Effect of Statistical Uncertainties on Predicted Extreme Wind Speeds

Á. Rózsás[1)] and M. Sýkora[2)]
[1)]*Department of Structural Engineering, Budapest University of Technology and Economics, Budapest 1111, Hungary, rozsas.arpad@epito.bme.hu*
[2)]*Department of Structural Reliability, Klokner Institute, Czech Technical University in Prague, Prague 16608, Czech Republic, miroslav.sykora@cvut.cz*

Abstract: Statistical uncertainties, arising from the uncertainty of parameter estimation and model selection, are often neglected in probabilistic assessment of engineering structures. However, few previous studies indicate that this might cause severe underestimation of extreme loads and lead to insufficient structural reliability. This contribution aims to qualitatively and quantitatively investigate the effect of this simplification on extreme values of wind speed that are commonly associated with design values. The probabilistic modelling of basic wind speeds is thoroughly investigated. Moderately high temporal resolution data – daily 10 min averaged maxima from three distinct one hour long measurement sessions – are obtained from the Carpatclim database, covering a 50-year observation period. Data for Budapest are taken into account as a representative example. Block maxima and peak over threshold approaches are applied to extract maxima and to fit associated distributions. Frequentist and Bayesian statistics are used to assess the effect of statistical uncertainties. The parameter estimation uncertainty is quantified by uncertainty intervals. Statistical model uncertainty is explored using different distribution types and taken into account by Bayesian model averaging. The conducted analyses imply that neglecting statistical uncertainties might yield to considerable underestimation of extreme values. Using the currently widespread annual maxima approach, the parameter estimation uncertainty can lead to underestimation of 1000-year return period values by about 20%. The commonly adopted Gumbel model yields 20% larger values with a return period of 1000 years than those based on the generalized extreme value distribution. The latter fits better to data though unambiguous, fully data-driven recommendation on model selection cannot be made. Bayesian posterior predictive distribution is recommended for accounting parameter estimation uncertainty. Moreover, if viable, smaller than one year block size, multiple maxima in a block, or peak over threshold methods are recommended to increase sample size and reduce statistical uncertainties. This leads to 70% reduction in the range of a 90% confidence interval for 1000-year extremes for the selected location.

Keywords: Bayesian statistics, design value, extreme load, generalized extreme value distribution, model selection, parameter estimation, statistical uncertainty, wind speed

1. Introduction

1.1. Motivation

Probabilistic models of extreme wind speeds are thoroughly studied by researchers from various disciplines, particularly by meteorologists. However, structural reliability requirements often substantially differ from interests of meteorologists, e.g. estimation of large (>1000 year) return period events and focus on prediction rather than best fit to observations. For instance, extreme events dominating structural reliability may have return periods of thousands years as is a common requirement for structures in nuclear industry. Since observation periods are only small fraction of these return periods, extrapolation to unobserved ranges is inevitable. This can yield to highly uncertain model estimates at the critical tail regions. Nevertheless it appears that this uncertainty is commonly neglected in civil engineering while drafting standards or conducting probabilistic analysis. The aim of this paper is to investigate the effect of this neglect on representative wind speed fractiles and to find out whether the current practice is (reasonably) conservative. Scarcity of observations leads to aforementioned uncertainty for all random variables; hereinafter it is referred to as statistical uncertainty. It is reasonable to make distinction between two classes: parameter estimation uncertainty and model selection uncertainty. The former is concerned with the uncertainty of parameter estimation for a given model, while the latter relaxes this constraint by considering a group of candidate models and quantifying their "goodness".

Previous studies show that the neglect of these uncertainties might cause severe underestimation of extreme loads and lead to insufficient structural reliability. For ground snow load – using the annual maxima for inference – a 1000-year return period value might be underestimated by 20% due to the neglect of parameter estimation uncertainty (Rózsás and Sýkora, 2015a), and the neglect of model selection uncertainty can lead to over an order of magnitude underestimation of failure probability (Rózsás and Sýkora, 2015b). Other studies found similar, often more significant effects in cases of extreme rainfalls (Coles and Pericchi, 2003), extreme hydrological hazards (Sisson et al., 2006), and slope reliability (Li et al., 2015).

In the following we critically review the current practice of extreme wind modelling in civil engineering and propose statistical approaches and a general rationale to quantify and incorporate statistical uncertainties. The applied methods can be utilized for other extreme climatic actions such as flooding and snow loads.

1.2. STATISTICAL MODELLING IN CIVIL ENGINEERING

Table I. reports a brief overview of current civil engineering practice in statistical modelling of extreme wind speeds. Features relevant for the present analysis are highlighted, e.g. applied method to extract extreme values and its parameters of the method, considered distribution type, treatment of statistical uncertainties and sources of information. More details are given in Section 2.

The suite of Eurocodes lacks commentary or references to background documents. However, based on some paragraphs and background documents, it seems that statistical uncertainty has been neglected while drafting EN 1991-1-4 on wind actions on structures. To our knowledge this approach is applied to other actions too, e.g. point estimates are used to derive representative ground snow loads (Sanpaolesi, 1998) and thermal actions as well (Sanpaolesi and Colombini, 2005). Since annual maxima are typically used to infer probabilistic models, and to derive representative fractiles and partial factors in Eurocodes, this neglect might have considerable impact.

The American load code ASCE 7-10 mentions the effect of statistical uncertainty – termed therein as sampling error – on estimating basic wind speed from regional climatic data. However, no general justification is provided and the provision seems to be rather arbitrary.

Table I. Summary of civil engineering practices in statistical modelling of extreme wind speeds.

Reference	Method	Distribution	Variable	Stat. uncertainty
EN 1991-1-4 (4.2)	Block method*	Gumbel	basic wind speed**	neglected
EN 1991-1-4 (E.10)	Block method*	Weibull	wind speed, wind induced cyclic loading	NA
JCSS Part 2.13	Block method*, annual maxima mentioned	Gumbel (Weibull)	wind speed maxima	neglected
(Niemann and Diburg, 2013)	Block method - annual maxima (9-102 observations)	Gumbel, GEV	wind speed annual maxima	neglected
(Kruger, 2010)	Block method with annual maxima and peak over threshold method with various thresholds	GEV, GP, Gumbel, etc.	wind speed maxima	upper bound of 75% confidence interval of parameters
ASCE 7-10 Ch.26.	Block method*	Gumbel	basic wind speed**	guidance for regional data

*Judgement of the authors; **EN 1991-1-4: obtained from 10 minutes mean wind speed at 10 m above ground level in open country terrain, modified to account for the direction of the wind being considered and the season (if required), characteristic value corresponding to 50 years return period; ASCE 7-10: three-second gust speed at 10 m above the ground in open terrain prevailing in the upwind direction (detailed definition in the standard), the value given in the standard corresponding to 300-1700 years return period depending on occupancy category.

