WIND ENERGY
Wind Energ. 2008; **11**:613–635
Published online in Wiley Interscience
(www.interscience.wiley.com) DOI: 10.1002/we.303

Research Article

Towards an Improved Understanding of Statistical Extrapolation for Wind Turbine Extreme Loads

Jeffrey Fogle, Puneet Agarwal and Lance Manuel*, Department of Civil, Architectural and Environmental Engineering, University of Texas, Austin TX 78712, TX, USA

Key words:
statistical loads extrapolation;
wind turbine loads;
simulation;
global maxima;
block maxima

One of the load cases that must be evaluated per the International Electrotechnical Commission standard for wind turbine design requires that characteristic loads associated with a 50-year return period be established. This is usually done by carrying out aeroelastic response simulations of the turbine. In order to estimate such rare loads, extreme loads data of adequate quantity and quality are required to facilitate robust predictions. Practitioners have expressed concerns about aspects of the load extrapolation—for instance, questions have arisen related to the minimum number of required ten-minute turbine response simulations, about whether only a single (global) maximum load from each simulation should be saved or whether, alternatively, several time-separated (block) maxima are preferred. Also, though turbine load types are not influenced by each wind speed between cut-in and cut-out to the same degree, focused simulation effort on winds that control the largest loads for each load type is not addressed. Using global and block maxima for four load measures from aeroelastic simulations on a 5 MW turbine model, we study short-term load distributions as a function of wind speed. Block maxima for different block sizes (time separations) are tested for independence and empirical load distributions for global and block maxima are compared. We present a proposal for addressing load extrapolation that focuses on efficiency, that spells out how to employ either global or block load maxima, and that provides convergence criteria for deciding on an adequate number of simulations that must be performed before attempting long-term load prediction using extrapolation.

Received 29 April 2008; Revised 27 September 2008; Accepted 30 September 2008

Introduction

Statistical extrapolation of wind turbine loads from limited simulations is required in order to predict rare long-term loads associated with an important design load case (DLC) specified in the International Electrotechnical Commission (IEC) standard for the design of wind turbines (IEC 61400-1, Edition 3, 2005).[1] A response simulation represents the stochastic response of a wind turbine to specified random

*Correspondence to: L. Manuel, Department of Civil, Architectural and Environmental Engineering, University of Texas, Austin TX 78712, TX, USA
E-mail: lmanuel@mail.utexas.edu
Contract/grant sponsor: Sandia National Laboratories (contract no. 743378). Contract/grant sponsor: National Science Foundation (grant nos. CMMI-0449128 and CMMI-0727989).

environmental conditions. Each DLC specifies the environmental conditions to be used for the aeroelastic simulations. One particular load case, DLC 1.1, will be the focus of our discussions; therefore, a brief background about it is appropriate here. DLC 1.1 deals with extreme loads that a wind turbine might experience during normal operations—when wind speeds range between cut-in and cut-out. DLC 1.1 also specifies that inflow conditions used in the simulations should represent those associated with near-neutral atmospheric conditions. A normal turbulence model (NTM) is prescribed in the IEC standard that must be used to represent the inflow turbulence fields pertinent to DLC 1.1. In this load case, the hub-height wind speed, V, averaged over 10 min may be treated as a single random variable representing the environment; in the IEC standard, turbulence intensity needed for the NTM is specified in terms of this wind speed, depending on the class for which the turbine design is considered. DLC 1.1 requires that aeroelastic simulations be conducted over the entire power-producing wind speed range from cut-in to cut-out. It is convenient, as the IEC standard permits, to carry out simulations over discrete wind speed intervals or bins; typically, bins of 2-m/s size for V are employed.

As DLC 1.1 relates to an ultimate limit state for design, it requires that a 'characteristic' load be established that has a low probability of occurrence. The standard states that this characteristic load must have a return period of 50 years or, equivalently, that this load may be exceeded on average only once every 50 years.[1] We will refer to this characteristic load as the 50-year load in the following. It is clear that prediction of the 50-year load needs to recognize the various wind speeds that will be encountered and their relative likelihoods. If load statistics or distributions are established separately for each wind speed bin, it is important that a sufficient number of simulations are carried out for each bin, and that aggregation and proper weighting of loads from each bin are also done correctly. Clearly, it is computationally infeasible to carry out the large number of 10-min simulations that would be needed to accumulate loads data, which account for the actual duration that would match the target return period. Instead, a limited number of simulations are generally carried out; effort, though, must be judiciously expended in running simulations most carefully for wind speed ranges that bring about the largest loads as well as the most variable ones. Careful statistical extrapolation from such limited simulations can then make it possible to derive the required 50-year loads.

In this study, we address some concerns that have been raised with regard to experiences practitioners have had with attempts to address DLC 1.1 in the IEC standard.[1] The standard requires that the 50-year load be established but it does not unambiguously provide a procedure that will lead to this load from simulations. The guidelines that are provided are vague at best, for example, when addressing the issue of what represents a sufficient number of simulations to run. There are also no clear indications of what constitutes a check that the 50-year load when derived is a robust or stable estimate. The standard does not explicitly suggest that effort might best be focused on the most important bins (usually at or around rated wind speeds for some load types and at or around cut-out wind speeds for others), although this would be prudent. Finally, the standard does not clearly describe what extreme load statistics may be saved from each 10-min simulation—the use of a single (global) maximum from each simulation needs to be considered against alternatives that utilize several time-separated (block) maxima from each simulation; in the latter, the question of what constitutes a set of independent block maxima is important as it fundamentally affects the derivation of the 50-year load. The standard allows use of multiple maxima by methods such as the peak-over-threshold procedure but details are missing with regard to robust tests for independence. In this study, we address each of these issues. In brief, we address: efficiency when we discuss which wind speed bins are design drivers for each load; convergence criteria that lead to approaches to quantify when an adequate number of simulations have been run that yield stable short-term empirical load distributions; and the issue of independence in block sizes and statistical tests for independence, and also discuss the difference in load predictions based on the use of global and block maxima. Throughout, insights are provided to guide the effort involved in carrying out statistical loads extrapolation as required for DLC 1.1. Although we have highlighted several issues related to loads extrapolation and suggested that these are limitations of the IEC standard, a fairer assessment needs to recognize that industry standards and guidelines are rarely sufficiently detailed so as to provide a recipe-like approach to what is required to derive design loads. Rather than viewing the present study as an attempt to address limitations, it might be better to characterize our work as an attempt to address known issues that practicing engineers are facing with one load case in the IEC standard.

This work represents our continuing effort to improve understanding of statistical loads extrapolation as it applies to wind turbine design. The load simulation data sets used in this study were provided to the authors as well as to other members of a loads extrapolation evaluation exercise (LE^3) working group that was formed at the request of the maintenance committee of the IEC 61400-1 turbine design standard. The simulated loads data sets were generated for a baseline 5MW wind turbine model developed at the National Renewable Energy Laboratory (NREL). All of the findings that are reported here are based exclusively on statistical studies on data from simulations with this turbine model alone.

The LE^3 Data Set

The LE^3 working group was constituted in order to address ongoing issues related to the IEC standard and, particularly, DLC 1.1 that deals with statistical loads extrapolation from limited simulation. The LE^3 working group approved a proposal to develop a database of simulated load time histories and summary statistics from a 5MW turbine model developed by NREL. This model is based on an onshore version of NREL's baseline turbine model developed to represent a utility-scale 5MW offshore wind turbine[2]; this onshore model has identical properties to the offshore turbine above the mudline. The turbine has a hub height of 90 m and a rotor diameter of 126 m. The machine is a variable-speed, collective pitch-controlled turbine with a rated wind speed of 11.5 m/s. The maximum rotor speed is 12.1 rpm. Moriarty[3] provides a detailed account of the turbine model and the various inflow conditions covered by the simulations. Inflow turbulence was simulated using TurbSim v12.0[4]; a Kaimal power spectrum, a shear exponent of 0.20 and a deterministic turbulence standard deviation (given V) were employed based on the NTM. The program, FAST v6.02b,[5] was used to carry out aeroelastic simulations for hub-height wind speeds, V, varying between cut-in and cut-out wind speeds. For the IEC Class I-B site assumed, the 10-min hub-height wind speed follows a standard Rayleigh distribution with mean equal to 10 m/s.

Two data sets were generated for use by the LE^3 working group. The first data set consists of 1200 10-min simulations for each of the 12 wind speeds ranging from 3 to 27 m/s, yielding a total of 14,400 different load time series (when running TurbSim to generate inflow turbulence time histories, the target wind speeds were set at discrete values of 3, 5 m/s, etc., up to 25 m/s, and were assumed to represent 2-m/s bins centered at the target values; realized 10-min average wind speeds varied slightly from the target values). A second data set was generated by representing wind speeds according to a Rayleigh distribution for five full years and then carrying out aeroelastic response simulations for those inflow conditions.

In the present study, four representative and contrasting loads were analyzed from the first LE^3 data set. These loads include the out-of-plane bending moment (OOPBM) at a blade root, the out-of-plane blade tip deflection (OOPTD), the fore-aft tower bending moment (FATBM) at the base and the in-plane blade bending moment (IPBM) at a blade root.

Statistical Load Extrapolation

The first step in statistical load extrapolation involves the identification of load extremes from the turbine simulations. Consider the case where the single largest (global) maximum is extracted from each 10-min time series for a wind speed bin, V_k. The probability, $P(L > l \mid V_k)$, that a given load of interest, L, will exceed any specified load level, l, in 10 min may be estimated by rank-ordering the N_k real-valued global maxima (X_i; $i = 1$ to N_k) that are obtained by running N_k simulations for wind speed bin, V_k. In practice, once the load level, l, is specified, if it lies within the range of the observed loads, one can obtain the empirical short-term

DOI: 10.1002/we

conditional distribution by finding two integers j and $j + 1$ that are such that where $X_j \leq l \leq X_{j+1}$ where $1 \leq j \leq (N_k - 1)$. As $P(L \leq X_j | V_k) = j/(N_k + 1)$ and $P(L \leq X_{j+1} | V_k) = (j + 1)/(N_k + 1)$, one can obtain $P(L > l | V_k) = 1 - n_k(l)/(N_k + 1)$ where $n_k(l)$ is obtained by interpolation. For values of l outside the range of the observed loads, extrapolation may be used. In summary, the empirical short-term load distribution for any bin, V_k, may be estimated as follows:

$$\begin{aligned} P(L > l|V_k) &= 1 - \frac{n_k(l)}{N_k + 1}; \quad \text{where } n_k(l) = j + \frac{l - X_j}{X_{j+1} - X_j}, \quad \text{if } X_j \leq l \leq X_{j+1} \text{ and } 1 \leq j \leq N_k - 1 \\ &= \frac{N_k}{N_k + 1}; \quad \text{if } l < X_1 \\ &= \frac{1}{N_k + 1}; \quad \text{if } l > X_{N_k} \end{aligned} \tag{1}$$

Figure 1 shows example empirical short-term distributions for two loads, OOPBM and OOPTD, estimated using global maxima from 200 simulations in each wind speed bin.

