Extreme Value Analysis of Multivariate High Frequency Wind Speed Data

Christina Steinkohl

Center of Mathematical Sciences, Technische Universität München, D-85748 Garching, Germany

Richard A. Davis

Department of Statistics, Columbia University, New York, United States

Claudia Klüppelberg

Center of Mathematical Sciences, Technische Universität München, D-85748 Garching, Germany

Summary. In this paper we analyze the extremal behaviour of wind speed with a measurement frequency of 8 Hz located on three meteorological masts in Denmark. In the first part of this article we set up a conditional model for the time series consisting of threshold exceedances from maxima per second for two consecutive days. The model directly captures the non-stationary nature of wind speed during the day. We assume that the conditional distribution of an exceedance given previous exceedances follows a generalized Pareto distribution. In addition, we analyze the dependence structure in extreme wind speeds between two masts using bivariate extreme value models. The initial motivation for this research was in the context of renewable energy. Specifically, the extremal dynamics of wind speed at small time scales plays a critical role for designing and locating turbines on wind farms.

Keywords: wind speed, generalized Pareto distribution, bivariate extreme value theory

1. Introduction

In the context of renewable energy, wind power has become the most dynamically growing energy source and, especially in the last few years, the installation of new wind turbines has expanded rapidly. According to the World Wind Energy Report World Wind Energy Association (2009), wind energy increased at a rate of 29% from 2007 to 2008. For the production of electrical power, a group of wind turbines in the same location, called wind farm, is used. To maximize the power output, wind farms are built offshore or in open fields far away from buildings and trees. Each wind turbine is usually equipped with instruments measuring wind speed, but these observations are disturbed from the turbulences produced from the big rotors. Before building a wind farm, one or more meteorological masts, known as met masts, are erected and characteristics of wind are measured at a targeted location. Since much, if not all of this data is proprietary, we were restricted to only publicly available data, which were in turn viewed as a proxy for the analysis one might carry out at a potential site.

E-mail: steinkohl@ma.tum.de
E-mail: rdavis@stat.columbia.edu
E-mail: cklu@ma.tum.de

One key objective in applying extreme value theory to wind speed data is often the determination of return levels corresponding to a specified return period. For a given return period the return level is the wind speed at which the probability of exceedance is 1/(return period). Often, engineers must design structures that can withstand extreme meteorological conditions for 50 or 100 years. In this case, it is vital for designers to have a good estimate of 50 or 100 year return levels for wind speed. The above described considerations are the motivation of many studies, including for instance Holmes and Moriarty (1999) and Walshaw and Anderson (2000). These previous studies mainly focus on large-scale weather events corresponding to macrometeorological fluctuations in wind with a time range of days down to hours. Our primarily interest lies in the analysis of velocity measurements on a finer time scale in the so-called micrometeorological range. Our data set consists of observations measured in the atmospheric boundary layer on two different days with a frequency of 1 Hz. The two days show different wind situations, including a windy day with wind speeds up to 25 meters per second, and a day, where the wind speed is decreasing during the day. We are dealing with time series from three different cup anemometers situated on measurement masts at a height of 30 meters above ground located in Denmark.

The main objective of this study is to model tail-behaviour of wind speed over a time-dependent threshold by using extreme value theory and in particular the Peaks-Over-Threshold approach. This method has received much attention and there exist standard textbooks, including Coles (2001), Embrechts et al. (1997) or Beirlant et al. (2004), where methods for inference are described. The non-stationary nature of the time series leads us to a conditional model, where we assume that, given previous wind speed values, the exceedances over a time-dependent threshold possess a Markov-like structure, where the conditional distribution follows the generalized Pareto distribution (GPD) with time-dependent parameters. From the fitted distributions we estimate one-step ahead quantiles, which predict the risk of an extreme wind speed value within the next second.

It is stated in Chapter 2 of the wind energy handbook by Burton et al. (2001) that "on still shorter time-scales of minutes down to seconds or less, wind speed variations can have a very significant effect on the design and performance of the individual wind turbines, as well as on the quality of power delivered to the network and its effect on consumers." Our models adjust for these intermittency effects present in the 1-second observations by allowing the scale parameter in the model to be a function of previous large values. We show that our technique works well for the Danish Lammefjord data set, especially on days in which there are extended periods with highly volatile wind speed.

The starting point in most extreme value models is to assume a GPD for the distribution of exceedances. The shape parameter ξ in the GPD family is perhaps the most interesting since it determines the tail behaviour of the exceedance distribution and is directly linked to the choice of extreme value distribution for the maximum. In many articles, including for example Coles and Walshaw (1994), who modeled hourly maximum wind gust speeds together with the wind gust direction, the shape parameter in the extreme value distribution is negative, which implies a distribution in the Weibull (Type III) domain of attraction, not to be confused with a standard Weibull distribution. This type of extreme value distribution has a finite right endpoint $x^* = \sup\{x \in \mathbb{R} : F(x) < 1\}$ of support and we show that the conditional distribution for the Lammefjord velocity data also lies in this domain of attraction. This is also supported by the studies of Holmes and Moriarty (1999), Walshaw and Anderson (2000) and Simiu and Heckert (1996).

In traditional wind engineering hourly mean wind speeds are usually modelled by the Weibull distribution (see for instance Burton et al. (2001)) which, in the extreme value

setting, is a distribution in the Gumbel domain of attraction. Since we are dealing with velocities on a finer time scale this is not in discrepancy to our approach. On finer time scales down to less than a second maxima are often modeled by the Rice distribution (Ronold and Larsen, 1999). This is based on the assumption that the underlying signal is a Gaussian process. Recent developments about turbulence and velocity show that the Gaussian process assumption may not be appropriate for micrometeorological measurements which might cast doubt on the suitability of the Rice distribution for maxima. For example see Barndorff-Nielsen and Schmiegel (2007).

We also analyze the extremal dependence in large velocities measured at different masts using bivariate extreme value theory. In order to account for the movement of wind, we investigate the time-lag giving the highest dependence at different masts.

A detailed description about the data set can be found in Section 2. Section 3 develops the modelling framework we use for the univariate wind speed time series, and we describe a procedure to measure extremal dependence between wind speed records at different masts. Summary comments are made in Section 4.

2. Description of wind speed data set

The data used in this study can be downloaded from a database on wind characteristics, that is supervised by Kurt S. Hansen from the Technical University of Denmark (Lyngby, north of Copenhagen) together with Gunner C. Larsen from the Risø National Laboratories in Roskilde (Denmark) (see `www.winddata.com`). The database provided by the Lammefjord Station in Denmark consists of 8 Hz (8 measurements per second) observations measured by cup anemometers and wind vanes situated 30 meters above ground. Thus, the velocity measurements are taken from the atmospheric boundary layer (ABL) of the troposphere. The ABL is directly influenced by the Earth's surface friction from vegetation and topography and velocities display rapid fluctuations. The terrain in Lammefjord is flat and homogeneous. The measurement system consists of 3 meteorological masts which are located in Lammefjord, a reclaimed fjord on the Danish island of Zealand. On each mast several instruments are erected at different heights, including cup anemometers and wind vanes. The heights are 10, 20 and 30 meters, respectively and we choose the 30 meters height data for our analysis. In terms of wind farm analysis the Lammefjord data set represents free stream conditions, since there are no wind turbines at the site.

Figure 1 shows the allocation of masts in Lammefjord together with the location of the instruments from which the data were collected. In the following we refer to the leftward most mast as mast 1, the middle mast as mast 2 and the mast on the right hand side as mast 3.