Kruger (2010) analyzed extreme wind speeds for South Africa and considered the effect of parameter estimation uncertainty. His approach is to upward adjust the estimated parameters by an appropriate confidence limit: 75% is recommended. This again seems to be unjustified, moreover a snow related study showed that this approach is not appropriate to account for statistical uncertainties and is usually non-conservative (Rózsás and Sýkora, 2015c). This brief review indicates that parameter estimation and model selection uncertainties are commonly neglected or inadequately addressed in probabilistic reliability studies involving wind actions.

1.3. Adopted approach and available database

Wind speed data from a representative location of the Carpathian Region are initially analyzed by accepting the widespread techniques and assumptions in civil engineering. Then advanced statistical analysis is carried out to quantify and incorporate statistical uncertainties and is considered to provide reference values. To explore the effect of statistical uncertainties, two different statistical paradigms are employed: frequentist and Bayesian statistics. Extreme value theory of mathematical statistics is used to extract extremes from observations and to select theoretically supported distribution functions. All models in this study are fully statistical, e.g. no physical arguments and principles are incorporated.

The wind data under study are obtained from the Carpatclim database (Szalai et al., 2013), covering a 50-year observation period. The climatological grid covers the region between latitudes 44°N and 50°N, and longitudes 17°E and 27°E. The data are gathered at 10 m height above the ground in horizontal direction from 270 stations with relatively homogeneous spatial distribution. These are homogenized and spatially interpolated using meteorological and statistical models; these post-processed data are available in the database and used as inputs in this study. The database has moderately high temporal resolution: for each day maxima of 10 min averaged wind speed from three distinct measurement sessions spanning over one hour, and about 10 km spatial resolution. Note that the experience from detailed analyses of wind speeds indicates that annual maxima obtained from continuous measurements exceed those inferred from stations with the three measurement sessions per day by about 5 %. Obviously, this difference needs to be taken into account while deriving representative values for structural design, but is neglected in this study focused on comparison of various statistical approaches to modelling of extreme wind speeds.

For the 50-year period more than 18000 observations for a grid point are thus available. The 10 min averaged value “is typically sufficiently long to incorporate most of the shorter period fluctuations in natural wind (turbulence) but is sufficiently short to be normally regarded as representing a period of near-constant background mean wind” (Harper et al., 2010; EN 1991-1-4).

Budapest (E 19.1, N 47.5°) is selected as a representative location to illustrate the effect of parameter estimation and model selection uncertainty. It is characterized by a single type of wind phenomena – strong wind generated by thunderstorms. Thus unimodal distributions (no mixed models) seem to suffice for statistical modelling.

2. Statistical Tools

2.1. CONSIDERED DISTRIBUTIONS

Extreme value theory offers two popular approaches to analyze extremes: block maxima and peak over threshold methods. These are applied to extract maxima from wind speed observations and select asymptotic distribution functions (Coles, 2001; Reiss and Thomas, 2007):

- If observations are divided into blocks and largest values are selected from each, the block maxima distribution asymptotically approaches the generalized extreme value family (GEV) under rather permissive conditions. This method is commonly referred to as block maxima method.
- If values over a selected threshold are considered, the distribution of these values asymptotically approaches the generalized Pareto family (GP) under rather permissive conditions (Coles, 2001; Reiss and Thomas, 2007).

For the block maxima approach with one year block size, two- (LN2) and three-parameter lognormal (LN3), and Gumbel distributions are used in addition to GEV. Gumbel is a special case of the GEV distribution with shape parameter converging to zero, and is commonly used in engineering to model extremes (Table I). Two-parameter lognormal distribution is found to describe well snow extremes, it is mainly applied in the US (ASCE, 2010). The LN3 distribution was successfully applied and propagated in the Czech Republic to model various types of extremes (Holický and Sýkora, 2015 and 2016). For block sizes different from one year, the GEV model is considered only. For multiple maxima in a block analyses, a multivariate generalized extreme value distribution (rGEV) is applied, considering the dependence between maxima in a block (Coles, 2001). In the peak over threshold approach, the GP distribution is used. The applied parametrization of these distributions are given in Table II. The parametrization has no effect on frequentist inference, however marginally affects the Bayesian estimates as priors generally differ for alternative parameterization of the model and flat (uniform) prior on scale and location parameters likely correspond to non-flat priors of moment characteristics.

Table II. Summary of considered distribution types, their application range, and parametrization.

Distribution	Application in this study	Parametrization	Reference
Gumbel	block method; single maxima; 1-year block size	scale and location	(Coles, 2001)
GEV	block method; single maxima; arbitrary block size	shape, scale, and location	(Coles, 2001)
rGEV	block method; multiple maxima; 1-year block size	shape, scale, and location	(Coles, 2001)
LN2	block method; single maxima; 1-year block size	shape and scale	(Singh, 1998)
LN3	block method; single maxima; 1-year block size	shape, scale, and threshold	(Singh, 1998)
GP	peak over threshold, threshold as a study parameter	shape, scale, and threshold	(Coles, 2001)

2.2. STATISTICAL INFERENCE

Frequentist and Bayesian statistical paradigms are selected to fit models and to quantify statistical uncertainties. Bayesian statistics treats parameters as random variables and assigns probability distribution to them. The latter is convenient when the inferred parameters are inputs and the full representation of their uncertainty is needed for instance in probabilistic reliability and risk analyses. Additionally it can handle complex problems with messy data and can combine information from different sources. These advantages distinguish it from the commonly used frequentist statistics that focuses on data variability given a parameter value (Spiegelhalter and Rice, 2009).

