Discussion of:

Graf, P. et al. "Adaptive stratified importance sampling: hybridization of extrapolation and importance sampling Monte Carlo methods for estimation of wind turbine extreme loads" [DOI:10.5194/wes-2017-30]

Lance Manuel
lmanuel@mail.utexas.edu

August 26, 2017

Overview and General Comments

This is a most interesting presentation of a vexing problem that has proven to be a challenge to wind turbine loads analysts for many years. The ideas developed by the authors and the narrative discussing the desire to "bridge" more conventional extrapolation methods and variance reduction techniques that go beyond brute-force Monte Carlo simulations are welcome. Casting the problem as an optimization problem, albeit without the usual formalisms, so as to adaptively improve estimates of long-term loads is done most effectively. Throughout, there are interesting insights and discussions that make for an illuminating reading and exposure to the essential issues.

Specific Section-by-Section Comments

Introduction

The description about IFORM on Page 2, Lines 12–13 should be clarified. More correctly, only with the environmental contour (EC) method which is the most commonly employed version of IFORM, one uncouples the "environment" from the "response," and global extremes of interest associated with a target return period are approximated by using the maximum response from among all response levels derived only for candidate environmental variables consistent with that return period (in other words, response variability given environmental conditions is neglected). That said, in cases such as the one described in this article, where the environment is described using only one random variable (wind speed), the EC method has limited use. The EC method is better suited when a pair of random variables (say, wind speed and turbulence intensity, or wind speed and wave height for offshore turbines) are included. When only one random variable defines the environment, the "environmental contour" for a target return period is reduced to a single wind speed. This has limited use because, for instance, in the present example, for both side-to-side and fore-aft tower bending moments, the method will suggest that only rare and high wind speeds around cut-out that are associated with a 50-year return period need to be considered in turbine aeroelastic simulations. This will clearly lead to inaccurate 50-year fore-aft tower moments.

Now, despite the preceding comment that the EC method reduces the conventional environmental contour to a degenerate point or single-valued wind speed to consider for turbine response simulations, in fact, in the present case, one could instead use IFORM in its more general form and use wind speed and response as two random variables and formally derive estimates of 50-year side-to-side and fore-aft tower bending moments. This is discussed later, along with comments offered in the context of Section 2.4 (IFORM).

Extrapolation

Page 4, Line 14: Strictly speaking, the 50-year return period event is that event that is exceeded "on average" once in 50 years. Even though it is not the same in general, sometimes the event is defined as one that is exceeded "on average" with a probability of 1/50 in one year.

Page 4, Line 26: 3.8^{-7} should be 3.8×10^{-7}.

Monte Carlo importance sampling for extreme loads

Page 6, Line 23: The comment that some form of accept-reject sampling can be used with importance sampling is an intriguing one. It is unclear how exactly this would be done given that $Y(x)$ is not known in closed form; any additional notes, even if included very briefly, regarding such sampling would help.

IFORM

As stated earlier, the EC method doesn't apply here as there is *no* environmental contour corresponding to the authors' example—such a contour is a degenerate single wind speed value obtained as $\mathrm{F}_V^{-1}(3.8 \times 10^{-7})$ and as such has limited value for, say, the fore-aft tower bending moment where the derived 50-year load will certainly be under-predicted. Indeed, in this single environmental random variable case, the degenerate single simulation needed for *any* response or load of interest would require simulations to be run for a single wind speed above cut-out, i.e., for V equal to 43.3 m/s. This would be meaningless.

The authors correctly point to the deficiencies of IFORM (on Page 7) but, in light of comments in the preceding paragraph, since there is no environmental contour at all that can be defined to describe their example study, much of the extended discussion regarding the EC method and environmental contours as presented in Section 2.4 is not relevant.

