (τ) of 60 min.

similarly from a statistical point-of-view. Moreover, the tail-index statistics exhibited minimal dependence with respect to height above the ground.

Another important aspect of these results showed that the tails of the wind ramp distributions do not follow a power-law distribution, rather modulated by a slowly varying function. We speculate this function to be an exponential. Therefore, in future work, we will use several types of pdfs from the generalized hyperbolic distribution family (e.g., GHSST, NIG) to determine the ideal candidate for capturing the tail characteristics of wind ramp distributions.

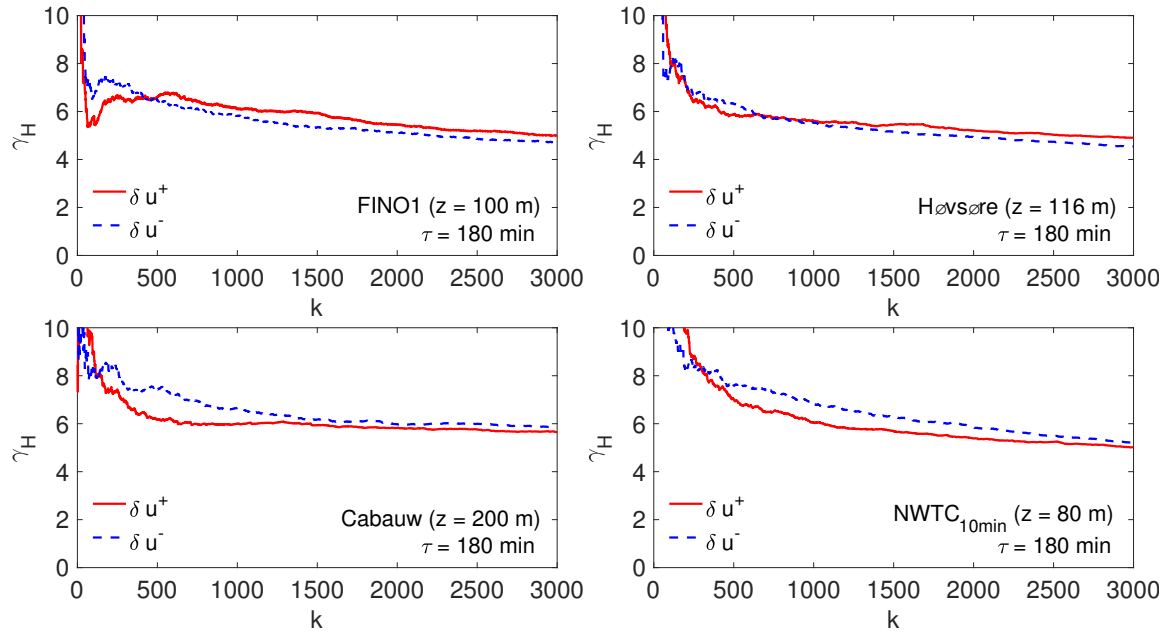


Figure 10. Same as Fig. 8, except for time increment (τ) of 180 min.

As the wind energy industry continues to flourish and wind ramp prediction becomes increasingly important, the results from this study should be utilized for model validations and improvement. It would be critical to find out if the state-of-the-art numerical weather prediction models and time-series forecasting tools are able to capture the extreme ramp behaviors accurately. It is also envisaged that the contemporary synthetic wind speed generators (e.g., [41, 42]), heavily relying on statistical information, will tremendously benefit from our findings.

A. GENERALIZED HYPERBOLIC SKEW STUDENT'S t (GHSST) DISTRIBUTION

The GHSST distribution is a subclass of the generalized hyperbolic (GH) family and its pdf is defined as [39, 43, 44]:

$$f(x; \nu, \beta, \mu, \delta) = \frac{2^{\frac{1-\nu}{2}} \delta^\nu}{\sqrt{\pi} \Gamma(\nu/2)} \left(\frac{y_x}{|\beta|} \right)^{-\frac{\nu+1}{2}} K_{\frac{\nu+1}{2}}(|\beta| y_x) e^{\beta(x-\mu)}, \quad \beta \neq 0, \quad (8)$$

where ν , β , μ , and δ are the four parameters of the GHSST distribution. $y_x = \sqrt{\delta^2 + (x - \mu)^2}$ and Γ is the gamma function. The parameters ν and β together control the degree of heavy-tailedness and skewness of the tails. μ is the location parameter and is slightly different from the mean of the distribution. δ is a scale or peakedness parameter and it controls the shape of the pdf near its mode. $K_\nu(x)$ is the so-called modified Bessel function [45].