The distribution given by equation (1) is termed a short-term distribution on the global maximum load, L, as it is a distribution conditional on wind speed. All the various wind speeds likely to be encountered need to be considered in order to yield the long-term distribution on L. In terms of a continuous random variable, V, the long-term distribution can be obtained as follows:

$$P(L > l) = \int_{V_{\text{in}}}^{V_{\text{out}}} P(L > l|V = v) f_{\text{V}}(v) dv \tag{2}$$

To evaluate equation (2) in order to obtain long-term distributions for turbine loads, one needs short-term distributions as well as the wind speed probability density function, $f_{\text{V}}(v)$; the latter is taken to be the Rayleigh density function for IEC Class I-B conditions in this study and only the mean value of V of 10 m/s is needed. One can use equation (2) to obtain the long-term load distribution for loads if parametric distribution fits are attempted to each empirical short-term distribution given by equation (1). This might be termed the 'fitting-before-aggregation' approach.

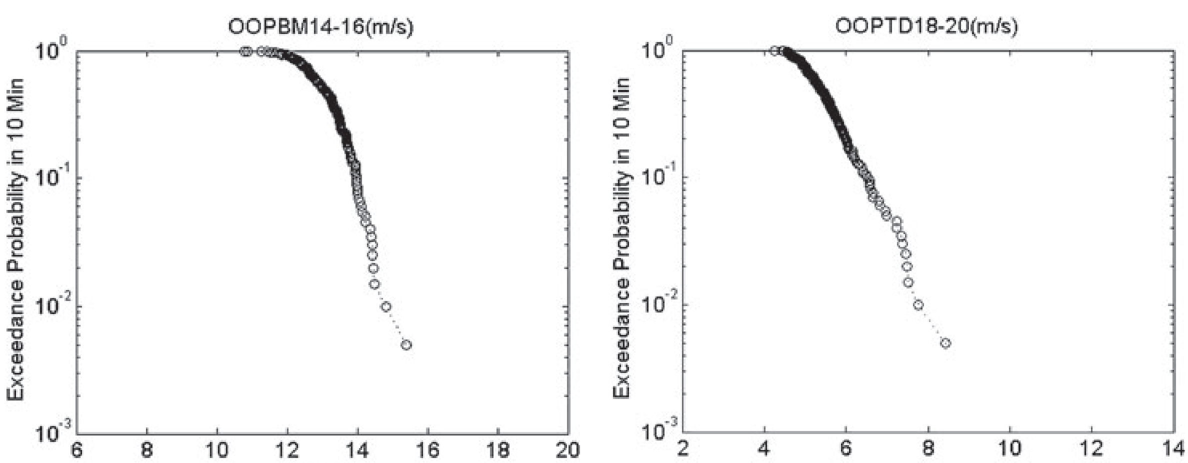


Figure 1. Example empirical short-term empirical distributions for OOPBM and OOPTD estimated using 200 simulations for each wind speed bin

Alternatively, one could obtain the long-term distribution by collecting data from all the wind speed bins together (this can be conceived of as putting all the data into one box). Assuming that the maxima in bin i are rank-ordered such that if there are total of N_i extremes, then $l_{1,1} \le l_{1,2} \le l_{1,3} \ldots \le l_{1,N}$, the notation used implies that $l_{i,k}$ is the k^{th} rank-order maximum from bin i. Note that the total number of load maxima from all bins, N, is equal to $\sum_{i=1}^{N_B} N_i$, where N_B is the number of wind speed bins. If the number of simulations run in each bin is proportional to the actual likelihood of that bin, then the empirical long-term distribution may be estimated simply as follows:

$$
\begin{aligned}
P(L \le l) &= \sum_{i=1}^{N_B} \left(\sum_{k=1}^{N_i} \frac{I[l_{i,k} \le l]}{N+1} \right) \\
\text{where } I[l_{i,k} \le l] &= 1, \quad \text{if} \quad l_{i,k} \le l \\
&= 0, \quad \text{otherwise}
\end{aligned}
\tag{3}
$$

Note that equation (3) provides an empirical expression for the long-term distribution of the global maximum of 10-min maximum load, L. This distribution can be used to derive the 50-year load by noting that this load has a return period of 50 years. If one assumes independence between global 10-min maxima, the desired 50-year load, l_{50}, must be such that the probability of its exceedance in 10 min is 10 min / (50 years $\times$ 365.25 d/year $\times$ 24 h/day $\times$ 60 min/h) = 1/2,629,800 = 3.8×10^{-7}. Clearly, in order to predict loads with such low probabilities of exceedance, statistical extrapolation will be necessary from the limited simulations that will be carried out.

It is worthwhile to note that in the 'aggregation-before-fitting' approach suggested by equation (3), as the data are more heterogeneous as they represent different wind speed bins, parametric fits can focus on the tails of the empirical data. Extrapolation to the desired 50-year return period level can follow directly with either of the two approaches. Ragan and Manuel[6] provided examples of the use of generalized extreme value distribution fits to field data on loads (global maxima) from a utility-scale wind turbine based on the 'fitting-before-aggregation' approach. The small amounts of data usually available in empirical short-term data often lead to fits of poor quality;[6] moreover, such fits are needed for all wind speed bins even where loads are not large. The alternative 'aggregation-before-fitting' approach involves fitting to distributions only on long-term loads; aggregated data are generally larger in number and given that large rare loads are of interest, fitting can be concentrated in the tail that is most useful for extrapolation.

The Relative Importance of Different Wind Speeds to Turbine Load Extremes

To compare the relative importance of different wind speed bins on load extremes, it is of interest to study load extreme statistics as a function of wind speed. With a little effort (i.e. limited simulations), it is often possible to identify which wind speeds can cause the largest turbine loads on average and, equally important, which ones show the greatest load variability. The largest loads are associated with the lowest probability of exceedance and, as such, are closest to the rare probability levels to which extrapolation is needed; the wind speeds where these largest loads occur are therefore of obvious interest. Empirical short-term distributions need to be well estimated in these bins in particular. Even if loads realized in some bins are not among the largest, if variability in extremes from simulations in those bins is large, they can have a significant influence on distribution tails and, hence, on extrapolation.

From Figure 2, it is possible to identify those wind speed bins associated with the largest loads and greatest short-term extreme variability, and also to identify those bins that most influence the tails of the long-term load distributions. This is useful as summary plots such as these make it possible to focus efforts in the most important bins. For OOPBM and FATBM, important controlling wind speed bins are in the 14–22 m/s range with perhaps the dominant winds being closest to the 14–16 m/s bin. The OOPTD is controlled by somewhat lower wind speeds; the largest loads occur between 10 and 16 m/s. The IPBM is clearly dominated by wind speeds close to the cut-out wind speed of 25 m/s. Although these controlling wind speed bins were identified by carrying out 200 simulations per bin, it is possible to identify important bins with considerably less simu-

DOI: 10.1002/we

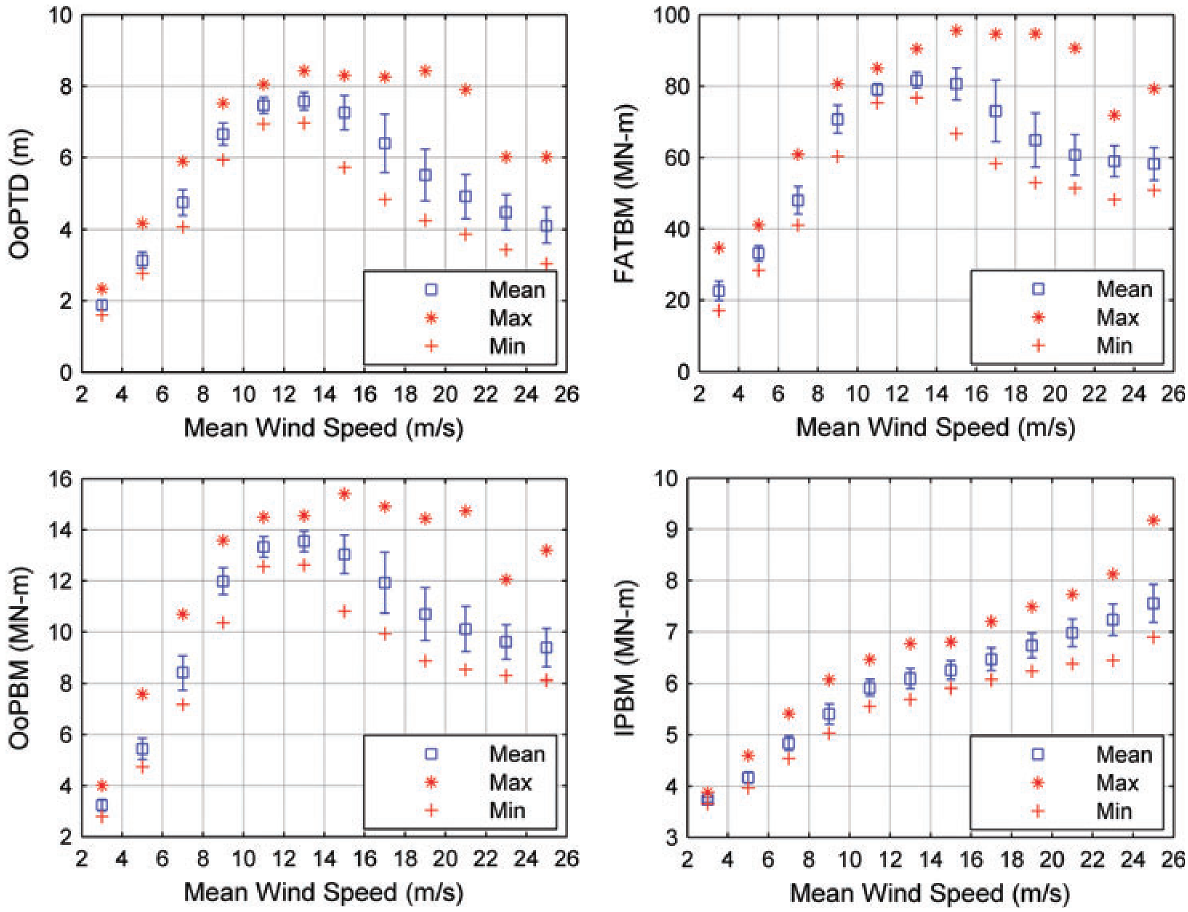


Figure 2. Distribution of short-term load maxima as a function of wind speed for four loads, OOPTD, FATBM, OOPBM and IPBM, based on 200 simulations per wind speed bin

lation effort. From Figure 2, it is clear that wind speeds below 10 m/s do not contribute large loads to any of the four loads types discussed; moreover, they also do not exhibit large variability in load extremes. In carrying out simulations with a view towards extrapolation, it is worthwhile to understand turbine load extremes as a function of wind speed in this manner to avoid excessive computational effort.