The main instrument of Bayesian statistics is Bayes' rule which incorporates the information conveyed by the data and prior knowledge through the likelihood function and prior distribution of parameters, respectively. The distribution of parameters obtained in this why is termed a posterior distribution. When future observations are to be predicted – as is typical in structural reliability studies – the posterior predictive distribution then serves this purpose by averaging over the posterior distribution of parameters (Aitchison and Dunsmore, 1980). Parameter estimation uncertainty can be expressed by providing the whole posterior distribution or its credible intervals (Gelman et al., 2003). Model selection uncertainty can be handled through Bayesian model averaging where weighted average over candidate models is calculated. A weight (b_i) is the probability of the i^{th} model given the data relative to the summed probability of all considered models (Hoeting et al., 1999). The weights favor parsimony (Occam's razor), i.e. penalizes model complexity. Thereby overfitting can be avoided and models with different complexity can be compared. In the Bayesian analyses, flat (uniform) priors are used for all the parameters in Table II with practically infinite range; multiple parameter settings are verified to confirm convergence.

In the frequentist paradigm, the maximum likelihood method is used to obtain point estimates. The uncertainty intervals – termed often as confidence intervals – are obtained by the delta method (Coles, 2001) and bootstrapping (Efron et al., 1994). A resampled empirical distribution function is created using linear interpolation among points within the data range.

In general Bayesian inference requires substantially more computational power than frequentist approaches. Yet these calculations can be carried out on personal computers. The computational burden is even larger for structural reliability applications where uncertain tails of distributions are of crucial importance and their estimation requires much larger simulation numbers. Taking advantage of the small dimension (less or equal to three) of probabilistic models in this study, direct numerical integration is used for Bayesian inference.

3. Results of Analysis

3.1. Block maxima

3.1.1. Annual maxima

Initially representative fractiles for different distributions are calculated and compared. The selected extremes correspond to 50, 100, 500, and 1000-year return period events. The 50-year event is the characteristic value of meteorological actions in Eurocode while the other are intended to indicate the design point coordinate of a structure subject to increasingly dominant wind action or with increasing structural reliability. They are comparable to representative basic wind speeds of ASCE 7-10 that associates 300, 700, and 1700-year return period values with risk categories I, II and III-IV, respectively.

The point estimates of selected extremes with 90% uncertainty intervals are shown in Figure 1. Maximum likelihood point estimates (solid circle) and 90% confidence intervals are obtained by delta method (dashed) and bootstrapping (solid). The bootstrap point estimate is the mean of the sample with symmetrical intervals.

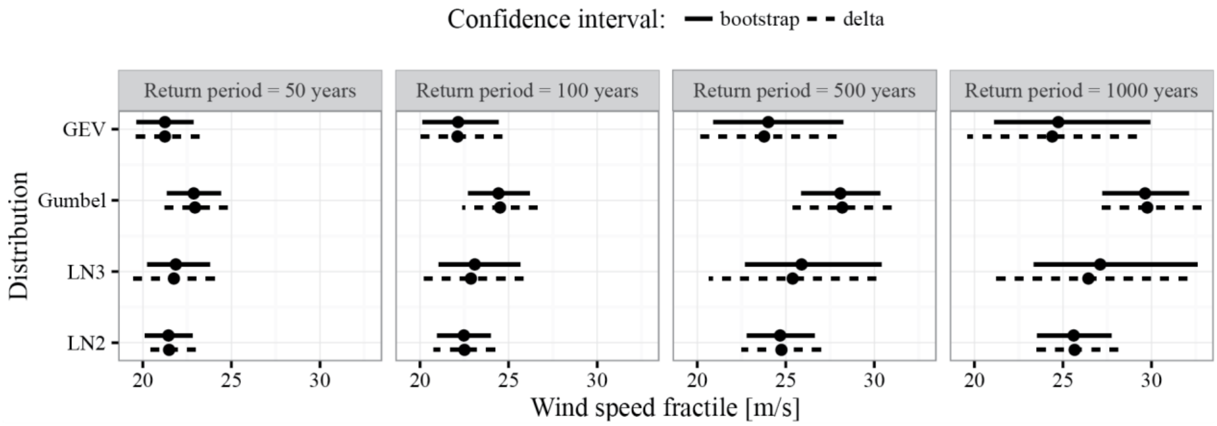

Figure 1. Summary of representative wind extremes with 50, 100, 500, and 1000 years return period for various distributions.

The inference is conducted with one year block size using maximum likelihood method and the confidence intervals are estimated using delta method and bootstrapping. Although the 50-year extremes differ subtly only, the difference is not negligible. For instance the GEV and LN3 distributions yield to 7% lower extreme values than that of Gumbel. The range of 90% confidence intervals is 15% of the associated point estimate.

The bootstrap and delta method confidence intervals match well. However, for larger return periods, their difference progresses particularly in respect of the lower interval endpoints: the delta method tends to provide more conservative estimates. This is partially due to the symmetric nature of delta method based confidence intervals. Confidence intervals are rapidly increasing with an increasing return period. The relative coverage of the 90% confidence intervals to point estimates is 35-40% for the three-parameter distributions; however, they are essentially the same: 15-20% for two-parameter ones even at 1000-year return period level. The difference among distributions in point estimates of 1000-year extremes reaches up to 18%. The results show that both distribution selection and parameter estimation uncertainty have substantial bearing on representative extreme values.

The return period-return level Gumbel plots enable visualization of the effects of distribution selection and parameter estimation uncertainty. Return level plots with confidence intervals are shown in Figure 2. In the figure the maximum likelihood point estimates (white solid line) are accompanied by 90% confidence intervals (gray) obtained by the delta method[1]. The point estimate and confidence interval of the characteristic value v_k along with the number of extreme observations are also displayed on each plot.