The tails of the distribution exhibit the following traits:

$$f_x(x) \sim C|x|^{-\nu/2-1} \exp(-|\beta x| + \beta x) \quad \text{as } x \rightarrow \pm\infty. \quad (9a)$$

Thus, the heavier tail decays as:

$$f_x(x) \sim C|x|^{-\nu/2-1} \quad \text{when } \begin{cases} \beta < 0 & \text{and } x \rightarrow -\infty, \\ \beta > 0 & \text{and } x \rightarrow +\infty, \end{cases} \quad (9b)$$

and the lighter tail behaves as:

$$f_x(x) \sim C|x|^{-\nu/2-1} \exp(-2|\beta x|) \quad \text{when } \begin{cases} \beta < 0 & \text{and } x \rightarrow +\infty, \\ \beta > 0 & \text{and } x \rightarrow -\infty. \end{cases} \quad (9c)$$

In this work, we generated GHSST distributed random variates following an algorithm described by Aas and Hobæk Haff [43]:

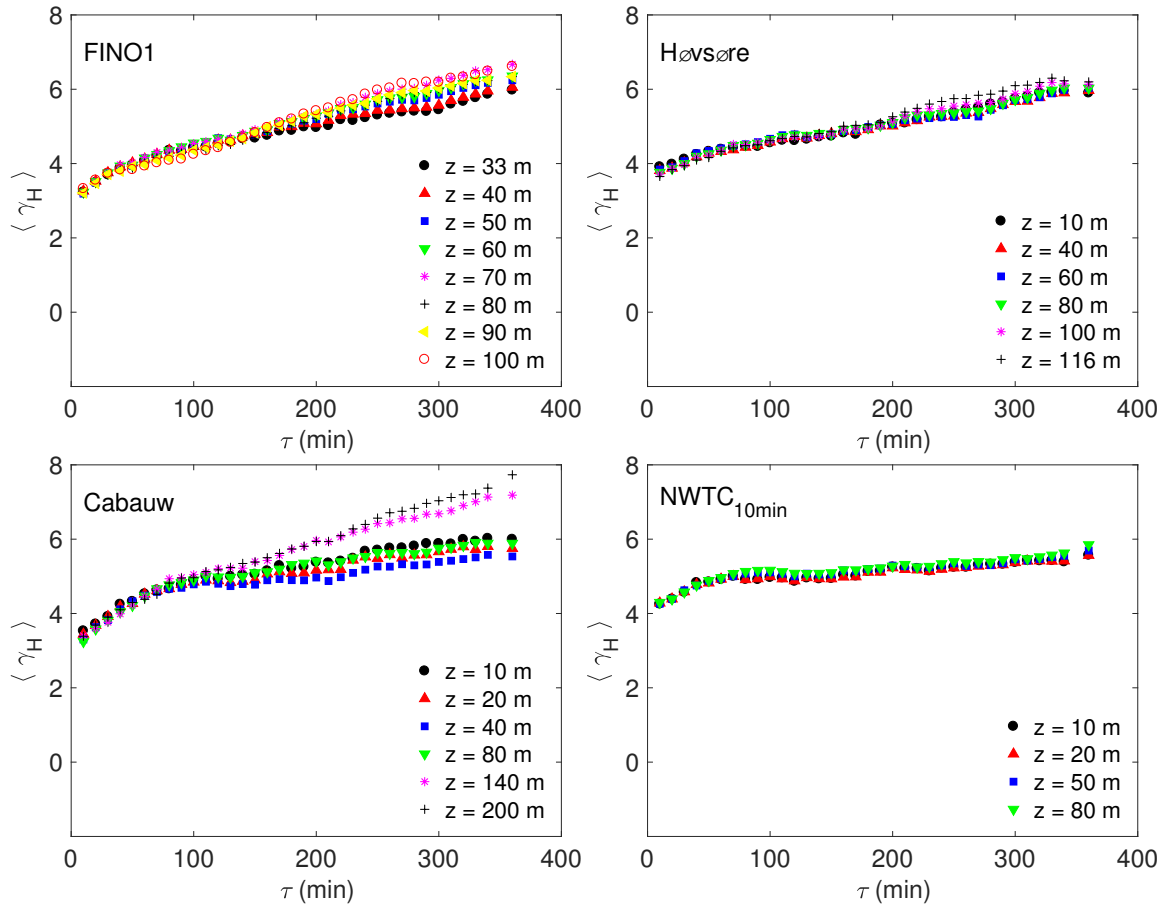


Figure 11. The averaged values of tail indices ($\langle \gamma_H \rangle$) as a function time-increment (τ). Wind ramp-up data from all the available sensors from all the four sites are utilized here.

- Generate Gamma distributed variates (G) with parameters $\nu/2$ and $\delta^2/2$.
- Obtain the Inverse Gamma variates by simply using $Z = G^{-1}$.
- Generate Gaussian random variates Y with zero mean and unit variance.
- Compute the GHSST random variates, $X = \mu + \beta Z + \sqrt{Z}Y$.

B. EFFECTS OF NONSTATIONARITY

For observational time-series, quantifying the effects of nonstationarity on the computed statistics is a challenging task. In meteorology literature, quite often time-varying mean and variance of a series are used to establish the (non-)existence of