Block Maxima and the Issue of Independence

An alternative to the use of only a single maximum (i.e. the global maximum) from each simulated 10-min time series is to extract several extremes from each time series in a systematic manner. Although this can be done by methods such as the peak-over-threshold procedure,[6] there are simpler methods as well. For instance, one could split or partition the time series into individual non-overlapping blocks of constant duration. From each of these blocks, then, a single largest value is extracted; together these extracted extremes constitute a set of block maxima. Figure 3 shows an example OOPBM load time series with 1-min block maxima indicated by circles and the a single 10-min global maximum indicated by an asterisk.

Figure 3 indicates that from a single load time series of 10-min duration, more extremes data may be extracted if block maxima are employed as opposed to global maxima in extrapolation. As long as the n block

DOI: 10.1002/we

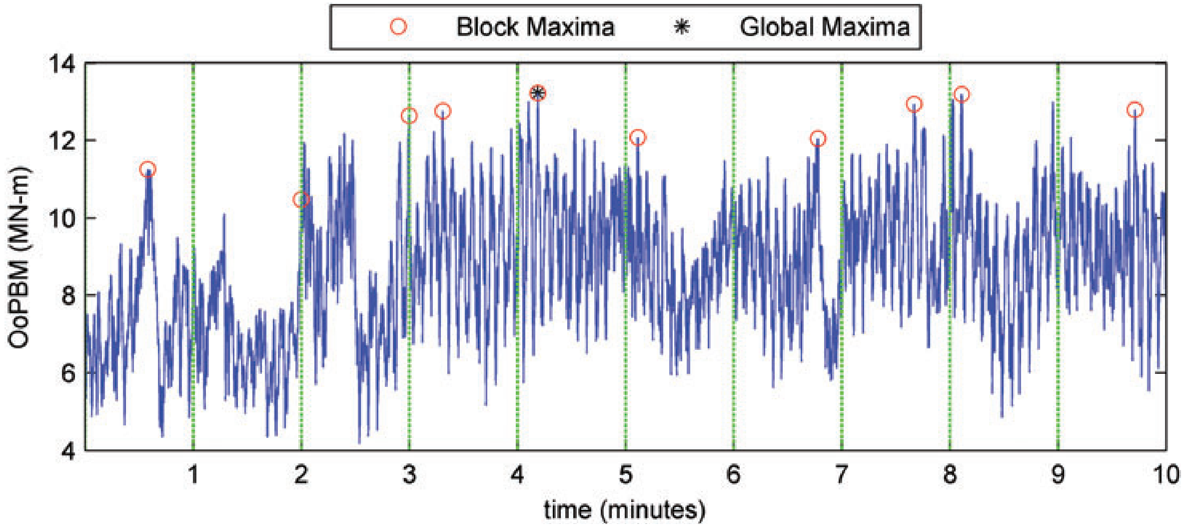


Figure 3. Example OOPBM load time series showing global and block maxima

maxima in a 10-min sample can be shown to be mutually independent, short-term global maxima (L) distributions as in equation (1) can be related to short-term block maxima (L_{block}) distributions as follows:

$$P(L < l|V_k) = [P(L_{block} < l|V_k)]^n \quad \text{or} \quad F_L(l|V_k) = [F_{L_{block}}(l|V_k)]^n \tag{4}$$

where $F_L(\ )$ and $F_{L_{block}}(\ )$ refer to the cumulative distribution functions for L and L_{block}, respectively.

In terms of probabilities of exceedance of any load level, l, one can also write:

$$P(L > l|V_k) = 1 - [1 - P(L_{block} > l|V_k)]^n \tag{5}$$

If one is interested in the p-quantile 10-min maximum load, l_p, defined such that $F_L(l|V_k) = p$, the adjusted load quantile in terms of the block maximum distribution must be adjusted as follows:

$$l_p = F_L^{-1}(p) = F_{L_{block}}^{-1}(p^{1/n}) \tag{6}$$

where n represents the number of blocks contained in 10 min. As might be expected, the non-exceedance probability, p, for global maxima needs to be adjusted to a rarer non-exceedance probability level, $p^{1/n}$, for block maxima, if it is to correspond to the same load. So, although a greater amount of extremes data are extracted when block maxima are employed and hence lower exceedance probability levels can be empirically estimated, the same p-quantile load needs to be sought farther in the tail of the block maxima distribution. For instance, if the 80th percentile 10-min maximum load is required (which corresponds to a non-exceedance probability of 0.80 for global maxima) when 1-min block maxima are used, the corresponding non-exceedance quantile for block maxima is $0.80^{1/10}$ or 0.978 that is considerably farther in the tail of the distribution of the block maxima. Figure 4 illustrates this effect for OOPBM load maxima extracted from six simulations of a single wind speed bin. The asterisks represent block maxima, and the circles represent the global maxima. To highlight the point that the global maxima are also block maxima, the specific global maxima extracted are shown twice to indicate where they appear in the block maxima distribution. Some extracted block maxima are higher than a few global maxima and arguably better defined tail trends are seen in the block maxima distribution. However, as can be seen, the 80th percentile 10-min maximum load corresponds to an exceedance probability level of (1 – 0.978) or 0.022 if the block maxima distribution is used. In this case, the actual 80th percentile load itself is read off at roughly the same level with either choice of distribution.

As was stated before, equations (4)–(6) are valid as long as block maxima selected from each time series are independent of each other. Intuitively, it may be expected that smaller block sizes will lead to greater

DOI: 10.1002/we

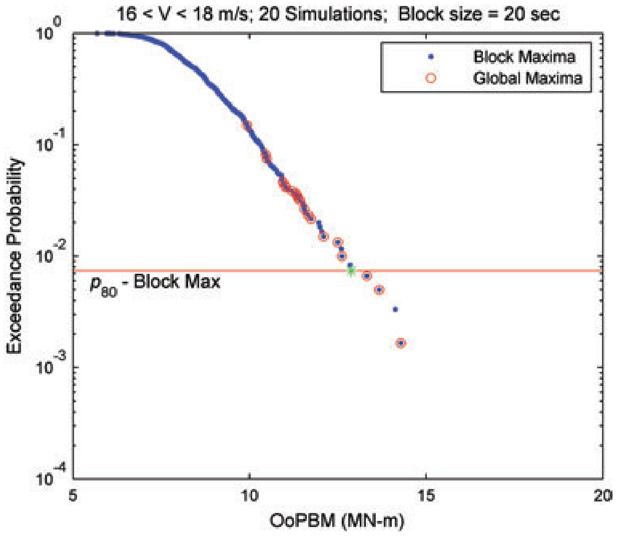


Figure 4. Comparison of block and global maximum probability levels associated with a given load quantile for OOPBM loads (in MN-m)

dependence among the extracted block maxima. Statistical tests for independence represent the only objective means of assessing the extent of independence or lack thereof in a sample of block maxima from load simulations.

Test for Independence

Several tests to evaluate independence between two random variables are available in the literature. We focus here on a test proposed by Blum *et al*.[7] Details related to this test along with examples may be found in Hollander and Wolfe;[8] that reference also provides a correction for a typographical error in an equation in Blum *et al*.[7] Blum's test has been used by Skaug and Tjøstheim[9] to test for independence in time series data, for which it was not originally developed.

Two random variables, X and Y, may be stated to be independent of one another if the product of their marginal probability distribution functions is equal to their joint distribution. According to Blum's test for independence, the null hypothesis, H_0, is that the two variables X and Y are independent. Thus, we have in terms of cumulative distribution functions:

$$H_0 : F_{X,Y}(x, y) = F_X(x)F_Y(y) \tag{7}$$

Blum's test makes use of a test statistic, B, that must be checked against a critical value, B_{cr}, at any specified significance level. This test statistic is computed as follows:

$$B = \frac{1}{2}\pi^4 N \sum_{j=1}^{N} \frac{(N_1(j)N_4(j) - N_2(j)N_3(j))^2}{N^5} \tag{8}$$

where N is the sample size for both X and Y. The quantities, $N_1(j)$ to $N_4(j)$, are computed for all values of j from 1 to N or effectively for all choices of $(X, Y) = (x_j, y_j)$ such that

- $N_1(j)$ is the number of (x, y) pairs such that $x \leq x_j$ and $y \leq y_j$.
- $N_2(j)$ is the number of (x, y) pairs such that $x > x_j$ and $y \leq y_j$.
- $N_3(j)$ is the number of (x, y) pairs such that $x \leq x_j$ and $y > y_j$.
- $N_4(j)$ is the number of (x, y) pairs such that $x > x_j$ and $y > y_j$.

DOI: 10.1002/we

If the value of B as computed by equation (8) is greater than B_{cr}, then the null hypothesis is rejected and the two variables, X and Y, are not independent at the specified significance level.

Useful illustrative examples of the use of this test involve examining this B statistic for paired data that are known to be either strongly dependent or independent. Note that for a bivariate Gaussian distribution, zero correlation implies independence (the distribution is completely defined by a correlation coefficient and the first two marginal moments of each variable). If Blum's test is carried out for two jointly distributed Gaussian random variables that are strongly correlated, the B statistic is likely to be large; the opposite is true if the correlation is weak.

Variables, X and Y, assumed jointly Gaussian with a correlation coefficient of 0.9, were simulated; a scatter plot of the data is shown in Figure 5(a). Note that for any (x_j, y_j) pair that is part of this data set (where by design, X and Y, are strongly correlated and, thus, dependent), the values of N_1 and N_4 are generally much larger than N_2 and N_3; hence, the computed B value is large. For this case, B is equal to 31.8, which is much larger than the critical value, B_{cr}, of 4.23 at a 1% significance level. Hence, the independence (null) hypothesis is rejected. Another extreme case is considered where the variables, X and Y, assumed jointly Gaussian with a correlation coefficient of 0.05, were simulated; a scatter plot of the data is shown in Figure 5(b). Note that for any (x_j, y_j) pair that is part of this data set, this time, the values of N_1 and N_4 are generally of similar magnitude to those of N_2 and N_3; hence, the computed B value is relatively smaller than in the previous case. For this case, B is equal to 2.10, which is smaller than the critical value, B_{cr}, of 4.23 at a 1% significance level. Hence, the null hypothesis of independence is not rejected. These two examples serve to illustrate the use of Blum's test for independence between two random variables in general.