Cup anemometers are mechanical instruments with a vertical axis of rotation, usually consisting of three or four hemispherical cups mounted, where the rate of rotation measures the wind speed. There is data available for year 1987 and we choose two days such that we have no missing values and that the days represent different wind situations. In the dataset we found two days with these properties, namely July 12th and July 13th. Day 1 (corresponding to July 12th) was a windy day with maximum wind speed of 23.84 meters per second measured at mast 1. In day 2 the wind speed is decreasing over time. Since we are mostly interested in modelling large wind speeds and for computational convenience, we calculate the maxima per second for each day and work with the six resulting univariate time series, each having a length of 86 400 measurements. Figure 2 shows the time series for day 1

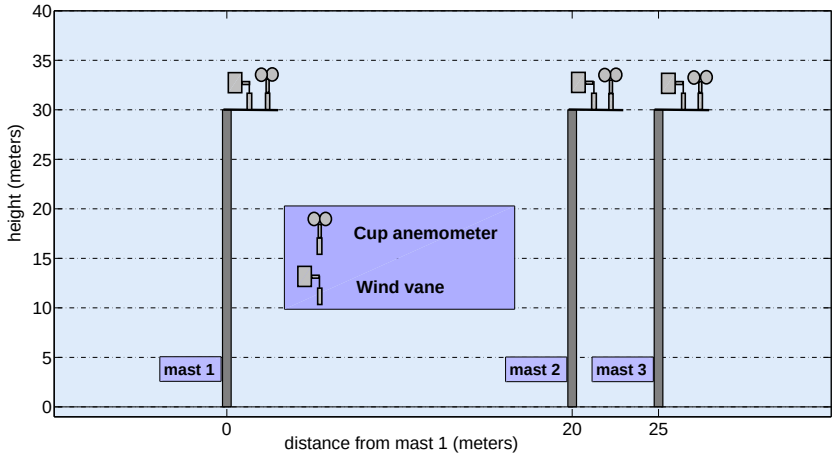

Fig. 1. Allocation of masts in Lammefjord

and 2 (mast 1). The dataset also provides information for the wind direction, from which we know, that the wind is mainly blowing from the east to the west on both days. This means, that the wind first passes mast 1 and then mast 2 and 3. Since we are dealing with one direction, only, the wind direction unfortunately is not a possible variable to describe the time-dependence of wind speed for consecutive observations. We also applied the following analyses to the velocity observations measured by the anemometers at 10 meters height and obtained very similar results.

3. Modelling wind speed threshold exceedances

3.1. Univariate modelling

In the following, we describe the modelling framework for the univariate wind speed time series. The main objective is to model large velocities over time. Based on this model, we estimate one-step ahead conditional quantiles. Short term extreme wind events are of importance, since they can cause extreme loading as described in the introduction.

There is an extensive literature available on univariate extreme value theory and modelling threshold exceedances by the generalized Pareto distribution. A detailed introduction can for instance be found in Embrechts et al. (1997) and Coles (2001) and we just describe the basic theory and introduce the notation needed. For $X_1, \ldots, X_n$ independent and identically distributed (i.i.d.) random variables with distribution function F we define

$$M_n = \max\{X_1, \ldots, X_n\}, \quad n \in \mathbb{N}.$$

The main objective of extreme value theory concerns the determination of the limiting distribution G of $(M_n - b_n)/a_n$, for $n \to \infty$, where (a_n) with $a_n > 0$ and (b_n) are sequences of constants. If G is a non-degenerate distribution function, the limiting distribution is given by the generalized extreme value distribution

$$G_\xi(x) = \begin{cases} \exp\left\{-\left(1+\xi\frac{x-\mu}{s}\right)^{-1/\xi}\right\}, & 1+\xi\frac{x-\mu}{s} > 0,\ \xi \neq 0, \\ \exp\left\{-e^{-(x-\mu)/s}\right\}, & x \in \mathbb{R},\ \xi = 0. \end{cases} \tag{1}$$

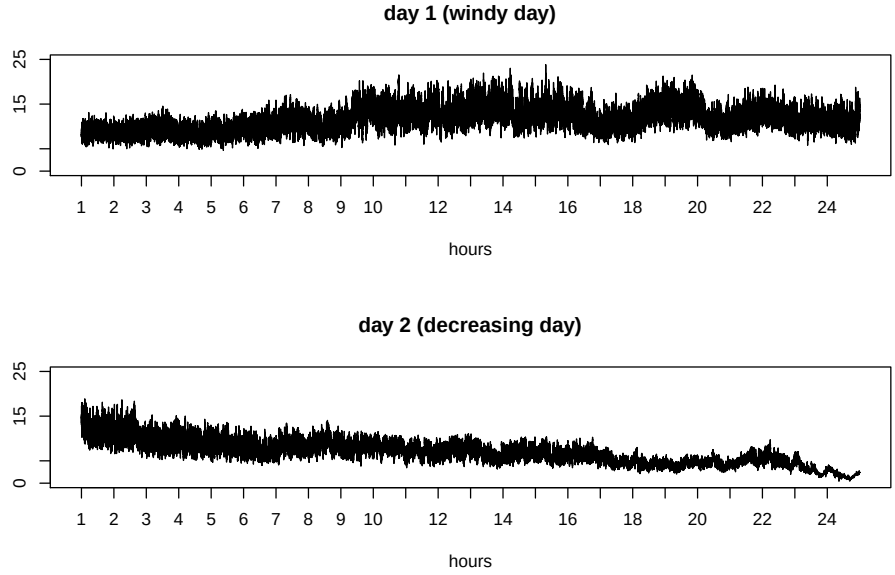

Fig. 2. Maxima per second wind speed time series for day 1 and day 2 (mast 1)

We say that F lies in the maximum domain of attraction of G_ξ, denoted by $F \in \mathrm{MDA}(G_\xi)$. The extension of this result from the i.i.d. case to a stationary time series satisfying some additional mixing conditions was shown by Leadbetter (1974, 1983). The shape parameter $-\infty < \xi < \infty$ determines the type of extreme value distribution according to the Gumbel family (Type II) for $\xi = 0$, the Fréchet family (Type I) for $\xi > 0$ and the Weibull family (Type III) for $\xi < 0$. The main objective of our analysis is the modelling of extreme wind speeds over a high threshold. The excess distribution for a random variable X with cumulative distribution function $F \in \mathrm{MDA}(G_\xi)$, which has right boundary point $x_* = \sup\{x \in \mathbb{R} : F(x) < 1\}$, is given for thresholds $u < x_*$ by

$$F_u(x) = P(X \leq u + x \mid X > u), \quad x \geq 0. \tag{2}$$

Pickands (1975) showed that this distribution can be approximated for high thresholds u by the generalized Pareto distribution (GPD) with distribution function

$$H_{\xi,\sigma}(x) = \begin{cases} 1 - \left(1 + \xi\frac{x}{\sigma}\right)_+^{-1/\xi}, & \xi \neq 0, \\ 1 - e^{-x/\sigma}, & \xi = 0, \end{cases} \qquad x > 0, \tag{3}$$

where $x_+ = \max\{x, 0\}$. The shape parameter ξ has a similar interpretation as in the generalized extreme value distribution and characterise the tail of the distribution.