nonstationarity (for example see [46] and the references therein). A far more stringent test of nonstationarity is usually employed in the dynamical systems literature. Quite often a specific higher-order statistic (e.g., correlation dimension, Lyapunov exponent, Shannon entropy) is computed for different segments of a series [47]. In the case of systematic trends and/or significant variations of this specific statistic, the series is flagged as nonstationary.

In this appendix, we use the $NWTC_{1min}$ time-series to probe into the issue of nonstationarity. Given the large sample size, we are able to divide it into one hundred non-overlapping and contiguous subsets. Each subset has a sample size of approximately 57,800. We perform Hill plot analysis on each subset in sequence. For each subset, $\langle \gamma_H \rangle$ is computed for $2000 \leq k \leq 3000$. The resultant $\langle \gamma_H \rangle$ series for ramp-up and ramp-down cases with various time-increment (τ) values are reported in Fig. 13.

A few observations can be readily made based on Fig. 13. First of all, the estimated $\langle \gamma_H \rangle$ values are much smaller than those reported in Figs. 6–7. This behavior is completely in line with our earlier findings related to the sample size dependency of γ_H (thus, $\langle \gamma_H \rangle$) values. Second, the average values of $\langle \gamma_H \rangle$ marginally increases with τ . This trend was also reported earlier in Section 4. Furthermore, similar to Fig. 7, we see that the variability of $\langle \gamma_H \rangle$ increases slightly with