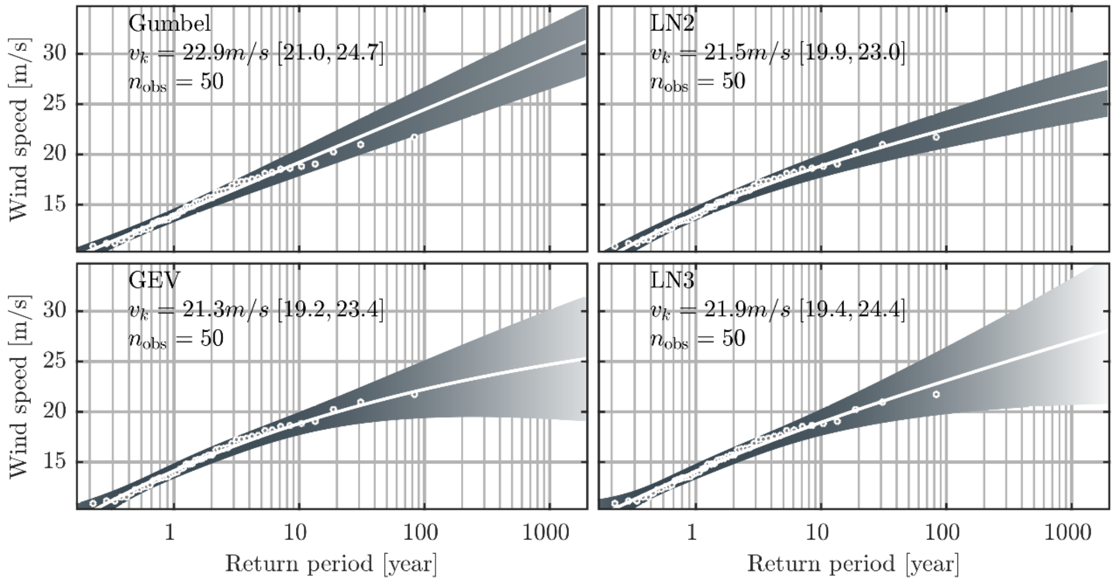

Figure 2. Return level plots of annual wind maxima for the selected distributions.

Although the difference between bootstrap and delta method based confidence intervals can be considerable (Figure 1), the latter is used only in further analysis for simplification. Estimates based on the delta method are deemed to be sufficiently accurate for this investigation. The confidence intervals are substantially widening as the cumulative distribution functions approach the regions with few data and regions of extrapolation. The difference between the models is remarkable, particularly the narrow confidence interval of the Gumbel distribution fails to include the largest observation.

Statistical tests and information theory based goodness-of-fit measures do not clearly support or reject any of the considered distributions even though the tail of the Gumbel model noticeably deviates from the observations. The other distributions – particularly GEV and LN2 – seem to capture the upper tail better. However, this observation needs to be confirmed by analysis of data from more stations.

However for the three-parameter models this is partially attributable to their greater flexibility that is traded-off by increasing parameter estimation uncertainty, i.e. they have wider uncertainty interval than the two-parameter distributions. For two-parameter models, available information allows better model identification.

[1] The interval coloring is 'ink-preserving', i.e. the same 'amount of ink' is used for every vertical section, hence creating a linear transition from the narrowest (dark gray) to the widest interval (white). In a particular 2×2 or 3×2 figure, equal vertical ranges have the same color on each subplot, thus the models are directly comparable based on coloring as well.

The independence of wind speed maxima separated by about four days can be reasonably assumed (Simiu and Heckert, 1996). The focus on annual maxima may thus lead to the loss of information as extreme values generated by less severe storms are discarded, though they can be informative. This additional information can be incorporated by (*i*) reducing block size; (*ii*) considering multiple maxima in a block; and (*iii*) using values above a selected threshold. The first two belongs to the block maxima approach and converge to GEV or GEV-like (rGEV) distributions. The third procedure is commonly referred to as the peak over threshold method and its asymptotic distribution family is the GP distribution (section 3.2). Another approach to reduce sampling variability is to combine wind speed measurements from several stations with similar climatic conditions (Holmes, 1998).

3.1.2. Block maxima with various block sizes

Initially the effect of block size is examined. There is no universal rule to select an optimal block size. Typically the balance between bias and variance is searched. Smaller block increases sample size and thus reduces statistical uncertainties (variance) while non-extreme values might contaminate the sample and introduce bias to the inference.

Figure 3 compares GEV distributions fitted to block maxima with a block size of 1, 1/2, 1/4 and 1/10 of year. The parameters are inferred using the maximum likelihood method. The point estimates (white solid line) are accompanied by 90% confidence intervals (gray) obtained by the delta method. The point estimate and confidence interval of the characteristic value v_k along with the number of extreme observations are also displayed on each plot.

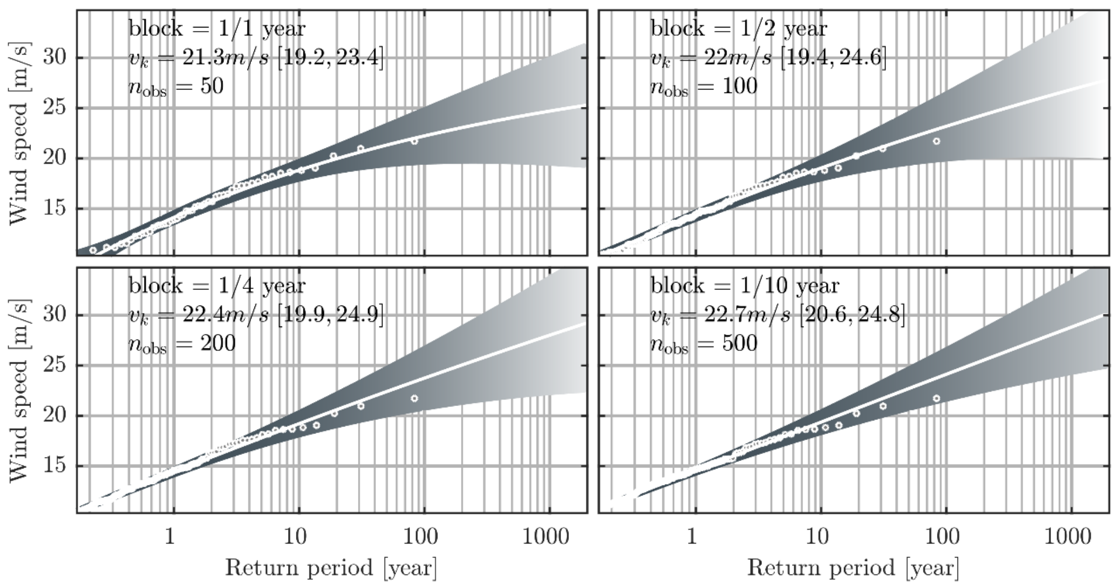

Figure 3. Return level plots of wind maxima with different block sizes (Gumbel plot).

With decreasing block size, smaller observations dominate the fit, yet the 90% confidence intervals still covers all the large observations at the right tail. It is interesting to observe that ten times more observations for the block size of 1/10 of year have no important effect on the confidence intervals of larger upper fractiles.