The independence of block maxima of wind turbine loads for different block sizes may be studied in a similar manner to that used in the preceding illustrative example. Blum's test statistic, the B value, for block maxima may be computed by forming lag-one vectors, X and Y, from all the block maxima in each 10-min time series. The B value may be computed for these lag-one extremes to test if they are independent. It is expected that these extremes will become more independent as the block size is increased. At a certain optimum block size, computed B values will fall below the critical value, B_{cr}. Although it is possible to study the B values for each simulation corresponding to a given wind speed, it is more instructive to study these B values (and, thus, independence) statistically as a function of block size by considering multiple simulations for each

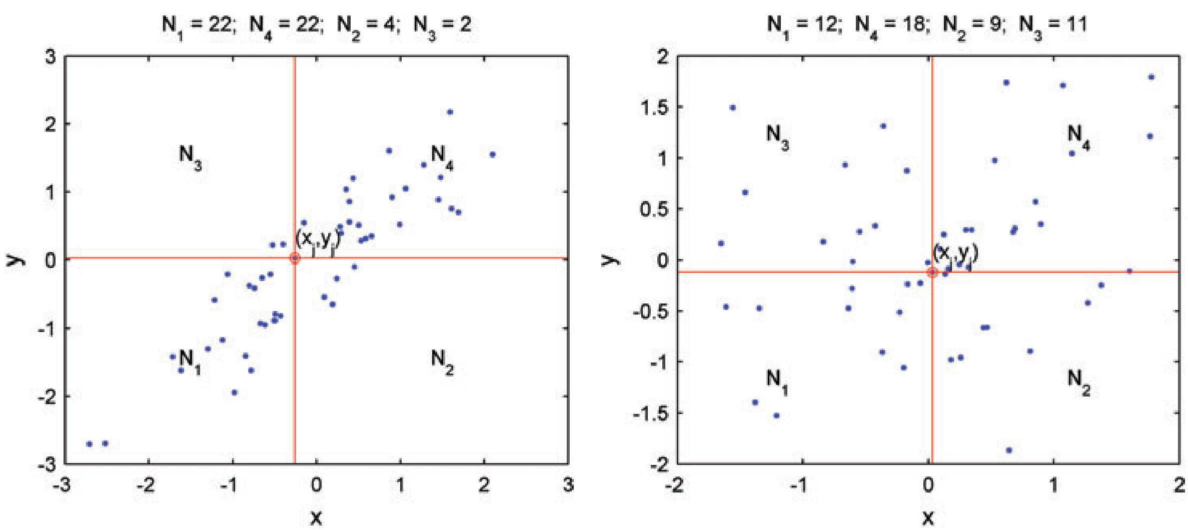


Figure 5. Scatter plot of simulated samples of two bivariate Gaussian random variables with correlation coefficients of (a) 0.9 and (b) 0.05. Also indicated are the values of N_1, N_2, N_3 and N_4 for (x_j, y_j) equal to (–0.260, +0.027) and (+0.034, –0.120), respectively, for correlation coefficients of 0.9 and 0.05, as computed while carrying out Blum's test for independence

DOI: 10.1002/we

wind speed; this makes it possible to account for scatter or uncertainty in the B values over different simulations. To this end, the mean, μ_B, and standard deviation, σ_B, of the B values from 200 simulations are computed for four different load measures: OOPBM, FATBM, OOPTD and IPBM. Note that even with a small number of simulations, on the order of 15 to 20, statistics of the B values are quite stable and 200 simulations are not really needed. The mean B values with error bars representing one standard deviation are shown in Figure 6 for the four load types and for three different wind speed bins: 10–12, 16–18 and 22–24 m/s. As expected, mean values of B decrease monotonically with increasing block size. Even if the more stringent ($\mu_B + \sigma_B$) level is checked against the critical value, B_{cr}, at the 1% significance level, independence of block maxima is virtually assured for block sizes longer than 30 s for all four load types and in all three wind speed bins. Summarized in Table I are the appropriate block sizes for independence based on criteria where either μ_B or ($\mu_B + \sigma_B$) values are compared with B_{cr} at the 1% significance level. Clearly, for a given block size, load maxima in some wind speed bins (e.g. the lower wind speed bins) exhibit greater dependence than in but it appears that—at least for this LE^3 loads data set—one could safely choose block sizes of around 40–60 s, extracting between 10 and 15 extremes (block maxima) from each 10-min time series and use these extremes to establish short-term load distributions.

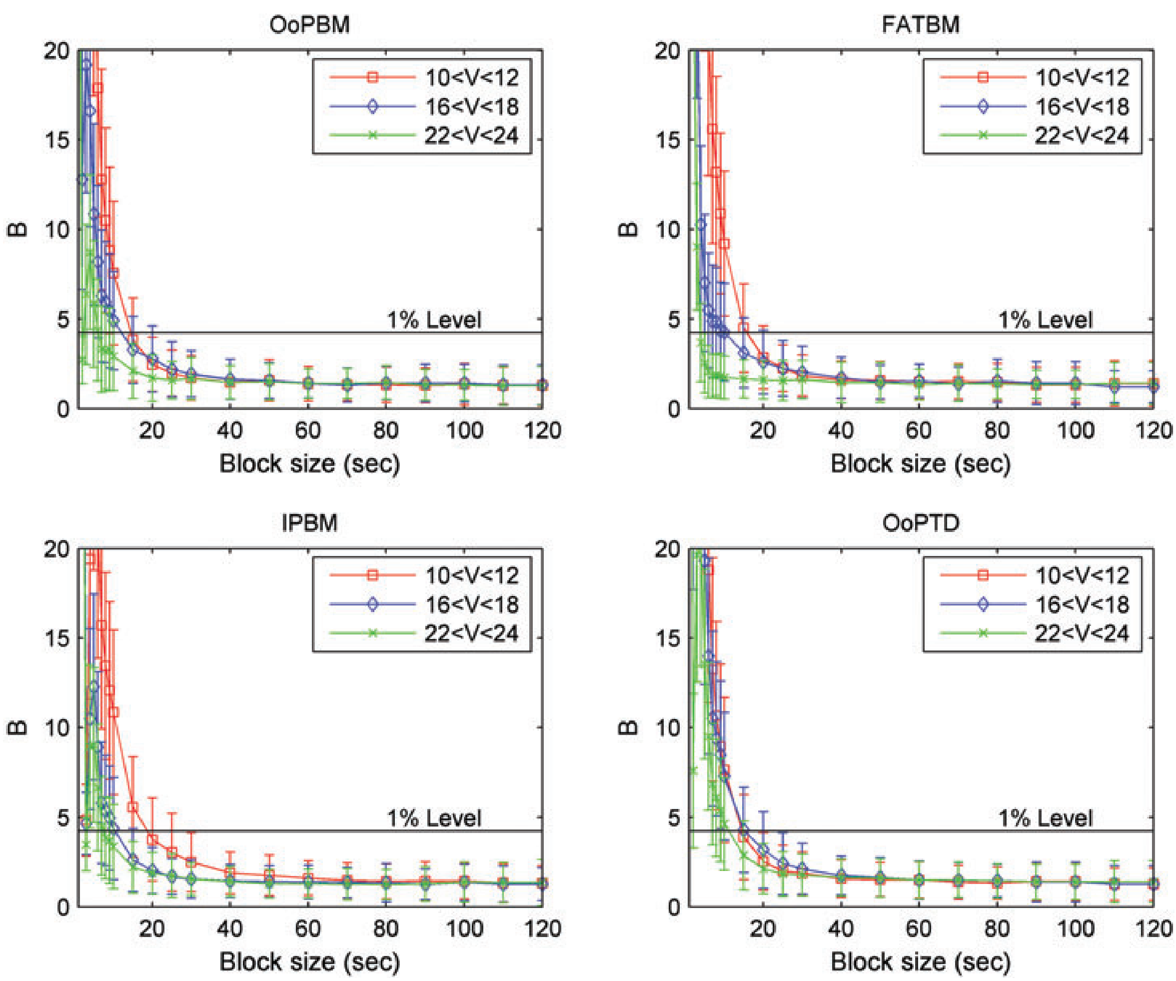


Figure 6. Variation of Blum's test B statistic for four loads as a function of block size (computed from 200 10-min time series for each load type and in three wind speed bins)

DOI: 10.1002/we

Table I. Suggested block sizes (in seconds) for independent block maxima based on mean (μ_B) and mean plus one standard deviation ($\mu_B + \sigma_B$) values from 200 simulations and tested at the 1% significance level.

Wind speed (m/s)	OOPBM		FATBM		IPBM		OOPTD	
	μ_B	$(\mu_B + \sigma_B)$	μ_B	$(\mu_B + \sigma_B)$	μ_B	$(\mu_B + \sigma_B)$	μ_B	$(\mu_B + \sigma_B)$
$2 < V < 4$	50	70	30	50	30	60	50	70
$4 < V < 6$	40	60	25	40	40	60	40	60
$6 < V < 8$	40	60	30	50	30	50	40	60
$8 < V < 10$	40	50	40	50	25	40	40	50
$10 < V < 12$	15	20	20	25	20	30	15	20
$12 < V < 14$	20	30	20	30	15	20	25	40
$14 < V < 16$	20	30	20	30	15	20	20	30
$16 < V < 18$	15	25	15	25	15	20	20	25
$18 < V < 20$	15	20	6	15	10	15	15	25
$20 < V < 22$	7	20	5	6	9	15	15	20
$22 < V < 24$	7	15	4	5	8	15	15	20
$24 < V < 26$	6	15	4	5	7	15	15	20

Table II. Average values of B computed using block maxima with different block sizes based on 20 simulations of four different loads for the most important wind speed bins

Block size (s)	OOPBM	FATBM	IPBM	OOPTD
	$16 < V < 18$	$16 < V < 18$	$24 < V < 26$	$16 < V < 18$
5	9.65	6.91	8.37	18.50
10	4.19	3.72	4.52	6.42
15	3.18	2.73	2.31	3.81
20	3.15	2.95	1.90	3.19
30	2.15	2.12	1.79	2.34
60	1.34	1.39	1.33	1.53

The critical value, B_{cr}, at the 1% significance level is 4.23.

Discussion on Independence

Returning to our earlier discussion on the matter of which wind speed bins bring about the largest loads, it was suggested that lower wind speeds contribute almost no useful information to the tails of long-term load distributions. As such, increasing block sizes to ensure independence in these low wind speed bins offer no benefit to our ultimate goal of statistical loads extrapolation. Table II shows that if only the most important wind speed bins are considered (the ones that cause the largest loads), maxima from block sizes as short as 10, 10, 15 and 15 s, respectively, may be considered acceptably independent for OOPBM, FATBM, IPBM and OOPTD (based on comparing μ_B levels from 20 simulations versus the 1% significance level B_{cr} value of 4.23).

It is instructive to see how a simpler statistical measure, such as the sample correlation coefficient between lag-one extremes, X and Y, used before in the independence test, varies as the block size is changed. Such measures are easier to use than the Blum's test for independence described above; this makes them appealing for studying independence. However, it is important to note that an indication of lack of correlation does not guarantee independence. Nevertheless, sample correlation coefficients of lag-one block maxima (averaged over 20 simulations) for four load types and for the most important wind speed bins for each load are presented in Table III. It is clear that sample correlation coefficient values on lag-one block maxima, as was the case with B values, generally decrease with increasing block size. However, no obvious acceptable correlation coefficient level on lag-one extremes can be claimed as a demarcation point for accepting lack of correlation among block

DOI: 10.1002/we

Table III. Averaged sample correlation coefficients from lag-one block maxima with different block sizes based on 20 simulations of four different loads for the most important wind speed bins

Block size (s)	OOPBM	FATBM	IPBM	OoPTD
	$16 < V < 18$	$16 < V < 18$	$22 < V < 24$	$16 < V < 18$
5	0.39	0.33	0.34	0.53
10	0.33	0.32	0.27	0.41
15	0.29	0.24	0.17	0.35
20	0.26	0.24	0.11	0.30
30	0.22	0.20	0.09	0.25
60	0.11	0.03	0.03	0.11

Table IV. Estimates of the 84th percentile global maximum load for four different load types as obtained from block maxima distributions using different block sizes

Block size (s)	OOPBM	FATBM	IPBM	OOPTD
	MN-m	MN-m	MN-m	m
	$16 < V < 18$	$16 < V < 18$	$24 < V < 26$	$16 < V < 18$
5	12.87	78.91	7.47	7.31
10	12.87	78.92	7.47	7.31
15	12.87	78.92	7.47	7.31
20	12.88	78.93	7.47	7.31
30	12.88	78.94	7.47	7.32
60	12.89	78.97	7.48	7.32

maxima. Although no attempt is made to address this issue, it can be seen that for the optimum block lengths needed to accept the independence criterion at the 1% significance level based on Blum's test (see Table II), correlation coefficients are estimated to be 0.33, 0.32, 0.17 and 0.35 for OOPBM, FATBM, IPBM and OOPTD, respectively. Thus, one could make a general observation, at least based on this LE^3 loads data set, that if correlation coefficients are about 30% or smaller between lag-one block maxima, then the block sizes selected lead reasonably well to independent block maxima. Note, however, that there is no implied direct relation—theoretical or empirical—between B values from Blum's test for independence and sample correlation coefficients.

