Pareto distribution for extreme loads on wind turbines

M. Hänler, U. Ritschel[1], and I. Warnke

Windrad Engineering GmbH, Querstr. 7, 18230 Zweedorf, Germany

Abstract: Extreme loads from 50-years wind and gusts play an important role in the design of wind turbines. Normally these loads are derived from simulated time series by identifying the absolute extreme values, leaving a high amount of uncertainty and chance in these design parameters.

Here we describe a method to determine statistically well founded extreme loads on the basis of the extreme value theory. Analyzing tower loads, we find that the extreme values of a given load component are described by a Pareto distribution, and non-correlated simultaneous components are normally distributed. Utilizing the peak-over-threshold method we obtain the optimal parameters of the Pareto distribution with the help of the maximum-likelihood estimation.

1. Introduction

In a recent publication [1] it was pointed out that turbulence is "a challenging problem for wind energy". This statement is true in two respects. Firstly, turbulence itself is still subject of research activities and its theoretical understanding is not yet complete. Secondly, for the turbine consists of flexible parts that tend to vibrate. As a result, there is a complex interaction between wind turbulence and turbine structure studied under the heading of aeroelasticity.

Normally structural loads caused by turbulence have been analysed with regard to fatigue life. Ultimate loads occurring in the structure have been typically investigated with deterministic wind, where the focus was put on functional effects caused by transient processes (like for example emergency braking) and on the effects of deterministic gusts. The transfer from the stochastic external forces to structural vibrations and loads is completely neglected in this approach. It is obvious, however, that extreme loads may also occur in load cases with turbulent wind, for example during normal power production at high wind speeds or during storms when the turbine is normally shut down. This was finally also taken into account by the standards. In the new edition of the IEC [2] the appendix "statistical extrapolation of loads" is added that deals with this subject.

An example for structural loads caused by the turbulent wind is shown in Fig.1. For the sake of simplicity, only the wind speed at hub height and the tilt moment at the tower base are displayed. The average wind speed in this example is about 38 m/s with a turbulence intensity of 11 %. The moment is shown in units of the maximum average tilt moment in the power production regime. The data are generated by a simulation code called SIWEC which was developed recently by the authors. For more information on SIWEC see Poster BL3.137.

It is straightforward to determine the extreme value in Fig. 1. The result is about 3.2, a value that might serve as the first candidate for the ultimate load at the tower base. Evaluating the loads with other wind files, however, leads to different answers for the ultimate load. In Fig. 2 the results from 80 independent time series are shown. The maximum is about 3.6 whereas the minimum is 2.6. Thus, there is a factor of almost 1.4 between lowest and highest extreme value, a span larger than would be taken into account by the safety factor of 1.35 commonly applied. Realizing that the safety factor to cover not only statistical uncertainties but (probably in the first place) model uncertainties as well, it is obvious that naïve

[1] E-mail: uritschel@windrad-engineering.de

determination of extreme values from a limited sample of time series does not lead to reliable results for ultimate loads.

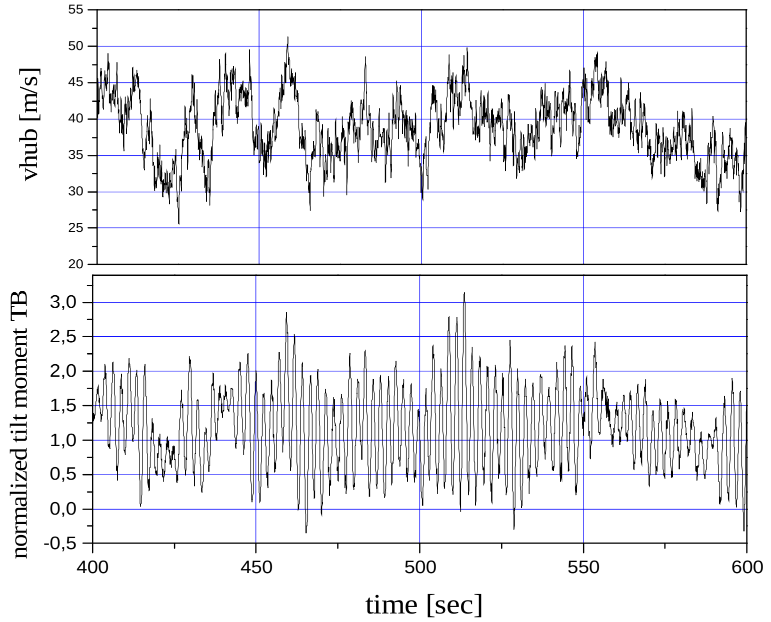

Fig. 1: Time series for wind speed at hub height and (rescaled) tilt moment at the tower base

