Wind Energ. Sci. Discuss., https://doi.org/10.5194/wes-2017-30
Manuscript under review for journal Wind Energ. Sci.
Discussion started: 25 July 2017

Adaptive stratified importance sampling: hybridization of extrapolation and importance sampling Monte Carlo methods for estimation of wind turbine extreme loads

Peter Graf[1], Katherine Dykes[1], Rick Damiani[1], Jason Jonkman[1], and Paul Veers[1]
National Renewable Energy Laboratory, Golden, CO, 80401 - USA

Correspondence to: Peter Graf (peter.graf@nrel.gov)

Abstract. Wind turbine extreme loads estimation is especially difficult because turbulent inflow drives nonlinear turbine physics and control strategies, so there can be huge differences in turbine response to essentially equivalent environmental conditions. The two main current approaches, extrapolation and Monte Carlo sampling, are both unsatisfying: extrapolation-based methods are dangerous because by definition they make predictions outside the range of available data, but Monte Carlo methods converge too slowly to routinely reach the desired 50-year return period estimates. Thus a search for a better method is warranted. Here we introduce an adaptive stratified importance sampling approach that allows for treating the choice of environmental conditions at which to run simulations as a stochastic optimization problem that minimizes the variance of unbiased estimates of extreme loads. Furthermore, the framework, built on the traditional "bin"-based approach used in extrapolation methods, provides a close connection between sampling and extrapolation, and thus allows the solution of the stochastic optimization (i.e., the optimal distribution of simulations in different wind speed bins) to guide and recalibrate the extrapolation. Results show that indeed this is a promising approach, as the variance of both the Monte Carlo and extrapolation estimates are reduced quickly by the adaptive procedure. We conclude, however, that due to the extreme response variability of turbine loads to the same environmental conditions, our method and any similar method quickly reaches its fundamental limits, and that therefore our efforts going forward are best spent elucidating the underlying causes of the response variability.

1 Introduction

Estimating extreme loads for wind turbines is made especially difficult by the nonlinear nature of the wind turbine physics combined with the stochastic nature of the wind resources driving the system. Extreme loads, such as those experienced when a strong gust passes through the rotor or when a turbine has to shut down for a grid emergency, can drive the design of the machine in terms of the material needed to withstand the events. The material requirements in turn drive wind turbine costs and overall wind plant cost of energy. Thus, accurate modeling and simulation of extreme loads is crucial in the wind turbine design process. This paper discusses the use of adaptive importance sampling in estimation of such loads. Importance sampling (IS) (Robert and Casella, 2004) is a well established method for using samples from one distribution to estimate statistics from another. Adaptivity in importance sampling has been introduced in (Karamchandani et al., 1989) and elsewhere, but does not appear to have broadly taken hold. Here we introduce an adaptive importance sampling method for extreme loads estimation.

Wind Energ. Sci. Discuss., https://doi.org/10.5194/wes-2017-30
Manuscript under review for journal Wind Energ. Sci.
Discussion started: 25 July 2017

The essential task in wind turbine extreme loads estimation is to evaluate the probability of exceedance (POE) integral

$$P(Y > l) = \int P(Y > l|x) f(x) dx. \tag{1}$$

Here $P(Y > l)$ is the probability of load Y exceeding target/threshold l, $P(Y > l|x)$ is the conditional probability of exceedance given wind speed x, and $f(x)$ is the distribution of wind speeds (or other environmental conditions). Because we are interested in extremely low probability events and $Y(x)$ is stochastic and only available via simulation, standard methods of integration do not apply.

Existing approaches fall generally into three main classes. The first is based on extrapolation: data is gathered in different wind speed “bins", extreme value distributions are fit to the empirical distribution function for each bin, and these are then integrated. The second is based on Monte Carlo methods: exceedance probabilities are written as expectations of indicator functions, samples are drawn from an assumed wind distribution, and unbiased estimates made by the usual Monte Carlo summation. A highly efficient third approach, the Inverse First Order Reliability Method (IFORM) (Winterstein et al., 1993), estimates global extremes by modeling their variation from median extremes. (We will discuss IFORM briefly below, but it is not the main focus of this paper.) Unfortunately, to date none of these methods is satisfactory. The crux of the difficulty is that on the one hand too many samples are required for converged *Monte Carlo* estimates, but on the other hand reliable *extrapolation* of nonlinear physics under uncertain forcing is extremely problematic, especially without knowledge of the form (e.g. quadratic, etc.) of the nonlinearity. Nevertheless, the computational expense of MC implies that except in rare cases, some sort of extrapolation will be necessary in order to reach the desired 50 year return period estimates. Perhaps we can at least use MC/IS to make sure extrapolations are accurate to the resolution of data we actually have, and to gather data in ways that accelerates their convergence.

Our goal is to therefore to develop methods that make unbiased estimates that *minimize variance* as a function of the number of samples/simulations, and to use these to dynamically update extrapolations. Our proposed method, Adaptive Stratified Importance Sampling (ASIS), is essentially a global stochastic optimization method where the search variables are the number of samples from each wind speed bin we use, and the objective function is the variance of our Monte Carlo estimates. The key tool here is importance sampling, which allows us to continually produce unbiased estimates of exceedance probabilities even as the distribution of bins changes. These quasi-optimal samples are then used to make the best possible extrapolations. Results below show that this is indeed a promising approach.

The organization of this paper is as follows. First we present the necessary background on the existing extrapolation method (e.g., as recommended in the IEC standard (IEC, 2005)), Monte Carlo methods, and importance sampling. Next we describe our ASIS algorithm. Then we present a brief study illustrating some of its properties. The paper provides a context for discussing extrapolation and importance sampling in the same framework. Our conclusion highlights the potential for this approach, reiterates some of the fundamental difficulties with the endeavor, and leads to suggestions for where we should next focus our efforts to crack this difficult problem.

Wind Energ. Sci. Discuss., https://doi.org/10.5194/wes-2017-30
Manuscript under review for journal Wind Energ. Sci.
Discussion started: 25 July 2017

2 Background on turbine simulation, extrapolation, Monte Carlo, and importance sampling

2.1 Turbine simulation

Throughout this paper, we use FAST, NREL's aeroelastic simulation tool (Jonkman, 2013). FAST is a widely-used industry and academic tool for wind turbine loads estimation. NREL's WISDEM software allows for the execution of FAST and its companion tool TurbSim (which generates turbulent wind fields for input to FAST) in a programmatic fashion from `python`, as has been reported previously (Graf et al., 2016).