Given a realization of an i.i.d. sample $X_1, \ldots, X_n$ from the unknown distribution function F, we define $Y_1, \ldots, Y_{N_u}$ to be the corresponding exceedances $Y_j = X_{T_j} - u$, $j = 1, \ldots, N_u$, where N_u is the number of values which exceed the threshold u and $T_1, \ldots, T_{N_u}$ are the exceedance times. The parameters of the GPD are estimated by maxi-

mizing the log-likelihood function

$$l(\xi, \sigma; Y_1, \ldots, Y_{N_u}) = -N_u \log(\sigma) - \left(\frac{1}{\xi} + 1\right) \sum_{i=1}^{N_u} \log\left(1 + \xi \frac{Y_i}{\sigma}\right).$$

The tail probability

$$\overline{F}(x) = P(X > x) = \overline{F}(u)\overline{F}_u(x - u), \quad x > u$$

can then be approximated by

$$\hat{\overline{F}}(x) = \frac{N_u}{n}\left(1 + \hat{\xi}\frac{x - u}{\hat{\sigma}}\right)^{-1/\hat{\xi}}, \quad x > u.$$

By inspecting the time series plots for the Lammefjord wind speed data in Figure 2, we clearly see the non-stationary nature of wind speed. There exist different approaches for handling the non-stationarity in time series. Davison and Smith (1990) were the first, who suggested to use the GPD as basis and enhanced the method to the non-stationary case by allowing the parameters to be modelled as functions of covariates. Another description of this approach can be found in Coles (2001) or Beirlant et al. (2004). In a more recent study by Eastoe and Tawn (2009), who studied daily maxima of hourly ozone concentrations, some approaches dealing with non-stationarity in the GPD- approach are discussed and summarized. As a starting point we estimate the parameters of the GPD using subperiod-samples. The length of each subperiod is chosen to be 600, since within this 10-minutes time range, the time series seem to be stationary. It is common agreement in wind engineering, that wind is stationary within a 10-minutes period. In each subperiod of length 600 we choose the threshold such that 60 values lie above the threshold. We assume that within each subperiod the exceedances form a stationary sequence of random variables. The results of this exploratory analysis showed that the shape parameter ξ stays almost constant for all 144 subperiods. The scale parameter σ varies over time and, based on the estimated values for σ, we tried several regression models with response variable σ_{T_j} and covariates Y_j, X_{T_j} and $X_{T_{j-1}}$, respectively, to obtain an appropriate structure for our model. In this way, we account for intermittency effects present in the univariate wind speed time series. In our further analysis we use a time-dependent threshold $u = u_t$, which is calculated by a rolling-window procedure with a window length of 600 seconds, where the threshold is chosen as the 98% empirical quantile in each window based on the previous 600 wind speed values. With such a high threshold, the assumption of a generalized Pareto distribution as an approximate distribution may be reasonable. We define the exceedance times by $T_1, \ldots, T_{N_u}$ and the corresponding exceedances by

$$Y_j = X_{T_j} - u_{T_j}, \quad j = 1, \ldots, N_u.$$

Conditional on previous wind speeds with recorded exceedances, we assume that the exceedances possess a Markov-like structure. The scale parameter is modelled through a generalized linear model with exponential inverse link function, where we make the scale parameter dependent on previous wind speed values with recorded exceedance. The shape parameter is set constant over time. For $j = 2, \ldots, N_u$, this leads to the following model:

$$\begin{aligned} Y_j \mid \mathcal{F}_{T_{j-1}} &\sim \text{GPD}(\sigma_{T_j}, \xi), \\ \log(\sigma_{T_j}) &= \alpha_0 + \alpha_1 X_{T_{j-1}}, \end{aligned} \tag{4}$$

Table 1. MLEs for the marginal GPD-fits and confidence intervals for the shape parameter ξ (CI(ξ)).

	$\hat{\alpha}_0$	$\hat{\alpha}_1$	$\hat{\xi}$	CI($\hat{\xi}$)
day 1, mast 1	-1.573	0.083	-0.125	$[-0.163, -0.091]$
	(0.120)	(0.007)	0.013	
day 1, mast 2	-1.689	0.084	-0.083	$[-0.128, -0.053]$
	(0.127)	(0.007)	(0.013)	
day 1, mast 3	-1.558	0.077	-0.109	$[-0.148, -0.075]$
	(0.126)	(0.007)	(0.018)	
day 2, mast 1	-1.854	0.110	-0.113	$[-0.149, -0.077]$
	(0.059)	(0.006)	(0.015)	
day 2, mast 2	-2.006	0.121	-0.071	$[-0.112, -0.037]$
	(0.064)	(0.006)	(0.018)	
day 2, mast 3	-1.937	0.115	-0.066	$[-0.097, -0.022]$
	(0.063)	(0.006)	(0.018)	

where $\mathcal{F}_{T_{j-1}}$ denotes the σ-algebra generated by $X_{T_1}, \ldots, X_{T_{j-1}}$ and contains all information up to time T_{j-1}. The conditional density for the exceedances, given the previous wind speed observations for which there is an exceedance, is given by

$$f_{Y_j|X_{T_{j-1}}}(y) = \frac{1}{\exp(\alpha_0 + \alpha_1 X_{T_{j-1}})} \left(1 + \xi \frac{y}{\exp(\alpha_0 + \alpha_1 X_{T_{j-1}})}\right)^{-1/\xi - 1}$$

for $\xi \neq 0$. The log-likelihood for observed exceedances $Y_1, \ldots, Y_{N_u}$ can, therefore, be expressed as

$$\begin{aligned} &l(\alpha_0, \alpha_1, \xi; Y_1, \ldots, Y_{N_u}) = \\ &-\sum_{j=2}^{N_u} \left(\alpha_0 + \alpha_1 X_{T_{j-1}}\right) - \left(\frac{1}{\xi} + 1\right) \sum_{j=2}^{N_u} \log\left(1 + \xi \frac{Y_j}{\exp(\alpha_0 + \alpha_1 X_{T_{j-1}})}\right). \end{aligned}$$

In Table 1 we list the estimates and the corresponding standard errors resulting from the maximum-likelihood estimation. In addition, we approximate confidence bounds using the limiting distribution of the maximum-likelihood estimates (MLEs) as calculated in Smith (1987, Section 2) leading to the formula

$$\hat{\xi} \pm z_{\alpha/2} N_u^{-1/2} (1 + \hat{\xi}), \tag{5}$$

where $z_{\alpha/2}$ is the $(1 - \alpha/2)$-quantile of a standard normal distribution. In all cases the shape parameters are slightly below zero and the confidence intervals do not contain the value zero, which corresponds to distributions in the Weibull domain of attraction with finite right endpoint of support. This result is consistent with other studies on wind speed data as already motivated in the introduction.

To test the goodness of fit we use probability-probability (pp-plots) and quantile-quantile plots (qq-plots). In order to apply such diagnostic checks, the observations have to be standardized as described in Coles (2001, Section 6.2). The transformation given for $j = 2, \ldots, N_u$ by

$$\tilde{Y}_j = -\log\left(1 - \hat{F}_u(Y_j)\right) = \frac{1}{\hat{\xi}} \log\left(1 + \hat{\xi} \frac{Y_j}{\exp(\hat{\alpha}_0 + \hat{\alpha}_1 X_{T_{j-1}})}\right)$$

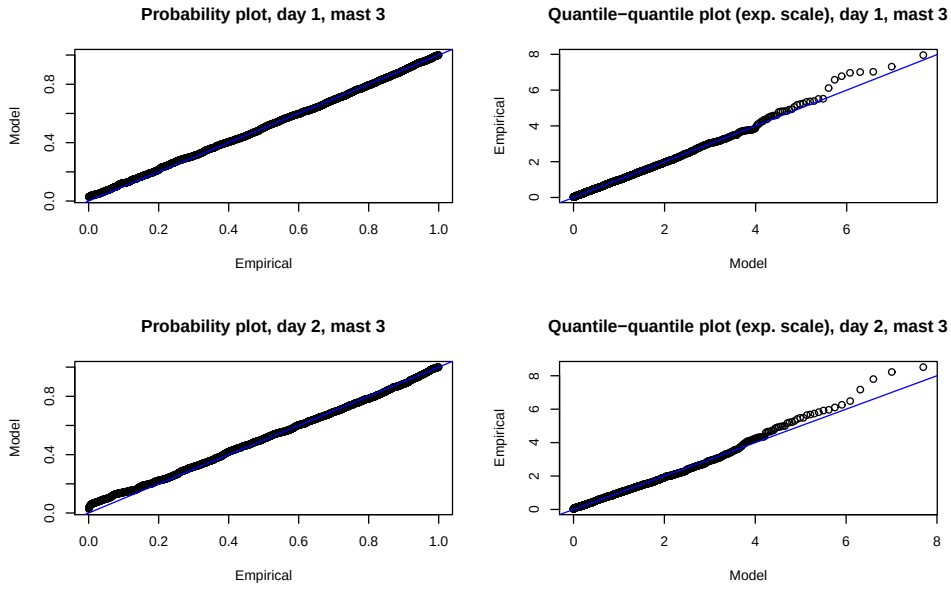

Fig. 3. Diagnostic plots for the marginal GPD fits. Left: pp-plots for day 1 (top) and day 2 (bottom). Right: qq-plots for day 1 (top) and day 2 (bottom).