It has already been established that independence of block maxima is required for the mathematical relationships expressed in equations (4)–(6) to hold. However, it is also important and perhaps more useful to discuss what effect, if any, the assumption of independence (whether or not it is justified) has on extrapolated loads. One can study, for example, what effect the assumption that a given block size assures independence has on the tails of short-term load distributions. Table IV shows estimates of the 84th percentile global maximum load for four load types obtained using block maxima distributions along with the fractile adjustment given by equation (6). Block sizes are varied from 5 to 60 s for the construction of this table.

Examining Table IV reveals that the selection of block size has almost no effect on short-term loads at fairly rare fractile levels. This suggests that although smaller block sizes do not guarantee independence, assuming otherwise does not lead to predictions of grossly inaccurate loads at the 84th percentile non-exceedance probability level for global maxima. This also suggests that one could use very small block sizes without concern for independence; however, no new information is gained by the significantly larger sample of extremes that would result with this small block size. Stated differently, the findings summarized in Table IV based on this LE^3 data set somewhat surprisingly suggest that not enforcing an independence assumption among the block maxima will

DOI: 10.1002/we

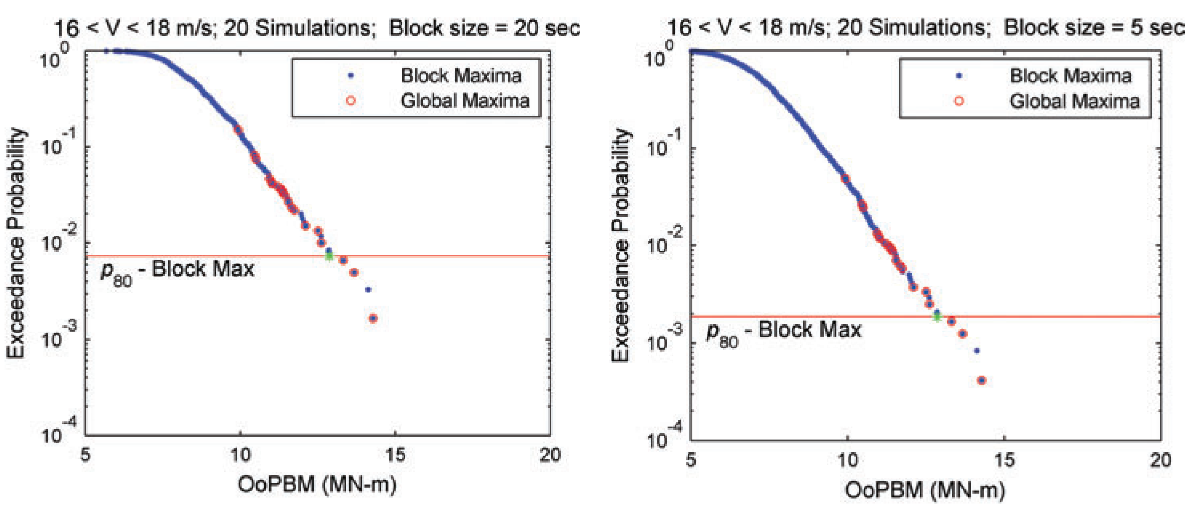
likely have no significant impact on short-term distributions and hence on extrapolation. Figure 7 illustrates this issue by showing short-term empirical distributions based on two different block sizes, one of a 20-s duration that guarantees independent block maxima and the other of a 5-s duration that almost certainly does not.

Short-term distributions on block maxima for the two block sizes are adjusted based on equation (5) so as to be directly represented in terms of exceedance probability in 10 min. From Figure 7, it is seen that even with a block length as short as 5 s the tail of the short-term distribution of OOPBM for this 16–18 m/s wind speed bin is almost identical to the one obtained by using 20-s block maxima. This confirms our statement that the assumption of independence among block maxima, even if not justified, has insignificant impact on predicted rare load fractiles. At the same time, this leads one to the conclusion that empirical distributions based on the use of larger samples of closely spaced block maxima offer no advantages over the use of smaller samples of well-separated block maxima or even global maxima. It should be pointed out that these findings are based on studies only with the LE^3 loads data sets.

Finally, we might state that instead of focusing on the question of independence among extremes, it may be more beneficial to focus on establishing stable tails of the short-term empirical loads distributions. This point has been emphasized several times where we have indicated a need to focus attention on the importance of short-term distribution tails and their propagation to the aggregated long-term distribution curve and its use in extrapolation. In the following, we discuss procedures for controlling the uncertainty associated with these rare empirical short-term load fractiles.

Convergence Criteria

From the preceding discussions, we have seen that adjusting the block size when extracting extremes data from load time series may have a limited effect on the tails of short-term distributions. Accordingly, now our focus shifts to whether or not increases in the number of simulations—with global maxima extracted from each simulation—can help to better define distribution tails and, hence, improve predictions of aggregated long-term distributions and of any extrapolated rare load. It is of interest to be able to estimate how many global maxima (or simulations) would be required to adequately define rare load fractiles. It is expected that by increasing the number of simulations, additional useful information about rare large loads can be gained. This should better



(a) (b)

Figure 7. Short-term distributions for OOPBM for (a) a 20-s block size where independence is verified (see Table II); and (b) a 5-s block size where block maxima are dependent

DOI: 10.1002/we

define short-term load distributions for two reasons. First, additional simulations will realize both more maxima and generally some large loads that can help to fill out and better define the tail of the distribution; this will result in reduced uncertainty in the aggregated long-term distribution as well and, ultimately, in an improved extrapolated load prediction. Second, with additional simulations, estimation of lower and lower probabilities of exceedance is possible leading to a reduction in the extent of extrapolation needed beyond the simulated loads data.

The real issue with running simulations, however, is one of practicality. Carrying out a large number of simulations can be time-consuming; hence, it would be beneficial to know what minimum number of simulations is needed to limit the uncertainty in load predictions to some specified level. A proposal is to enforce a convergence criterion on the tail fractiles of the empirical short-term distributions, not on the long-term distributions nor on the extrapolated long-term (50 years) load itself; for instance, one could require that the uncertainty in the p-quantile load of the short-term load distribution be no larger than some specified value. This could be prescribed by enforcing a maximum limit on confidence intervals on the p-quantile load. We propose such a convergence criterion that may be expressed mathematically in the following manner:

$$\frac{\hat{L}_{\alpha,p} - \hat{L}_{(1-\alpha),p}}{\hat{L}_p} < \frac{q}{100} \tag{9}$$

where the denominator in equation (9) represents the empirical estimate of the p-quantile load (similar to l_p defined in equation (6) earlier), whereas the numerator represents the $(2\alpha\text{-}1)\%$ confidence interval on the p-quantile load. The right-hand side of the inequality contains the variable, q, which represents the maximum acceptable percent error permitted on the normalized confidence interval (where normalization is with respect to the p-quantile load itself). The convergence criterion, as stated, implies that if the normalized confidence interval based on a certain number of simulations exceeds the specified maximum acceptable error, additional simulations will need to be run to reduce this normalized confidence interval. The rationale behind specifying the convergence criterion in this manner is that if the p-quantile is chosen to be reasonably far in the tail of the short-term distribution, uncertainty in its estimate can be controlled or limited. As the tail quite directly influences the long-term aggregated distribution and the extrapolated load, the convergence criterion mentioned earlier—although it does so only indirectly—aids in the overall purpose of statistical loads extrapolation. We need to next discuss how the confidence interval in the numerator of equation (9) can be estimated from data and how the maximum acceptable percent error, q, may be selected. We need to be cognizant of the excessive level of effort that may be needed if q is specified very small or if p is specified as a quantile level that is too far in the tail of the short-term load distribution.

Confidence Intervals Based on Bootstrapping

Estimates of the confidence interval of the p-quantile load are related to the number of simulations or data points used in estimating that load. The bootstrap technique proposed by Efron and Tibshirani[10,11] makes it possible to estimate such confidence intervals by a process that involves randomly resampling data, with replacement, many times. Efron and Tibshirani[11] point out that bootstrapping is a computer-intensive procedure as the data set may need to be resampled thousands of times in order to accurately estimate confidence intervals on statistics such as the quantiles of a distribution. Henderson[12] provides an excellent review of the bootstrap method and offers many examples of its implementation and subtleties involved in its use.

Using the bootstrap procedure to form confidence intervals begins with taking the initial set of data on, say, n global maxima (m_1, m_2, m_3, m_4, $m_5 \ldots m_n$) and randomly resampling these data with replacement to form each time a new set (m_1*, m_2*, m_3*, m_4*, m_5* . . . m_n*) or a bootstrap resampling of the same size as the original sample. Note that bootstrap resamplings will be composed of repeated values from the original sample because, for each resampling, data are sampled randomly with replacement. The process is repeated so as to form a large number, N_b, of bootstrap resamplings. From each of these sets of n data, individual estimates of the p-quantile can be obtained. From these N_b estimates, confidence intervals can be found in the usual manner by rank-ordering the N_b p-quantile estimates. These can then be used for the numerator of equation (9). The estimate of the p-quantile that is obtained from the original data represents the denominator of equation (9).

DOI: 10.1002/we

It is worthwhile to discuss the need to perform a sufficiently large number of bootstrap resamplings to obtain reliable estimates of confidence intervals. The literature suggests that sometimes as few as 25 bootstrap resamplings may be sufficient to estimate some statistics; however, it has also been pointed out that many more resamplings are required to estimate confidence intervals.[10] Lunneborg[13] has suggested that bootstrap resamplings be increased incrementally until some stability results in estimates of the standard error. However, as pointed out earlier, bootstrapping is a computer-intensive method and the question of how many bootstrap resamplings are required is perhaps not so important as it is just as easy to carry out 5000 resamplings as it is carry out 25; the benefit of a larger number is that it will lead to more reliable estimates of the desired statistic (see Chernik[14] for further discussion on this issue). In the present study, 5000 bootstrap resamplings were used to form confidence interval estimates on load quantiles and these were found to be adequately stable. Figure 8 illustrates differences between confidence interval estimates on the 0.84-quantile OOPBM load when a small number of bootstrap resamplings is used (25 in the top panel of the figure) compared with that when a large number is used (5000 in the bottom panel of the figure). Variability in estimates of the normalized 90% confidence interval width is obvious with the smaller number as what is evident by running the same bootstrap procedure 10 times. With the smaller number of bootstrap resamplings, 90% confidence intervals on a 0.84-quantile load from different runs can vary greatly from the stable estimates obtained with the larger number; deviations are smaller with the larger number of bootstrap resamplings.