It seems that the one year blocks capture all the largest values and the enlargement of the sample does not reduce parameter estimation uncertainty in this case. The difference in characteristic value point estimates is subtle, at maximum 6% for 1/10 year block size compared to the one year model. However, the difference between these two models for the extremes with 1000-year return period progresses to 20%. A notable effect of block size reduction is the straightening of the point estimate of the distribution (white solid line), i.e. the shape parameter is approaching zero and the distribution approaches Gumbel distribution.

3.1.3. Multiple maxima in a block

Another technique aimed to increase a sample size considers multiple maxima in a block. This leads to an asymptotic multivariate rGEV distribution. It takes into account the dependence between largest values within the same block, i.e. the second largest must be smaller than the largest. Figure 4 summarizes the results of distribution fitting with number of largest values from 1 to 120 per block. The parameters are inferred using the maximum likelihood method. The point estimates (white solid line) are accompanied by 90% confidence intervals (gray) obtained by the delta method. Also shown in the figure are the point estimates and confidence intervals of the characteristic value v_k along with the number of extreme observations.

Though the number of observations is increased by 120-times, no effect on trends nor on bias in the extreme value-return period plots is observed. However, parameter estimation uncertainty appears to be significantly reduced; the uncertainty intervals for characteristic values reduce by 50% when a number of maxima in a block increases from 1 to 120. The reduction is more pronounced for extremes with longer return periods, e.g. it is 70% for 1000-year return period. The results show that valuable additional information can be extracted from the observations and the uncertainty intervals can be significantly reduced by increasing the number of maxima in a block.

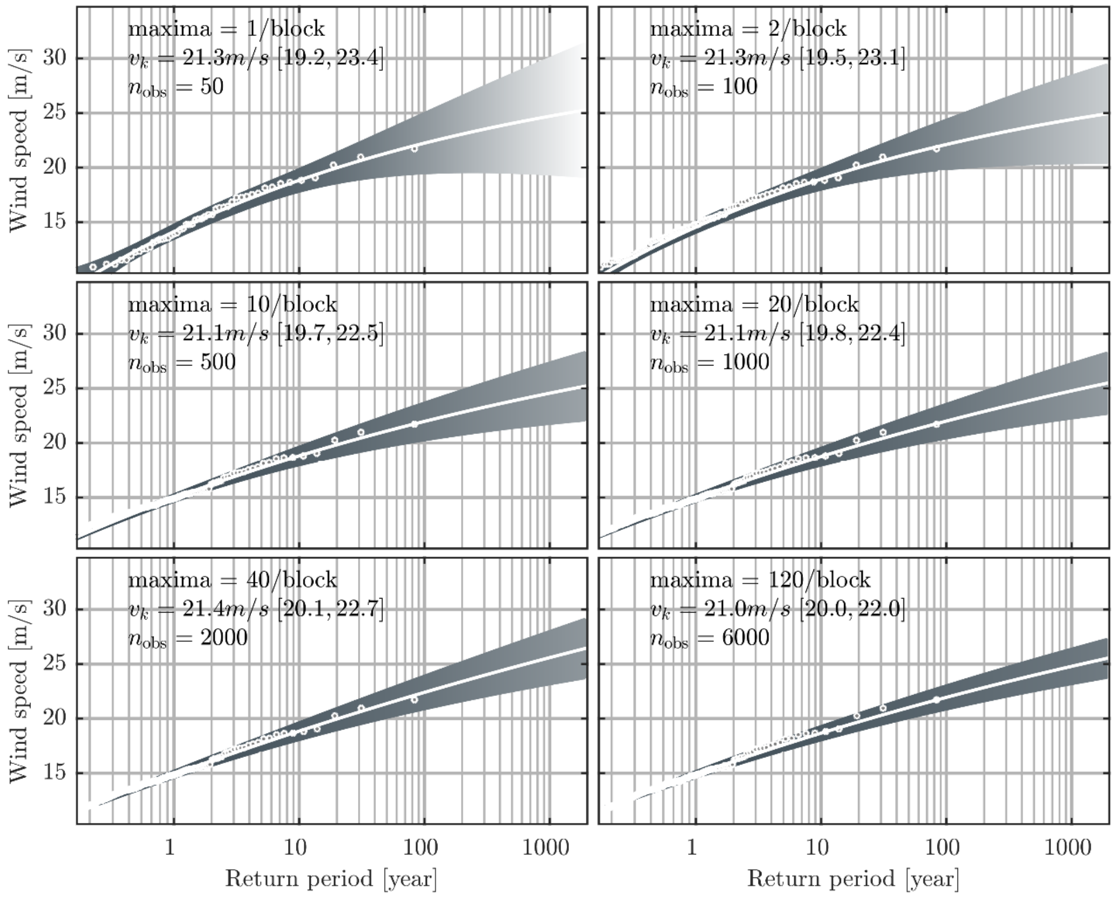

Figure 4. Return level plots wind maxima for various number of maxima in a block; rGEV is used for each plot.

3.2. PEAK OVER THRESHOLD

In this section the peak over threshold method is used to extract extreme wind speeds for Budapest. Since it has different asymptotic distribution than the block maxima, the return value plot transformation is also different, though it has similar characteristics such as logarithmic horizontal scale and linear cumulative distribution if the shape parameter is approaching zero. The threshold selection is complex due to the trade-off between bias and variance as is similar with the block size and multiple largest value selection for block maxima. Again lower threshold allows more observations, however they might not be representative for extremes and can introduce bias.

Mean residual plot and parametric analyses (Coles, 2001) are utilized to assess the stability (variance and bias) of parameter estimates by changing the threshold level. These diagnostic plots suggest an appropriate threshold in the range from 5 to 15 m/s. Return value plots with six different thresholds from this region are presented in Figure 5. The thresholds are chosen to allow comparison with multiple maxima in a block (Figure 4). The GP distribution supported by the extreme value theory is considered for the analyses only. The parameters are inferred using the maximum likelihood method. The point estimates (white solid line) are accompanied by 90% confidence intervals (gray) obtained by the delta method. The point estimate and confidence interval of the characteristic value v_k along with the number of extreme observations are also displayed on each plot.

