Received: 5 July 2018 | Revised: 15 February 2019 | Accepted: 24 May 2019
DOI: 10.1002/we.2382

WILEY

RESEARCH ARTICLE

Load case reduction for offshore wind turbine support structure fatigue assessment by importance sampling with two-stage filtering

Lars Einar S. Stieng | Michael Muskulus

Department of Civil and Environmental Engineering, Norwegian University of Science and Technology NTNU, Trondheim, Norway

Correspondence
Lars Einar S. Stieng, Department of Civil and Environmental Engineering, NTNU, Trondheim 7491, Norway.
Email: lars.stieng@ntnu.no

Funding information
Norges Forskningsråd, Grant/Award Number: 193823; Strategiske Forskningsråd, Grant/Award Number: 1305-00020B

Abstract

A complete fatigue assessment for operational conditions for offshore wind turbines involves simulating thousands of environmental states. For applications such as optimization, where this assessment needs to be repeated many times, that presents a significant computational problem. Here, we propose a novel way of reducing the number of simulated environmental states (load cases) while maintaining an acceptable accuracy. From one full fatigue analysis of a base design, the OC3 monopile (with the NREL 5MW turbine), the distribution of fatigue damage per load case can be used to estimate the lifetime fatigue damage of a range of modified designs. Using importance sampling and a specially adapted two-stage filtering procedure, we obtain pseudo-optimal sets of load cases from which the fatigue damage is estimated. This is applied to seven different designs that have been modified to emulate iterations of an optimization loop. For several of these designs, sampling less than 1% of all load cases can give damage estimates with median errors of less than 2%. Even for the most severe cases, using 3% of the environmental states yields a maximum error of 10%. While further refinement is possible, the method is considered viable for applications within design optimization and preliminary design.

KEYWORDS

fatigue, importance sampling, load case reduction, offshore wind turbines, support structures

1 | INTRODUCTION

One of the main challenges for the design of offshore wind turbine support structures is the complexity of the offshore environment and the corresponding loads on the structure. The structural response in the time domain to any given loading situation produced by the wind, the corresponding motion of the rotor, the waves, and the interaction of all of these, can be reasonably accurately captured by modern simulation tools. However, the modeling of both the structure and the loading is quite involved, and the computational effort (as measured in either computational time or machine resources) for any single simulation is non-negligible. Furthermore, to faithfully represent all the different environmental conditions that are likely to occur during a year of normal operation, as dictated by the standards (eg, IEC[1]), requires literally thousands of different simulations. Recent advances in computational capacity have somewhat alleviated this issue in the context of single-design assessments. However, the problem remains where assessments of a large number of designs to the same environment is needed, for example, in the case of design optimization.[2-4] This is the main motivation and setting for the methodology proposed in this study.

The reduction of computational effort for simulation-intensive problems is a much more general topic than its applications to (offshore) wind energy, see, eg, Sacks et al[5] and Santner et al.[6] However, given the offshore wind energy-specific problem addressed here, as well as the extent to which our proposed methodology is tailored to this problem, our literature review will be mostly contained to this area. In general, one can say that previous work on this topic has obtained various degrees of success with various degrees of more general applicability. Zwick and Muskulus[7] investigated the use of two different linear approximations to the fatigue damage of a jacket support structure when varying the wind speed.

From a total of 21 different operational wind speeds, using lumped sea states and a fixed direction for the incoming wind and waves, the authors were able to reduce the number of required environmental states down to sets of between three and six. A maximum error in the estimated lifetime fatigue damage of around 6% was achieved in the case of sets with three load cases. Häfele et al[8,9] studied a much more comprehensive set of load cases, in total 2048, for multiple jacket designs. Using random sampling from probability distributions derived from the probability of occurrence of each load case, they obtained a hierarchy of reduced load case sets from which the total fatigue damage could be approximately estimated. The errors in the estimates for the smaller sets were quite substantial, though the method still showed the possibility of significant reduction in the load case set while achieving acceptable accuracy.

Studies have also been made of various sampling-based approaches to simplify fatigue assessment for floating support structures. Graf et al[10] used both grid-based and Monte Carlo sampling on a very large environmental data set. Müller, Dazer, and Cheng[11] used response surface modeling based on Latin Hypercube Sampling and artificial neural networks. Müller and Cheng[12] used Sobol sequences to select which environmental subsets to sample allowing for more rapid convergence compared with conventional Monte Carlo sampling. The latter study was able to ensure a maximum error of about 10%. Kim et al[13] used a frequency domain approach based on an artificial neural network to modify the stress transfer function. A more general study on selection of subsets of environmental conditions for offshore wind turbines using a maximum dissimilarity algorithm was conducted by Guanche et al.[14] Finally, Choe, Byon, and Chen[15] used an extended type of importance sampling to estimate the reliability in a framework of stochastic simulation models. Summarizing these previous studies as they apply to the current work, we can say that they all demonstrate the power of various sampling-based techniques (and in some cases, surrogate models) to simplify fatigue damage evaluation for various types of offshore wind turbine support structures and more general frameworks. However, for the most part, the applied methods are meant to simplify the damage assessment of single designs, focusing on simplifying the environmental states without considering the specific fatigue damage caused by each state. This is reasonable for the applications likely intended by the authors, but leaves a gap as far as applications to (eg) optimization are concerned. In optimization, repeated fatigue analyses subject to the same environmental states are performed, meaning that information about the specific fatigue damage in each state is available in the subsequent analyses. Hence, we seek methods that can use this information to generate reduced sets of load cases to be used during the optimization.

In this study, we resolve some of the issues left open by previous work with regard to how these sampling-based approaches for simplified fatigue assessment could be utilized within design optimization. To do this, we employ a scheme based on importance sampling. While this is a slightly different sampling-based approach than those used by the cited previous studies, the most important idea in the current work is how we can tailor this type of approach to be used to predict the fatigue damage of modified support structure designs. Hence, with some modifications, it should be possible to extend these more practical ideas to many of the methods referenced above. For example, while Häfele et al[8,9] also used a type of importance sampling, they considered only the distribution of environmental states and did not fine-tune the sampling based on the actual fatigue response of the structure. The fundamental assumption of our approach is that the fatigue distribution for the set of environmental states should be approximately invariant when the design is modified. By further combining the fatigue distributions at all relevant locations along the height of the structure, one can then obtain a sampling distribution that can be used to obtain an importance sampling estimate of the total fatigue damage of a modified design. We have furthermore applied an additional two-stage filtering procedure in order to reduce both the variance of the estimates and to reduce the error of the most severe outliers seen when the sampling procedure is repeated many times.

2 | BACKGROUND AND METHODOLOGY

If an offshore wind turbine support structure is to be considered viable and safe with respect to structural performance under cyclic loading, it needs to satisfy the design criteria for fatigue[16] for each design load case required by the standards.[1] This is in general a detailed procedure consisting of many checks, but a significant amount of computational time is spent on the load cases, together forming the design load case (DLC), representing operational conditions, since this is where the largest number of environmental conditions need to be checked. In the following, we restrict the analysis to the DLC representing operational conditions (DLC 1.2 in the IEC standard[1]). Each environmental condition (from now on referred to as *load cases*) E, representing a particular realization of wind and waves (including different incoming directions), has a certain probability of occurrence, $P(E)$, during any given year of operation at a particular site. To estimate the total fatigue damage in a single location of the structure that is expected to occur during operational conditions during a lifetime T_{lt}, the designer must first estimate the fatigue damage $D(E)$ expected to occur for each load case. This can be done and will be done in the following, by a time domain structural analysis (simulation) resulting in a time series for the normal stress σ_n at the location in question, where the stress is evaluated at eight points i around the circumference of each location:

$$\sigma_n^i = \frac{F_x}{A} + \frac{M_y r}{I_y}\cos(\theta_i) - \frac{M_z r}{I_z}\sin(\theta_i), \tag{1}$$

Where F_x is the axial force, A is the area of the section, M_y and M_z are the in-plane and out-of-plane moments, I_y and I_z the corresponding second moments of area, r the radius of the section, and the angles θ_i represent the position of the point i around the circumference of the section. For fatigue analysis, the contribution of the axial force is small and could in principle be neglected, but we include it here for the sake of completeness. Applying the conventional methodology of rainflow counting[17] and damage accumulation, via application of SN-curves (the T curves in tables 2-1 and 2-2 in DNV-RP-C203[18] were used) and the Palmgren-Miner rule, then results in the desired value for $D(E)$, where the maximum of the eight damage values found around the circumference of each location in the structure is used. The total fatigue damage D_{tot} is then found as:

$$D_{\text{tot}} = T_{\text{lt}} \cdot \sum_E P(E) \cdot D(E). \tag{2}$$

We note that, for reasons of simplicity, we will ignore the usual safety factor that would be applied to (2). For similar reasons, the fact that it only changes the result by a fixed constant, we will, in the following, also neglect the factor T_{lt} scaling up the damage from simulation time to lifetime. In particular, since we will present our main results as relative errors, any shared constant factors automatically cancel out.