Now, in a most interesting way, the very issue that doesn't allow for a critique of the EC method—namely, that the authors choose only V as an environmental random variable—actually allows the more general IFORM procedure to be used with the authors' own simulation results and will lead to reasonable results (how this can be achieved is presented here very briefly). The idea is as follows: Consider that there are two random variables—wind speed, V, and the response or load, Y, whose statistics are derived from 10-min simulations . We will assume that we know the probability distribution for V (for instance, here, the authors use a Weibull V with shape and scale parameters equal to 2 and 11.28 m/s, respectively); we establish conditional distributions for Y given V based on simulations. We can use IFORM, though *not* the EC method, to find the required quantiles of $Y|V$ for any V of interest. This is purely a geometry problem (involving mapping of V and Y to two independent standard normal random variables). To illustrate this, because the results presented in Figure 1 are the easiest to read off and learn from without great effort, one would find using IFORM that for TwrBsMxt, the 24 m/s bin would require that the desired 50-year response must have a probability of exceedance in 10 minutes of 5.95×10^{-6}. Given the 5th, 25th, 50th, 75th, and 95th percentile loads in Figure 1, a 2-parameter Weibull fit to these data leads to a 50-year TwrBsMxt value of 29,700 kN-m. Other (lower) wind speeds occur more often and associated load levels to be checked for those wind speeds using the IFORM procedure must be rarer, i.e., with exceedance probabilities in 10 minutes that are smaller than 5.95×10^{-6}. Given the data, these wind speeds do not lead to TwrBsMxt values at the desired probability levels that exceed what was found for V equal to 24 m/s. In a similar manner, for TwrBsMyt, selecting V equal to 16 m/s, the desired 50-year response for IFORM must have a probability of exceedance in 10 minutes of 7.20×10^{-7}. Again, from the data in Figure 1, a 2-parameter Weibull fit to these data read off easily, leads to a 50-year TwrBsMyt value of 94,300 kN-m. Again, note that other wind speeds and associated (different) response quantiles need to be checked, as part of the IFORM procedure, to ensure that the largest load quantile across all the wind speed bins is then claimed as the 50-year load. Details regarding all the calculations are not presented here but IFORM computations are based on the Weibull V and loads data from Figure 1.

In sum, the authors' comment regarding searching on the environmental contour (or just inside it) is not pertinent here. There is also no need to discuss above-median response levels in this context. Both the preceding comments would have been appropriate if, in addition to wind speed, another environmental random variable were included such as turbulence intensity. As illustrated, IFORM in its general form can be easily employed here and response variability can be directly accounted for—as the authors state correctly, this variability is ignored by the EC method. It is *not* ignored by IFORM in general and, as shown in the previous paragraph, even with the limited data presented in Figure 1, the method works quite efficiently in deriving 50-year loads. By running additional simulations at critical wind speeds, the resulting loads data and subsequent distribution fits to the same will lead to reduced uncertainty in derived 50-year loads.

Adaptive stratified importance sampling (ASIS)

Reference to POE, on Page 8, Line 20, should really be to the probability of non-exceedance.

In Step 2 of the algorithm on Page 10, why not simply obtain new samples (with rounding) in proportion to $\Delta_N J(N)$? It is not clear but it seems that the most important bin (where $\Delta_N J(N)$ is largest) is allocated some samples and the remaining (of 20) are then randomly allocated to other bins. Why randomly?

The algorithm, as presented, appears not to state what is the criterion for stopping or convergence. The discussion regarding the "umbrella" concept that suggests a minimal superset of sample distributions across bins is exactly what is needed. It is the only way to guarantee adequate samples of response extremes to meet very distinct response characteristics such as between TwrBsMxt and TxrBsMyt. The ASIS algorithm as presented doesn't explicitly state this but assumes convergence when all the response measures are adequately sampled in all bins so as to yield unbiased 50-year response values, presumably with some specified confidence level on these predictions. If it helps, in an offshore wind turbine application, Sultania and Manuel [3], employed bootstrap-based confidence intervals for specific sea states (akin to bins here) to arrive at the appropriate number of simulations for accuracy in response probabilty distributions, conditional on the environment. Given the results presented in this article, it appears that convergence on long-term loads for each bin is indeed achieved by the authors in their examples by examining the variance of loads associated with low POE levels.

It was not clear, upon reading, what was the reason for using the 5 largest loads. The largest loads will automatically drive the tail of long-term loads distributions and, as such, the 50-year load, when these largest loads are included along with *all* smaller loads; so, why retain only the 5 largest? Some clarification would help here.

Results

The results in Figures 3–5 are very interesting and suggest that the methodology proposed by the authors offers a robust and efficient means of deriving long-term turbine loads for design. The rapid reduction in variance in estimates of long-term loads by using the ASIS algorithm is convincing.