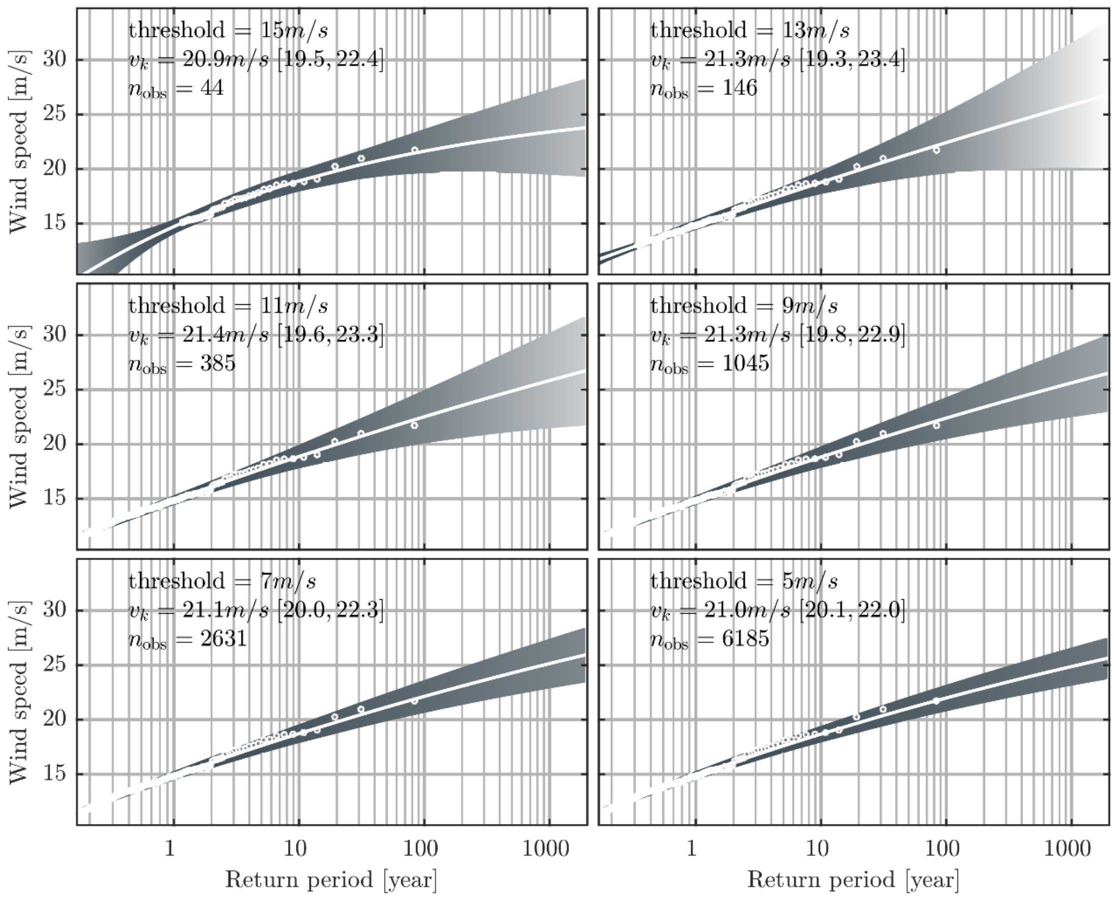

Figure 5. Return level plots wind maxima over various thresholds; GP is used for each plot.

The results show that lower thresholds lead to narrower uncertainty intervals of the fractiles. However, this observation is difficult to generalize as for instance the uncertainty intervals for a threshold of 13 m/s are wider than those for 15 m/s. This discrepancy is attributable to the shift of a shape parameter from negative to positive values. The latter corresponds to a distribution without an upper bound which substantially increases the effect of parameter estimation uncertainty on larger fractiles. For smaller thresholds, this is gradually counterbalanced by the information conveyed by the additional data. Looking at the point estimates of characteristic values, only small differences (up to 2%) are found, thus this representative fractile is sufficiently stable in the investigated range of thresholds. The associated 90% confidence intervals are reduced considerably – by about 50% when the threshold decreases from 15 m/s to 5 m/s. This reduction is the results of a 150-times increased sample size. For larger return periods, the difference of point estimates is larger, it can reach 15% for a 1000-year return period compared to the 15 m/s threshold model. In respect of bias, the 5 m/s threshold model still seems to be appropriate as all observations are within the confidence interval and the cumulative distribution function is rather stable. Note that these statements apply for the dataset under consideration, other locations might have more consistently varying shape parameter or larger bias at lower thresholds, thus the gain in variance reduction might not be attainable.

Point estimates and confidence intervals obtained for the peak over threshold and multiple maxima block method are in good agreement. The outcomes indicate that parameter estimation uncertainty can be significantly reduced by using the peak over threshold method and incorporation of a great amount of additional data seems not to introduce considerable bias into the model.

3.3. EFFECT OF PARAMETER ESTIMATION UNCERTAINTY

Uncertainty intervals provide only a visual insight into the effect of parameter estimation uncertainty. This is valuable, but cannot be directly used in probabilistic reliability and risk analyses where all uncertainties need to be taken into account. Therefore, it is more useful to infer the probabilistic distribution of model parameters and capture related uncertainties that can be further propagated and integrated in reliability analysis. This can be readily achieved by Bayesian analysis that treats parameters as random variables. Instead of giving the whole distribution function of uncertain parameters, the parameter estimation uncertainty can be directly incorporated into the distribution of the variable of interest (such as wind speed) by its posterior predictive distribution function. This approach is taken here in conjunction with popular analysis focused on annual maxima. The maximum likelihood, posterior mean, and posterior predictive distributions for Gumbel, LN2, LN3 and GEV are plotted in Figure 6. The posterior predictive is unique for each distribution type.