Below, we will present a scheme for estimating D_{tot} from a small subset of load cases. It is based on the idea of importance sampling, now taken to a discrete setting (sums). While we also present further modifications to the method that decrease the variance at the cost of some added complexity, a lot of the power of this approach is contained within the basic setup.

2.1 | Importance sampling

Importance sampling[19] is an extension of the traditional Monte Carlo sampling method to estimate integrals I of the form

$$I = \int f(x)\mathrm{d}x, \tag{3}$$

where f is some function, x is some variable, and, for simplicity, we assume that the domain of f is such that $\int \mathrm{d}x = 1$. Conventional Monte Carlo estimates I by sampling f at N random points drawn from a uniform distribution on the whole domain and estimates the integral as the average value of f over all the samples $\{x_i\}$:

$$\hat{I} = \frac{1}{N}\sum_{i=1}^{N} f(x_i), \tag{4}$$

Where $\hat{I}$ is the Monte Carlo estimate of the integral I. Note that here, x_i is just a number, but later it will be used as an identifier for load cases, which are in a sense vectors of parameter values. The Monte Carlo estimate has the property that as N tends to infinity, $\hat{I}$ tends to I. However, the convergence of $\hat{I}$ can be very slow, typically proportional with the inverse square root of N. Hence, one often seeks to speed up the convergence by modifying the approach. Importance sampling is based on the idea of introducing a new function g, called the *sampling distribution*, to be used instead of the original distribution when drawing samples. By design, samples drawn from this new distribution tend to be in the part of the domain where f has the highest contribution to the integral, the region with the highest "importance." Sampling from g instead of a uniform distribution should then ensure that, when g is chosen well, $\hat{I}$ converges to I much faster. To proceed, we start by multiplying and dividing by $g(x)$ in (3):

$$\begin{aligned} I &= \int f(x)\frac{g(x)}{g(x)}\mathrm{d}x \\ &= \int \frac{f(x)}{g(x)}g(x)\mathrm{d}x. \end{aligned} \tag{5}$$

Using the same ideas as when going from (3) to (4), the importance sampling estimate of I, can then be found as follows:

$$\hat{I} = \frac{1}{N}\sum_{i=1}^{N}\frac{f(x_i)}{g(x_i)}, \tag{6}$$

Where the N samples $\{x_i\}$ have now been drawn from the probability density $g(x)$. Note that importance sampling is just a special case of the situation, where the variable x is distributed according to some probability distribution $h(x)$, which would then be substituted for g in the above expressions. In the general case, the random samples are also drawn from h, and $\hat{I}$ is estimated in the same way as in (6), but the convergence of $\hat{I}$ to I may be much slower than when the distribution is chosen with convergence in mind, which is the main idea of importance sampling.

2.2 | Importance sampling for fatigue estimation

A similar idea as above holds for approximating the value of a sum. Suppose we want to evaluate some property S, defined as a sum of N_{tot} terms:

$$S = \sum_{i=1}^{N_{\text{tot}}} f(x_i). \tag{7}$$

Just as when deriving (6) from (3), we can then obtain an importance sampling estimate, $\hat{S}$, of S, from a reduced set $\mathcal{J}$ consisting of N samples drawn from a (now discrete) sampling distribution $g(x_k)$ as follows:

$$\hat{S} = \frac{1}{N}\sum_{i\in\mathcal{J}}\frac{f(x_i)}{g(x_i)}, \tag{8}$$

Where g is defined such that it generates samples that give as much information about f as possible with respect to the evaluation of the sum. Note that the indices $i \in \mathcal{J}$ no longer run from 1 to N like in (6), since we are now selecting among an already discrete (and finite) set of points. Hence, the indices (points) are not consecutive.

The above is exactly the type of situation we want to address for fatigue estimation. From (2), letting $E = x_i$ and ignoring the scale factor, we get

$$D_{\text{tot}} = \sum_{i=1}^{N_{\text{tot}}} P(x_i) \cdot D(x_i), \tag{9}$$

Where N_{tot} now refers to the total number of load cases. Using (7) and (8) and letting $f(x_i) = P(x_i) \cdot D(x_i)$ and $g(x_i) = \frac{f(x_i)}{D_{\text{tot}}}$, we arrive at the following importance sampling estimate of the total fatigue damage:

$$\hat{D}_{\text{tot}} = \frac{1}{N} \sum_{i \in \mathcal{J}} \frac{f(x_i)}{g(x_i)}. \tag{10}$$

We should stress at this point that this definition of g is of course the optimal choice if we actually know D_{tot}, but this value is exactly what we are trying to estimate. This can be seen by simplifying the above to obtain the following:

$$\begin{aligned} \hat{D}_{\text{tot}} &= \frac{1}{N} \sum_{i \in \mathcal{J}} \frac{f(x_i)}{\frac{f(x_i)}{D_{\text{tot}}}} \\ &= \frac{1}{N} \sum_{i \in \mathcal{J}} D_{\text{tot}} \\ &= D_{\text{tot}}. \end{aligned} \tag{11}$$

So this estimate is in fact exact (even for $N = 1$) in the idealized case, where we know the true value of D_{tot} already. A conventional Monte Carlo sampling with a uniform distribution would require on average $N = N_{\text{tot}}$ samples to attain a similar accuracy. The motivation for the definition of g should hence be clear: If the fatigue damage distribution changes very little when the support structure design is changed, or if the changes for each load case are all the same, then importance sampling with g will give an essentially exact estimate from a minimal sample size (in principle a sample size of 1).

The problem that we actually aim to solve is the following: Suppose we already know the fatigue damages for each load case, $D(x_i)$, and hence the total fatigue damage, D_{tot}, of some initial support structure design. We want to estimate the total fatigue damage, $D_{\text{tot}}^{\text{new}}$, of some new (similar) design, from the corresponding fatigue damages per load case, $D^{\text{new}}(x_i)$. If we knew $D^{\text{new}}(x_i)$ for every load case, the exact value would then be calculated as

$$D_{\text{tot}}^{\text{new}} = \sum_{i=1}^{N_{\text{tot}}} P(x_i) \cdot D^{\text{new}}(x_i). \tag{12}$$

However, we want to estimate $D_{\text{tot}}^{\text{new}}$ without having to perform simulations for every load case. Defining g using the information from the initial design, $g = \frac{P(x_i)D(x_i)}{D_{\text{tot}}}$, and generating a set $\mathcal{J}$ of N load cases from this distribution, we may then estimate the new total fatigue damage as follows:

$$\hat{D}_{\text{tot}}^{\text{new}} = \frac{1}{N} \sum_{i \in \mathcal{J}} \frac{f(x_i)}{g(x_i)}, \tag{13}$$

Where $f(x_i) = P(x_i) \cdot D^{\text{new}}(x_i)$. For clarity, we may simplify this expression to obtain the following:

$$\hat{D}_{\text{tot}}^{\text{new}} = D_{\text{tot}} \cdot \frac{1}{N} \sum_{i \in \mathcal{J}} \frac{D^{\text{new}}(x_i)}{D(x_i)}. \tag{14}$$

Hence, the estimate for the new total fatigue damage is essentially a scaling factor times the old total fatigue damage, where the scaling factor is the average ratio between the new and old fatigue damage incurred for each load case.

2.3 | Sampling distribution for multiple locations

Using importance sampling in practice requires some further refinement. If all we were interested in was the total fatigue damage at a single location of the structure, (13) or (14) would suffice. However, one often wants the damage in several locations along the height of the structure. This raises an immediate question: Is the sampling distribution induced by the fatigue damage the same for all locations in the structure? Intuitively, one would expect that there are no major differences, but because of, for example, the effect of misalignment, which might be dependent on location, there is no reason why they should be exactly the same. In fact, an initial investigation showed that there were some differences, which motivated a further refinement of the sampling distribution. Especially since for a small set of samples, even minor differences can lead to significant differences in performance when the above method is applied to estimate the damage in each location. Hence, we need to redefine the sampling distribution such that it samples in a balanced way. There is not really a way to accommodate each location perfectly, nor some ready-made state-of-the-art approach to follow, so in order to maintain balance between the locations we propose to define a new distribution $\tilde{g}$ in the following way: For each location j, define a sampling distribution g_j as in the procedure leading to (13) by using the fatigue damages at that location. We can then obtain $\tilde{g}$ as a normalized, weighted sum, of the distributions g_j, taking care to normalize both the individual distributions and the final new distribution:

$$\tilde{g}(x_i) = \frac{\sum_j w_j \cdot \frac{g_j(x_i)}{\max\{g_j\}}}{\sum_k \sum_j w_j \cdot \frac{g_j(x_k)}{\max\{g_j\}}}, \tag{15}$$

Where w_j are the weights, and we have normalized each individual distribution by their respective maxima in order to make sure that distributions with higher overall numerical values contribute the same (up to the given weight) as ones with lower overall numerical values. If there is no reason from a design point of view to favor one location over another (or to *disfavor* one location over another), then the weights may all be set to the reciprocal of the number of locations N_p, making the new distribution simply the average of the others. Note that in the latter case, by a quick inspection of (15), we may actually set the weights to 1. This is because when all the weights are the same, there is perfect cancellation.