Confidence Intervals Based on the Binomial Distribution

As an alternative to the bootstrap procedure discussed earlier, the binomial distribution may also be used to obtain confidence interval estimates on the p-quantile load.[15] This can limit the computational effort necessary when evaluating the numerator in equation (9). The theoretical development is presented here.

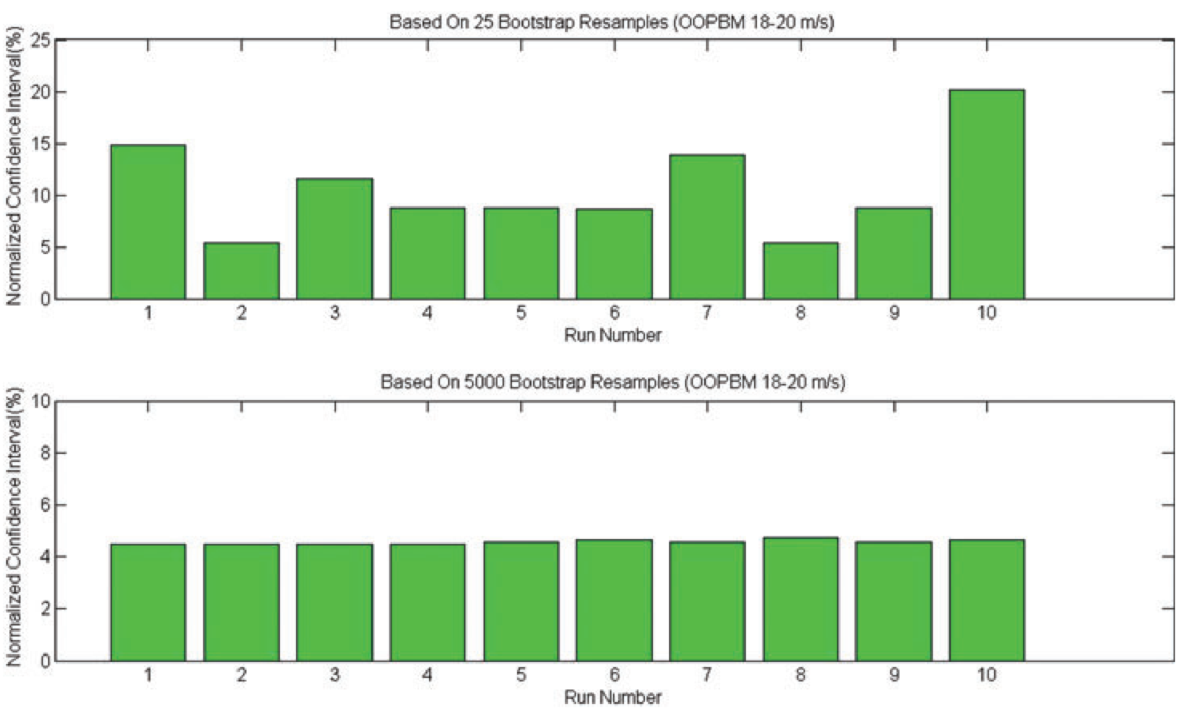


Figure 8. Normalized 90% confidence interval estimates on the 0.84-quantile global maximum OOPBM load in the 18–20 m/s wind speed bin based on 30 simulations followed by 25 bootstrap resamplings (top figure) and 5000 resamplings (bottom figure)

DOI: 10.1002/we

We start by writing the formula for the binomial probability mass function, $B(i; m, p)$, which expresses the probability of i occurrences of an event of interest in m Bernoulli trials when the probability of occurrence of the event in any single trial is p.

$$B(i; m, p) = \frac{m!}{i!(m-i)!} p^i (1-p)^{m-i}; \quad i = 0, 1, 2 \ldots m \tag{10}$$

In the present context, let the event of interest refer to the non-exceedance of the estimated p-quantile load and let m refer to the number of aeroelastic simulations run as well as to the number of global maxima extracted. From the definition of the probability mass function in equation (10), a cumulative distribution function, $C(j; m, p)$, may be written as follows:

$$C(j; m, p) = \sum_{i=0}^{j} B(i; m, p); \quad j = 0, 1, 2 \ldots m \tag{11}$$

To form the confidence interval on the p-quantile load needed for the numerator of equation (9), two load levels, x_k and x_l, need to be found such that the following is true (where X refers to the p-quantile load):

$$P(x_k < X < x_l) = 2\alpha - 1 \tag{12}$$

To simplify notation, the quantities, $\hat{L}_{(1-\alpha),p}$ and $\hat{L}_{\alpha,p}$, in equation (9) have been replaced by x_k and x_l, respectively, in equation (12). The load levels, x_k and x_l, need to be obtained by searching the rank-ordered extremes, x_{1*}, $x_{2*}, \ldots x_{m*}$, from the m simulations and then interpolating. It is generally possible to find two integer values, k^* and l^*, where k^* is the largest integer such that

$$C(k^*; m, p) \le (1-\alpha) \tag{13}$$

and where l^* is the largest integer such that

$$C(l^*; m, p) \le \alpha \tag{14}$$

and $1 \le k^* < l^* \le m - 1$.

The integers k^* and l^* are such that x_{k^*} and x_{l^*} bound the desired load levels, x_k and x_l, from below. This in turn means that $x_{k^*} \le x_k \le x_{(k+1)^*}$ and $x_{l^*} \le x_l \le x_{(l+1)^*}$. Once the integers k^* and l^* are found using equations (13) and (14); the load levels x_k and x_l may be found by interpolation as follows:

$$x_k = \frac{x_{(k+1)^*} - x_{k^*}}{C_{(k+1)^*} - C_{k^*}} [(1-\alpha) - C_{k^*}] + x_{k^*} \tag{15}$$

$$x_l = \frac{x_{(l+1)^*} - x_{l^*}}{C_{(l+1)^*} - C_{l^*}} [\alpha - C_{l^*}] + x_{l^*} \tag{16}$$

where, to simplify notation, we have set $C(k^*; m, p) = C_{k^*}$, $C[(k+1)^*; m, p] = C_{(k+1)^*}$, $C(l^*; m, p) = C_{l^*}$, and $C[(l+1)^*; m, p] = C_{(l+1)^*}$. The desired $(2\alpha$-$1)$% confidence interval required for equation (9) is simply equal to $x_l - x_k$.

Normal Approximation to the Binomial Distribution

Note that the binomial cumulative distribution function is needed to find the integer values of k^* and l^* as well as to obtain the values of x_k and x_l using equations (15) and (16). It is possible to replace the binomial distributed integer random variable representing the number of occurrences in m trials (where the probability of event occurrence in a single trial is p) by a normally distributed real random variable with mean equal to mp and variance equal to $mp(1 - p)$. If this is done, the confidence interval developed following the steps indicated by equations (11)–(16) remains valid; the only difference is a simplification of equation (11) that may be approximated as follows:

$$C(j; m, p) \cong \Phi(z_{\mathrm{j}}) \quad \text{where} \quad z_{\mathrm{j}} = \frac{(j+0.5)-mp}{\sqrt{mp(1-p)}} \tag{17}$$

In equation (17), Φ( ) refers to the cumulative distribution function of a standard normal random variable. The reduced effort in steps involving determination of the values of k^* and l^* in equations (13) and (14) represents the most significant advantage as it is much easier to find k^* and l^* by using the inverse cumulative distribution of a standard normal random variable than it is to evaluate the binomial cumulative distribution function using equations (10) and (11). Note that the normal approximation to the binomial involves a continuity correction of 0.5 in the definition of z_{j} in equation (17) as we are replacing an integer random variable with a real one. Although the normal approximation to the binomial is most accurate when the values of mp and $m(1 - p)$ are greater than 5, in other cases as well the approximation is reasonably accurate as we shall see.

Binomial Confidence Bounds Simplified for Wind Turbine Applications

The confidence intervals based on the binomial distribution as developed in equations (10)–(16) are less computationally intensive than those computed using the bootstrap procedure. It is possible, though, to simplify the binomial-based confidence intervals to an even greater extent for applications to statistical extrapolation of wind turbine extreme loads. Essentially, this is done by tabulating values of k^* and l^* that will result for most common situations where the number of simulations is in the order of 15 to 35 for each wind speed bin. This number of simulations will be shown to be reasonable if, in equation (9), the 90% confidence interval on the 84th percentile load is computed, and the maximum error on the normalized confidence interval of equation (9) is to be less than 15% (i.e. q = 15). Note that the simplified approach to the binomial-based confidence interval can only be reasonably tabulated for specific values of p and α. For instance, for p equal to 0.84 and α equal to 0.95, Table V provides values of k^* and l^* as well as two other values, A and B, needed for interpolating as done in equations (15) and (16).

Table V. Parameters needed to establish binomial-based confidence intervals (for $\alpha = 0.95$ and $p = 0.84$)

	Number of simulations	k^*	l^*	A	B
For 90% confidence interval on the 84th percentile load	15	9	14	0.50	0.32
	16	10	15	0.27	0.19
	17	11	16	0.10	0.03
	18	11	16	0.87	0.96
	19	12	17	0.58	0.90
	20	13	18	0.35	0.83
	21	14	19	0.16	0.76
	22	14	20	1.00	0.69
	23	15	21	0.69	0.60
	24	16	22	0.45	0.50
	25	17	23	0.25	0.39
	26	18	24	0.08	0.26
	27	18	25	0.85	0.12
	28	19	25	0.58	0.98
	29	20	26	0.36	0.91
	30	21	27	0.18	0.83
	31	22	28	0.02	0.75
	32	22	29	0.75	0.66
	33	23	30	0.51	0.56
	34	24	31	0.31	0.44
	35	25	32	0.13	0.32

DOI: 10.1002/we

Table V works in conjunction with a design equation that is tailored to be used with it. This design equation and Table V give the 90% confidence interval for the 84th percentile 10-min maximum [i.e. $\alpha = 0.95$ and $p = 0.84$ in equation (9)]. The design equation can be written as follows:

$$(x_l - x_k) = (x_{l^*} - x_{k^*}) + B(x_{(l+1)^*} - x_{l^*}) - A(x_{(k+1)^*} - x_{k^*}) \tag{18}$$

where l^*, k^*, A and B are given in Table V as a function of the number of simulations run, and x_{l^*}, $x_{(l+1)^*}$, x_{k^*} and $x_{(k+1)^*}$ are obtained from the rank-ordered simulated extremes.