The particular turbine on which we are testing these methods is the NREL 5 MW reference turbine, often used for such studies (Jonkman et al., 2015; Choe et al., 2016), in an onshore configuration. The "environmental conditions" are thus described by hub height mean wind speed (modeled by a Weibull distribution with scale and shape parameters of 11.28 m/s and 2, respectively). Additional environmental parameters of turbulence intensity, spectrum, coherence function, and wind shear are kept fixed at nominal values. It should be noted that the stochasticity in the combined TurbSim/FAST simulation comes from the "random seed" that is input to TurbSim. This seed governs the exact starting conditions for the generation of the turbulent inflow wind field. Because this is an ergodic system, a long enough simulation would eventually cover all possible wind conditions. However, for finite simulations, multiple runs using different random seeds allow us to "sample" the space of all possible turbulent flow field snapshots with the same mean wind speed, turbulence intensity, etc. This is common practice in extreme loads analysis. We can regard random seed as a proxy for sampling over a uniform distribution of turbulent inflows for each set of environmental conditions.

For this study, we selected two output channels of interest, tower base side-side bending ("TwrBsMxt" in FAST nomenclature) and tower base fore-aft bending ("TwrBsMyt"), which provide contrast because the wind speeds where their highest loads occur overlap differently with the typical wind speed distribution. The side-side moments grow with hub-height wind speed, making their extremes hard to estimate with traditional MC sampling because they do not overlap well with typical wind distributions. The fore-aft counterparts do overlap quite closely with the typical wind distributions.

2.2 Extrapolation

The current standard for estimating extreme loads relies on extrapolation. We refer the reader to the relevant literature for a detailed exposition of the extrapolation method (Ragan and Manuel, 2008; Moriarty, 2008; Toft et al., 2011; Graf et al., 2017). Here we present a concise statement of the method and discuss one or two subtleties. The protocol to construct exceedance curves using binning and extrapolation is as follows:

1. Run Turbsim/FAST N_i times per wind speed x_i at center of bin i (Typically $N_i \sim 6$).

2. For each bin i, concatenate the data from each seed and extract peaks (see below r.e. peak extraction and time scales). For future reference we refer to the resulting dataset as $\{Y_{i,k}\}$ where i indexes over wind speed bins and k indexes over the peaks we have extracted at that wind speed

3. For each bin i, form empirical cumulative distribution functions (CDFs) and fit a chosen distribution $F_i(x) = P(Y < l|x)$ to them. In this paper we use a 3 parameter Wiebull distribution.

4. (optional) For each bin: convert each fitted distribution to desired time scale (see below r.e. peak extraction and time scales).

5. Finally, $P(Y < l) = \int P(Y < l|x)f(x)dx \sim \sum_i P(Y < l|x_i)f(x_i)\Delta x_i$, where $f(x)$ is pdf of wind distribution, and Δx_i is the width of bin i. The extrapolated estimate of the probability of exceedance is then $P(Y > l) = 1 - P(Y < l)$.

The distributions chosen to fit the bin-wise CDFs are the theoretically appropriate extreme value distributions (Generalized Extreme Value, 3 parameter Weibull, etc.). However, this does not mean they accurately represent the behavior of the particular FAST loads in a specific context. The optimality properties of extreme value distributions are asymptotic properties, but we are doing “intermediate asymptotics": long term–but not infinitely long term–trends. Nevertheless, these distributions are the appropriate starting point.

It is important to be clear regarding various time spans at play here. First, there is the ultimate time of interest, typically in wind studies the “50 year return period" (note this does not mean that in 50 years the event in question happens with probability 1; typically it is *defined* as an event having probability $1/50$ of happening in one year). Next there is the simulation time, i.e., the length of each FAST run. This is almost always 10 minutes in the literature. Relatedly, there is the time span over which our estimates of exceedance probability apply. These are also traditionally 10 minutes, i.e. reported probabilities of exceedance are probabilities of exceedance *in 10 minutes*, but there is nothing in principle to make this fixed. Finally, there is the length of time between independent peaks. This time can be estimated empirically by examining the autocorrelation of the data; values as low as 4 seconds have been justified in previous studies (Ragan and Manuel, 2008), and 10 seconds seems to be more than adequate.

The various time spans come into play as we extract peaks and make estimates. The rule that connects them is the simple “AND" rule of probability: If $Y < l$ for time T, and $T = Kt$, then $Y < l$ for time t for K times in a row, so $P_T(Y < l) = P_t(Y < l)^K$. For probabilities of exceedance $P(Y > l)$, we write $P(Y > l) = 1 - P(Y < l)$ and use the same idea, resulting in the familiar expression $P_T(Y > l) = 1 - (1 - P_t(Y > l))^K$. For example, the former is typically used to convert peak distributions collected on a 10 second, 1 minute, or peak-over-threshold basis to a 10 minute basis, while the latter is used when we derive the ubiquitous value of 3.8^{-7} that represents the probability of the 50 year return period event happening in 10 minutes.

In this paper we adopt the following setup: All our simulations are 11 minutes long, where we discard the first minute as a transient and retain the final 10 minutes for our studies (we will occasionally be loose with the terminology and refer to these as “10 minute simulations"). To gather peaks, we take the maximum of each 1 minute segment in our simulations. This provides exactly 10 peaks per simulation, which allows for building 1-minute empirical cumulative distribution functions (i.e. probability of exceedance in 1 minute) in a consistent manner. The resulting 1 minute empirical POEs are converted to 10 minute POEs as described above (i.e. $K = 10$) to conform to standard practice. Below we experiment with how many of these peaks should be used in the fitting of the extreme value distributions used in the extrapolation method. This alternative to the peak-over-threshold method makes comparison and correspondence between extrapolation and sampling easier, because we

Wind Energ. Sci. Discuss., https://doi.org/10.5194/wes-2017-30
Manuscript under review for journal Wind Energ. Sci.
Discussion started: 25 July 2017

eawe european academy of wind energy
Open Access WIND ENERGY SCIENCE DISCUSSIONS

have a fixed number of peaks per FAST run. We have divided the wind speed range into 5 bins centered at 8, 12, 16, 20, and 24 m/s.

To acquire a sense of the basic variability of the response, we have run 20 independent sequences of the simulations described above for the extrapolation method (6 random seeds per bin). Figure 1 consists of box and whiskers plots of the peaks from the first 6000 peaks (10 peaks per run X 6 seeds X 20 repetitions X 5 bins). For both the side-side and fore-aft loads, the variability within each bin is large. For example, the difference between 95th and 5th quantile for the 24 m/s side-side bin is comparable to its median value. For the side-side load the dependence on wind speed is clearly also very strong, whereas for the fore-aft load, the variability *within* the bins is as large as it is *between* the bins.