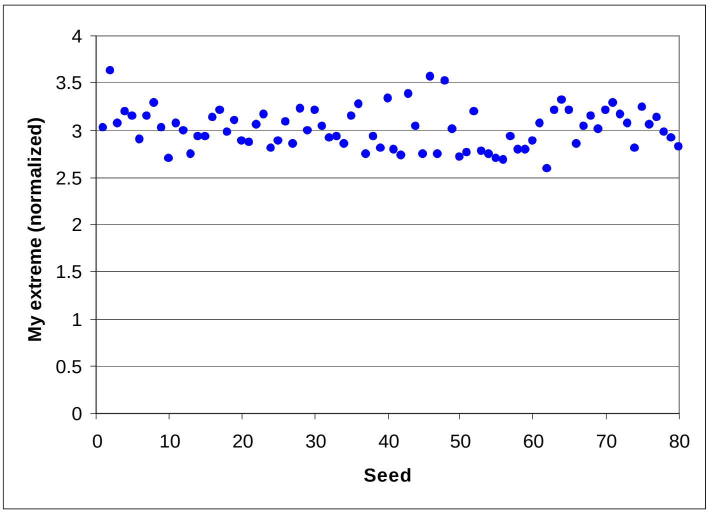

Fig.2: Extreme tilt moment at tower base from 80 time series.

The answer to the question, how to determine ultimate load levels with well-defined probabilistic properties is given by extreme value theory (EVT). It has been realized by Gumbel [3] that for a large class of stochastic processes and corresponding probability distributions, the statistics of the extreme events takes a relatively simple form. In other words, one does not have to know the precise rules for the original process, it suffices to count large events and fit a simple curve to the obtained frequencies. With that curve one may assign a certain probability statement to a certain load, as for example that the level is exceeded on average once within a year or once within fifty years.

EVT is employed in many other fields as well like for example quantitative financial analysis, and it was developed in these contexts during the last decades. Especially the peak-over-threshold (POT) method, a very powerful tool to analyze the extreme events, has been described in Ref. [4]. EVT has previously been used to analyze loads on wind turbines in [5-8].

The rest of this paper is organized as follows. Firstly, a brief introduction to the theoretical background is given. For all mathematical details we refer to the literature [3,4]. Next the procedure is described how to practically apply EVT to ultimate loads on wind turbines. Finally as an example the tower loads in 50 years wind are analyzed with the method.

2. Theoretical background – extreme value theory for the practitioner

The new edition of the IEC standard states "it is necessary to analyse the extreme values of the loading on a statistical basis in order to determine a suitable characteristic load" [2]. For that purpose let us consider a typical load for which we calculate or measure a time series *M(t)* like the one from Fig. 1. The local peak values of *M(t)*, let us call them M_i, can be viewed as the events of a stochastic process.

Of interest with regards to extreme loads is the number of events M_i above a certain threshold M_0 within a certain reference time interval *T*. The basic probabilistic formula for this quantity is given by

$$\mathrm{Prob}\left(M_i > M_0 \,\middle|\, T\right) \equiv \mathrm{P}_e\left(M_0, T\right) = \int_V dV\, \mathrm{Prob}\left(M_i > M_0 \,\middle|\, T, V\right) \rho\left(V\right). \tag{1}$$

It contains the wind-speed distribution $\rho(V)$. The integration runs over all wind speeds that potentially contribute to the ultimate load. The probability $\mathrm{Prob}\left(M_i > M_0 \,\middle|\, T, V\right)$ is given by

$$\mathrm{Prob}\left(M_i > M_0 \,\middle|\, T, V\right) = 1 - \left[F_{\max}\left(M_0 \,\middle|\, V\right)\right]^{n(V,T)} \tag{2}$$

where $F_{\max}\left(M \,\middle|\, V\right)$ is the (integral) probability function of local peaks and $n(V,T)$ is the number of peaks in *T*. On the right hand side of (2), the time dependence comes in only via $n(V,T)$. The probability F_{max} for the peak values itself is time-independent.