As an illustration of how equation (18) and Table V can be used, consider a situation where 20 simulations have been carried out. Then, $l^* = 18$, $k^* = 13$, $A = 0.35$ and $B = 0.83$ according to Table V, and if the rank-ordered extremes, x_{13^*}, x_{14^*}, x_{18^*} and x_{19^*} are found, equation (18) can be used to compute the confidence interval needed for equation (9). If the convergence criterion is not met according to equation (9), additional simulations may be run and new values of l^*, k^*, A and B obtained again from Table V. This may be repeated until the convergence criterion is met.

Application of the Convergence Criteria to the LE^3 Loads Data Set

Convergence criteria for four load measures, OOPBM, FATBM, OOPTD and IPBM are studied using the LE^3 loads database. Based on equation (9), we are interested in computing the percent error in terms of the normalized 90% confidence interval of the 84th percentile 10-min global maximum for each load type for different numbers of simulations. If the maximum allowable percent error, q, is specified, the appropriate number of simulations can be ran. Table VI shows computed percent errors when 30 simulations are ran for each wind speed bin and for all of the four loads. The results presented in these tables suggest that the convergence criterion is adequately met if the maximum error permitted is 15% (i.e. if q is equal to 15). For IPBM loads, even a 10% maximum error criterion would be met when 30 simulations are ran. For the OOPBM and OOPTD, slowest convergence is seen in the 16–18 m/s wind speed bin but even there, 30 simulations lead to normalized 90% confidence intervals on the 84th percentile load that are smaller than 15%.

Short-term load distributions for the four load types are summarized in Figure 9 for the single wind speed bin that in each case showed slowest convergence based upon the normalized 90% confidence interval on the 84th percentile load. The 90% confidence intervals are shown for each empirical distribution at the 1–0.84 exceedance probability level. The 84th percentile load is shown along with the confidence interval based on the bootstrap method as well as the binomial method.

The preceding discussion suggests that, for the LE^3 data set, adequately stable tails for the short-term distributions in all wind speed bins can be obtained if 30 simulations are ran. Enforcing the convergence criterion of a maximum percent error of 15% on the normalized 90% confidence interval on the 84th percentile load in the short-term distributions leads us to state that these distributions are reasonably well-estimated. The next and more important issue to study with regard to extrapolation is whether or not the controlled uncertainty in short-term distributions propagates to stable aggregated long-term distributions as well. Tail stability of aggregated long-term distributions for each load type can be evaluated using bootstrap procedures in a similar manner to that employed for the short-term distributions. However, as the aggregated distribution involves weighting based on the Rayleigh wind speed distribution, it is not possible to bootstrap the data in this long-term distribution directly. Instead, the short-term distributions can be each bootstrapped separately and then for each wind speed bin's bootstrap resampling, the aggregation can be carried out leading to multiple long-term distribution curves. From these multiple long-term probability distributions, it is possible to evaluate confidence intervals on any desired long-term load quantile. All the wind speed bins are aggregated in the long-term distribution but are weighted to different degrees according to the Rayleigh distribution. The worst wind speed for convergence at the short-term level is diminished (in relative terms) in importance at the long-term level. As a result, long-term distributions often appear fairly stable and have low uncertainty, once convergence criteria for short-term load distributions have been enforced.

In closing our discussion on convergence criteria, we note that when a maximum error of 15% was imposed for the short-term distributions (in conjunction with 90% confidence intervals of the 84th percentile load), 90%

DOI: 10.1002/we

Table VI. Estimates of the 84th percentile load, x_k, x_l, and the binomial-based normalized 90% confidence interval for OOPBM, FATBM, OOPTD, and IPBM based on 30 simulations

OOPBM	x_{84} (MN-m)	x_k (MN-m)	x_l (MN-m)	$(x_l - x_k)/x_{84}$ (%)
2 < *V* < 4	3.50	3.36	3.55	5.5
4 < *V* < 6	5.82	5.54	5.90	6.2
6 < *V* < 8	9.09	8.52	9.47	10.5
8 < *V* < 10	12.45	12.28	12.63	2.8
10 < *V* < 12	13.52	13.38	13.60	1.6
12 < *V* < 14	13.84	13.59	13.90	2.3
14 < *V* < 16	13.87	13.57	14.02	3.2
16 < *V* < 18	13.33	12.17	13.84	12.5
18 < *V* < 20	11.58	11.08	11.66	5.0
20 < *V* < 22	10.83	10.33	10.95	5.8
22 < *V* < 24	10.05	9.78	10.57	7.9
24 < *V* < 26	9.80	9.54	10.12	5.9
FATBM	x_{84} (MN-m)	x_k (MN-m)	x_l (MN-m)	$(x_l - x_k)/x_{84}$ (%)
2 < *V* < 4	24.75	23.56	25.52	7.9
4 < *V* < 6	35.82	35.52	36.00	1.3
6 < *V* < 8	51.51	48.94	53.59	9.0
8 < *V* < 10	73.30	72.18	74.97	3.8
10 < *V* < 12	80.51	79.27	80.72	1.8
12 < *V* < 14	85.04	83.20	86.78	4.2
14 < *V* < 16	84.06	82.82	85.49	3.2
16 < *V* < 18	81.16	78.20	83.07	6.0
18 < *V* < 20	71.89	67.24	72.69	7.6
20 < *V* < 22	62.05	60.55	64.95	7.1
22 < *V* < 24	63.45	59.23	64.80	8.8
24 < *V* < 26	61.84	59.09	62.59	5.7
OOPTD	x_{84} (m)	x_k (m)	x_l (m)	$(x_l - x_k)/x_{84}$ (%)
2 < *V* < 4	2.08	1.93	2.14	10.2
4 < *V* < 6	3.35	3.24	3.47	6.8
6 < *V* < 8	5.05	4.80	5.25	9.0
8 < *V* < 10	6.93	6.87	7.04	2.5
10 < *V* < 12	7.73	7.48	7.78	3.9
12 < *V* < 14	7.74	7.65	7.77	1.6
14 < *V* < 16	7.69	7.45	7.73	3.7
16 < *V* < 18	7.52	6.56	7.69	15.0
18 < *V* < 20	6.04	5.71	6.36	10.7
20 < *V* < 22	5.52	5.17	5.65	8.8
22 < *V* < 24	4.94	4.52	4.95	8.6
24 < *V* < 26	4.45	4.18	4.78	13.5
IPBM	x_{84} (MN-m)	x_k (MN-m)	x_l (MN-m)	$(x_l - x_k)/x_{84}$ (%)
2 < *V* < 4	3.79	3.75	3.81	1.7
4 < *V* < 6	4.25	4.21	4.28	1.8
6 < *V* < 8	4.96	4.90	5.05	2.9
8 < *V* < 10	5.67	5.60	5.82	3.9
10 < *V* < 12	6.03	5.96	6.06	1.6
12 < *V* < 14	6.35	6.28	6.45	2.6
14 < *V* < 16	6.45	6.33	6.54	3.2
16 < *V* < 18	6.70	6.61	6.75	2.0
18 < *V* < 20	7.03	6.72	7.26	7.6
20 < *V* < 22	7.37	7.16	7.42	3.6
22 < *V* < 24	7.62	7.43	7.90	6.2
24 < *V* < 26	7.69	7.44	7.85	5.2

DOI: 10.1002/we

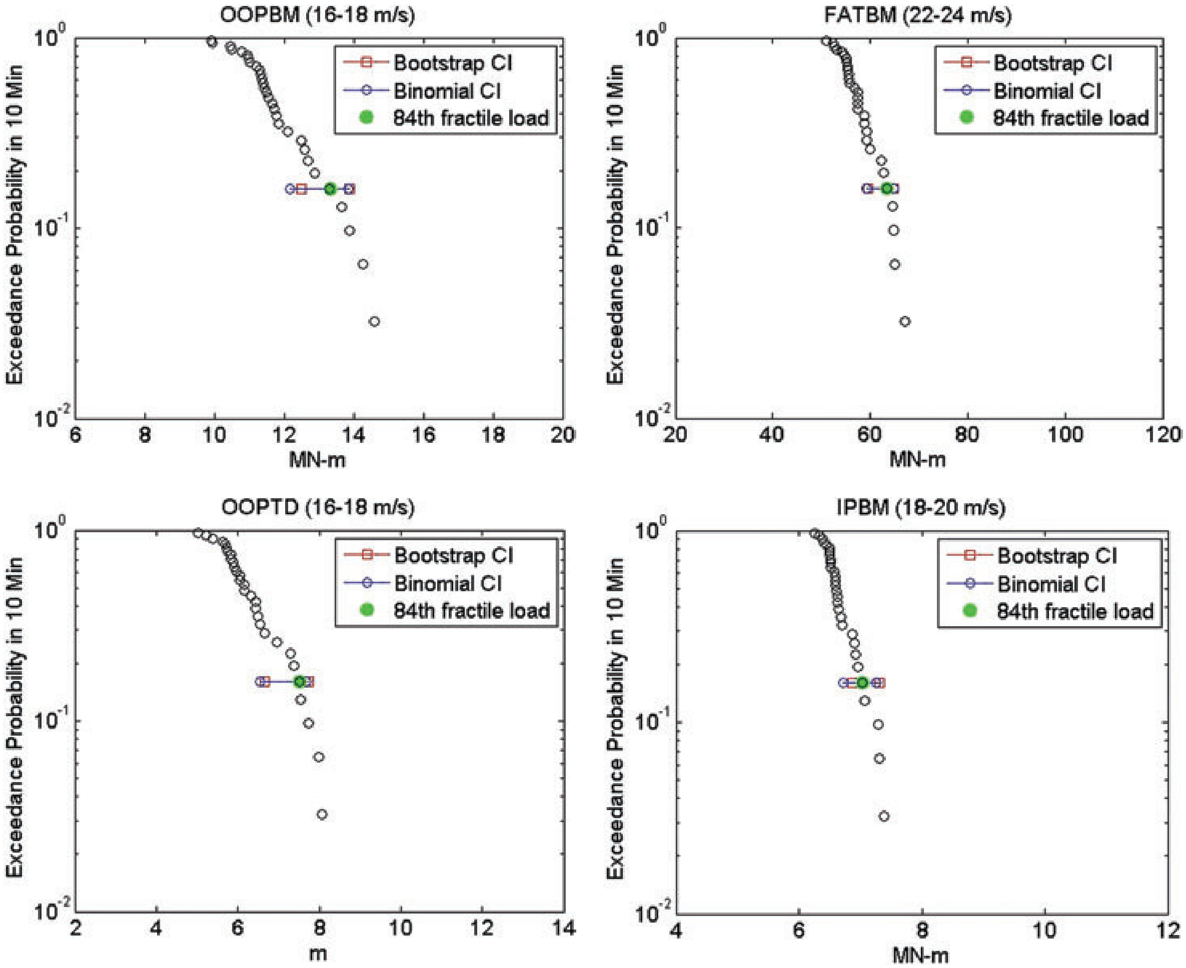


Figure 9. Short-term distributions for OOPBM, FATBM, OOPTD and IPBM based on 30 simulations. Bins selected have largest 90% relative confidence bounds (RCBs) on the 84th percentile load. The 84th percentile load is shown as are binomial- and bootstrap-based confidence intervals (CIs)

confidence bounds long-term distributions are also small. It is important to note, though, that these results were based on studies of four loads and came from one wind turbine model alone. If simulated loads data from other turbines are considered, results might be different; therefore, it is suggested that the value of q in equation (9) be adjusted as appropriate so as to not cause excessive amount of simulations. The convergence criterion triad ($\alpha = 0.95$, $p = 0.84$, $q = 15$) for short-term distributions has been demonstrated to yield stable tails of the short-term distributions as well as stable long-term distributions. It was found that if the short-term distributions are verified through convergence checks, uncertainty in aggregated long-term distributions is acceptably low as well.