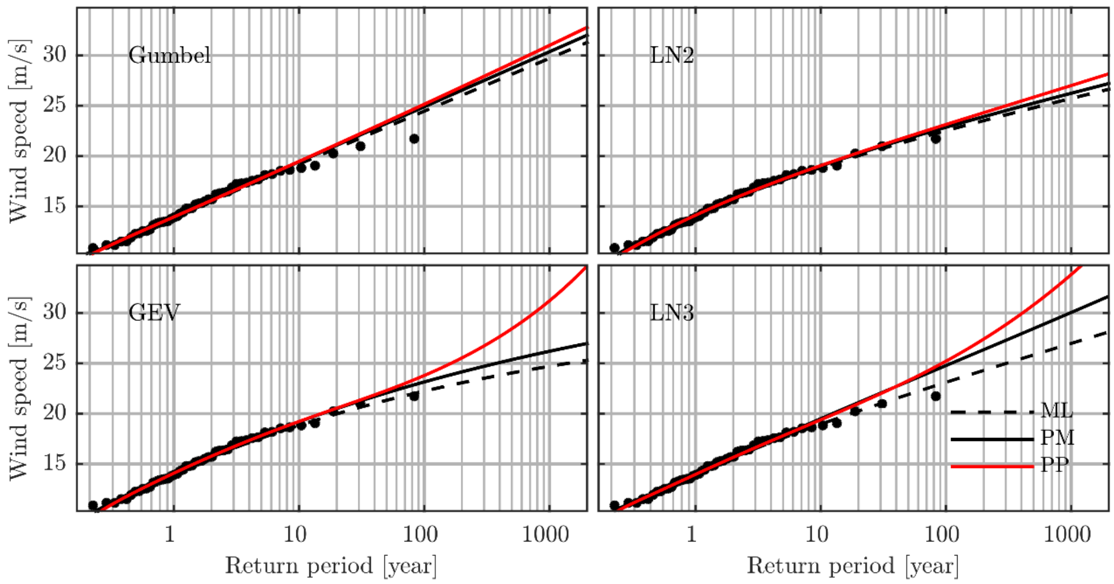

Figure 6. Return level plots of annual maxima with maximum likelihood (dashed black, ML), posterior mean (solid black, PM), and posterior predictive (solid red, PP) distributions.

For the two-parameter models, the effect of parameter estimation uncertainty is negligible even for large return periods (less than 3%). Likewise the effect on characteristic values is negligible for all the models. However, the difference between posterior mean and posterior predictive 1000-year extremes is 12% and 19% for LN3 and GEV, respectively. The results show that parameter estimation uncertainty has significant bearing on larger return period extremes for three-parameter distributions and non-conservative estimates are obtained when its effect is neglected. The plots also show the considerable difference in frequentist and Bayesian point estimates that is a result of the skewed posterior distribution of the fractiles.

3.4. Effect of model selection uncertainty

Although parameter estimation uncertainty seems to have substantial effect on large return period extremes for three-parameter distributions, this uncertainty might not be of concern when the data can be described by Gumbel or other two-parameter models with sufficient confidence. This section examines this by calculating Bayes weights for each distributions under consideration. The weights express the relative goodness of models; i.e. the relative probability that a distribution is an appropriate underlying model for the observed variable from the considered pool of models.

The calculated weights are 0.22, 0.30, 0.25, and 0.23 for Gumbel, LN2, GEV, and LN3 distributions, respectively. Hence, no distribution is clearly favored. Note that the weights and model averaging is conditional on the pool of candidate models. Figure 7 shows that Gumbel (max. 7%) and LN3 (max. 8%) distributions overestimate the model averaged one while LN2 (max. 8%) and GEV (max. 6%) distributions provide lower estimates. Both types of statistical uncertainty can be incorporated by averaging over the posterior predictive distributions of selected distributions. Note that the model averaged posterior mean is same for each distribution type and associated Bayes weights are indicated in Figure 7.

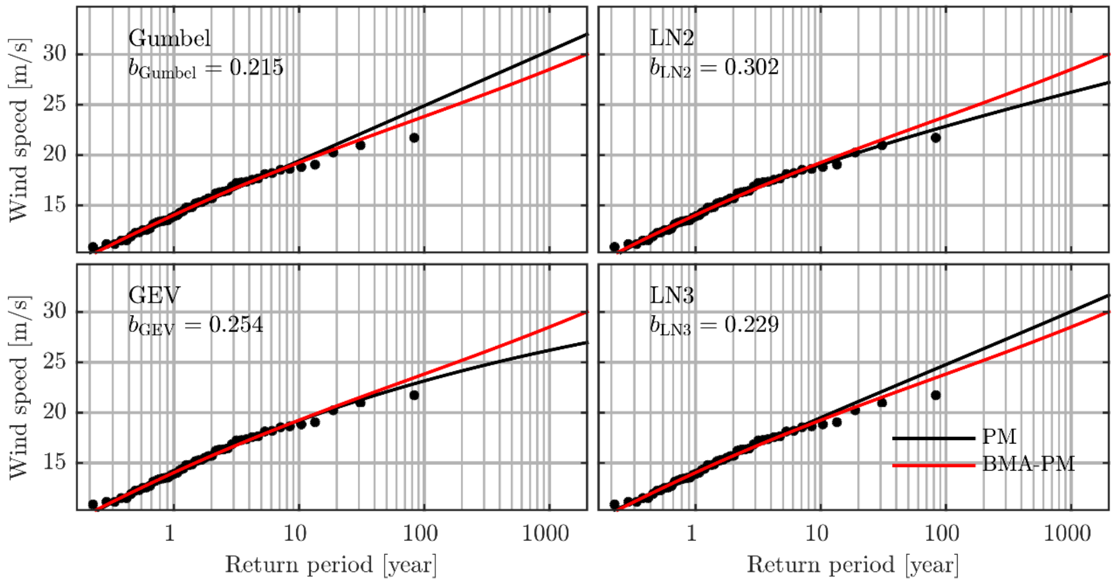

Figure 7. Return level plots of annual wind maxima with Bayesian posterior mean (solid black, PM) and model averaged posterior mean (solid red, BMA-PM) distributions.

4. Conclusions

This contribution compares the current practice of extreme wind speed modelling in civil engineering with more advanced statistical techniques. The former neglects while the latter is capable of quantifying statistical uncertainties. Analysis of extreme wind speeds for the representative location of Budapest reveals that:

- Considering the annual maxima approach, the parameter estimation uncertainty has negligible effect on extremes with long return periods for two-parameter distributions, but it is substantial for three-parameter models. Underestimation can reach up to 20% and the current design practice seems to be non-conservative.
- The popular Gumbel distribution yields about 20% larger 1000-year return period values than those based on the generalized extreme value distribution. The latter fits better to data, though unambiguous recommendation concerning distribution selection cannot be provided due to a limited amount of data.
- Gumbel confidence intervals seem to be deceptively narrow and the largest observation is outside of a 90% uncertainty interval. The generalized extreme value and generalized Pareto distributions imply that the wind maxima have an upper bound, however the 90% confidence intervals overlap with the unbounded region as well.
- If viable, smaller than one year block size, multiple maxima in a block or peak over threshold methods are recommended to increase sample size and reduce statistical uncertainties. This leads to 70% reduction in the range of a 90% confidence interval for 1000-year extremes for the selected location.
- Bayesian posterior predictive distribution is recommended for accounting parameter estimation uncertainty. Furthermore, Bayesian model averaging can be used to account for model selection uncertainty.