To use the method in practice then means to perform simulations and corresponding fatigue analyses for all load cases for one fixed support structure design. Once this has been done, the sampling distribution can be defined by (15), and one can generate a reduced set of load cases to analyze for any new design, under the assumption that the structural changes have little impact on the overall distribution of fatigue damage across the total set of load cases.

2.4 | Two-stage filter

Importance sampling, like any sampling procedure drawing from a probability distribution, is inherently random and therefore has an associated variance. In anticipation of the need to reduce this variance, we will now propose an addition to the importance sampling procedure as explained so far. The overall idea is to first reuse the information we already have from the initial set of simulations to construct an optimization scheme that further conditions our samples on properties that increase the accuracy of the fatigue estimate beyond what is possible to encode in the sampling distribution directly. Then, we identify an issue that can cause large errors in the estimate when the support structure is changed significantly and define a criterion to alter samples affected by this behavior. Put together, this extended procedure may be called importance sampling with a two-stage filter.

2.4.1 | Stage 1: Optimal sampling through simulated annealing

First, we want to further condition the sample on the fatigue distribution. When defining the fatigue distribution based on several locations, as in (15), the property that the importance sampling gives the exact answer for the original design, discussed beneath (10), no longer holds. However, there should still be a sample set, which is *optimal* for all locations. One that would give the best possible importance sampling estimate when applied to the original design. Finding the optimal sample set can be done using only the information we already have when defining the sampling distribution, but we need a measure for what it means for a sample set to be optimal. A straightforward definition of such a measure at each location j is the absolute value of the relative estimation error:

$$\delta_j = \left| 1 - \frac{\hat{D}^j_{\text{tot}}}{D^j_{\text{tot}}} \right|. \tag{16}$$

In order to find a sample set that will be optimal when considering all N_p locations, we then define the total performance measure, Δ, as follows:

$$\Delta = \sum_{j=1}^{N_p} \delta_j. \tag{17}$$

We could also have used a quadratic error measure, like the sum of squares. However, this puts too much emphasis on larger errors. For instance, $\sqrt{0.095^2 + 0.005^2} = 0.0951$, whereas $0.095 + 0.005 = 0.1$. The former alternative would pass a 0.1 tolerance check, but the latter would not. However, there is another modification of Δ that makes the objective more robust against overfitting. If we denote a given sample set by $\{x_k\}$, we can make the following definition of the final performance measure:

$$\tilde{\Delta} = \frac{\Delta}{\sum_j \tilde{g}(x_j)}. \tag{18}$$

Dividing by the sum of the sampling distribution evaluated at each sample point, corresponding to the sum of the isolated probabilities of obtaining each sample point means that sample sets that are very unlikely to be picked out of the population according to $\tilde{g}$ (even ones with small errors) will have larger values of $\tilde{\Delta}$. Hence, such sample sets are unlikely to be selected when minimizing $\tilde{\Delta}$. This helps us avoid overfitting our proposed sample set to the initial design. For a simplified, and more general, explanation of how such overfitting might occur, see Figure 1. Finally then, what we want to find is the following:

$$\min_{\{x_k\} \in \Omega_k} \tilde{\Delta}, \tag{19}$$

Where Ω_k is the sample space for the set of all load cases and a given sample set size k.

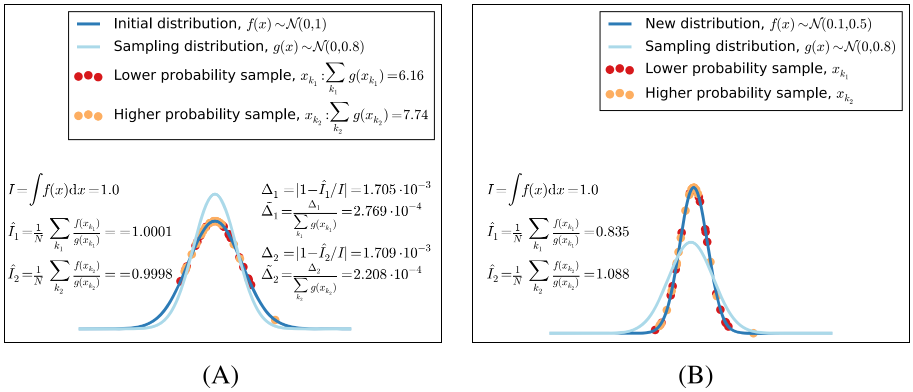

FIGURE 1 The effect of overfitting, for simplicity illustrated with various normal distributions ($\mathcal{N}(\mu, \sigma)$). Note how the two different samples (x_{k_1}, x_{k_2}), with noticeably different probability sums, give about the same estimate for the integral of the initial distribution f, which we are using to select the optimal sample for future estimation (panel A). In fact, using the uncorrected version of the performance measure Δ would lead us to select the sample with the smaller value of the probability sum (expected to be less robust when estimating the integrals of other distributions). We see, however, that it is indeed the higher probability sample, which is more suitable for such estimation (panel B). Using Δ rather than $\tilde{\Delta}$, we would have relied too much on information from the initial f, which is an example of overfitting to the training data [Colour figure can be viewed at wileyonlinelibrary.com]

We now have an objective to optimize, but we still need a way to solve (19). This is a discrete optimization problem and one that is tricky to rephrase in a way that would make it suitable for conventional gradient-based methods. The application of gradient-based methods for discrete problems would require additional work outside the scope of the current study. Hence, we instead look to heuristic optimization approaches. There are a lot of choices here, like a naive random search, particle swarm, or a genetic algorithm, but through testing, we have found simulated annealing[20,19] to perform quite well for this problem. The specific details of the implementation will vary, for one thing simulated annealing usually requires a bit of parameter tweaking, but one thing to note here in general is that we are really only seeking local minima. From a practical standpoint, the size of Ω_k even for very small k makes finding a global solution impractical. However, it is also likely that finding such a global solution, or even something close to it, will tend to produce an overfitted sample set. Hence, the termination criterion should be set to a small but reasonable number. Doing so usually makes the algorithm converge in a few seconds.

2.4.2 | Stage 2: Removal of outliers: Motivation

The second stage of the filter is meant to safeguard against very severe outliers that can occur because significant systematic changes in the design can induce substantial nonuniform changes in the fatigue distribution. To understand this, let us consider a simplified model for a wind turbine supported by a monopile structure. If we model this system as a uniform cantilever beam with one fixed end and a mass placed at the other end, then it can be shown[21,22] that the first eigenfrequency, f_1, is proportional to the following:

$$f_1 \propto \sqrt{\frac{k_{\mathrm{eq}}}{M_{\mathrm{RNA}} + \frac{33}{140} m_{\mathrm{t}}}}, \tag{20}$$

Where k_{eq} is the equivalent stiffness of the monopile tower, M_{RNA} is the mass of the rotor-nacelle assembly, and m_{t} is the mass of the monopile tower. Note that the constant factor in front of m_{t} is derived under specific assumptions and will in general depend on structural properties. However, this dependence is small enough[23,24] that it can be neglected for the present purpose. If the dimensions of the support structure (meaning its element diameters and thicknesses) are scaled uniformly by a factor s, then the stiffness and mass will also be scaled. The stiffness depends on the dimensions via the second moment of area, which depends on the dimensions to the fourth power. Similarly, the area depends on the dimensions to the second power. Hence, if the dimensions are scaled by s, the first eigenfrequency will change as follows:

$$f_1 \propto \sqrt{\frac{k_{\mathrm{eq}} s^4}{M_{\mathrm{RNA}} + \frac{33}{140} m_{\mathrm{t}} s^2}}. \tag{21}$$

For small values of s, this is a quadratic dependence; for large values of s, this is a linear dependence; and for intermediate values, the dependence is somewhere between quadratic and linear. Hence, if the dimensions of the structure are scaled, the first eigenfrequency can change dramatically. The reason this is important for the importance sampling is because changes to the first eigenfrequency can lead to dynamic amplification for some (but not all) wind speeds. Depending on the design of the turbine and the base monopile model, an increase or decrease in the first eigenfrequency can mean a shift towards either the 1P or 3P frequency of the turbine. However, such a shift will only lead to