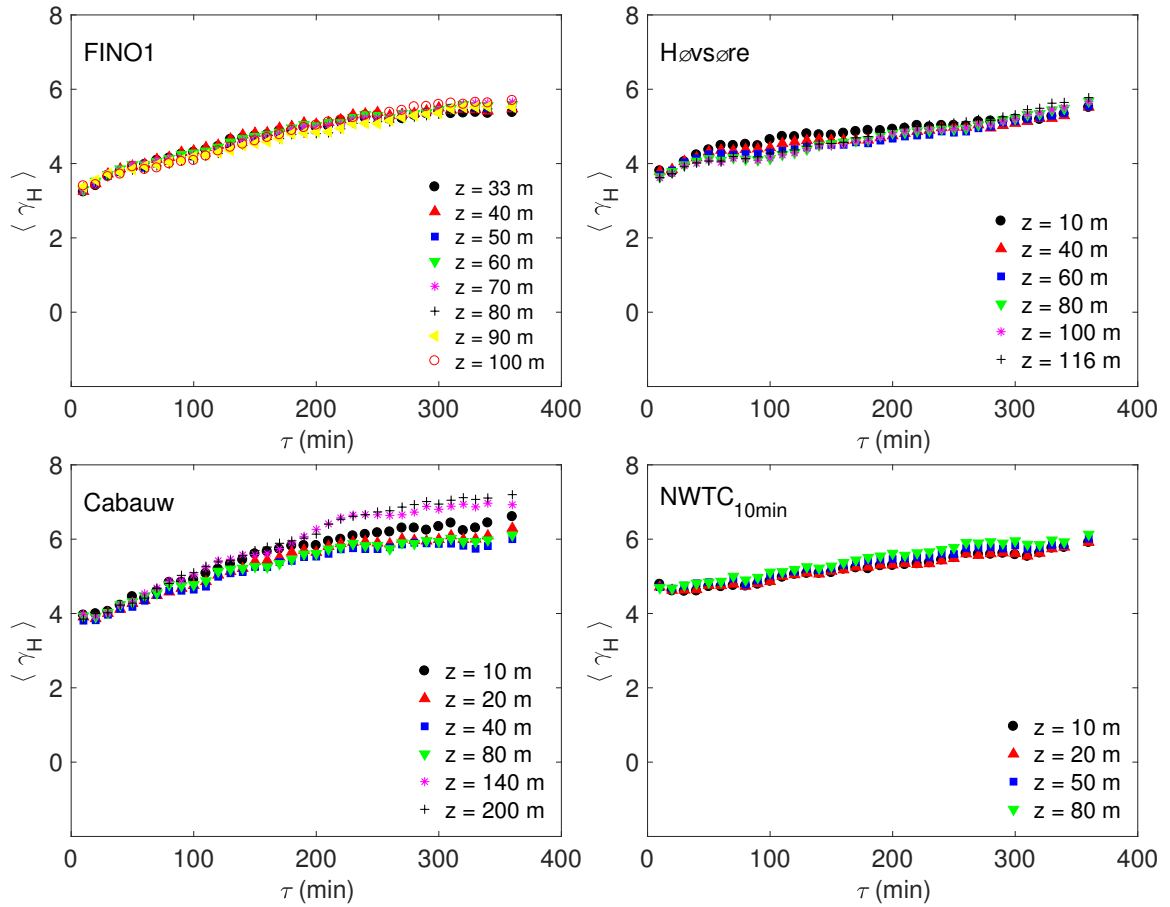


Figure 12. Same as Fig. 11, except for ramp-down events.

larger τ values. Most importantly, we do not see any trend or significant variations of $\langle \gamma_H \rangle$ with time (proxied by subset #). In other words, there is no tell-tale sign of nonstationarity that we can detect in the NWTC dataset.

C. EFFECTS OF CORRELATION

Rigorously speaking, the tail-index analysis should be performed on independent and identically distributed (i.i.d) random variates. However, observational data often possess intrinsic correlation. In this appendix, we report the autocorrelation functions of wind ramp time-series. We make use of both the NWTC_{1min} and NWTC_{10min} time-series from $z = 80$ m level. In Fig. 14, the results are shown for $\tau = 10$ min (left panel) and $\tau = 60$ min (right panel). Clearly, the autocorrelation function decreases rapidly for all the cases. In the case of $\tau = 10$ min, the minimum autocorrelation occurs around lag = 10 min; in contrast, the minimum occurs at lag = 60 min for $\tau = 60$ min. As a matter of fact, for all the other cases that we analyzed (not shown), the autocorrelation function always passed through a local minimum close to their respective τ values. Several years ago, [48] reported identical results in the context of fully developed turbulence.

Last, we would like to point out that most of our analyses reported in Section 4 were based on 10-min averaged data from several sites. Like the NWTC time-series, the autocorrelation function in other cases were also close to zero for $\tau = 10$ min (not shown); for these cases, the samples could be considered independent. For higher values of τ , the samples were correlated.