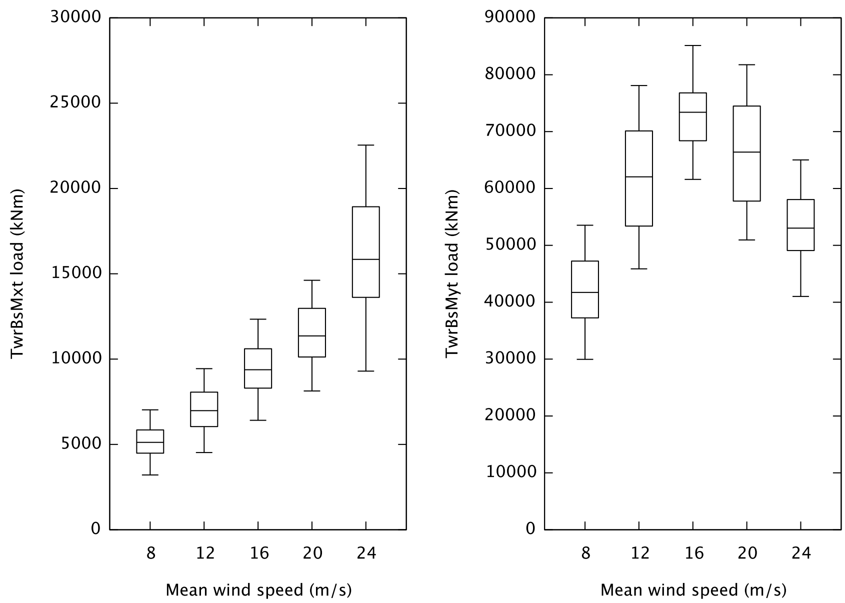

Figure 1. Box and whiskers plots of the distribution of raw response (specifically, all 1-minute maxima) of the combined TurbSim/FAST simulation as a function of wind speed bin for side-side (left) and fore-aft (right) loads. The difference in general trend between side-side and fore-aft loads is clearly evident. The variability within each bin is extreme (esp. for the fore-aft load), which puts an upper bound on the utility of sampling methods (e.g., importance sampling) targeting certain wind speeds. The boxes show the median, 25th, and 75th percentile. The whiskers are positioned at the 5th and 95th percentile. The data are the absolute maxima in 1 minute segments of 120 separate 10 minute simulations per bin (1200 total peaks for each bin).

2.3 Monte Carlo importance sampling for extreme loads

Monte Carlo (MC) methods are widely used to estimate expectations of quantities calculated by stochastic simulations (Robert and Casella, 2004). Importance Sampling (IS) is an MC method in which an auxiliary *importance distribution* $q(x)$ is used to focus sampling on areas of the target distribution $f(x)$ that are most relevant with respect to the functions of interest (e.g.,

Wind Energ. Sci. Discuss., https://doi.org/10.5194/wes-2017-30
Manuscript under review for journal Wind Energ. Sci.
Discussion started: 25 July 2017

the loads $Y(x)$). Often "relevant" means *minimal variance* (see below). The broad applicability of the method arises from the so-called Importance Sampling Identity:

$$E_f[Y(x)] = \int Y(x)f(x)dx \quad (2)$$

$$= \int Y(x)\frac{f(x)}{q(x)}q(x)dx \quad (3)$$

$$= E_q[Y(x)\frac{f(x)}{q(x)}] \quad (4)$$

where E_f and E_q represent the expectation with respect to f and q, respectively. This means, from an MC standpoint, that both

$$E_f[Y(x)] \sim \frac{1}{M_{tot}}\sum_i^{M_{tot}} Y(x_i) \qquad \text{with } x_i \text{ drawn from } f, \quad (5)$$

and

$$E_q[Y(x)\frac{f(x_i)}{q(x_i)}] \sim \frac{1}{M_{tot}}\sum_i^{M_{tot}} Y(x_i)\frac{f(x_i)}{q(x_i)} \qquad \text{with } x_i \text{ drawn from } q, \quad (6)$$

(where M_{tot} is the total number of samples we have of the quantity we are estimating) are *unbiased* estimates of the same quantity $E_f[Y(x)]$.

For us M_{tot} will be the number of peaks gathered from the FAST runs. In what follows we will use M to represent numbers of peaks, and N to represent numbers of FAST runs. With our convention of taking the maximum over 1 minute spans of 10 minute simulations, we will have $M = 10N$ throughout. The subscript "$_{tot}$" will be the total number (over all bins), whereas index subscripts (e.g. "$_i$"), etc., will refers to the peaks or runs within the corresponding bin. Thus as above N_i is the number of FAST runs (i.e. random seeds) in the ith bin, and $M_i = 10N_i$ is the number of peaks extracted from the N_i runs, and N_{tot} is the total number of FAST runs.

Although the above estimates (5) and (6) are both unbiased, they could have drastically different *variance*, which means that they may converge at drastically different rates. The minimal variance importance distribution can be derived and is

$$q^*(x) = \frac{Y(x)f(x)}{E_f[Y(x)]} \quad (7)$$

From a practical standpoint there are two obvious problems with this result. First, it depends on the expectation we wanted to calculate in the first place. This objection we can overcome using some form of accept-reject sampling that does not depend on the normalization constant. A more significant problem, however, is that q^* also depends on the function $Y(x)$ whose expectation we are trying to calculate with as few evaluations as possible. Finally, in our case, an even worse problem is that $Y(x)$ is not a deterministic function of x, so, as stated, q^* is also a stochastic function. Nevertheless, Eqn. (7) provides a guide for our quest to find the minimal variance importance distribution: it should be as close as possible to proportional to the product of load $Y(x)$ and wind probability $f(x)$.

Wind Energ. Sci. Discuss., https://doi.org/10.5194/wes-2017-30
Manuscript under review for journal Wind Energ. Sci.
Discussion started: 25 July 2017

For our purposes, finally, note that Eqn. (1) can be written as an expectation of the so-called “indicator" function that is 1 if $Y(x) > l$ and 0 otherwise:

$$P(Y > l) = E_f[I(Y > l)] \sim \frac{1}{M_{tot}} \sum_i I(Y(x_i) > l) \frac{f(x_i)}{q(x_i)} \qquad \text{with } x_i \text{ drawn from } q. \tag{8}$$

Eqns. (1) and (8) form the basis of the mathematical bridge between extrapolation and MC/IS methods described in section 3.

2.4 IFORM

Here we summarize the Inverse First Order Reliability Method (IFORM) in the present context. IFORM was introduced by Winterstein (Winterstein et al., 1993) and addresses the estimation of extreme loads from a different perspective. Instead of directly computing the integral in Eqn. (1), IFORM seeks to find just the conditions that *cause* the extreme load. IFORM works best if the response of interest is a completely deterministic function of environmental conditions that themselves have known probability. Then one can directly search the *environmental contour* (e.g. all wind/wave/turbulence/etc. combinations that have probability 1-in-50 years of occurring) to find the highest load. IFORM introduces the important notion of *response variability*, which refers to the magnitude of the variation of the nonlinear stochastic response for *fixed* environmental conditions. This notion helps to explain why it is problematic to rely on IFORM for the purposes of wind turbine extreme loads estimation. The approaches to handling response variability in IFORM generally involve modeling or estimating the variation of the response away from its "median extreme" value (e.g. median of all the maximum values, here over each 1 minute window, over several simulations). But in our case the response variability (for example, plotted directly in Figure 1) is simply too large to be credibly modeled as "median plus an approximation" (the true/maximal extreme is too far from the median extreme). Also, the strategy of starting from the desired extreme (e.g. 50 year return period) environmental contour, or just inside it, is not feasible here. Such wind speeds, for example, are not even in the operating range of the turbine; the wind speed where the extreme response will occur is in fact not that improbable. IFORM may not work well for wind turbine extreme loads because the 50-year loads may not be associated with the 50-year environmental contour. The present task, that of estimating extreme loads with wind speed as the only environmental variable, is governed mostly by the response variation. Therefore we will not estimate extreme loads by IFORM in this paper. However, IFORM is critical for conceptual understanding: where possible, our goal should be to "convert" response variation (intractable) to environmental variation (tractable) through better understanding of its physical cause. It is the extreme response variability (different random seeds for the same environmental conditions can cause very different FAST output) that makes extreme loads estimation a difficult problem. We return to this subject in section 5.