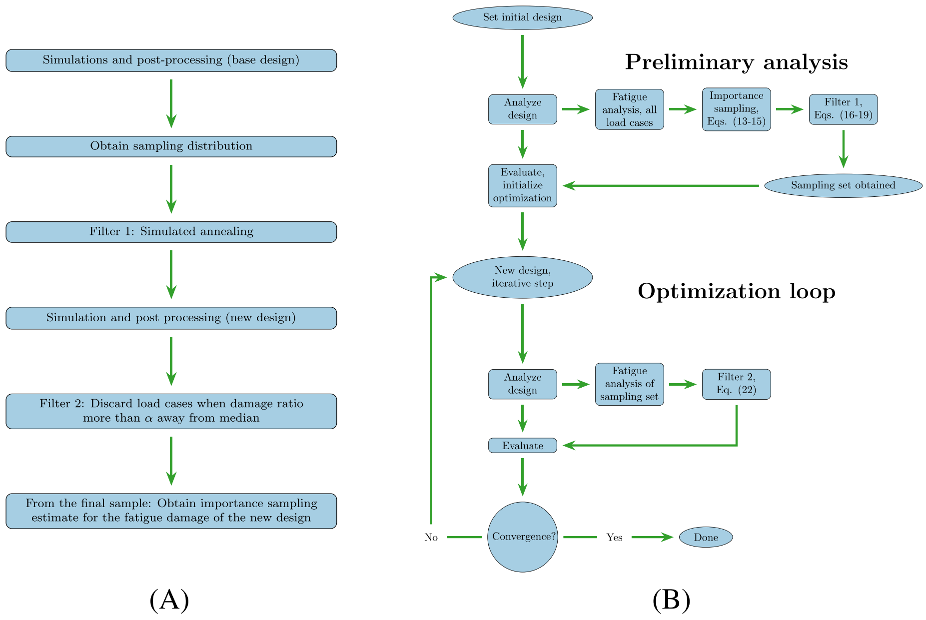

FIGURE 2 Workflow of the load case reduction method (A) and the workflow when used specifically for structural optimization (B) [Colour figure can be viewed at wileyonlinelibrary.com]

dynamic amplification for low or high wind speeds, since this is where the initial eigenfrequency will tend to be the closest to either the 1P or 3P frequency. If dynamic amplification (leading to higher fatigue damage) occurs for only some wind speeds, meaning only some subset of load cases, then the basic assumption, that the fatigue distribution changes uniformly, will break down. The idea of the second stage of the filter is to modify the importance sampling in such a way that the impact of a scenario like the one just described will be minimal.

2.4.3 | Stage 2: Removal of outliers: Implementation

From (14), we know that the importance sampling estimator is determined by the average of the ratios of incurred fatigue damage in the new design to those of the old design. If the ratio changes a lot for only certain load cases, the resulting estimate can be quite off from the true overall average ratio. This motivates the formulation of a secondary filtering approach. The core idea is to recognize that these outliers occur with low probability and that the cause of an inaccurate estimate in these scenarios is usually just one or a few load cases with abnormal ratios. Hence, we can formulate a criterion to see if any given sample set contains such load cases. We look for this by considering the difference of each fatigue ratio from the median ratio. That is, for each load case x_i, we check whether

$$\left|1 - \frac{D_{\text{new}}(x_i)/D(x_i)}{\text{M}\left(D_{\text{new}}/D\right)}\right| > \alpha, \tag{22}$$

Where M refers to the median, and α is the threshold for what constitutes an acceptable difference from the median. As with the termination criterion for the optimization in the first filter, the best choice for α will vary a bit with the size of the sample set and the extent to which one will risk false positives. We have found values in the range 1.0 to 5.0 to be suitable. When one or more such outliers are found, we simply discard the fatigue values associated with these load cases and compute the importance sampling estimate from the remaining ones. The anomalous proportion of any sample set is relatively small, especially for intermediate or large sample set sizes, so the loss of information in discarding these load cases is minimal.

A summary of the complete load case reduction procedure, as well as how it would be used for structural optimization, can be found in Figure 2.

2.5 | Setup and testing procedure

The turbine and support structure models used in this study are the NREL 5MW turbine[25,26] and the OC3 monopile,[27] respectively. The time domain analysis has been performed using the fully integrated aero-elastic software FEDEM Windpower.[28] The monopile was clamped at the seabed. Clamping the monopile has the effect of increasing the global stiffness and the eigenfrequency, which also has some effect on the distribution of loads and hence the fatigue, though this is expected to have a very limited effect on the results, which measure relative changes. A drawing of the system is shown in Figure 3. To ensure that the results are as applicable as possible to realistic scenarios, a substantial amount of load cases have been included. These represent different wind speeds, different sea states (peak period and significant wave height), and different wind and wave misalignment. These data have been taken from the Ijmuiden Shallow Water site.[29] The load cases used represent wind speeds from 4 to 24 m/s, in increments of 2 m/s, with a corresponding turbulence intensity for each wind speed (giving a total of 11 different cases). Note that for the NREL 5MW turbine, the cut-in and cut-out wind speeds are 3 and 25 m/s, respectively, and the rated wind speed is 11.4 m/s (12 m/s being the closest wind speed to rated studied here). For each wind speed, a number of different sea states have been used, with between 21 and 42 combinations of significant wave height and peak period. For each combination of wind speed and sea state parameters, the misalignment

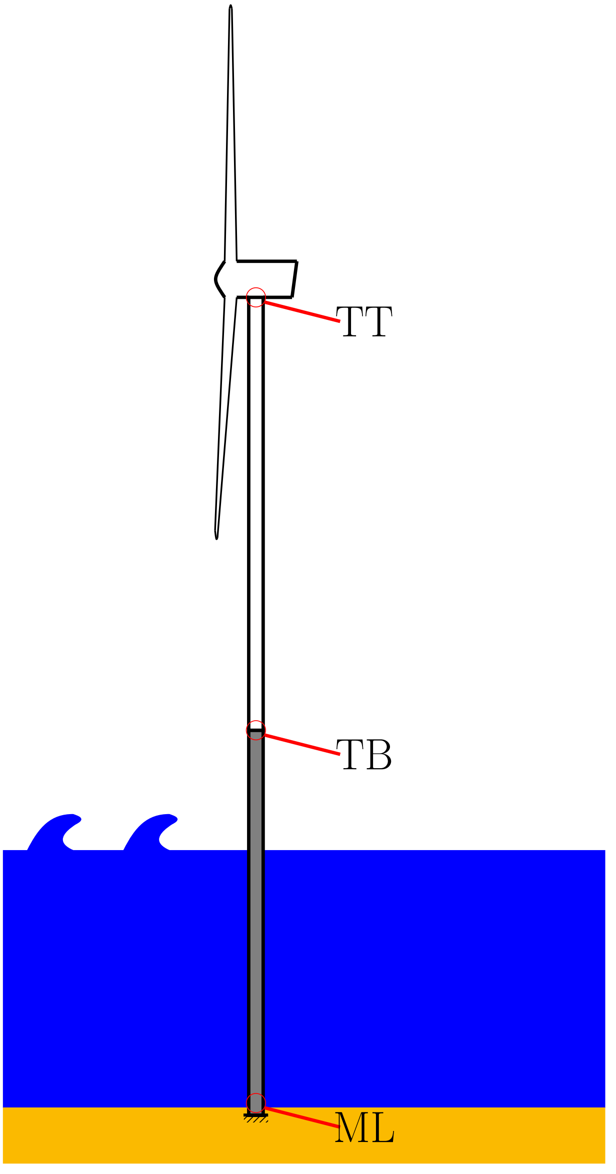

FIGURE 3 The monopile model used in this study, drawn with wind turbine on top. The three locations studied are indicated: the tower top (TT), the tower bottom (TB), and the mudline (ML) [Colour figure can be viewed at wileyonlinelibrary.com]

between incoming wind and waves has been varied between 0° and 330° in increments of 30° (giving a total of 12 possible directions). The number of sea state parameter combinations and misalignment angles used for each wind speed varies, since some combinations were not observed at the site, and these, therefore, have zero occurrence probability. In total, 3647 load cases were used. Each load case yields a 10-minute time series of usable data after deleting transients, with a different random seed for the wind and wave input for each realization. Normally, a full analysis would need to use 60 minutes of data per load case. However, if the method works for fatigue estimates based on a 10-minute time series, the increased stability of a 60-minute time series should be even better. Hence, it is reasonable to test the method with shorter simulation lengths than those that should be used in practice. The variance caused by using 10-minute simulation lengths is generally expected to be larger than the accuracy we aim for with our approach, see, eg, Zwick and Muskulus[30] for more details about this issue. Though this is generally less of an issue for optimization, where one is interested in repeated evaluations of the response to a fixed set of realizations of the loads.