A few observations are offered.
1. It would be very useful and insightful to see how $g(x)$ or the bin-wise importance sampling changes with iteration for the the two load measures, TwrBsMxt and TwrBsMyt, separately, and what the ultimate umbrella sampling ends up being, after convergence, or as the number of iterations changes—given the contrasting characteristics of the two load measures, one might expect a bi-modal sampling with one mode around or slightly above rated and another closer to cut-out will result.
2. It is not completely clear what M_{pks} refers to. For instance, does $M_{pks} = 40$ imply that, after accounting for samples from all bins, only the largest 40 are used or does it mean that for each bin, the largest 40 loads are retained and then combined with 40 from the other bins before extrapolation?
3. The non-monotonic upward trending cooefficient of variation on the 0.05 POE load seen in Figure 3 for a very large number of peaks might be caused by dependence among the peaks that results when too many peaks are extracted from each 10-min simulation (see Fogle et al [2]).
4. Is it possible that the x axis units in Figure 4 are 10,000's of KN-m. rather than 1,000's of kN-m as stated? This would be consistent with Figures 1 and 2 which appear to show TwrBsMxt loads an order of magnitude higher. It would also be consistent with the IFORM-based estimates that were computed (above) as 29.7×10^3 kN-m and 94.3×10^3 kN-m, respectively, for TwrBsMxt and TwrBsMyt. The caption for Figure 5 is correct; the one for Figure 4 might need to be corrected.
5. As presented, Figures 4 and 5 do not show loads for the low POE level of 3.8×10^{-7} associated with the target 50-year return period. It would have been useful to see those results. Indeed, ASIS-based convergence at the lowest POE levels that are presented, suggest that 25 iterations

and the use of 40 peaks is very good.

In the caption for Figure 3, 5^{-2} should be 5×10^{-2}.

Conclusions

The closing discussion is most helpful in setting this work in the context of other studies regarding the derivation of long-term loads for wind turbine design. The ASIS algorithm, as presented, will prove useful to loads analysts. While the authors have demonstrated its efectiveness using 100 independent tests in their numerical studies, in practice, it may be useful to suggest use of the ASIS algorithm followed by bootstrapping, after each iteration, and confirmation that the coefficient of variation on some load quantile (as in Figure 3, for example) is acceptably low—for example, 5%. It is clear that ASIS can achieve this target with a far smaller number of aeroleastic simulations than extrapolation based on ordinary or conventional sampling methods. The authors do, in fact, recognize in their closing discussions, that bootstrapping with a single data set could be used with ASIS and thus reduce computational effort.

Only to provide a contrast with the authors' work, it should be noted that IFORM-based approaches, referred to in this article, can be very efficient. This has been demonstrated in this discussion using the authors' own data (crudely derived from their Figure 1); note that in the illustration presented response variability is *not* ignored as is done with the environmental contour method. Accounting for response variability in IFORM is not difficult—it is conceivable that an ASIS-like formulation for sampling could prove far more efficient than even IFORM but this needs to be demonstrated for situations where more than one random variable defines the "environment."

Finally, related to this last point, the ASIS algorithm would be especially important to employ along with the introduction of stochastic turbulence and with the treatment of turbulence intensity or turbulence standard deviation explicitly as a random variable. The role of gusts and turbulence in extreme loads is known to be significant and the need then for bivariate importance sampling distributions could present challenges (Bos et al. [1], van Eijk et al. [4]); at the same time, the benefits to be derived from approaches such as ASIS, if its efficiency and accuracy is demonstrated in such cases, would then add to its appeal.

On Page 15, Line 3, "stratified adaptive" should be "adaptive stratified" to be consistent with the ASIS acronym.

References

[1] R. Bos, W. Bierbooms, and G. van Bussel. Importance sampling of severe wind gusts. *11th EAWE PhD Seminar on Wind Energy in Europe*, 2015.

[2] J. Fogle, P. Agarwal, and L. Manuel. Towards an improved understanding of statistical extrapolation for wind turbine extreme loads. *Wind Energy*, 11(6):613–635, 2008.

[3] A. Sultania and L. Manuel. Reliability analysis for a spar-supported floating offshore wind turbine. *Wind Engineering, (DOI: 10.1177/0309524X17723206)*, 2017.

[4] S. F. van Eijk, R. Bos, and W. A. A. M. Bierbooms. The risks of extreme load extrapolation. *Wind Energy Science*, 2(2):377–386, 2017.