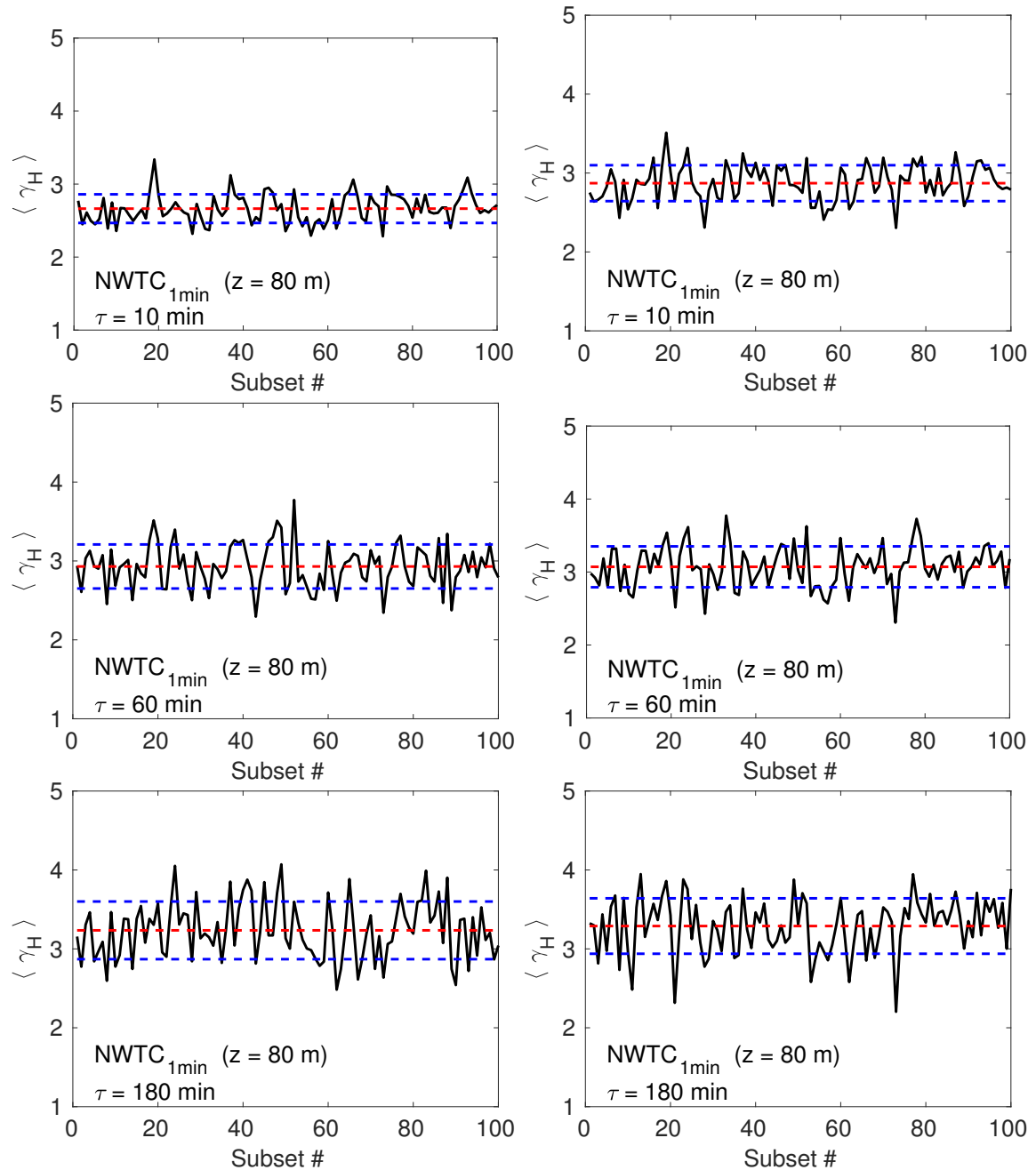


Figure 13. The averaged values of tail indices ($\langle \gamma_H \rangle$) based on the non-overlapping and contiguous subsets of NWTC_{1min} time-series. The left and right panels correspond to ramp-up and ramp-down events, respectively. Time increments (τ) of 10 min, 60 min, and 180 min are utilized in top, middle, and bottom panels, respectively. For visual aid, the temporal mean values of $\langle \gamma_H \rangle$ and one standard deviation around the mean are depicted with the red and blue dashed lines, respectively.