Testing the method involves running one full fatigue analysis on the initial support structure design, training the method on this data, and seeing how well we can approximate a full fatigue analysis of altered designs by using importance sampling and the two-stage filtering procedure. To verify the performance, we perform full analyses using the same wind and wave data on the modified designs as well, to compare with the exact values. Since the main motivation for this simplification method is design optimization, we select the modified designs in ways that emulate certain kinds of designs one might encounter during an optimization loop, including some particularly selected design configurations that could lead to significant loss of accuracy for the method. The type of optimization we emulate here is mass optimization, which means modifying the diameters and thicknesses of the structural elements to find the minimal mass (defined as the sum of element volumes times the material density). Hence, designs where the diameters and thicknesses have been scaled either up or down, either systematically or randomly, should then be sufficiently realistic emulations of an optimization process while also probing designs where the method might be expected to fail. To this end, we have generated seven new designs. These are summarized in Table 1.

TABLE 1 Names of modified support structure designs and how they have been modified

Design name	Design Description
MD5	Element sizes decreased by 5%
MI5	Element sizes increased by 5%
MR5	Element sizes randomly increased or decreased by 5%
MD10	Element sizes decreased by 10%
MI10	Element sizes increased by 10%
MR10	Element sizes randomly increased or decreased by 10%
MRU10	Element sizes randomly increased or decreased by up to 10%, using a uniform probability distribution

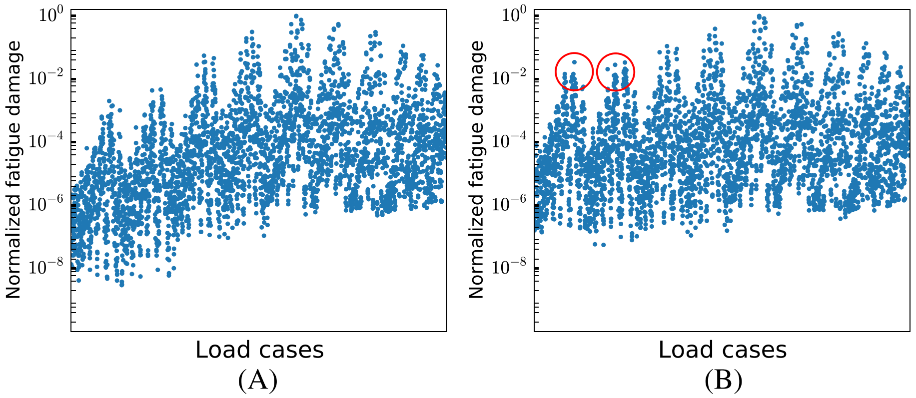

FIGURE 4 Normalized fatigue distributions for the base design (A) and for design MI10 (B) with changes for lower wind speeds marked in red [Colour figure can be viewed at wileyonlinelibrary.com]

For the purposes of testing in this study, a set of locations in the structure thought to be representative have been selected. These locations are the tower top, the tower bottom, and the mudline. See Figure 3 for the placement on the structure. The sampling distribution is hence defined according to (15) using these three locations, with weights set to 1.

3 | RESULTS

From the displayed initial fatigue distribution in the left panel of Figure 4, it is clear that there are collections of peaks for each wind speed. The wind speeds with the most impact are, in descending order, 12, 14, 16, and 10 m/s. The load cases corresponding to the peaks within each wind speed bin tend to be the ones having the highest probability of occurrence (that is, usually wind and wave misalignment of 0°, 30°, or 330° and relatively moderate sea state parameters).

For the designs that have been uniformly scaled by 5% or 10%, it would be pertinent to ask whether any of them are likely to experience the kind of eigenfrequency changes that motivated the formulation of the second stage of the filter. Applying (21) to our model, with values of s corresponding to changes of 5% or 10%, gives that the first eigenfrequency should change by almost 10% or 20%, respectively (a quadratic scaling). Performing more detailed eigenfrequency analysis with our software confirms these results. Checking the Campbell diagram for the turbine[26] indicates that there may be dynamic amplification for lower wind speeds when the structure is scaled up. Investigating the corresponding fatigue distributions shows significant changes for low wind speeds when the design is scaled up by 10% (design MI10 from Table 1), see the right panel of Figure 4.

In the following, we will show how the method performs when applied to all the models listed in Table 1. To measure the performance, the absolute value of the relative error as defined in (16) seems most suitable. The use of the absolute value is mostly for convenience and hides no information that could otherwise have been useful. For example, there seems to be no tendency to either overbe no tendency to either or underpredict the fatigue damage. Furthermore, since the method contains randomness, we seek to elucidate both its mean behavior and variance when the procedure is repeated many times. The results are hence displayed as boxplots. The boxplots used here are defined in such a way that the median of the data is represented by the horizontal line within each box, the top edge of the boxes represent a value not exceeded by 75% of the data, the ends of the "whiskers" (lines extending vertically from the boxes) represent 1.5 times the interquartile range, and the remaining points (shown as dots) are outliers. Note that the number of repetitions used below (1000 and 10 000, respectively) is not based on any type of convergence criterion for the Monte Carlo method. The idea of the repetition is simply to illustrate the statistical behavior of the method, since the importance sampling (even with the two-stage filter) can generate a variety of different sample sets. The specific numbers chosen have been set as a compromise between achieving representative statistics and practical computational effort. With seven different designs, three different locations in the structure at which the fatigue damage is evaluated and multiple sample set sizes tested, there is in principle a lot of data that could have been shown. Usually, no additional insight is provided by displaying the results for multiple locations. Hence, within any given figure, all the estimates used to make the boxplots are from the location with the largest recorded error for that design (usually the tower bottom or the mudline).

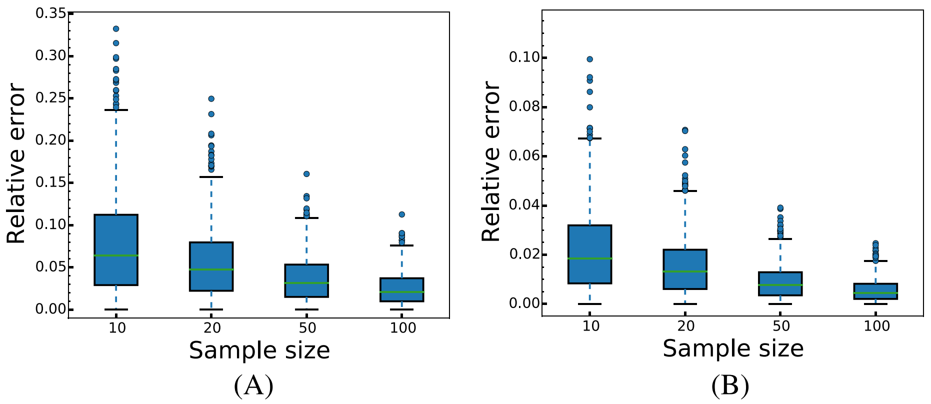

FIGURE 5 Importance sampling for design MD5, without (A) and with (B) filtering [Colour figure can be viewed at wileyonlinelibrary.com]

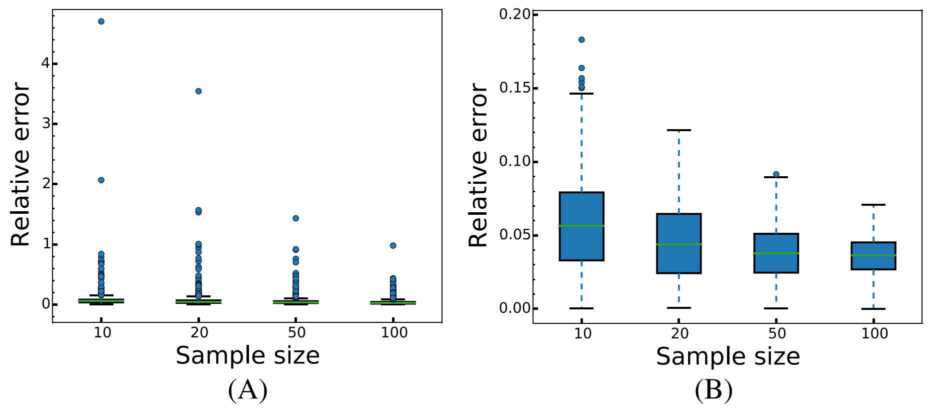

FIGURE 6 Importance sampling for design MI10, without (A) and with (B) second stage filtering [Colour figure can be viewed at wileyonlinelibrary.com]

3.1 | The effect of using the two-stage filter

With the added effort in computation and implementation required by the two-stage filter, it seems prudent to ask whether this additional effort is even required. What are the capabilities of the importance sampling method on its own? Boxplots of the estimates obtained from design MD5, using 1000 repetitions of the importance sampling procedure, are shown in the left panel of Figure 5. The results are certainly reasonable, with the estimate using 100 samples (a bit less than 3% of the size of the total set) obtaining a maximum error of about 12%, a median error of about 2% and with 75% of all estimates at less than 4% error. Still, there is room for improvement. Especially for the estimates using smaller sample set sizes and when considering that a higher number of repetitions could mean even larger outliers. Applying the first stage of the filter (in this case the filter's second stage did not trigger) when generating the estimates for the same model yields the results shown in the right panel of Figure 5. This represents a substantial improvement, with none of the estimates with sample set size 10 having an error exceeding 10%. The outliers, if unlikely to occur, still have substantial errors in this case, so higher sample set sizes are needed to be more robust. Going to 20 samples, the performance is already a marked improvement compared with 100 samples without the filter. Using 100 samples in this case gives a maximum error of less than 3% and 75% of the estimates do not exceed an error level of about 1%. Some of these numbers are expected to inflate slightly if the procedure is repeated even more times, but the tendencies remain the same.