should lead to exponentially distributed random variables, if the fitted distribution is appropriate. Figure 3 shows pp-plots and qq-plots, assuming exponential distributions for $(\tilde{Y}_j)$ for day 1 and 2 (mast 3). The approximate straight line pattern in all plots leads us to the conclusion that the estimated model provides a plausible statistical fit for the given data. The plots for the other masts (not shown) look very similar. For further testing of the significance according the shape parameter, we estimate the parameters α_0 and α_1 in a setting with the shape parameter ξ forced to be zero leading to a Gumbel distribution. By comparing the two resulting models via inspecting the probability and quantile-quantile plots and calculating the Akaike-Information criterion (AIC), we conclude that ξ is below zero.

In the following we determine conditional quantile plots for the original data. Let n denote the total number of observations and define $\tilde{X}_t := \max\{X_t, u_t\}$ for $t = 1, \ldots, n$. The conditional quantiles are then defined for $t = 2, \ldots, n$ by

$$Q_t(p) = \inf\left\{ z \in \mathbb{R} : P(\tilde{X}_t \leq z \mid \tilde{X}_{t-1}) \geq p \right\}, \quad \text{for } p \in [0,1].$$

The conditional distribution for $\tilde{X}_t$ given $\tilde{X}_{t-1}$ for $t = 2, \ldots, n$ can be calculated as

$$\begin{aligned} P\left(\tilde{X}_t \leq z \mid \tilde{X}_{t-1} = \tilde{x}_{t-1}\right) &= P\left(\max\{X_t, u_t\} \leq z \mid \tilde{X}_{t-1} = \tilde{x}_{t-1}\right) \\ &= P\left(X_t \leq z,\ u_t \leq z \mid \tilde{X}_{t-1} = \tilde{x}_{t-1}\right) \\ &= \left(1 - \tfrac{N_u}{n}\left(1 + \xi \tfrac{x - u_t}{\exp\{\alpha_0 + \alpha_1 \tilde{x}_{t-1}\}}\right)^{-1/\xi}\right) \mathbf{1}_{\{z \geq u_t\}}, \end{aligned}$$

where $\mathbf{1}_B$ denotes the indicator function for some set B. For z large enough, we can calculate

Table 2. Validation for one-step ahead conditional 99%-quantile estimates: absolute and relative number of excesses over the quantile estimates.

	number of excesses	proportion above quantile estimate
day 1, mast 1	872	1.01%
day 1, mast 2	873	1.01%
day 1, mast 3	865	1.00%
day 2, mast 1	885	1.02%
day 2, mast 2	865	1.00%
day 2, mast 3	879	1.02%

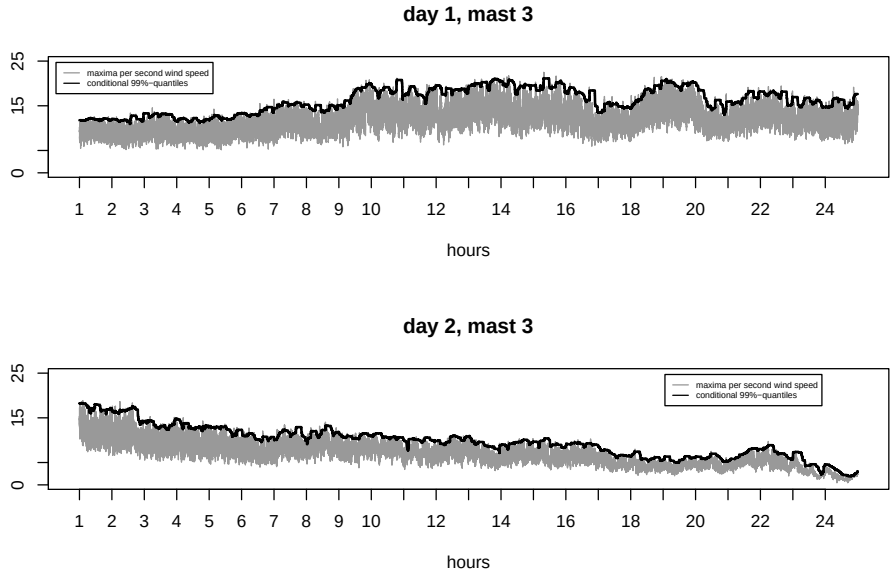

Fig. 4. Conditional quantile estimates: one-step ahead prediction for day 1 (top) and day 2 (bottom).

the one-step ahead conditional quantiles as

$$Q_t(p) = u_t + \frac{1}{\xi}\exp\left\{\alpha_0 + \alpha_1 \tilde{X}_{t-1}\right\}\left(\left(\frac{n}{N_u}(1-p)\right)^{-\xi} - 1\right), \quad p \in [0,1].$$

The quantile estimates are obtained by plugging in the MLEs. The main objective of extreme value theory is to model very large values and, hence, we also want to obtain large quantile estimates. Since the threshold was chosen such that 2% of the data lie above the threshold in each subperiod, reasonable estimates for the conditional quantiles are given for p larger than 0.98. In order to validate our quantile estimates, we estimate the 99% quantile and then compute the proportion of observations that exceed these quantiles. As seen in Table 2, the frequencies of these exceedances match the targeted exceedance probability of 1% reasonably well.

Table 3. Pickands dependence functions for the parametric models.

Model	Pickands dependence function
logistic	$A(u) = (u^{1/r} + (1-u)^{1/r})^r$,
asymmetric logistic	$A(u) = (\theta^r u^r + \phi^r (1-u)^r)^{1/r} - (\theta - \phi)u + 1 - \phi$,

3.2. Dependence structure for extreme wind speed values

The next step in our analysis is to estimate the extremal dependence between velocities measured at the investigated observation points. Based on the fitted distributions for the marginals, we determine estimates for the joint distribution function for bivariate threshold exceedances of wind speed observations from two different masts. The Pickands dependence function is a commonly used tool for measuring the extremal dependence. It can be shown (Beirlant et al., 2004, Chapter 8) that a bivariate extreme value distribution has the general representation

$$G(x,y) = \exp\left\{(v_1+v_2)A\left(\frac{v_2}{v_1+v_2}\right)\right\}, \quad x,y \in \mathbb{R}, \tag{6}$$

where

$$v_1 = \log(G_X(x)), \quad v_2 = \log(G_Y(y)),$$

G_X and G_Y are the marginal distributions of G, and A is the Pickands dependence function. Since G is a bivariate extreme value distribution, G_X and G_Y are univariate extreme value distributions. The dependence function A satisfies the following properties:

- $A(0) = A(1) = 1$ and $A(u) \geq 1/2$ for all $u \in [0,1]$.
- A is convex on $[0,1]$.
- The marginal components are independent if and only if $A(u) = 1$ for all $u \in [0,1]$, and
- If $A(u) = \max\{u, 1-u\}$ for all $u \in [0,1]$, the components are completely dependent.