A concrete example for the design value would be the level that is exceeded on average once every 50 years. In that case the level $M_0=M_{50}$ is determined from the implicit equation

$$\mathrm{P}_e\left(M_{50}, T\right) = \frac{T}{50\,years} = 3.8 \times 10^{-7} \quad \text{for } T = 10 \text{ min} \tag{3}$$

The problem with the above equations is that the probability function F_{max} is not known. Thus even if the number of peak values is counted carefully, it is hardly possible to derive a reliable estimate for M_{50}.

As a consequence, the formulae in the IEC standard [2] can not be utilized directly for the actual work. This is where the EVT comes in. For ultimate state analyses only the extreme events are of interest. EVT states that for a large class of underlying probability distributions, the distribution of the extreme values takes a simple form, a function with a small number of parameters that can be easily fitted to the sample data obtained from the analysis of the time series.

A very powerful method based on the above statement is the POT method [4]. In the POT scheme a threshold M_t is set and the conditional probability that $M_i < M_0$ subject to the condition that $M_i > M_t$ is considered. It can be shown that the conditional probability approaches the function

$$P_{\sigma,\xi}(x) = \begin{cases} 1-(1+\dfrac{\xi}{\sigma}x)^{-\frac{1}{\xi}} & \xi \neq 0,\ 1+\dfrac{\xi}{\sigma}x > 0 \\ 1-e^{-\frac{x}{\sigma}} & \xi = 0 \end{cases} \tag{4}$$

called the Pareto distribution, where x $=M_i - M_t$ and ζ and σ are parameters that determine the form of the function and can be obtained by fitting to the data.

The M_i, used for the analysis have to be statistically independent realizations of a stochastic process. In the case of the wind turbine the time dependence of the loads is a result of stochastic forces acting on an oscillating system. As a result, a particular high peak event tends to be surrounded by other high values that are deterministically related to each other. As will be discussed in the next section, the POT count has to be supplemented by a procedure that assures statistical independence of the data.

Fig. 3 shows the graphs of typical examples of the Pareto function. The black line represents the $\zeta=0$ case, where probability one is approached exponentially and which is called a thin-tailed distribution. The lower blue line represents the case $\zeta>0$, with an approach to one slower than exponential. This case is called fat-tailed. Finally in the case $\zeta<0$ the value one is reached at a finite value of x, i.e. there is a finite limit to the quantity considered. This case is called finite-tail distribution.

The virtue of the POT method is that the simple form (4) can be fitted to data obtained by counting the peaks over a certain threshold. Since the functional form is given, standard methods from statistics like the maximum-likelihood method can be used to determine the unknown parameters, and tests of the goodness of fit like the χ^2-test can be employed straightforwardly.

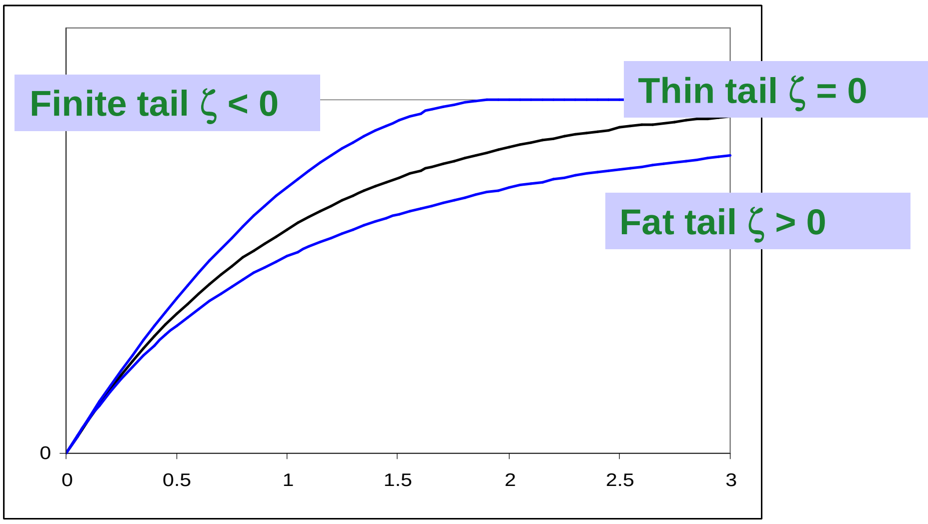

Fig. 3: Examples for the Pareto-distribution

3. Application in load calculations for wind turbines

In this section the application of the POT method to ultimate loads is described. It may be used to determine ultimate levels for loads, i.e. sectional forces and moments, and other quantities as for example the blade-tip-to-tower distance. Further, it may be used to analyze both measured and simulated time series.