3 Adaptive stratified importance sampling (ASIS)

3.1 Bin-wise empirical CDFs as the bridge between extrapolation and MC

When we perform the bin-wise simulations used in the extrapolation methods, we are performing *stratified* sampling. Recognizing that these samples can be described as a probability distribution provides a bridge to using them in an IS context, as

Wind Energ. Sci. Discuss., https://doi.org/10.5194/wes-2017-30
Manuscript under review for journal Wind Energ. Sci.
Discussion started: 25 July 2017

discussed in (Graf et al., 2017). The basic idea is as follows. Assume a set of samples $\{x_i\}$ from an arbitrary distribution $g(x)$ of wind speeds (g may be either $f(x)$, $q(x)$, or an empirical one derived from binning the data). By running FAST, the set $\{Y_i\}$ of corresponding loads can be generated and sorted from the lowest to the highest. Then for any given load Y_j, the IS estimate is:

$$P(Y < Y_j) = E_f[I(Y < Y_j)] \tag{9}$$

$$= \int I(Y(x) < Y_j) f(x) dx \tag{10}$$

$$= \int I(Y(x) < Y_j) \frac{f(x)}{g(x)} g(x) dx \tag{11}$$

$$\sim \frac{1}{M_{tot}} \sum_i I(Y(x) < Y_j) \frac{f(x_i)}{g(x_i)} \quad \text{with } x_i \text{ drawn from } g \tag{12}$$

$$= \frac{1}{M_{tot}} \sum_i \begin{cases} \frac{f(x_i)}{g(x_i)} & \text{if } i < j \\ 0 & \text{otherwise} \end{cases} \tag{13}$$

$$= \frac{1}{M_{tot}} \sum_{i<j} \frac{f(x_i)}{g(x_i)}. \tag{14}$$

Letting $w_i \equiv \frac{f(x_i)}{g(x_i)}$, we can concisely write the empirical CDF for *all* the loads $\{Y_j\}$ as:

$$P(Y < Y_j) = \frac{1}{M_{tot}} \sum_{i<j} w_i. \tag{15}$$

To apply this formula to data from the binning method, we need the appropriate $g(x)$, i.e. the appropriate weights w_i of each sample. Here we denote the dataset as $\{Y_{i,k}\}$ where i indexes over wind speed bins and k indexes over the peaks we have extracted at that wind speed. We rewrite the integral over wind speeds as a sum of integrals, and then approximate each separate integral by the M_i peaks derived from the N_i runs (in our case, over TurbSim random seeds) at the fixed wind speed x_i for bin i:

$$\int I(Y(x) < Y_j) f(x) dx = \sum_{i=1}^{N_{bins}} \int_{x_i}^{x_{i+1}} I(Y(x) < Y_j) f(x_i) dx \tag{16}$$

$$\sim \sum_{i=1}^{N_{bins}} \frac{1}{M_i} \sum_{k=1}^{M_i} \begin{cases} f(x_i)\Delta x_i & \text{if } Y_{i,k} < Y_j \\ 0 & \text{otherwise} \end{cases} \tag{17}$$

From this, we see that the "weight" contributed by sample i,k to the POE for Y_j is $\frac{1}{M_i} f(x_i)\Delta x_i$ for all i,k s.t. $Y_{i,k} < Y_j$. To calculate the empirical CDF from the bin data, then, we assign weight $w_i = \frac{1}{M_i} f(x_i)\Delta x_i$ to all samples from bin i and apply Equation (15). Table 1 summarizes the exact correspondence between extrapolation and IS/MC. The importance density corresponding to stratified sampling is seen to be $g(x_i) = \frac{M_i}{M_{tot}} \frac{1}{\Delta x_i} = \frac{N_i}{N_{tot}} \frac{1}{\Delta x_i}$.

Thus we have a "bridge" between fitting and sampling. Bin-wise empirical CDFs can be directly compared with fitted distributions (in fact, they are what we fit to). But the estimate over all wind speeds can then be expressed generically in the

Wind Energ. Sci. Discuss., https://doi.org/10.5194/wes-2017-30
Manuscript under review for journal Wind Energ. Sci.
Discussion started: 25 July 2017

Table 1. Bin-wise empirical cumulative distribution functions provide a bridge from extrapolation, which builds POE from bin-wise fitted distributions F_i, and importance sampling, which builds an unbiased estimate from the appropriately defined distribution g.

Method	$P(Y < l) = \int P(Y < l\|x) f(x) dx$	Remarks
Empirical bin-wise CDF	$\sum_i \frac{1}{N_i} \sum_k I(Y_{i,k} < l) f(x_i) \Delta x_i$	$P(Y < l\|x_i) \sim \frac{1}{N_i} \sum_k I(Y_{i,k} < l)$
Extrapolation	$\sum F_i(l) f(x_i) \Delta x_i$	$P(Y < l\|x_i) \sim F_i(l)$, fitted to above
Importance Sampling	$\frac{1}{M_{tot}} \sum_{i,k} I(Y_{i,k} < l) \frac{f(x_i)}{g(x_i)}$	sampling from $g(x_i) = \frac{N_i}{N_{tot}} \frac{1}{\Delta x_i}$

IS language suited to comparison with MC and IS estimates. IS/MC do not provide any “bin-wise" information (there are no bins), so it is otherwise impossible to “debug" their divergence from extrapolation. This formulation allows us to see, first, that error accrues from lack of convergence of empirical CDFs, for both methods. Additionally though, for extrapolation, the error is compounded by lack of fit between the chosen extreme value distribution and the empirical CDFs, which is the price we pay for being able to extrapolate to arbitrarily low POEs.