The utility of the first stage of the filter is hence clear, but what about the second stage? In the above example, the first stage was enough to ensure high performance, and the second stage did not even trigger, but for some of the models, this is not the case. In the left panel of Figure 6, results for design MI10 are shown. Here, the first stage of the filter has been applied but not the second. The performance is a lot worse than that seen previously. Even if one were to disregard the most severe outliers (with errors of about 500%-100%), there are a lot of other outliers with unacceptably high errors. If we apply the second stage of the filter for this design, the results in the right panel of Figure 6 clearly demonstrate the viability of this approach. There is a substantial overall improvement and a reduction of the maximum errors of about a factor 10. There is still an outlier with an error of about 18% when using a sample set size of 10, but the overall performance is reasonable for such a small sample set (75% of the estimates do not exceed an error of 8%, 99% of the estimates do not exceed an error of 14%). With 100 samples, the maximum error is about 7%. Clearly, this model is harder to deal with than the rest. This seems to be related to changes in the eigenfrequency, as discussed when the second stage of the filter was defined and at the beginning of this section. Again, some of the numbers reported here will be inflated when a larger number of repetitions is used.

3.2 | Importance sampling with the two-stage filter

The results for each design, using both stages of the filter and 10 000 repetitions, are shown in Figures 7 and 8. The overall performance is on the same level or better than those cases shown previously. Note in particular that the designs with randomly scaled element sizes is where the

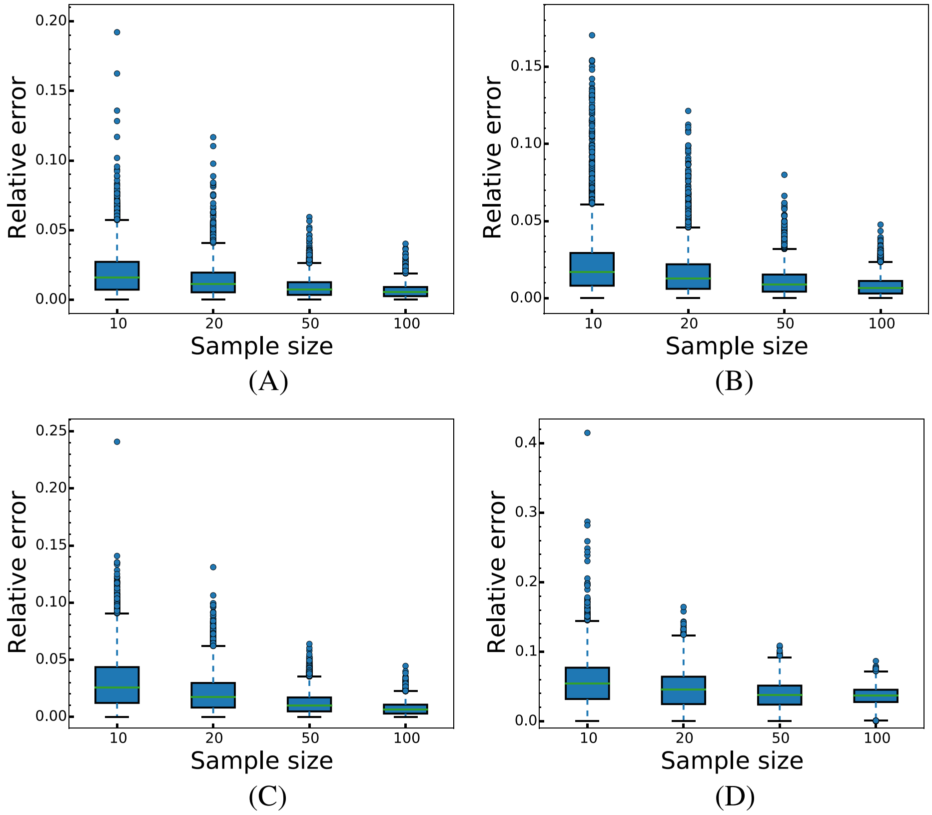

FIGURE 7 Importance sampling results with two-stage filter for systematically scaled models MD5 (A), MI5 (B), MD10 (C), and MI10 (D) [Colour figure can be viewed at wileyonlinelibrary.com]

method attains the lowest errors. Using just 10 samples (less than 0.3% of the original set, a computational reduction of more than a factor 360) is viable in several cases, at least when noting that though outliers with errors of 15%-25% may occur, they do so with very low probability. Increasing the sample set size to 50 or 100 gives satisfactory, or more than satisfactory, performance in every case, with maximum errors less than 2% for some designs. The median and expected performance (defined either as the extent of the boxes or the extent of the whiskers, depending on how strict one wishes to be) are in some cases contained within 1% error. The median error for MR5 with 100 samples is about 0.2%, and the corresponding standard deviation is also about 0.2%. With the exception of model MR5, where the errors generally are so low that the deviation seems insignificant, we observe a steady convergence of the maximum errors as the sample set size increases. While the same is also true for the median error and the variance (relatively speaking illustrated by the size of the boxes), the effect there is less dramatic. Hence, the main motivation for increasing the sample set size in most cases is the reduction of the error of the outliers. There are visible differences between the estimates for the randomly scaled models. For example, the performance for MR10 is significantly reduced compared with the other two. This is most likely due to the fact that when the changes in element sizes become large enough, the random distribution yields less protection against systematic changes (eg, if elements that are already larger are scaled one way, and smaller elements are scaled the opposite way), which introduces more changes in the underlying "true" fatigue distribution. Models MR5 and MRU10 perform very similarly, except for the single, 1 in 10 000, 14% error outlier for MRU10. This is most likely because, while the element scaling is expected to be somewhat different, neither model changes in ways that have a significant impact on the fatigue distribution.

Additionally, we note that the actual load cases sampled by the procedure follows the basic fatigue distribution fairly closely. A selection of the most severe load cases are usually included in the sample sets, and the rest of any given sample set tends to be made up of high to medium severity load cases. One visible effect of the two-stage filter in this regard is that the distribution of the sampled load cases is even more sharply peaked around the high impact load cases, nearly doubling the probability of sampling the most severe load case compared with when the filter is not used. So the additional conditioning of the first stage of the filter is behaving as expected.

4 | FURTHER DISCUSSION

4.1 | The two-stage filter

The overall results clearly show the power of importance sampling from the empirical fatigue distribution as a method for estimating the total fatigue of modified support structure designs, using only a small number of load cases. Though it is also clear that importance sampling alone, while clearly superior to conventional Monte Carlo sampling,[31] is not robust enough. The benefit of the first stage of the filter is clear from comparing the left and right panels of Figure 5. While not shown here for the sake of brevity, similar results hold for almost all the models. The exception is model MI10, where it seems the performance with and without the first stage of the filter is almost the same. Most likely, this is