Real Uncertainty in Short-Term Loads

We have shown that confidence intervals on any load quantile at the short-term level may be estimated using bootstrap- and binomial-based methods. These estimates rely on a limited single set of simulated data. It is possible, though, to examine confidence intervals and evaluate convergence criteria by using the entire LE^3

loads data set. If all the 1200 global maxima available for each wind speed bin are utilized, real uncertainty can be estimated directly without resorting to methods such as bootstrapping. For example, this may be done by subdividing the set of 1200 global maxima per bin into 40 individual sets of 30 maxima. Then, real confidence intervals of the *p*-quantile load may be found by extracting 40 estimates of that load. The convergence criterion based on equation (9) can then be checked using this confidence interval. As all the data here are real wind turbine maxima, not statistical estimates, the 90% confidence interval is normalized by the mean of these values. Figure 10 shows a comparison of 90% confidence intervals for the 84th percentile levels of four loads (OOPBM, FATBM, OOPTD and IPBM) based on 30 simulations, except in the 'real' case that is based on 1200 simulations. The figure shows no consistent trend or difference between confidence intervals based on real data versus those based on bootstrapping and the binomial approach. Note that the figure also shows that the normal approximation to the binomial is generally quite good for estimation of confidence intervals.

Conclusions

Using the LE^3 loads data, several questions related to the theory and practical implementation of statistical loads extrapolation have been addressed. First, we have emphasized the need to understand which wind speeds tend to cause largest loads of different types. This information is useful to have when determining where greater simulation effort is needed.

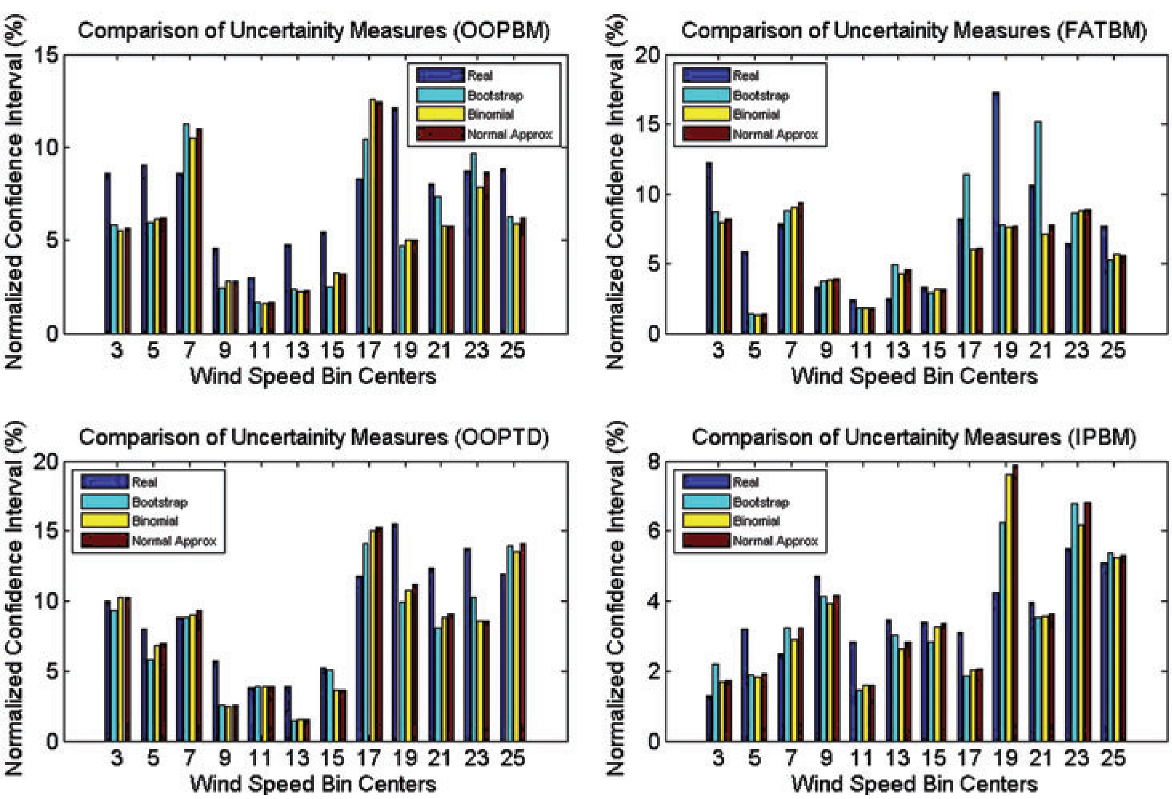


Figure 10. Normalized 90% confidence intervals on the 84th percentile load based on 1200 simulations (real data), as well as based on 30 simulations followed by bootstrapping, a binomial-based method and a normal approximation to the binomial. Results are shows for four load types (OOPBM, FATBM, OOPTD and IPBM) and for all wind speed bins

DOI: 10.1002/we

We have presented details related to several steps involved in predicting a long-term load of interest (such as the 50-year return period load) by defining short-term distributions first and then aggregating these to long-term distributions.

We have compared the use of global and block maxima in the context of statistical loads extrapolation. Issues related to independence of block maxima have been addressed and statistical tests of independence have been formulated. Ignoring unimportant wind speed bins, it was found that block sizes of around 40–60 s for four loads (OOPBM, FATBM, OOPTD and IPBM) led to independent block maxima when checked at the 1% significant level based on a well-accepted statistical test. Finally, with the LE^3 data, it was demonstrated that there is no advantage gained from using block maxima over global maxima when short-term loads are estimated. Moreover, even if block sizes are very small so as to exhibit dependence, ignoring this again leads to small error in estimation of short-term loads.

In order to assure stable or robust short-term distributions especially in the tails of these distributions, a convergence criterion was developed that relies on computing confidence intervals on rare load quantiles. Based on studies with the LE^3 data, the convergence criterion that was proposed is that the normalized 90% confidence interval on the 84th percentile load (normalization is with respect to the 84th percentile load estimate from the simulations) may not exceed 15%. The 90% confidence interval that forms part of the convergence criterion may be estimated based on bootstrap methods as well as the binomial distribution. Both procedures have been developed here. Additionally, for the binomial method, an approximation using the normal distribution was presented. Finally, a design equation and additional tabulated parameters were presented that enable quick computation of the normalized 90% confidence interval on the 84th percentile load as long as the number of simulations run is between 15 and 35. The convergence criteria were applied for four loads (OOPBM, FATBM, OOPTD and IPBM), and with 30 simulations, the normalized 90% confidence interval on the 84^{th} percentile load never exceeded 15%. Bootstrap- and binomial-based confidence intervals were reasonably similar. The convergence criteria applied to the short-term loads distributions were verified to lead to smaller uncertainty in aggregated long-term load distributions. Small uncertainty in long-term distributions is expected to lead to good, robust predictions of extrapolated rare loads although that was not a focus of this study.

In closing, it is important to note that all of these conclusions were derived based on findings from the LE^3 data set and the simulated loads data are for the LE^3 5MW turbine model alone. It is possible that the study of loads from other turbines will lead to different conclusions. Nevertheless, the focus of this study was to develop several new ideas that—it is hoped—will aid in statistical loads extrapolation for practicing engineers.

Acknowledgements

The authors are pleased to acknowledge useful discussions with members of the LE^3 working group over the last year. In particular, they thank Dr. Pat Moriarty of the Natural Renewable Energy Laboratory for the simulated loads data, and Dr. Paul Veers of Sandia National Laboratories for continued encouragement and advice on various aspects of this work. Finally, the authors wish to express their gratitude for the financial support received from Sandia National Laboratories (Contract No. 743378) and from the National Science Foundation (Grant Nos. CMMI-0449128 and CMMI-0727989).

References

1. International Electrotechnical Commission. *IEC 61400-1: Wind Turbines—Part 1: Design Requirements*. Edition 3. IEC: Geneva, Switzerland, 2005.
2. Jonkman JM, Butterfield S, Musial W, Scott G. *Definition of a 5MW Reference Wind Turbine for Offshore System Development*. NREL/TP-500-38060, National Renewable Energy Laboratory: Golden, CO, 2007.
3. Moriarty P. Database for validation of design load extrapolation techniques. *Wind Energy* 2008; **11**(6): 559–576. DOI:10.1002/we.305.

4. Jonkman BJ, Buhl ML Jr. *TurbSim User's Guide*. NREL/TP- 500-41136. National Renewable Energy Laboratory: Golden, CO, 2007.
5. Jonkman JM, Buhl ML Jr. *FAST User's Guide*. NREL/EL- 500-38230. National Renewable Energy Laboratory: Golden, CO, 2005.
6. Ragan P, Manuel L. Statistical extrapolation methods for estimating wind turbine extreme loads. *Proceedings of the 45th AIAA Aerospace Sciences Meeting and Exhibit*. AIAA-2007-1221, Reno, Nevada, January 2007.
7. Blum JR, Kiefer J, Rosenblatt M. Distribution free tests of independence based on the sample distribution function. *Annals of Mathematical Statistics* 1961; **32**: 485–498.
8. Hollander M, Wolfe DA. *Nonparametric Statistical Methods*. Wiley: New York, 1999.
9. Skaug HJ, Tjøstheim D. A nonparametric test of serial independence based on the empirical distribution function. *Biometrika* 1993; **80**: 591–602.
10. Efron B, Tibshirani RJ. *An Introduction to the Bootstrap*. Chapman and Hall: New York, 1993.
11.Efron B, Tibshirani RJ. Statistical data analysis in the computer age. *Science* 1991; **253**: 390–395.
12. Henderson AR. The bootstrap: a technique for data-driven statistics. Using computer-intensive analyses to explore experimental data. *Clinica Chimica Acta* 2005; **359**: 1–26.
13. Lunneborg CE. *Data Analysis by Resampling: Concepts and Applications*. Duxbury Press: Pacific Grove, CA, 2000.
14. Chernik MR. *Bootstrap Methods: A Practitioner's Guide*. John Wiley and Sons: New York, 1999.
15. Hogg RV, Craig A. *Introduction to Mathematical Statistics*, 5th Edition. Prentice Hall: Englewood Cliffs, NJ, 1995.

DOI: 10.1002/we