3.2 ASIS as stochastic optimization

The discussion above indicates that the samples from the bin based methods can alternatively be used to make empirical estimates via their implied importance distributions. This orientation suggests, also, that there is no barrier to changing the distribution of samples as we go. So we can think of the estimation procedure as an optimization problem: find the distribution of bins (number of samples per bin) that results in unbiased estimates with minimal variance. In (Graf et al., 2017) we have used a heuristic algorithm that looked for “gaps" in the empirical peak distribution. Here instead we introduce a gradient-based approach. As above, let N_i be the number of FAST runs performed in the ith wind speed bin, and let N be the *vector* of bin-counts. Now, the variance of the estimate using importance distribution g is

$$J(N) \equiv var_g[P(Y > l)] = E_g[\frac{I(Y > l)^2 f(x)^2}{g(x)^2}] - E_g[\frac{I(Y > l) f(x)}{g(x)}]^2. \tag{18}$$

Note the second term does not depend on g (in the corresponding integral, the g in the denominator cancels out). So

$$\frac{\partial J}{\partial N_j} = \frac{\partial}{\partial N_j} E_g[\frac{I(Y > l)^2 f(x)^2}{g(x)^2}] \tag{19}$$

$$\sim \frac{\partial}{\partial N_j} \sum_{i,k} I(Y_{i,k} > l)^2 f(x_i)^2 \frac{N_{tot}^2 \Delta x^2}{N_i^2} \tag{20}$$

$$= -2 \sum_{i,k} I(Y_{i,k} > l)^2 f(x_i)^2 \frac{N_{tot}^2 \Delta x^2}{N_i^3}. \tag{21}$$

Here we have used the fact that $\frac{M_i}{M_{tot}} = \frac{N_i}{N_{tot}}$ to write the expression in terms of bin counts instead of peak counts.

Our algorithm begins by running the standard 6 seeds per bin from the extrapolation method (i.e., N_i is initialzed to 6 for all i). Then we perform the following steps in an iterative fashion:

Wind Energ. Sci. Discuss., https://doi.org/10.5194/wes-2017-30
Manuscript under review for journal Wind Energ. Sci.
Discussion started: 25 July 2017

1. Compute $\nabla_N J(N)$.
2. Allocate a target number of new samples (e.g. 20 per iteration) to bins in two ways:
   (a) allocate some percentage of the new samples in proportion to $\nabla_N J$,
   (b) recognizing this is a *global* optimization problem, allocate the rest to other bins randomly.
3. Run TurbSim/FAST for the new batch.
4. Update our empirical estimates of POEs *and* our extrapolation estimates.

Note the algorithm as stated does not explicitly recognize the stochasticity of the underlying quantity $Y(x)$. Because we are using an unbiased estimate of the gradient of the variance, our approach, naive as it is, is known to converge "almost surely" to a local optimum (Robbins and Monro, 1951). Now that we have cast the problem in this form, though, we can take advantage of ongoing research in this area.

Also, the N_i need to be integers, which is accomplished by rounding. The resulting error is likely subsumed into the general convergence of the stochastic optimization procedure.

Next, as stated J only concerns one load type (e.g. tower base side-side bending moment). But previously (Graf et al., 2017) and above (i.e. Fig.1) we have seen that side-side and fore-aft loads favor different importance distributions. Since it defeats the purpose of the method to have to repeat it for every load, we have adopted an "umbrella" concept; we compute the desired bin distribution for all the loads of interest, and form the minimal superset of bins that includes them all.

Another issue, even for a single load type, is how many of its peaks Y_i are considered in computing the gradient of the variance. In fact we have to decide some number of "large peaks" we will use to evolve N. The results here use the 5 largest loads. This is an important issue to keep in mind and one of several that makes the application of IS theory tricky in practice. (E.g., (Choe et al., 2015) have derived an optimal importance distribution, but it is "peak-dependent", i.e. only strictly optimal for estimating the exceedance probability of the particular load value for which it was derived.)

4 Results

In this section we demonstrate the basic mechanics of the algorithm in the context of a study of the effect of the number of peaks, M_{pks}, used to fit the bin-wise extrapolation distributions ($F_i \sim P(Y < l|x_i)$, above). Figure 2 illustrates the variability of the extrapolated exceedance probability as a function of M_{pks}. Here we have fit the extreme value (Weibull 3 parameter) distribution to the peaks from the wind speed bin centered at 20 m/s for different values of M_{pks}. Based on the data, we need a way of saying what the "correct" number of peaks to use is. For this, we study the variance of the resulting extrapolation, which we can estimate simply by repeating the entire sampling, simulating, fitting, and extrapolation procedure 100 times. The results are summarized in Figures 3, 4, and 5.

For each of the 100 independent tests, we ran extrapolation and ASIS for 25 iterations (an iteration of extrapolation is simply re-performing the extrapolation procedure with the current ASIS bin data), which adds a varying number of new

Wind Energ. Sci. Discuss., https://doi.org/10.5194/wes-2017-30
Manuscript under review for journal Wind Energ. Sci.
Discussion started: 25 July 2017

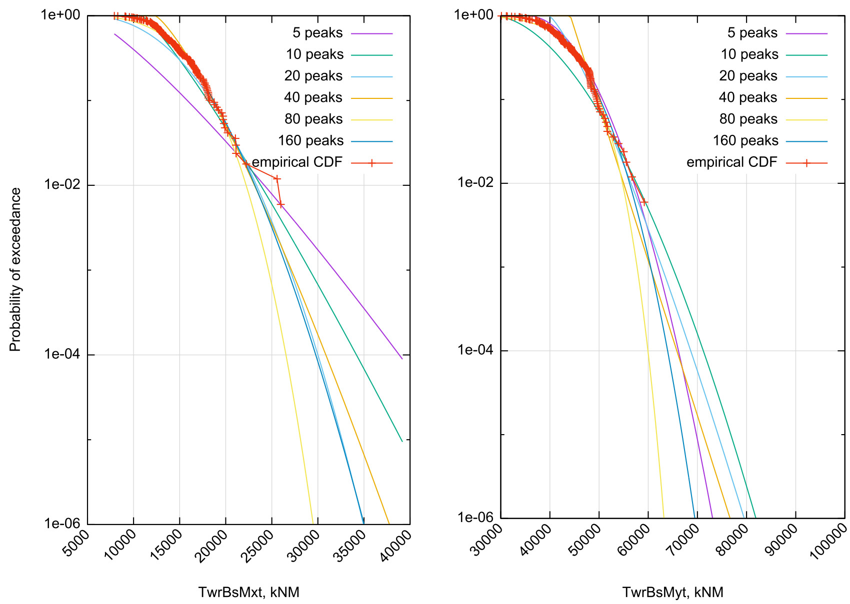

Figure 2. Weibull 3 parameter fits to empirical CDF for wind speed bin centered at 20 m/s for a variety of choices for how many peaks M_{pks} we use for the fit for side-side (left) and fore-aft (right) tower base loads. In all cases we are fitting the analytical CDF to the *highest* M_{pks} loads of the empirical CDF (a lower values of M_{pks} is like a higher threshold in the peak-over-threshold method). Especially for the side-side moment, the extrapolated POE values depend heavily on how many peaks are used.

samples to each bin at each iteration. The most obvious observation is that indeed the variance of the estimates is decreasing quickly as a function of iteration. ASIS reliably drives the variance of the estimates of POE down, and simply recalculating the extrapolations to "keep up" with ASIS drives the variance of the extrapolation estimates down as well. There is a slight dependence on M_{pks}, but it does appear there is a "sweet spot" around 40 peaks that is good for both loads (the detailed study of the optimal M_{pks} is beyond the scope of this paper; it is akin to the study of the optimal threshold in the peak-over-threshold method). The standard deviation drops by roughly a factor of 3 after only about 100 FAST runs, (compared to 30 for the original extrapolation (iteration 0)). This is closer to a "$\frac{1}{N}$" rate of convergence than the theoretical "$\frac{1}{\sqrt{N}}$" convergence. Thus our "adaptive extrapolation" approach appears capable of reducing variance somewhat dramatically with minimal additional computation. The two approaches maintain correspondence even while adapting the bin distribution, which allows for leveraging the variance reduction of the empirical ASIS estimate to reduce the variance of the extrapolation estimate. And the latter is the estimate of real importance, because that is what will be used in practice.