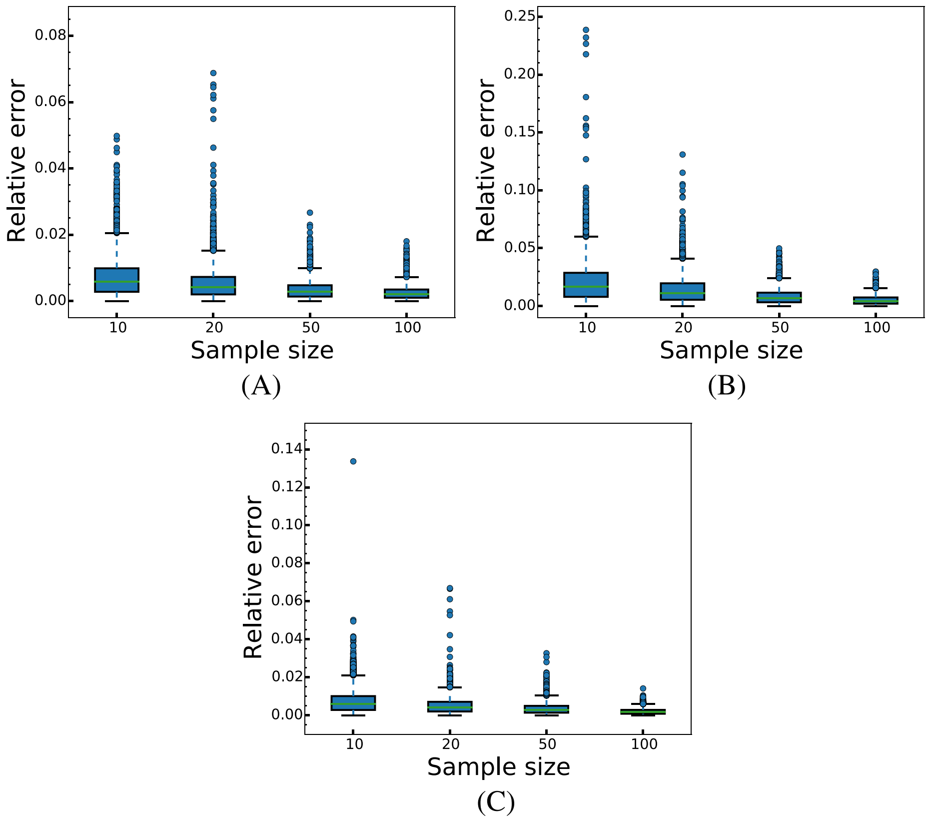

FIGURE 8 Importance sampling results with two-stage filter for randomly scaled models MR5 (A), MR10 (B), and MRU10 (C) [Colour figure can be viewed at wileyonlinelibrary.com]

because the procedure is designed to further condition the sample sets on the original distribution. Since the distribution in this case, as previously discussed, has changed significantly, the underlying problem, that the damage ratios change for only some load cases, remains unresolved.

As we have shown, the second stage of the filter is able to handle what the first stage is not, namely, the problem presented by the change in the fatigue distribution of model MI10. There were also some improvements using the second filter with model MI5 (the unfiltered case for which was not shown). This procedure is arguably a bit ad-hoc and could perhaps have been more stringently defined and/or justified, but the results do in large speak for themselves. Since the underlying problem is caused by somewhat predictable physics (changes in eigenfrequency), it would be possible to define criteria for when such a correction is likely to be needed. On the other hand, the procedure requires very little additional computation and has not been observed to modify the estimates for models where it is not needed. Furthermore, since the computational effort of the first stage of the filter is less than that of a single simulation, the total additional computational time of the two-stage filter compared with conventional importance sampling is in practical terms very small. Hence, using both stages of the filter always is a simple and cheap way of ensuring more robust results. More generally, the ideas used here is an example of how physics and engineering knowledge of the systems studied can help improve computational procedures that on paper rely more on abstract mathematical concepts.

4.2 | Overall viability of the results

While the applicability of the overall results here depends a lot on the requirements of individual situations, the method should be able to accommodate most such needs. The clear relationship between sample set size (corresponding to the number of simulations required) and error means that there is a lot of flexibility in how to use the approach. If very high accuracy is needed, eg, a maximum error of less than 1%, then perhaps several hundred samples are needed. This might seem like a comparative defeat when the suggestion of 10 or 20 samples is on the table, but the presence of strict requirements anyway likely means that the alternative is to use the whole load case set or possibly employ more conservative assumptions about the accuracy of the results. In that case, several hundred samples still represent a reduction of computational effort by a factor 10 or so. In general, it seems prudent to ask what these percentages really mean. A 5% error in the estimate for the total fatigue damage translates to predicting a lifetime of 19 or 21 years when we should have predicted 20. For a lot of applications, this is probably acceptable. Considering the many uncertainties involved in even a conventional full fatigue analysis, this error might not even be that significant. It is common practice to multiply these fatigue estimates by safety factors of 2 to 3, so an additional 5% or even 10% error would likely be subsumed into the already present uncertainty and make little difference. Alternately, since the results presented above give reasonably robust estimates for the maximum error, an additional safety factor could be implemented for those wishing to be on the safe side, eg, multiplying the estimate by 1.2 if the expected maximum error is 10% to 15%. In most cases, the results show that the risk of significant errors seems to be quite small (only a few out of 10 000), with the expected error lying safely within an acceptable regime. Especially using sample set sizes of 50 to 100 should give a high degree of confidence in the resulting estimates for most applications. In this study, we have investigated the performance of the method for monopile support structures, but it could also have been applied to other types of support structures. There might be some increased local behavior, eg, for jackets, decreasing the performance compared with what we have seen here because of differences between the local and global sampling distributions. However, we do not expect this to be a major issue. Similarly, though floating support structures need to be checked with respect to an even more comprehensive set of environmental states, the methodology should be viable for simplifying fatigue assessment for these structures as well. Furthermore, should one wish to include fatigue analyses of the rotor blades, it should be possible to augment the sampling distribution to include also one or more locations from blade structural components.

4.3 | Specific applications

The models used for testing have been intended to emulate various states likely to be encountered during design optimization. The randomly modified models are perhaps closest to what one expects to encounter during more conventional optimization loops, whereas the systematically modified models represent cases, where the initial design is either very under- or overdesigned with respect to fatigue resistance. The latter designs are also relevant for applications to preliminary design, where rough scaling of the elements can be employed to get a quick idea about system performance in various design regimes. Because the optimization itself is often time consuming, it has been a common practice to ignore the wider set of load cases when investigating optimization algorithms for support structure design, using only a few load cases for testing purposes.[23] Clearly, the presented method allows for approaches that use a small number of load cases without incurring high errors in the resulting fatigue estimates. While significant outliers were observed for large systematic design changes because of shifting eigenfrequencies, in the context of optimization, such situations could be avoided by enforcing eigenfrequency constraints. In addition to covering a wider set of environmental states than otherwise, the method could also allow for using longer simulation lengths. Even at 60 minutes of simulation time, there is a certain level of variance in the fatigue estimate because of the stochastic nature of the external loads. The large reduction in the number of simulations would make it possible to reduce this variance significantly by extending the length of each simulation, without making the analysis computationally infeasible.

4.4 | Comparison to previous work

The method presented in this study is distinct in both approach and intended application from most of the cited previous studies. The majority of previous work has focused on deriving optimal sets of load cases to sample for a single design, either as a way of simplifying single fatigue assessments or as a way of deriving optimal load case sets to be used in general (or methods for how to generally derive them). This work is, on the other hand, intended for applications involving repeated fatigue evaluations for the same overall set of load cases, given specific site data, for related designs. It is therefore difficult to directly compare the performance of this method with those used in the other studies. In particular, those other methods derive the results without having access to as much information as we do (namely, one full fatigue analysis). On the other hand, it could be that some of the approaches that reduce the number of environmental states by binning or clustering (eg, the work of Sanner et al[32]) might be combined with our importance sampling approach, simplifying the general environmental conditions before relying on a more detailed fatigue analysis for the final sample selection.

Additionally, one might want to compare the current method with methods where the structural analysis itself is simplified, rather than the set of environmental states. There is essentially a fundamental philosophical difference between our method and these approaches. Specifically, one has to choose between small, but accurate, data sets and large, but less accurate, data sets. We prefer the former. More quantitatively, a comparison with, eg, the QuLa method of Schløer et al,[33] shows that while this approach is more computationally efficient, the maximum error is larger or equal to the maximum error seen in our results when using 10 load case samples. For all but the most time sensitive, preliminary analyses, the small additional time required for a few full simulations is likely worth the gain in accuracy.

Furthermore, we should stress that the main motivation for using importance sampling as the sampling method in our approach—compared, for example, with the sampling methods used in the previous studies—is precisely because we base the methodology on having access to one full set of simulation results for every load case. Having this information makes importance sampling particularly useful, because we then have a very good approximation to the actual optimal sampling distribution. Importance sampling also compares favorably with many other sampling methods in terms of both theoretical and practical complexity. Finally, note that there is nothing particular about monopiles that motivates the use of this method. Just as this method can be applied to other structures, so too could other sampling methods be used for monopiles. It is the nature of the fatigue estimation problem in general, irrespective of support structure type, as well as the motivation of applications to design optimization, that prompts the use of this specific methodology.