We use a parametric and a non-parametric approach to estimate A from the data. We shortly introduce the formulas used in our analysis. A more detailed description of the estimates can be found in the Appendix. We focus on two parametric examples of the Pickands dependence function given by the logistic (Gumbel, 1958) and the asymmetric logistic model (Tawn, 1988) (see Table 3 for $0 < r \leq 1$, $0 \leq \theta \leq 1$ and $0 \leq \phi \leq 1$). The parameters are estimated using a censored likelihood approach as described in Coles (2001, Section 8.3.1), which accounts for the different combinations of a bivariate pair of points, where one component could lie below the threshold and the other component is above the threshold.

A non-parametric estimate for A, based on pseudo-polar coordinates, is given as follows:

$$\hat{A}(u) = \frac{2}{k}\sum_{t=1}^{n} \mathbf{1}_{\{\hat{R}_t > \hat{R}_{(n-k)}\}} \max\{u\hat{\omega}_{X,t}, (1-u)\hat{\omega}_{Y,t}\},$$

where

$$\hat{R}_t = X_{*t} + Y_{*t} \quad \text{and} \quad \hat{\omega}_{X,t} = \frac{X_{*t}}{\hat{R}_t},\ \hat{\omega}_{Y,t} = \frac{Y_{*t}}{\hat{R}_t}, \quad t = 1,\ldots,n,$$

$$X_{*t} = \begin{cases} \frac{1}{1-\hat{F}_{X_t}(X_t)}, & X_t > u_{X,t}, \\ 1, & X_t \leq u_{X,t}, \end{cases} \qquad Y_{*t} = \begin{cases} \frac{1}{1-\hat{F}_{Y_t}(Y_t)}, & Y_t > u_{Y,t}, \\ 1, & Y_t \leq u_{Y,t}, \end{cases}$$

k denotes the number of exceedances in each individual component and $\hat{R}_{(1)} < \cdots < \hat{R}_{(n)}$ are the order statistics of $(\hat{R}_t)$. A convex modification of this estimator, which ensures the properties mentioned above, is given by

$$\tilde{A}(u) = \max\left\{u, 1-u, \hat{A}(u) + 1 - (1-u)\hat{A}(0) + u\hat{A}(1)\right\}.$$

The motivation for this estimate is given in the Appendix. Figure 5 shows the resulting parametric and non-parametric estimates for all combinations of masts. When inspecting the plots we recognize the placement of the masts in Pickands dependence function. As seen from the plots in Figure 1 the dependence decreases with the distance between masts. An analysis based on various parametric models leads to the same conclusion. A useful summary measure for extremal dependence is the so-called tail dependence coefficient, which goes back to Geffroy (1959) and Sibuya (1960) and a detailed description can be found for instance in Falk et al. (2000). For a bivariate random vector (X, Y) with marginal distributions F_X and F_Y the tail dependence coefficient is defined by

$$\chi = \lim_{u \to 1} P(F_X(X) > u \mid F_Y(Y) > u),$$

provided that the limit exists. This value corresponds to the probability of one variable being extreme given that the other is extreme. The values $\chi = 0$ or $\chi = 1$ correspond to the two extreme cases of asymptotic independence and complete dependence, respectively. It can be shown that χ is related to A by

$$\chi = 2\left(1 - A\left(\frac{1}{2}\right)\right), \tag{7}$$

so that we can estimate the tail dependence coefficient from the estimate of the Pickands dependence function. Table 4 shows the resulting tail dependence coefficient based on the parametric and non-parametric estimates of the Pickands function. Again, we see how the tail dependence coefficients decrease with the distance of the masts (see Figure 1). In addition, the values on day 2 are lower than on day 1 due to the lower wind speed on the second day.

To discriminate between the parametric models we use likelihood ratio tests as suggested by Tawn (1988), which can be done since the two models are nested. The null-hypothesis, corresponding to the logistic model with $\theta = \phi = 0$, is rejected, if

$$-2\log\lambda(X) = -2(\log L(0, 0, \hat{r}; X) - \log L(\hat{\theta}, \hat{\phi}, \hat{r}; X)) > \chi^2_{1-\alpha,2},$$

where L denotes the likelihood function evaluated at the MLEs, which are shown for the parametric models in Table 6. Table 5 shows the resulting test statistic and the corresponding p-values of the likelihood ratio test for all combinations of masts. In all cases the p-values are above 0.05 leading to the conclusion that the symmetric logistic model cannot be rejected at the 5% significance level. This agrees with Figure 5. We also tried other parametric models, including for instance the mixed model (Tawn, 1988) and the bilogistic model (Joe et al., 1992), but these did not yield an improvement over the logistic and the non-parametric models.

Table 4. Estimated tail dependence coefficients for the parametric models and the non-parametric estimates.

$\hat{\chi} = 2(1 - \hat{A}(1/2))$	logistic	asymmetric logistic	non-parametric
day 1, masts 1 & 2	0.2083	0.2078	0.2093
day 1, masts 1 & 3	0.1407	0.1408	0.1374
day 1, masts 2 & 3	0.3577	0.3568	0.3612
day 2, masts 1 & 2	0.1986	0.1978	0.1953
day 2, masts 1 & 3	0.0987	0.0973	0.1027
day 2, masts 2 & 3	0.2721	0.2695	0.2719

Table 5. Test statistic and corresponding p-values for the likelihood ratio tests to discriminate between the logistic and the asymmetric logistic model.

	test statistic	p-value	test statistic	p-value
	day 1		day 2	
mast 1 & mast 2	4.281	0.1176	3.065	0.2159
mast 1 & mast 3	4.164	0.1247	1.774	0.4118
mast 2 & mast 3	1.300	0.5220	8.077	0.0189

The estimated parameters are given in Table 6. The parameter r of the logistic model can be interpreted as a dependence parameter, which equals one, if the two components are independent, and zero, if the marginals are completely dependent. In our case, the estimated values for r are below one, but the dependence is not as strong as one might think between masts which only have a distance apart of 20, 25 and 5 meters, respectively. The strongest dependence for extreme values is given, when the wind speed is high and the masts have a distance of 5 meters. The corresponding value on day 2 is lower and the dependence parameter for mast 1 and 3 gets close to one.

3.3. Cross-tail dependence

We now consider extremal dependence between masts at different time lags. In particular, we estimate the tail dependence coefficients for extreme velocities arising from the following time series

$$(X_{t+s,i})_{t\in\mathbb{Z}} \quad \text{and} \quad (X_{t,j})_{t\in\mathbb{Z}}, \quad \text{for } i,j = 1,2,3.$$

Figures 6 and 7 show the resulting estimates of the cross-tail dependence coefficient at lags from -8 to 8 seconds. For the Lammefjord data set, the wind is mainly moving in one direction first through mast 1 and then past masts 2 and 3. The diagonal plots show the cross-tail dependence within each time series and the three plots in the upper triangular part show the estimates for the combination of masts in direction of the wind. The results appear consistent with our prior results and the layout of masts. For instance, the distance between mast 1 and 2 is 20 meters and the median value of the exceedances for day 1 is around 16 m/s, from which we conclude that the highest dependence for the time series $(X_{t,1})_{t\in\mathbb{Z}}$ and $(X_{t,2})_{t\in\mathbb{Z}}$ should be around $20/16 = 1.25$. In particular, the estimated tail dependence coefficient for lags $s = -1$ and $s = -2$ are given by 0.23 and 0.21, respectively and are the largest among all other lags. Similarly the highest tail dependence coefficient for the time series $(X_{t,1})_{t\in\mathbb{Z}}$ and $(X_{t,2})_{t\in\mathbb{Z}}$ on day 2 (with a median value around 9 m/s for the exceedances) should be around $20/9 = 2.22$. When inspecting the plots in the upper