Wind Energ. Sci. Discuss., https://doi.org/10.5194/wes-2017-30
Manuscript under review for journal Wind Energ. Sci.
Discussion started: 25 July 2017

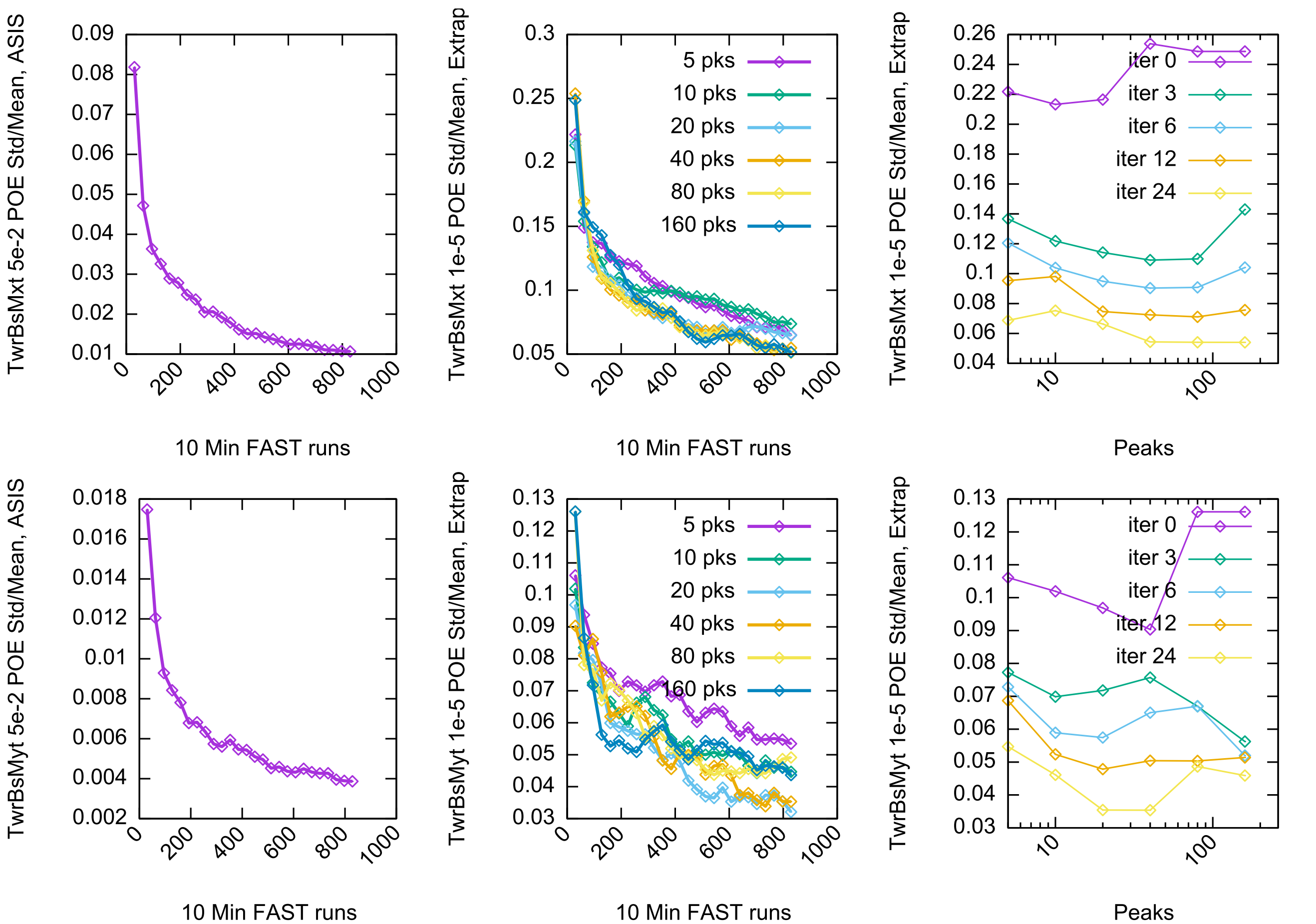

Figure 3. Summarizing the convergence of tower base load estimates using extrapolation and ASIS estimates over 100 independent runs. The top row is side-side, the bottom is fore-aft. The y-axis units are the ratio of the standard deviation to the mean estimate (relative standard deviation, measuring convergence). The x-axis is the number of FAST runs (measuring computational expense). The target POE for the empirical ASIS estimate (left) is 5^{-2} and for the extrapolation estimate (middle and right) is 10^{-5}. Adaptively selecting samples in a way designed to accelerate the convergence of the empirical estimate (ASIS, left) also accelerates the convergence of the extrapolation estimates (middle, right). There is a somewhat weak dependence on the number of peaks used for extrapolation, but $M_{pks} \sim 40$ appears robust for both loads. The side-side load estimate has larger initial relative variance, because (as shown above in Fig. 1) its extremes occur at high winds, but its relative variance is reduced more quickly by the adaptive procedure than the fore-aft load.

Wind Energ. Sci. Discuss., https://doi.org/10.5194/wes-2017-30
Manuscript under review for journal Wind Energ. Sci.
Discussion started: 25 July 2017