ACKNOWLEDGEMENTS

We are grateful to several people and multiple agencies for making the following long-term tower data available for this research: BMU (Bundesministerium für Umwelt), Federal Ministry for the Environment, Nature Conservation and Nuclear Safety for the FINO 1 dataset; Yoram Eisenberg from Technical University of Denmark – DTU Wind Energy for the Høvsøre dataset; Fred Bosveld (Royal Netherlands Meteorological Institute) for the Cabauw tower data; and The

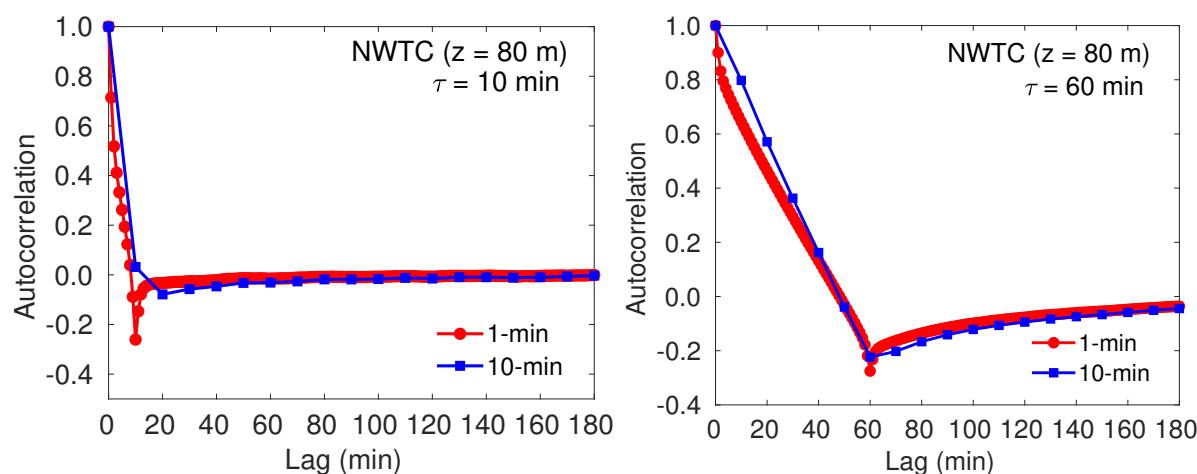


Figure 14. Autocorrelation functions based on NWTC_{1min} (red lines with circles) and NWTC_{10min} (blue lines with squares) time-series. The left and right panels correspond to time increments (τ) of 10 min and 60 min, respectively.

National Renewable Energy Laboratory (NREL) for the NWTC dataset. This work was partially funded by the National Science Foundation (AGS-1632679).