The effect of statistical uncertainties on extreme values mainly depends on the available information for probabilistic models. This is particularly important for projecting environmental loads with few observations in a long-term perspective. The recommended procedure can be utilized for other extreme climatic actions such as flooding and snow loads.

Acknowledgements

This work was supported by the Ministry of Education, Youth and Sports of the Czech Republic under Grants LG14012 and LD15037. The numerical analyses are completed using Matlab (Matlab, 2015) and R (R Core Team, 2015), the work and commitment of the developers of these applications are highly appreciated. All codes/ scripts, processed data, and results can be obtained from the authors.

References

ASCE. Minimum Design Loads for Buildings and Other Structures. ASCE/SEI 7-10. Reston, Virginia: American Society of Civil Engineers; 2010.

Aitchison J. and I. R. Dunsmore. *Statistical Prediction Analysis*. New York, USA: Cambridge University Press, 1980.

CEN. Eurocode 1: Actions on structures. Part 1-4: General actions - Wind actions, 2005.

Coles, S. *An Introduction to Statistical Modeling of Extreme Values*. Springer-Verlag, London, 2001.

Coles, S. and L. Pericchi. Anticipating catastrophes through extreme value modelling. *Journal of the Royal Statistical Society. Series C (Applied Statistics)*, 52(4):405–416, 2003.

Efron, B. and R. Tibshirani. *An Introduction to the Bootstrap*. Chapman and Hall/CRC, 1994.

Gelman, A., J. Carlin, H. Stern and D. Rubin. *Bayesian Data Analysis*. CRC Press, USA, 2nd edition, 2003.

Harper B.A., J. D. Kepert and J. D. Ginger. Guidelines for converting between various wind averaging periods in tropical cyclone conditions. World Meteorological Organization. WMO/TD-No. 1555, 2010.

Hoeting, J. A., D. Madigan, A. E. Raftery and C. T. Volinsky. Bayesian model averaging: A tutorial. *Statistical Science*, 14(4):382–417, 1999.

Holický, M. and M. Sýkora. Probabilistic Models for Climatic Actions. In *Proc. 13th International Probabilistic Workshop (IPW 2015)*, pages 281-290, 2015.

Holický, M. and M . Sýkora. Probabilistic Models for Wind Actions. In *Proc. SMRLO'16 The Second International Symposium on Stochastic Models in Reliability Engineering, Life Sciences and Operations Management*, pages 172-175, 2016.

Holmes, J. D. Wind loading of structures: application of probabilistic methods. *Progress in Structural Engineering and Materials*, 1(2):193–199, 1998.

Kruger, A. C. Wind climatology of South Africa relevant to the design of the built environment. (Ph.D.). Stellenbosch University, South Africa, 2010.

Li, D.-Q., X.-S. Tang and K.-K. Phoon. Bootstrap method for characterizing the effect of uncertainty in shear strength parameters on slope reliability. *Reliability Engineering & System Safety*, 140(0):99–106, 2015.

Matlab. version 7.10.0 (R2015a). The MathWorks Inc. Natick, Massachusetts, 2015.

Niemann, H.-J. and S. Diburg. Statistics of extreme climatic actions based on the Gumbel probability distributions with an upper limit. *Computers & Structures*, 126:193–198, 2013.

R Core Team. R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna, Austria, 2015.

Reiss, R.-D. and M. Thomas. *Statistical Analysis of Extreme Values with Applications to Insurance, Finance, Hydrology and Other Fields*. Birkhäuser Basel, 3rd edition, 2007.

Rózsás, Á. and , M. Sýkora. Effect of statistical uncertainties in ground snow load on structural reliability. In *IABSE Conference: Structural Engineering: Providing Solutions to Global Challenges*, pages 220–227, 2015a.

Rózsás, Á. and M. Sýkora. Model comparison and quantification of statistical uncertainties for annual maxima of ground snow loads. In Podofillini, L., Sudret, B., Stojadinovic, B., Enrico, Z., and Kröger, W., editors, *Safety and Reliability of Complex Engineered Systems: ESREL 2015*, p. 2667–2674. CRC Press, 2015b.

Rózsás, Á. and M. Sýkora. Neglect of parameter estimation uncertainty can significantly overestimate structural reliability. Transactions of the VŠB – Technical University of Ostrava, Civil Engineering Series, 15(2):10, 2015c.

Sanpaolesi, L., and S. Colombini. Implementation of Eurocodes. Handbook 3. Action effects for buildings. Chapter IV - Thermal actions on buildings. Aachen, Germany, 2005.

Sanpaolesi, L., D. Currie, P. Sims, C. Sacré, U. Stiefel, S. Lozza, B. Eiselt, R. Peckham, G. Solomos, , I. Holand, R. Sandvik, M. Gränzer, G. König, D. Sukhov, , R. Del Corsoand and P. Formichi. Scientific support activity in the field of structural stability of civil engineering works: Snow loads. Final report phase I. Report, Commission of the European Communities. DGIII-D3, 1998.

Simiu, E. and N. Heckert. Extreme wind distribution tails: A 'peaks over threshold' approach. Journal of Structural Engineering, 122(5):539–547, 1996.

Singh, V. *Entropy-Based Parameter Estimation in Hydrology*. Water Science and Technology Library. Springer Netherlands, 1998.

Sisson, S. A., L. R. Pericchi and S. G. Coles. A case for a reassessment of the risks of extreme hydrological hazards in the Caribbean. *Stochastic Environmental Research and Risk Assessment*, 20(4):296–306, 2006.

Spiegelhalter, D. and K. Rice. Bayesian statistics. *Scholarpedia*, 4(8):5230, 2009.

Szalai, S., I. Auer, J. Hiebl, J. Milkovich, T. Radim, P. Stepanek, P. Zahradnicek, Z. Bihari, M. Lakatos, T. Szentimrey, D. Limanowka, P. Kilar, S. Cheval, G. Deak, D. Mihic, I. Antolovic, V. Mihajlovic, P. Nejedlik, P. Stastny, K. Mikulova, I. Nabyvanets, O. Skyryk, S. Krakovskaya, J. Vogt, T. Antofie and J. Spinoni. Climate of the greater Carpathian Region. Report, 2013.