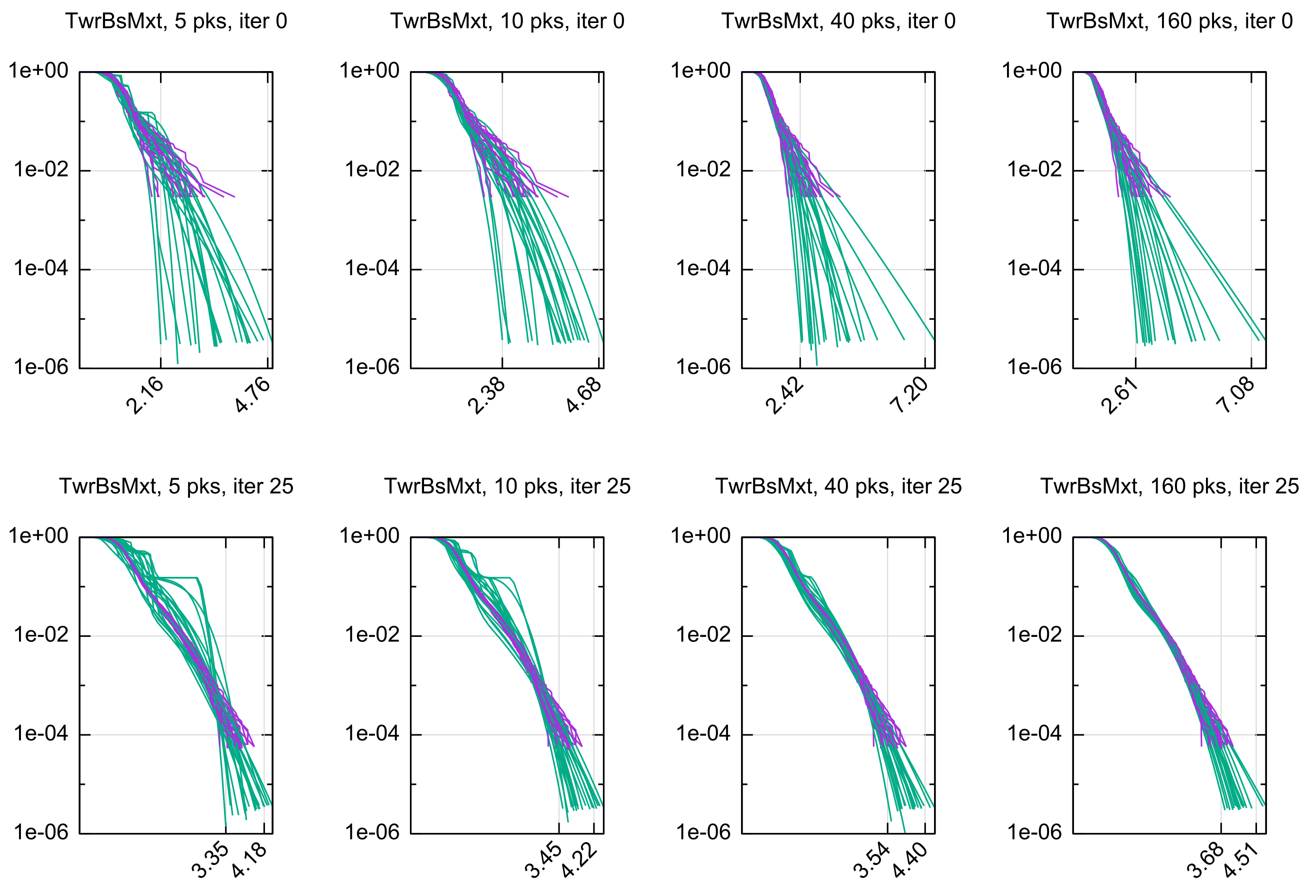

Figure 4. Behavior of ASIS and iterative updating of extrapolation for side-side tower base bending load over 20 separate runs as a function of M_{pks}. The x-axis unit are 1000s of kN-m, the y-axis is probability of exceedance in 10 minutes. The top row shows the results of both ASIS and extrapolation at iteration 0 (i.e. just based on the initial set of bin-wise samples) as a function of the number of peaks used for fitting the extrapolation distributions. The bottom row shows the estimates after 25 ASIS iterations. (Note the ASIS results are the same across each row because they are independent of M_{pks}.) Clearly the variance of the estimates is tightened. It is not clear from visual inspection if one choice of M_{pks} is better than any other.

Wind Energ. Sci. Discuss., https://doi.org/10.5194/wes-2017-30
Manuscript under review for journal Wind Energ. Sci.
Discussion started: 25 July 2017