REFERENCES

1. Greaves B, Collins J, Parkes J, Tindal A. Temporal forecast uncertainty for ramp events. *Wind Engineering* 2009; **33**:309–319.
2. Ferreira C, Gama J, Matias L, Botterud A, Wang J. A survey on wind power ramp forecasting. *Technical Report*, Argonne National Laboratory (ANL) 2011.
3. Yang Q, Berg LK, Pekour M, Fast JD, Newsom RK, Stoelinga M, Finley C. Evaluation of WRF-predicted near-hub-height winds and ramp events over a Pacific Northwest site with complex terrain. *Journal of Applied Meteorology and Climatology* 2013; **52**:1753–1763.
4. Gallego-Castillo C, Cuerva-Tejero A, Lopez-Garcia O. A review on the recent history of wind power ramp forecasting. *Renewable and Sustainable Energy Reviews* 2015; **52**:1148–1157.
5. Kamath C. Understanding wind ramp events through analysis of historical data. *Transmission and Distribution Conference and Exposition, 2010 IEEE PES*, IEEE, 2010; 1–6.
6. Freedman J, Markus M, Penc R. Analysis of West Texas wind plant ramp-up and ramp-down events. *AWS Truewind, LLC* 2008; .
7. Zheng H, Kusiak A. Prediction of wind farm power ramp rates: A data-mining approach. *Journal of Solar Energy Engineering* 2009; **131**:031 011.
8. Turner R, Zheng X, Gordon N, Uddstrom M, Pearson G, de Vos R, Moore S. Creating synthetic wind speed time series for 15 New Zealand wind farms. *Journal of Applied Meteorology and Climatology* 2011; **50**:2394–2409.
9. Bossavy A, Girard R, Kariniotakis G. Forecasting ramps of wind power production with numerical weather prediction ensembles. *Wind Energy* 2013; **16**:51–63.
10. Schmitt F, Schertzer D, Lovejoy S, Brunet Y. Empirical study of multifractal phase transitions in atmospheric turbulence. *Nonlinear Processes in Geophysics* 1994; **1**:95–104.
11. Ragwitz M, Kantz H. Indispensable finite time corrections for Fokker-Planck equations from time series data. *Physical Review Letters* 2001; **87**:254 501.
12. Boettcher F, Renner C H, Waldl, H-P, Peinke J. On the statistics of wind gusts. *Boundary-Layer Meteorology* 2003; **108**:163–173.
13. Böttcher F, Barth St, Peinke J. Small and large scale fluctuations in atmospheric wind speeds. *Stochastic Environmental Research and Risk Assessment* 2007; **21**:299–308.
14. Muzy JF, Baïle R, Poggi P. Intermittency of surface-layer wind velocity series in the mesoscale range. *Physical Review E* 2010; **81**:056 308.
15. Liu L, Hu F, Cheng SLL X-L. Probability density functions of velocity increments in the atmospheric boundary layer. *Boundary-Layer Meteorology* 2010; **134**:243–255.