Let us come back to the time series of Fig. 1. In Fig. 4 part of this time series is shown once more. To apply the POT method the following steps have to be taken:

- Choose a threshold and carry out the POT count. In Fig. 4 the threshold was chosen at the load level 2 for example.
- Make sure that the peaks used to determine the Pareto parameters are statistically independent. Here one has to find a compromise between long-time spans between the selected peaks and a large enough sample from a limited number of time series. Fig. 4 shows a practical procedure to guarantee statistical independence. After choosing a peak event, other events in the marked neighbourhood, a certain number of tower oscillation periods earlier and later, are excluded from the analysis.
- Fit the Pareto distribution to the selected POT events, for example with the maximum likelihood method.

- Carry out quality checks to assure statistical independence and quality of the fit on the basis of the mean excess function and the χ^2-test.
- Obtain the desired ultimate value, for example the 50 years load, with the equation

$$P(L_{50},T) = \sum_V [1-(P_{\sigma,\varsigma}(L_{50}-L_t))^{n(T,V)}]\cdot \rho(V) = 3.8\cdot 10^{-7} \qquad (5)$$

where $n(T,V)$ is the average number of peaks over the threshold L_t.

- For loads the synchronous components have to be determined as well. It turns out that the synchronous components can be described as a sum of a term that is linearly correlated to the analyzed quantity and a normally distributed contribution. This opens the possibility to determine the synchronous components equipped with a certain probability.

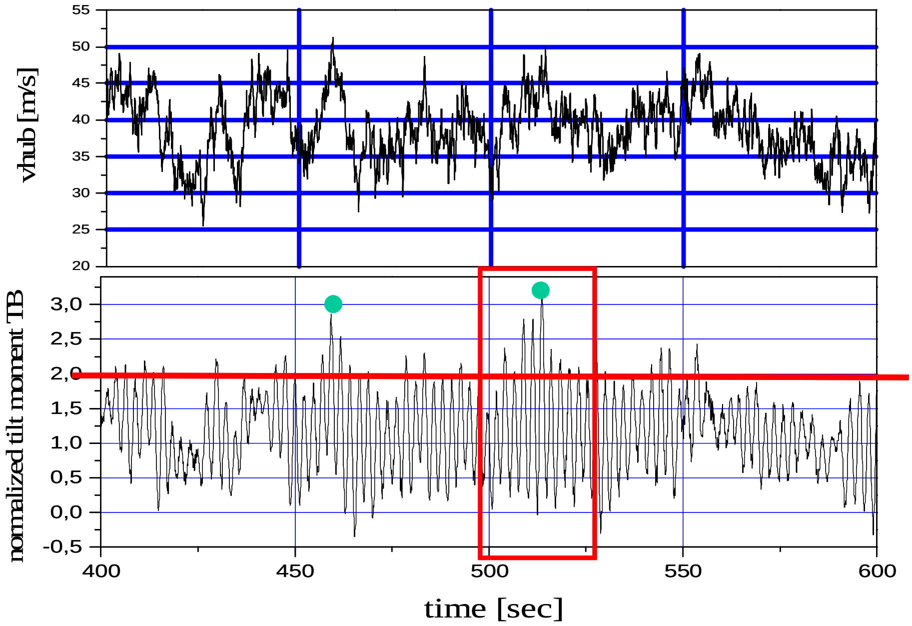

Fig. 4: Procedure to determine statistically independent peak values. Further explanations are given in the text.

The procedure described above determines the ultimate value of a certain load component and the synchronous components on a well defined probabilistic basis. A concrete example for the application is given in the next section.

4. Example: 50-years tower loads

In order to describe the tilt moment and synchronous components, the local coordinate system at the tower base is defined as follows: x-direction along the tower axis downwards, y-direction horizontal and perpendicular to the nacelle axis, and z-direction horizontal along the axis of the nacelle.

To determine the 50-years value for the tilt moment M_y, we have analyzed 80 time series according to the rules described in the previous section. The absolute extreme values of each time series were shown in Fig. 2. The result of the POT analysis is shown in Fig. 5, where the graph represents the probability of exceedance as a function of the load level. In this case the 50 years value of the load is 3.28. The highest observed level lies about 10 % above. The reason for this is simply that with 10 min series we have simulated much more time with extreme wind conditions than would occur within 50 years.