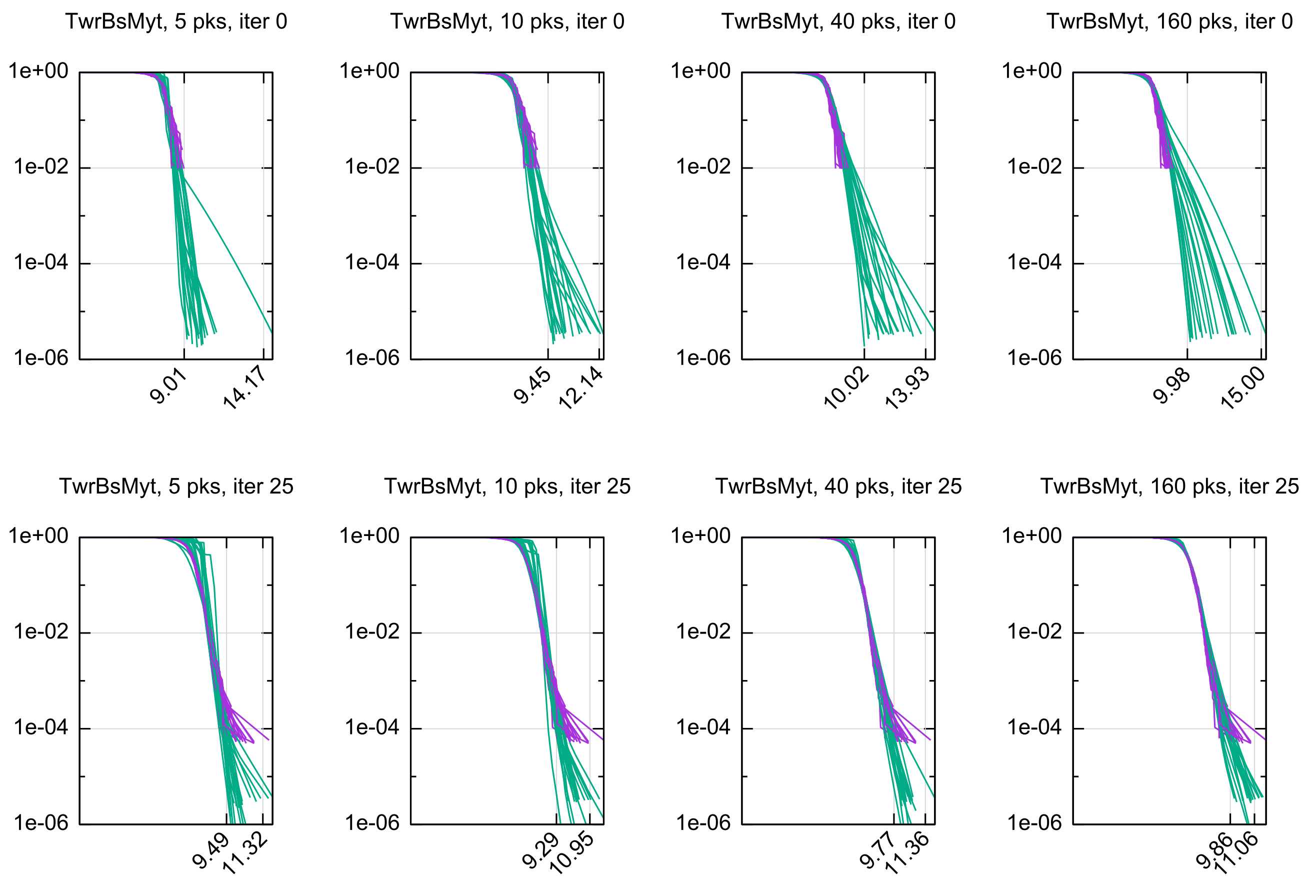

Figure 5. Behavior of ASIS and iterative updating of extrapolation for fore-aft tower base load over 20 separate runs. The x-axis unit are 10,000s of kN-m, the y-axis is probability of exceedance in 10 minutes. The format is the same as Fig. 4. Again it is interesting to compare the visual representation with the statistics presented in Fig. 3. The single extremely large load discovered by ASIS was also seen in (Graf et al., 2017) (Fig 5, panel (f)). Though it is beyond the scope of the present paper to do so, one of our main conclusions is that the statistical methods have come to a point where the best course forward will be to pursue the exact causes of such loads and integrate a statistical description of such situations into our methods.

Wind Energ. Sci. Discuss., https://doi.org/10.5194/wes-2017-30
Manuscript under review for journal Wind Energ. Sci.
Discussion started: 25 July 2017

5 Conclusions

In this paper we have built a bridge between bin-based extrapolative methods and sample-based importance sampling Monte Carlo methods. With this, we proposed a stratified adaptive importance sampling (ASIS) algorithm that is both more efficient than existing Monte Carlo approaches and maintains contact with the extrapolation methods and thereby allows for iteratively increasing the extrapolation accuracy. This is important, because only the extrapolations are able to routinely make estimates of extremely long return period load exceedance probabilities.

There are several specific questions we are now in a better position to tackle. First, given a set of peaks, what is the best way to use them to fit an extreme value distribution? In particular, using the largest peak values should more closely capture the extreme loads; but these are also the loads that have the largest variation (because they are the most rare), so this part of the empirical CDF is the least reliable. We can investigate the variation in fitted extreme value distributions as a function of schemes used to select peaks (the simplest of which, as explored above, is just taking the M_{pks} highest peaks).

Next, the importance distribution optimization problem we have introduced is clearly a *stochastic* optimization problem. As stated above, this is in fact a convergent algorithm. But this is an active area of research, and more sophisticated algorithms in the field of stochastic optimization can be used to improve our approach. We need to keep in mind, however, that the optimization problem is a means to an end. The real goal is minimal variance estimates with the smallest amount of effort. We want to *use* the optimal importance distribution at the same time as we are discovering it.

In the end we need to also keep in mind that we have the dual mission of both efficiently estimating the load POEs *and* accurately estimating their *variance*. We can use the peaks we sample to make unbiased estimates of variance just as we do expectation, but these are only *estimates*, and they themselves suffer from lack of convergence. The resampling method of *bootstrapping* offers a way to leverage a single data set to estimate statistics *and their variance*, and in a practical setting this would be recommended (as opposed to the completely separate runs we have described above).

Finally, it appears that this problem may be ripe for a machine learning approach: The physics *is* in the solver; to the extent it is possible, we should be able to learn from increasing amounts of data. For this, we need accurate variance estimation methods that can build a “loss function" for learning algorithms that examine data and decide how to best process it to make the best next estimate, and to choose the best next places to sample; here we have presented a framework for extrapolating from such data that allows for learning the best extrapolation strategy from the variance minimization algorithm.

On the other hand, we should realize there *is* a *physical* source of extreme response variation, which is the combination of turbulent inflow and nonlinear turbine response. By “opening up the black box", i.e., circling back to the original physics, we hope to *transfer* what in the present setup is response variability into the realm of environmental variability, at which point we can use its probability distribution (which, like wind speed and other environmental variables, we have a better chance of knowing) in an IFORM-like method to hone in on just the loads of interest (i.e., the extreme loads) more quickly. Further studies into the root causes of extreme response variation of wind turbine loads for fixed environmental conditions and their ultimate incorporation into more efficient statistical extreme loads estimation are ongoing.

References

Choe, Y., Byon, E., and Chen, N.: Importance Sampling for Reliability Evaluation With Stochastic Simulation Models, Technometrics, 57, 351–361, doi:10.1080/00401706.2014.1001523, http://dx.doi.org/10.1080/00401706.2014.1001523, 2015.

Choe, Y., Pan, Q., and Byon, E.: Computationally Efficient Uncertainty Minimization in Wind Turbine Extreme Load Assessment, J. of Solar Energy Engineering, 138, 041 012–041 012–8, 2016.

Graf, P., Stewart, G., Lackner, M., Dykes, K., and Veers, P.: High-Throughput Computation and the Applicability of Monte Carlo Integration in Fatigue Load Estimation of Floating Offshore Wind Turbines, Wind Energy, 19, 861–872, 2016.

Graf, P., Damiani, R., Dykes, K., and Jonkman, J.: Advances in the Assessment of Wind Turbine Operating Extreme Loads via More Efficient Calculation Approaches, in: AIAA SciTech 2017 - 35th Wind Energy Symposium, National Harbor, MD, 2017.

IEC: 61400-1. Wind turbines—Part 1: Design requirements, 2005.

Jonkman, J.: New Modularization Framework for the FAST Wind Turbine CAE Tool, in: Proceedings of the 51st AIAA Aerospace Sciences Meeting, AIAA, Dallas, TX, 2013.

Jonkman, J., Butterfield, S., Musial, W., and Scott, G.: Definition of a 5-MW Reference Wind Turbine for Offshore System Development, Tech. Rep. NREL/TP-500-38060, NREL, Golden, CO, 2015.

Karamchandani, A., Bjerager, P., and Cornell, C.: Adaptive Importance Sampling, in: Proc. ICOSSAR, International Conference on Structural Safety and Reliability, pp. 855–862, 1989.

Moriarty, P.: Database for Validation of Design Load Extrapolation Techniques, Wind Energy, 11, 559–576, 2008.

Ragan, P. and Manuel, L.: Statistical Extrapolation Methods for Estimating Wind Turbine Extreme Loads, J. Solar Energy Engineering, 130, 03101 111–03101 115, 2008.

Robbins, H. and Monro, S.: A stochastic approximation method, Ann. Math. Statistics, 22, 400–407, 1951.

Robert, C. and Casella, G.: Monte Carlo Statistical Methods, 1431-875X, Springer-Verlag New York, 2 edn., 2004.

Toft, H., Sørensen, J., and Veldkamp, D.: Assessment of Load Extrapolation Methods for Wind Turbines, J. of Solar Energy Engineering, 133, 021 001–1–8, 2011.

Winterstein, S. R., Ude, T. C., Cornell, C. A., Bjerager, P., , and Haver, S.: Environmental Parameters for Extreme Response: Inverse FORM With Omission Factors, in: Proc. ICOSSAR, International Conference on Structural Safety and Reliability, pp. 551–557, 1993.