5 | CONCLUSIONS

In this study, we have presented a method for simplifying the fatigue analysis of offshore wind turbine support structures. Specifically, a simplification method relevant for applications to optimization and preliminary design. The main idea being that there is a lot of information contained within the fatigue damage incurred by a small number of load cases and that, once identified, the role these load cases play in determining the total fatigue damage, represented formally by the distribution of fatigue damage over all load cases, remains approximately invariant when the design is modified. By using importance sampling of this fatigue distribution, adapted to be used for several locations in the structure, it is possible to obtain good estimates of the total fatigue damage from a small number of samples. Furthermore, by applying an additional two-stage filter, which both strengthens the conditioning of the samples on the distribution and identifies and removes severe outliers, significant improvements on the basic method was obtained. Using 100 samples (out of 3647), a maximum error of about 10% was achieved even for designs where significant shifts in the shape of the fatigue distribution occurred. For other designs, especially those closely mimicking configurations likely encountered during an optimization loop, samples set sizes as low as 10 or 20 were viable, with maximum errors of about 7% and median errors of about 1% or less. In general, the method allows for specific tailoring to individual needs, and our results provide indications for the expected maximum error that can be used to define safety factors that ensure the method could also be applied conservatively. While several similar studies into reduction of load cases have been performed in recent years, most of them differ in the purpose for which the methods are intended. Specifically, the applications tend to be deriving optimal load case sets to be used for single design assessments, without using information about the specific fatigue damage per load case for the designs in question. Hence, no direct comparison in terms of performance is relevant. On the other hand, we think this means that our approach fills a significant gap for the most part previously unexplored.

Though the overall results make the method viable for many applications, further refinements can be envisioned. For example, the development of updating procedures that use the samples obtained for each iteration (and corresponding design configuration) in an optimization loop to redefine or modify the sampling distribution. Another possibility would be utilizing the derived limitations of the approach (how the error decreases with sample set size and how the error increases for increasing overall design modifications) to obtain an adaptive scheme that sets the sample set size based on how far away the design is from the initial one and/or defines a cutoff point for the initial sampling distribution, a "radius" in configuration space beyond which the approximation to the fatigue distribution no longer is deemed to hold. The further development and application of such ideas, as well as simply using the method in a realistic design optimization study, would be relevant and interesting continuations of the work presented herein.

ACKNOWLEDGMENTS

This work has been partly supported by NOWITECH FME (Research Council of Norway, contract no. 193823) and by the Danish Council for Strategic Research through the project "Advancing BeYond Shallow waterS (ABYSS) - Optimal design of offshore wind turbine support structures."

ORCID

Lars Einar S. Stieng https://orcid.org/0000-0002-9167-0071
Michael Muskulus https://orcid.org/0000-0001-8826-8164

REFERENCES

1. International Electrotechnical Commision. Wind turbines - part 3: Design requirements for offshore wind turbines. International Standard, IEC 61400-3; 2009.
2. Chew KH, Tai K, Ng EYK, Muskulus M. Analytical gradient-based optimization of offshore wind turbine substructures under fatigue and extreme loads. *Mar Struct*. 2016;47:23-41.
3. Oest J, Sørensen R, Overgaard LCT, Lund E. Structural optimization with fatigue and ultimate limit constraints of jacket structures for large offshore wind turbines. *Struct Multidiscip Optim*. 2017;55:779-793.
4. Muskulus M, Schafhirt S. Design optimization of wind turbine support structures - a review. *J Ocean Wind Energy*. 2014;1:12-22.
5. Sacks J, Welch WJ, Mitchell TJ, Wynn HP. Design and analysis of computer experiments. *Stat Sci*. 1989;4:409-423.
6. Santner TJ, Williams BJ, Notz WI. *The Design and Analysis of Computer Experiments*. Second. New York: Springer; 2018.
7. Zwick D, Muskulus M. Simplified fatigue load assessment in offshore wind turbine structural analysis. *Wind Energy*. 2016;19(2):265-278.
8. Häfele J, Hübler C, Gebhardt CG, Rolfes R. Reconsidering fatigue limit state load sets for jacket substructures utilizing probability distributions of environmental states. In: Proceedings of the Twenty-Seventh (2017) International Ocean and Polar Engineering Conference International Society of Offshore and Polar Engineers; 2017; San Francisco, CA, USA:266-273.
9. Häfele J, Hübler C, Gebhardt CG, Rolfes R. A comprehensive fatigue load set reduction study for offshore wind turbines with jacket substructures. *Renew Energy*. 2018;118:99112.
10. Graf PA, Stewart G, Lackner M, Dykes K, Veers P. High-throughput computation and the applicability of Monte Carlo integration in fatigue load estimation of floating offshore wind turbines. *Wind Energy*. 2016;19:861-872.
11. Müller K, Dazer M, Cheng PW. Damage assessment of floating offshore wind turbines using response surface modeling. *Energy Procedia*. 2017;137:119-133.
12. Müller K, Cheng PW. Application of a Monte Carlo procedure for probabilistic fatigue design of floating offshore wind turbines. *Wind Energy Sci*. 2018;3:149-162.
13. Kim HJ, Jang BS, Park CK, Bae YH. Fatigue analysis of floating wind turbine support structure applying modified stress transfer function by artificial neural network. *Ocean Eng*. 2018;149:113-126.
14. Guanche Y, Guanche R, Camus P, Mendez F, Medina R. A multivariate approach to estimate design loads for offshore wind turbines. *Wind Energy*. 2013;16:1091-1106.
15. Choe Y, Byon E, Chen N. Importance sampling for reliability evaluation with stochastic simulation models. *Technometrics*. 2015;57:351-361.
16. Det Norske Veritas. Design of offshore wind turbine structures. Offshore Standard, DNV-OS-J101; 2014.

17. Rychlik I. A new definition of the rainflow cycle counting method. *Int J Fatigue*. 1987;9(2):119-121.
18. Det Norske Veritas. Fatigue design of offshore steel structures. Recommended Practice, DNV-RP-C203; 2005.
19. Press WH, Teukolsky SA, Vetterling WT, Flannery BP. *Numerical Recipes: The Art of Scientific Computing*. 3rd ed. New York: Cambridge University Press; 2007.
20. Kirkpatrick S, Gelatt CD Jr, Vecchi MP. Optimization by simulated annealing. *Science*. 1983;220:671-680.
21. W. Weaver Jr., Timoschenko SP, Young DH. *Vibration Problems in Engineering*. Fifth. New York: John Wiley & Sons; 1990.
22. Arany L, Bhattacharya S, Macdonald JH, Hogan SJ. Closed form solution of Eigen frequency of monopile supported offshore wind turbines in deeper waters incorporating stiffness of substructure and SSI. *Soil Dyn Earthq Eng*. 2016;83:18-32.
23. Gürgöze M. On the representation of a cantilevered beam carrying a tip mass by an equivalent spring-mass system. *J Sound Vib*. 2005;282:538-542.
24. Kim JE. On the equivalent mass-spring parameters and assumed mode of a cantilevered beam with a tip mass. *J Mech Sci Technol*. 2017;31:1073-1078.
25. Jonkman J, Butterfield S, Musial W, Scott G. Definition of a 5-MW reference wind turbine for offshore system development. *Technical Report*. NREL/TP-500-38060; 2009.
26. Jonkman JM, Jonkman BJ. FAST modularization framework for wind turbine simulation: full-system linearization. *J Phys: Conf Ser*. 2016;753(8):082010.
27. Jonkman J, Musial W. Offshore code comparison collaboration (OC3) for IEA task 23 offshore wind technology and deployment. *Technical Report*. NREL/TP-5000-48191; 2010.
28. Fedem Technology. Fedem user's guide. Release 7.2.1. Fedem Technology AS: Trondheim, Norway; 2016.
29. Fischer T, de Vries W, Schmidt B. Upwind design basis. *Technical Report*, Endowed Chair of Wind Energy (SWE) at the Institute of Aircraft Design, Stuttgart University; 2010.
30. Zwick D, Muskulus M. The simulation error caused by input loading variability in offshore wind turbine structural analysis. *Wind Energy*. 2015;18:1421-1432.
31. Stieng LES, Muskulus M. Sampling methods for simplified offshore wind turbine support structures load case assessment. In: Proceedings of the Twenty-Eight (2018) International Ocean and Polar Engineering Conference. International Society of Offshore and Polar Engineers; 2018; Sapporo, Japan.
32. Kanner S, Aubault A, Peiffer A, Yu B. Maximum dissimilarity-based algorithm for discretization of metocean data into clusters of arbitrary size and dimension. In: 37th International Conference on Ocean, Offshore and Arctic Engineering. ASME; 2018; Madrid, Spain.
33. Schløer S, Castillo LG, Fejerskov M, Stroescu E, Bredmose H. A model for quick load analysis for monopile-type offshore wind turbine substructures. *J Phys: Conf Ser*. 2016;753:092008.

How to cite this article: Stieng LES, Muskulus M. Load case reduction for offshore wind turbine support structure fatigue assessment by importance sampling with two-stage filtering. *Wind Energy*. 2019;1–15. https://doi.org/10.1002/we.2382