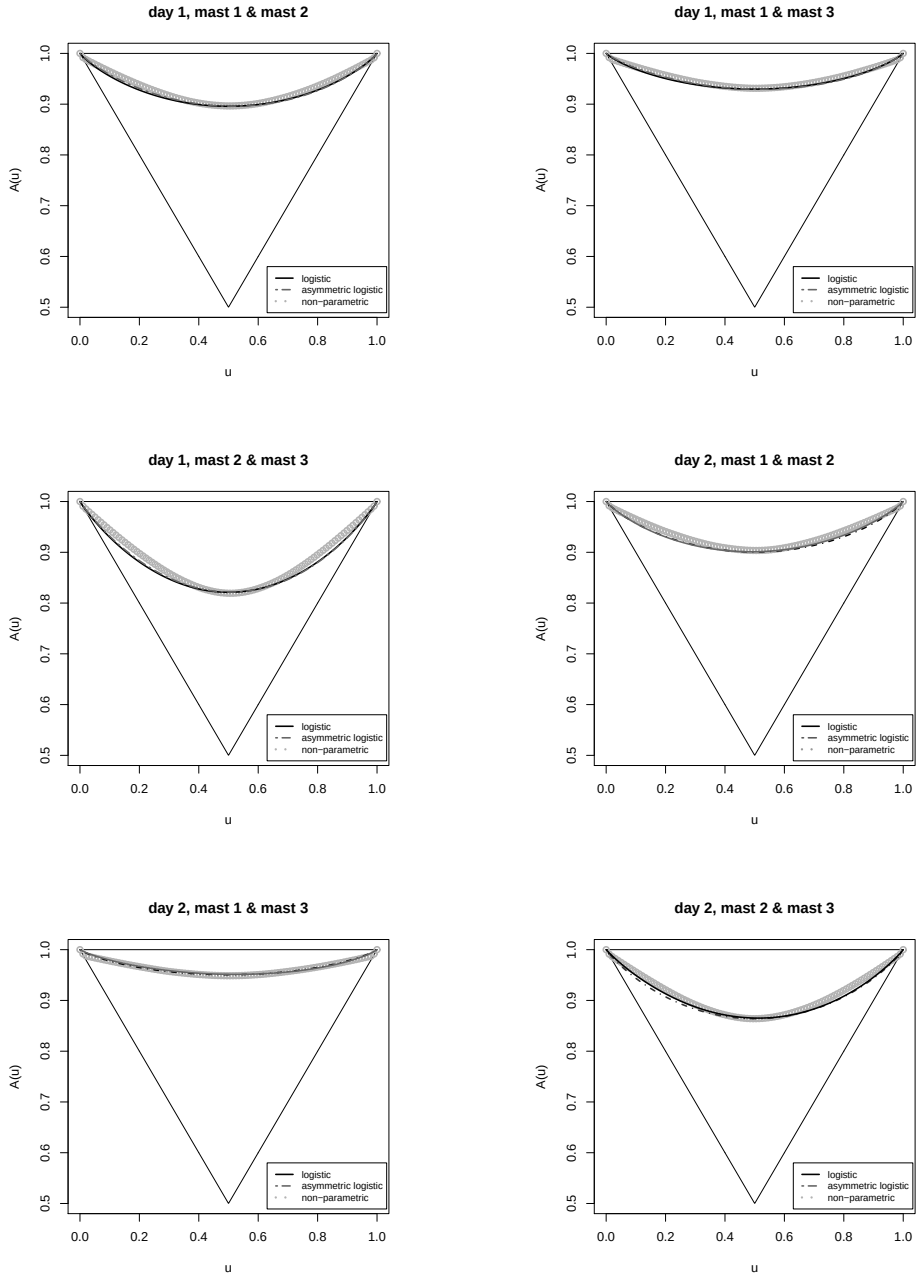

Fig. 5. Parametric and non-parametric estimates for Pickands dependence function for day 1 (left) and day 2 (right). Top: masts 1 & 2. Middle: masts 1 & 3. Bottom: masts 2 & 3.

Table 6. *Estimated parameters for the logistic and the asymmetric logistic model with standard errors in brackets.*

	logistic	asymmetric logistic		
	$\hat{r}$	$\hat{\theta}$	$\hat{\phi}$	$\hat{r}$
day 1, masts 1 & 2	0.8414	0.7914	0.8584	0.7994
	(0.0056)	(0.1102)	(0.1303)	(0.0304)
day 1, masts 1 & 3	0.8948	0.6182	0.7518	0.8423
	(0.0047)	(0.1458)	(0.1752)	(0.0378)
day 1, masts 2 & 3	0.7157	0.9405	0.9981	0.7062
	(0.0069)	(0.0132)	(0.0021)	(0.0242)
day 2, masts 1 & 2	0.8491	0.8451	0.7486	0.8082
	(0.0055)	(0.1460)	(0.1257)	(0.0330)
day 2, masts 1 & 3	0.9269	0.7654	0.9944	0.9170
	(0.0039)	(0.1188)	(0.0189)	(0.0081)
day 2, masts 2 & 3	0.7890	0.6864	0.7621	0.7021
	(0.0062)	(0.0525)	(0.0612)	(0.0223)

triangular part, we clearly see that the estimated tail dependence coefficients are higher for lags smaller than zero. The estimates for day 2 are shifted to the left due to lower velocities, where the wind flow needs more time to reach the next mast.

4. Conclusion

In this article we established models for extreme velocity measurements in the atmospheric boundary layer observed at a station in Denmark. The models present in principle techniques for analyzing large wind speeds on small time scales.

In Section 3 we established a conditional model for exceedances over a time-dependent threshold. The shape parameters in the generalized Pareto distribution were below zero for all time series indicating conditional distributions in the Weibull maximum domain of attraction. This is consistent with other studies about wind speed data and implies that the distributions have a finite boundary point of support. Based on the estimated distributions, we determined one-step ahead conditional quantiles predicting the risk of extreme wind speeds within the next second. The models adjust for intermittency effects present in wind speed by allowing the scale parameter in the GPD to vary over time. In the context of design and performance of wind turbines, our technique can be used to model wind speed exceedances over a variable threshold. The model for the scale parameter accounts for short term fluctuations which can have a significant effect on extreme loading (Nielsen et al., 2003, Introduction). In the second part of Section 3 we built a bivariate model for joint threshold exceedances based on the marginal fitted distributions. The dependence parameters of the logistic model and the tail dependence parameters clearly show that higher wind speeds lead to higher dependence between the extreme measurements of different masts. In addition, we analyzed the cross tail dependence by estimating the tail dependence parameters for the logistic model for temporally shifted observations. This gives a better understanding of the horizontal movement of air mass. According to different wind situations we obtained different time-lags for largest extremal dependence between the investigated observation points.