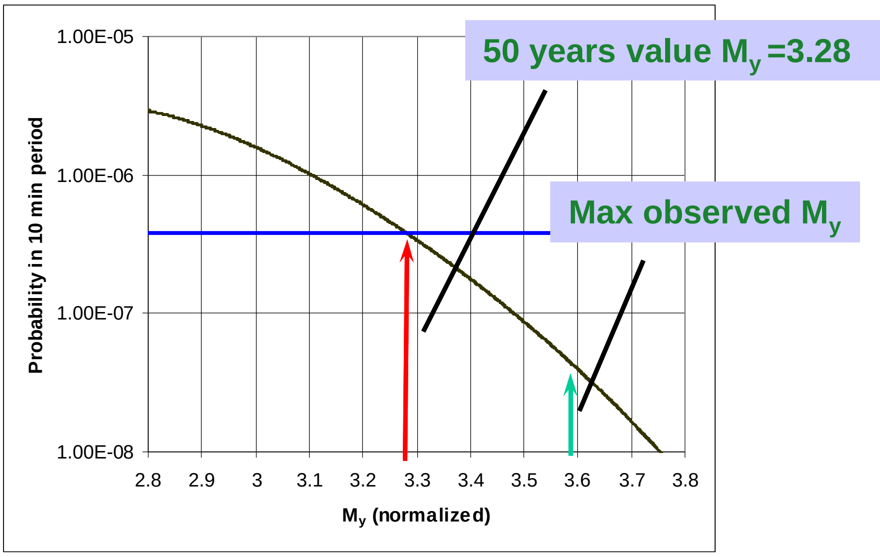

Fig. 5: Results for probability function and 50-years value of M_y

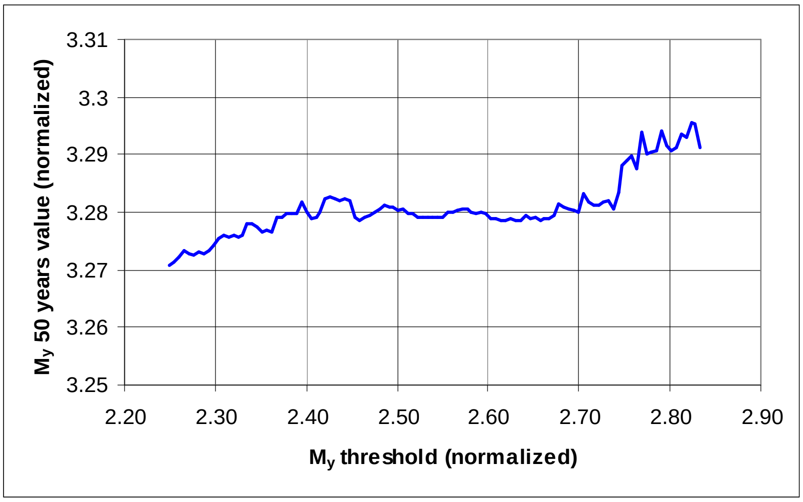

Fig. 6: Dependence on result for M_{50} on threshold

It is instructive to plot the dependence of the result on the threshold. Fig. 6 shows this dependence for thresholds between 2.3 and 2.9. For low thresholds where several hundred events are analyzed there is a dependence of the result on the threshold signalling that for these values the condition of the EVT are not fulfilled, the values analyzed are not extreme enough. On the other hand, for thresholds between 2.4 and 2.7 the result for M_{50} is stable. For thresholds above 2.7, when the number of POT values becomes too small, the results start to scatter.

Other quality checks for the obtained distribution are the χ^2-test and the mean excess function. In the material presented here the condition was imposed that the χ^2 value was within the 95% quantile.

Concerning the synchronous components, Fig. 7 shows the results for correlation coefficient for all load components with M_y. F_x and F_z are correlated with M_y. F_y and M_z are uncorrelated. Subtracting from each component the correlated part, one finds that the remainder can be described well by a normal distribution centred at zero. Fitting the width of the normal distribution one finally can determine synchronous load components with a certain probability.