16. Liu L, Hu F. Cascade-like and scaling behavior of wind velocity increments in the atmospheric surface layer. *Physica A* 2013; **392**:5808–5816.
17. Kiliyanpilakkil VP, Basu S, Ruiz-Columbié A, Araya G, Castillo L, Hirth B, Burgett W. Buoyancy effects on the scaling characteristics of atmospheric boundary layer wind fields in the mesoscale range. *Physical Review E* 2015; **92**:033 005.
18. Kiliyanpilakkil VP, Basu S. Extended self-similarity of atmospheric boundary layer wind fields in mesoscale regime: Is it real? *Europhysics Letters* 2016; **112**:64 003.
19. Neumann T, Nolopp K, Strack M, Mellinghoff H, Söker H, Mittelstaedt E, Gerasch W, Fischer G. Erection of German offshore measuring platform in the North Sea. *DEWI Magazin* 2003; **23**:32–46.
20. Türk M, Grigutsch K, Emeis S. The wind profile above the sea – investigations basing on four years of FINO-1 data. *DEWI Magazin* 2008; **33**:12–16.
21. Ernst B, Seume JR. Investigation of site-specific wind field parameters and their effect on loads of offshore wind turbines. *Energies* 2012; **5**:3835–3855.
22. Jørgensen HE, Mikkelsen T, Gryning SE, Larsen S, Astrup P, Sørensen PE. Measurements from Høvsøre met mast. *Technical Report*, Technical Report Risø 2008.
23. Peña A, Floors R, Sathe A, Gryning SE, Wagner R, Courtney MS, Larsén XG, Hahmann AN, Hasager CB. Ten years of boundary-layer and wind-power meteorology at Høvsøre, Denmark. *Boundary-Layer Meteorology* 2016; **158**:1–26.
24. Monna WAA, Van der Vliet JG. *Facilities for research and weather observations on the 213 m tower at Cabauw and at remote locations*. KNMI De Bilt, The Netherlands, 1987.
25. van Ulden AP, Wieringa J. Atmospheric boundary layer research at Cabauw. *Boundary-Layer Meteorology* 1996; **78**:39–69.
26. Verkaik JW, Holtslag AAM. Wind profiles, momentum fluxes and roughness lengths at Cabauw revisited. *Boundary-Layer Meteorology* 2007; **122**:701–719.
27. Kelley ND, Scott GN, Jonkman BJ. The great plains turbulence environment: its origins, impact and simulation. NREL/CP-500-40176, Golden, CO, 2006.
28. Drees H, De Haan L, Resnick S. How to make a Hill plot. *The Annals of Statistics* 2000; **28**:254–274.
29. Gumbel EJ. *Statistics of Extremes*. Columbia Univ. Press, New York, 1958.
30. Embrechts P, Klüppelberg C, Mokosh T. *Modelling Extremal Events: for Insurance and Finance*. Springer-Verlag, Heidelberg, 1997. 648 pp.
31. De Haan L, Ferreira A. *Extreme Value Theory: An Introduction*. Springer Science+Business Media, New York, 2007.
32. Crovella ME, Taquu MS, Bestavros A. Heavy-tailed probability distributions in the World Wide Web. *A Practical Guide to Heavy Tails: Statistical Techniques and Applications*, Adler RJ, Feldman RE, Taquu MS (eds.), Birkhäuser, Boston, 1998; 3–26.
33. Sornette D. *Critical Phenomena in Natural Sciences: Chaos, Fractals, Self-organization and Disorder: Concepts and Tools*. Springer-Verlag, Heidelberg, 2006.
34. Zipf GK. *Human Behavior and the Principle of Least Effort: an Introduction to Human Ecology*. Addison-Wesley Press, Cambridge, 1949.
35. Pickands III J. Statistical inference using extreme order statistics. *The Annals of Statistics* 1975; **3**:119–131.
36. Hill BM. A simple general approach to inference about the tail of a distribution. *The Annals of Statistics* 1975; **3**:1163–1174.
37. Dekkers ALM, Einmahl JHJ, De Haan L. A moment estimator for the index of an extreme-value distribution. *The Annals of Statistics* 1989; **17**:1833–1855.
38. Bibby BM, Sørensen M. Hyperbolic processes in finance. *Handbook of Heavy Tailed Distributions in Finance*, Rachev ST (ed.), Elsevier Science, Philadelphia, 2003; 211–248.
39. Aas K, Hobæk Haff I. The generalized hyperbolic skew students t-distribution. *Journal of Financial Econometrics* 2006; **4**:275–309.
40. DeMarco AW, Basu S. Estimating higher-order structure functions from geophysical turbulence time series: Confronting the curse of the limited sample size. *Physical Review E* 2017; **95**:052 114.
41. Negra NB, Holmstrøm O, Bak-Jensen B, Sørensen P. Model of a synthetic wind speed time series generator. *Wind Energy* 2008; **11**:193–209.
42. D'Amico G, Petroni F, Prattico F. First and second order semi-Markov chains for wind speed modeling. *Physica A* 2013; **392**:1194–1201.
43. Aas K, Hobæk Haff I. NIG and skew student's t: Two special cases of the generalised hyperbolic distribution 2005. SAMBA/01/05, Norwegian Computing Center, Oslo.
44. Paoletta MS. *Intermediate Probability: A Computational Approach*. John Wiley & Sons, Ltd., Chichester, 2007. 415 pp.

45. Abramowitz M, Stegun IA. *Handbook of Mathematical Functions: with Formulas, Graphs, and Mathematical Tables*. Dover, New York, 1972.
46. Mahrt L. The influence of nonstationarity on the turbulent flux–gradient relationship for stable stratification. *Boundary-Layer Meteorology* 2007; **125**:245–264.
47. Schreiber T. Detecting and analyzing nonstationarity in a time series using nonlinear cross predictions. *Physical Review Letters* 1997; **78**:843–846.
48. Huang YX, Schmitt FG, Lu ZM, Liu YL. Autocorrelation function of velocity increments time series in fully developed turbulence. *EPL* 2009; **86**:40 010.