We also applied the techniques established in Section 3 to the velocity measurements taken at the 10 meters height for comparison. The marginal estimated parameters are

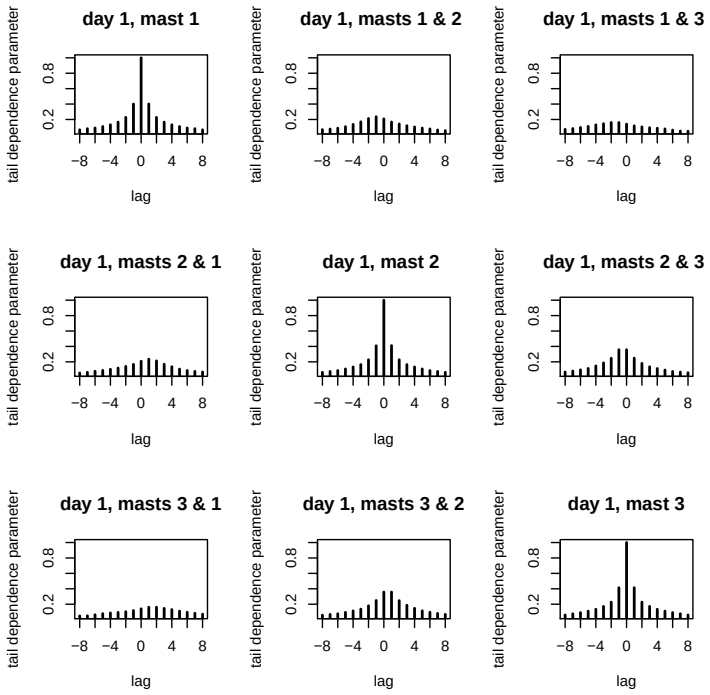

Fig. 6. *Day 1: Cross tail dependence parameters for large wind speeds coming from different masts. Each plot depicts the tail dependence coefficient from lags (in seconds) -8 to 8.*

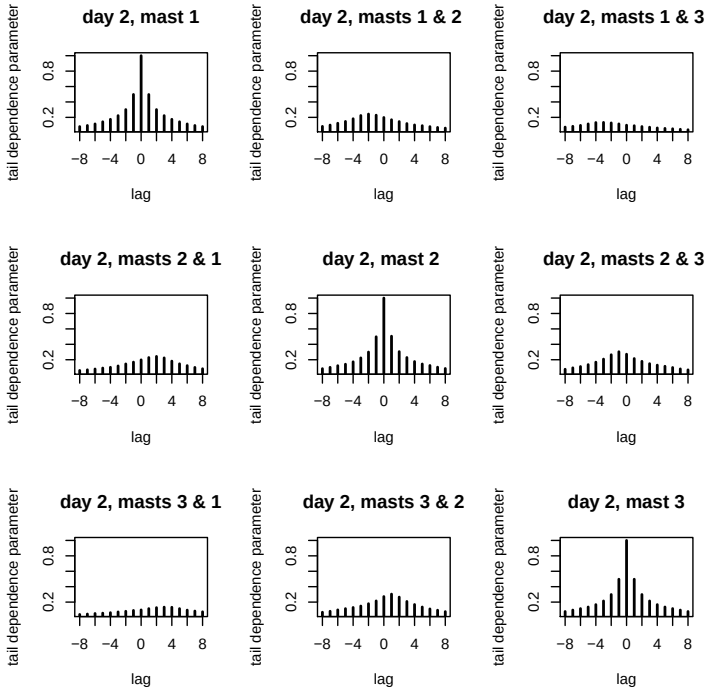

Fig. 7. *Day 2: Cross tail dependence parameters for large wind speeds coming from different masts. Each plot depicts the tail dependence coefficient from lags (in seconds) -8 to 8.*

almost the same as for the measurements from 30 meter above terrain and the dependence parameters are slightly lower for the 10 meters observations.

This technique allows for an advanced analysis of the dependence between two observations points and is based on the raw velocitiy time series rather than averaged wind speeds. Especially, the anaylsis of extremal dependence is of great important, since it allows for conclusions based on extremal wind speeds at one observation point.

Acknowledgment
The authors wish to thank Martin Greiner (Aarhus University) and Gunnar C. Larsen (Risø National Laboratory - Technical University of Denmmark) for sharing their insight into wind engineering within discussions.

The first author gratefully acknowledges the support by the International Graduate School of Science and Engineering (IGSSE) of the Technische Universität München.

References

Barndorff-Nielsen, O. and J. Schmiegel (2007). Time change, volatility, and turbulence. In A. Sarychev, A. Shiryaev, M. Guerra, and M. Grossinho (Eds.), *Mathematical Control Theory and Finance*, pp. 29–53. Springer, Berlin.

Beirlant, J., Y. Goegebeur, J. Segers, and J. Teugels (2004). *Statistics of Extremes, Theory and Applications*. Wiley Series in Probability and Statistics, John Wiley & Sons Ltd, Chichester.

Burton, T., D. Sharpe, N. Jenkins, and E. Bossanyi (2001). *Wind Energy Handbook*. John Wiley & Sons, Ltd, Chichester.

Coles, S. (2001). *An Introduction to Statistical Modeling of Extreme Values*. Springer Series in Statistics, Springer, New York.

Coles, S. and D. Walshaw (1994). Directional modelling of extreme wind speeds. *Journal of the Royal Statistical Society. Series C (Applied Statistics) 43*, 139–157.

Davison, A. and R. Smith (1990). Models for exceedances over high thresholds (with discussion). *Journal of the Royal Statistical Society. Series B 52*, 393–442.

Eastoe, E. and J. Tawn (2009). Modelling non-stationary extremes with applications to surface level ozone. *Journal of the Royal Statistical Society. Series C (Applied Statistics) 58*, 25 – 45.

Embrechts, P., C. Klüppelberg, and T. Mikosch (1997). *Modelling Extremal Events*. Springer, Berlin.

Falk, M., J. Hüsler, and R. Reiss (2000). *Laws of Small Numbers: Extremes and Rare Events* (2nd ed.). Birkhäuser Verlag.

Geffroy, J. (1958, 1959). Contributions à la théorie des valeurs extrème. *Publ. Inst. Stat. Univ. Paris 7, 8*, (7) 36–123; (8) 3–52.

Gumbel, E. (1958). *Statistics of Extremes*. Columbia University Press, New York.

Holmes, J. and W. Moriarty (1999). Application of the generalized Pareto distribution to extreme value analysis in wind engineering. *Journal of Wind Engineering and Industrial Aerodynamics 83*, 1–10.

Joe, H., R. Smith, and I. Weissman (1992). Bivariate threshold methods for extremes. *Journal of the Royal Statistical Society B 54*, 171–183.

Leadbetter, M. (1974). On extreme values in stationary sequences. *Zeitschrift für Wahrscheinlichkeitstheorie und verwandte Gebiete 28*, 289–303.

Leadbetter, M. (1983). Extremes and local dependence in stationary sequences. *Zeitschrift für Wahrscheinlichkeitstheorie und verwandte Gebiete 65*, 291–306.

Nielsen, M., G. Larsen, J. Mann, S. Ott, K. Hansen, and B. Pedersen (2003). Wind simulation for extreme and fatigue loads. Technical report, Risø National Laboratory.

Pickands, J. (1975). Statistical inference using extreme order statistics. *Annals of Statistics 3*, 119–131.

Resnick, S. (1987). *Extreme values, Regular variation, and Point processes.* Springer, New York.

Ronold, K. and G. Larsen (1999). Variability of extreme flap loads during turbine operation. In E. Peterson (Ed.), *Wind Energy for the Next Millenium: Proceedings of the European Wind Energy Conference*, pp. 224–228. James & James (Science Publishers) Ltd, London.

Sibuya, M. (1960). Bivariate extreme statistics. *Annals of the Institute of Statistical Mathematics 11*, 195–210.

Simiu, E. and N. Heckert (1996). Extreme wind distribution tails : A "peaks over threshold" approach. *Journal of Structural Engineering 122*, 539–547.

Smith, R. (1987). Estimating tails of probability distributions. *The Annals of Statistics 15*, 1174–1207.

Tawn, J. (1988). Bivariate extreme value theory: models and estimation. *Biometrika 75*, 397–415.

Walshaw, D. and C. Anderson (2000). A model for extreme wind gusts. *Journal of the Royal Statistical Society. Series C (Applied Statistics) 49*, 499–508.