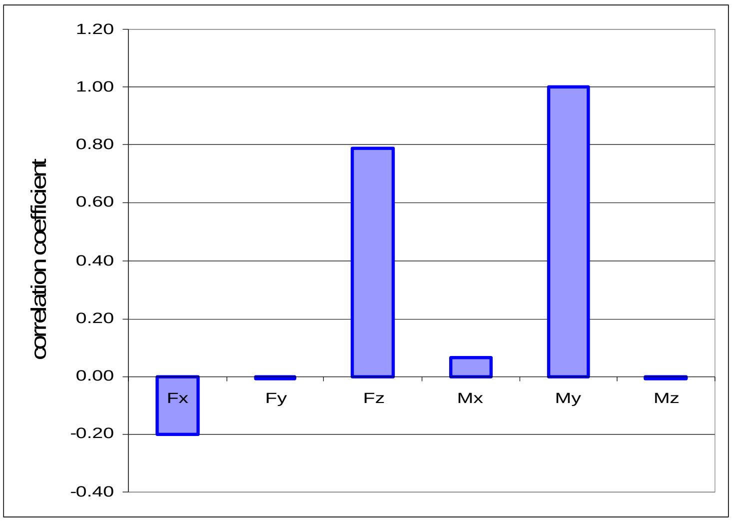

Fig. 7: Correlation coefficients of other load components

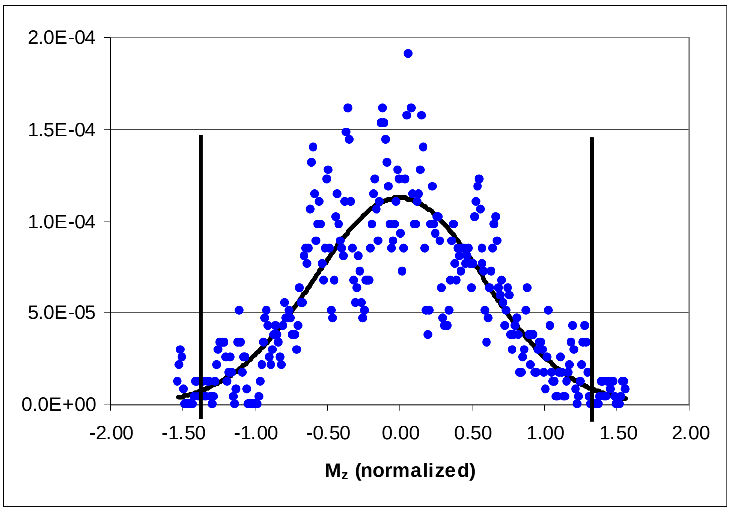

Fig. 8: Distribution of synchronous values of M_z

5. Summary and Outlook

In this paper we have described a procedure to determine extreme values for loads using methods from extreme value theory. Utilizing the peak-over-threshold method, we were able to design a computer code that allows

- to determine a ultimate load level with a certain probability, as for example that the level is exceeded on average once within 50 years,
- to carry out quality checks in order to guarantee that the conditions for the application of extreme value theory are satisfied.
- to determine the values of the synchronous load components.

As an example we analyzed the tower base moment from simulated data in the situation with 50 years wind. This analysis was carried out on the basis of 80 time series. Other analyses showed that about 30-40 time series are sufficient in order to obtain enough statistically independent peak values to carry out the analysis with sufficient quality with respect to the χ^2-test.

The method can be viewed as a post-processing tool for load data comparable for example with the rain-flow-count procedure used for fatigue calculations. It eliminates the statistical uncertainty of the naïve determination of extreme values from a limited sample. As a consequence it justifies a reduction of the safety factor, as suggested also in the IEC standard [2].

The method can be straightforwardly applied to other quantities like for example the tip-to-tower distance. It could also be used for the evaluation of measured data.

References

[1] J. Peinke et al., Physica A 338 (2004) 187-193.

[2] Int. Standard IEC 61400-1, Wind turbines – Part 1: design requirement, 3. ed. 2005.

[3] E. J. Gumbel, Statistics of Extreme, Columbia Univ. Press, New York, 1958.

[4] J. I. Pickands, Statistical inference using extreme value order statistics, Annals of Statistics, (1975):119–131

[5] P. H. Madsen, *Predicting ultimate loads for wind turbine design*, Report AIAA-99-0069.

[6] P. S. Veers and S. Butterfield, *Extreme Load Estimation for Wind Turbines: Issues and Opportunities for Improved Practice*, ReportAIAA-2001-0044.

[7] M.D. Pandey and H.J. Sutherland, *Probabilistic Analysis of LIST data for the estimation of extreme design loads for wind turbine components*, Report AIAA-2003-0866, 2003.

[8] P. J. Moriarty et al., *Effects of turbulence variation on extreme loads prediction for wind turbines*, Report AAIA-2002-0050