World Wind Energy Association (2009). World wind energy report 2008. URL: `www.wwindea.org`.

A. Bivariate extreme value theory - further details

In the following, we give a short introduction to bivariate extreme value theory and the modelling of joint threshold exceedances. As already described, a bivariate extreme value distribution has the general representation:

$$G(x,y) = \exp\left\{(v_1+v_2)A\left(\frac{v_2}{v_1+v_2}\right)\right\}, \quad x,y \in \mathbb{R},$$

where

$$v_1 = \log(G_X(x)), \quad v_2 = \log(G_Y(y)),$$

and G_X and G_Y are the marginal distributions of G. The main objective is to estimate the Pickands dependence function A which can be represented in terms of a spectral measure H,

$$A(u) = \int_{\mathfrak{B}} \max\{u\omega_1, (1-u)\omega_2\}\, dH(\omega_1, \omega_2), \ u \in [0,1], \tag{8}$$

where $\mathfrak{B} = \{(\omega_1, \omega_2)' \in [\mathbf{0}, \infty) \setminus \{\mathbf{0}\} : \omega_1 + \omega_2 = 1\}$, see for instance Beirlant et al. (2004, Chapter 8).

The non-parametric estimate for A used in our models is mainly based on this representation and the fact that the spectral measure is uniquely determined by a so-called exponent measure $\nu(\cdot)$

$$H(\cdot) = \nu\left(\left\{(x,y)' \in [\mathbf{0}, \boldsymbol{\infty}) \setminus \mathbf{0} : x + y \geq 1, \frac{(x,y)'}{x+y} \in \cdot\right\}\right),$$

for which the following convergence result holds.

$$tP(t^{-1}(X_*, Y_*) \in \cdot) \xrightarrow{v} \nu(\cdot), \quad \text{as } t \to \infty,$$

where v denotes vague convergence. We use the same number of exceedances in each marginal distribution and set $k = N_{u_X} = N_{u_Y}$. The observations $X_1, \ldots, X_n$ and $Y_1, \ldots, Y_n$ from two different masts with a certain distance are transformed using the marginal fitted distributions, so that the corresponding random variables are all standard Pareto distributed.

$$X_{*t} = \begin{cases} \frac{1}{1-\hat{F}_{X_t}(X_t)}, & X_t > u_{X,t}, \\ 1, & X_t \leq u_{X,t}, \end{cases} \qquad Y_{*t} = \begin{cases} \frac{1}{1-\hat{F}_{Y_t}(Y_t)}, & Y_t > u_{Y,t}, \\ 1, & Y_t \leq u_{Y,t}. \end{cases}$$

To obtain an estimate for the spectral measure H we build pseudo-polar coordinates

$$\hat{R}_t = X_{*t} + Y_{*t} \quad \text{and} \quad \hat{\omega}_{X,t} = \frac{X_{*t}}{\hat{R}_t}, \ \hat{\omega}_{Y,t} = \frac{Y_{*t}}{\hat{R}_t}, \ \ t = 1, \ldots, n.$$

The non-parametric estimate for the spectral measure H is given by

$$\hat{H}(\cdot) = \frac{2}{k} \sum_{t=1}^{n} \mathbf{2}_{\left\{\hat{R}_t > \hat{R}_{(n-k)}, (\hat{\omega}_{X,t}, \hat{\omega}_{Y,t})' \in \cdot\right\}},$$

where $\hat{R}_{(1)} < \cdots < \hat{R}_{(n)}$ denote the order statistics of $(\hat{R}_t)$. An estimate for Pickands dependence function can then be obtained from the representation (8)

$$\hat{A}(u) = \frac{2}{k} \sum_{t=1}^{n} \mathbf{1}_{\left\{\hat{R}_t > \hat{R}_{(n-k)}\right\}} \max\left\{u\hat{\omega}_{X,t}, (1-u)\hat{\omega}_{Y,t}\right\}.$$

To obtain an estimate which satisfies the characteristics of Pickands dependence function, namely that $\tilde{A}$ is convex and that $\max\{u, 1-u\} \leq \tilde{A}(u) \leq 1$, for all $u \in [0,1]$, we use the modification as proposed in Beirlant et al. (2004), given by

$$\tilde{A}(u) = \max\left\{u, 1-u, \hat{A}(u) + 1 - (1-u)\hat{A}(0) + u\hat{A}(1)\right\}.$$

The non-parametric estimate of Pickands dependence function can be used as a basic guideline for the choice of a parametric model.

In addition to the non-parametric estimate, we give a short theoretical overview of modelling joint threshold exceedances in the parametric approach. Let $(X_1, Y_1), \ldots, (X_n, Y_n)$ denote realizations of the bivariate distribution function F and let u_X and u_Y be specified high thresholds. For simplification, we skip the time-index t in the following. The distribution functions $F_X(\cdot)$ and $F_Y(\cdot)$ are approximated by

$$\tilde{F}_X(x) = 1 - \frac{N_{u_X}}{n}\left\{1 + \xi_X \frac{x - u_X}{\sigma_X}\right\}^{-1/\xi_X}, \quad x > u_X,$$

$$\tilde{F}_Y(y) = 1 - \frac{N_{u_Y}}{n}\left\{1 + \xi_Y \frac{y - u_Y}{\sigma_Y}\right\}^{-1/\xi_Y}, \quad y > u_Y,$$

as described in Subsection 3.1. Next, we transform the marginal distributions so that they are all extremal Weibull distributed according to the probability integral transform.

$$v_1 = \log\left(\tilde{F}_X(x)\right) \quad \text{and} \quad v_2 = \log\left(\tilde{F}_Y(y)\right), \quad x > u_X,\ y > u_Y \tag{9}$$

induce a random vector (V_1, V_2) with distribution function F_V, which has margins that are approximately standard Weibull distributed for $x > u_X$ and $y > u_Y$. It then follows that

$$F_V(v_1, v_2) = (F_V^n(v_1, v_2))^{1/n} \approx G^{1/n}\left(\frac{v_1 - b_{1,n}}{a_{1,n}}, \frac{v_2 - b_{2,n}}{a_{2,n}}\right),$$

where G is an extreme value distribution, $a_{1,n},\ a_{2,n} > 0$ and $(b_{1,n}),\ (b_{2,n})$ are sequences of constants. The approximating sign comes from the general definition of a bivariate extreme value distribution and indicates that we approximate F_V^n for large n by the extreme value distribution G. The max-stability property of extreme value distributions (Resnick, 1987, Chapter 5) provides sequences of constants $\alpha_{1,n} > 0,\ \alpha_{2,n} > 0$ and $\beta_{1,n} \in \mathbb{R},\ \beta_{2,n} \in \mathbb{R}$ such that

$$G^{1/n}(v_1, v_2) = G\left(\alpha_{1,n}\frac{v_1 - b_{1,n}}{a_{1,n}} + \beta_{1,n}, \alpha_{2,n}\frac{v_2 - b_{2,n}}{a_{2,n}} + \beta_{2,n}\right) =: \tilde{G}(v_1, v_2).$$

G and $\tilde{G}$ only differ in scale and location, but not in the shape parameter or in the dependence structure. Therefore, $\tilde{G}$ is an extreme value distribution and, since $F_V(v_1, v_2) = F(x, y)$, it follows that F can be approximated by

$$\tilde{F}(x, y) = \exp\left\{(v_1 + v_2)A\left(\frac{v_2}{v_1 + v_2}\right)\right\}, \quad x > u_X,\ y > u_Y \tag{10}$$

and $v_1,\ v_2$ as in (9). The marginal parameters (ξ_X, σ_X), (ξ_Y, σ_Y) and the dependence parameters (according to the parametric model) can be estimated by using a censored likelihood approach as described in Coles (2001, Section 8.3.1).