Author response to interactive comments on "Reducing the number of load cases for fatigue damage assessment of offshore wind turbine support structures by a simple severity-based sampling method"

Lars Einar S. Stieng Michael Muskulus

First and foremost, we wish to offer a general thanks to both referees for carefully reading and offering constructive criticism of our manuscript. Their feedback will help to improve the quality of the work and of the writing.

Please note that a version of the revised manuscript with marked up changes has been appended at the end of this document.

In the following, statements by the referees have been italicized and specific references by the authors to updated material in the revised manuscript have been written in boldface.

Response to Referee #1, Jan Häfele

This manuscript addresses a common problem in structural design of structures for offshore wind turbines, where the computational costs for FLS structural code checks are high. In general: Good work! The proposed approach is straightforward and the paper is well-written. I also believe that this work is relevant to practical applications.

Our sincere thanks to Jan Häfele for taking the time to read our manuscript and for offering some useful advice for improvements. The feedback is highly appreciated!

Page 2, line 2: "Furthermore, a fundamental assumption for this method is that the relative fatigue response to each load case remains approximately constant for an extended family of related support structure designs" - This is indeed a fundamental assumption and it is shown that it is valid under the given boundaries for the given (NREL 5MW) turbine. However, it is important to highlight that this may be invalid for a different turbine (due to severe resonance effects, for instance).

This is indeed true. The response of the structure is highly dependent on the turbine, both in terms of the loads and the overall dynamics induced by the turbine (1P and

3P frequencies for instance). One could argue that since any support structure must be designed with this in mind, any design that comes close to exhibiting severe resonance effects would likely be ruled out by other means (in optimization, constraints on allowable values for the first eigenfrequency are common). Some resonance effects were already seen in the results (when the design was scaled up by 10%) and the effect on the method was noticeable if not too severe. All in all this means that the tolerance for such effects within the method is enough that it mostly works (with some reduced efficiency), at least up to a point where the design being studied would still be viable under a more general set of criteria. All that being said, we agree that this point could have been stated more explicitly in the text.

We have added a paragraph covering this issue in section 4.1 (page 17, line 1 to 10, line 4 in the revised manuscript; page 17, line 22 to 31 in the marked up version).

Subsection 2.4: Needs (minor) improvement concerning description of load assumptions, i.e., how does your wave spectrum look like or how do you model the current?

Subsection 2.4: Can you elaborate a bit more on your "elements" or your structural model, respectively? I am actually not familiar with Fedem and I guess I am not the only one, so can you provide some more details?

These details were originally omitted for brevity, but we agree that some readers may find them of interest. Especially for the sake of reproducibility.

We have added some additional information covering these issues to section 2.4 (page 6, line 16 to page 7, line 6 in the revised manuscript; page 6, line 25 to page 7, line 3 in the marked up version).

It may increase the quality of the paper, when you use the same font style in all figures.

We have tried to be consistent with details like these where possible. However, it may be possible to improve consistency a bit more with a second look.

The font used within Fig. 1(a) (top of page 4 in the revised manuscript; top of page 5 in the marked up version) was made to conform with the font used in the other figures. Otherwise the font type (and for the most part the font sizes) should be the same for each figure.

Page 8, line 4: "has been quantified"?.

Done (page 9, line 11 in the revised manuscript; page 9, line 5 in the marked up version).

Page 14, line 27: "state-of-the-art approaches".

Done (page 20, line 8 in the revised manuscript; page 21, line 1 in the marked up version).

In your references list, try being consistent: Either "Jason Jonkman" or "Jason M.

Jonkman".

Done (page 21, line 16 in the revised manuscript; page 22, line 18 to 19 and line 24 to 25 in the marked up version).

References from DNV GL: Particularly the first one is antiquated. Take these: https://rules.dnvgl.com/docs/pdf/DNVGL/RP/2016-04/DNVGL-RP-C203.pdf (RPC203), http://rules.dnvgl.com/docs/pdf/DNV/codes/docs/2016-04/Os-J101.pdf (OS-J101).

Done (page 21, line 4 to 5 in the revised manuscript; page 22, line 4 to 7 in the marked up version).

Thank you for pointing these out.

Response to Referee #2

General comments The authors have proposed a method to decrease the number of load cases to be evaluated during fatigue analyses in wind turbine support structures. The method is interesting and could be very useful especially for optimization applications. However, some comments and suggestions are given below with the aim of clarifying the advantages and possible limitations of the method as well as to improve the quality of the document itself.

Our sincere thanks to Referee #2 for the careful reading of and comprehensive response to our manuscript. The comments will serve well to improve the quality of the paper. The feedback is highly appreciated!

P1. L11-L12. I would suggest rewriting the sentence "The method as is can be used without further modification" because it sounds like the method can not be improved and there is always a possibility of improvement.

The intention with this sentence was to convey the sense that no further improvement was strictly necessary before the method could be applied, but of course there is always room for improvements in one way or another. It certainly was not our intention to imply otherwise. We appreciate that the way it was written is not completely clear on this point.

This sentence was re-written in the revised manuscript (page 1, line 11 to 12 in the revised manuscript; page 1, line 11 to 12 in the marked up version).

P2. L6-L7. What would be the effect of considering other design situations besides the power production, such as parked conditions? I would suggest adding a short clarification about this.

Since we have not attempted this in our study, we cannot say for sure. However, it would most likely depend on how sensitive the fatigue damage in these design situations is to changes in the design. Or, specifically, whether or not the fatigue damage in these cases changes in ways that are proportional to the changes for power production. If the changes are proportional, then the method would still work as reported here. If not, then the results would be weakened, though it is hard to say by how much. We agree that, although the method has been designed specifically to reduce the very large computational effort associated with analyzing the power production DLC, a clarification of this point would be both instructive generally speaking and could also point the way for further studies.

A paragraph acknowledging and discussing this issue was added to section 4.4 (page 19, line 14 to 21 in the revised manuscript; page 20, line 5 to 12 in the marked up version).

P3. L22. Considering normal stresses means that the damage is estimated assuming under uniaxial stress states. How real is this assumption for these type of structures which are normally subjected to multiaxial stress states? What would be the effect of considering multiaxial stress states in the proposed model?

It is true that these structures are in reality subject to multiaxial stress. However, it has long been standard practice in the industry, and hence also in a lot of research aiming for industrial applications, to consider normal stress only. The main reason for this is that it allows the use of beam models and therefore a much less involved structural analysis when calculating stress and fatigue damage. To properly calculate multiaxial stress means having to use shell elements instead of beam elements in the finite element analysis, which requires more effort both on the modeling side and on the computational side. Additionally, the calculation of fatigue damage from multiaxial stress is much more involved, especially when it comes to the choice of methodology for identification and counting of stress cycles (see e.g. Stephens 2001, Metal fatigue in Engineering, for more on the multiaxial approach). Based on decades of experience from the oil and gas industry (see e.g. Lotsberg 2016, Fatigue design of marine structures) it has been seen that the normal stress approach, while not as accurate, gives reasonable (and often conservative) estimates for the fatigue damage in marine structures. While it would certainly be valuable to understand more about the behavior of the structure under multiaxial stress and how it would affect methods like the one proposed in this paper, especially since the more involved calculations in this case would be an even higher incentive to simplify the load case analysis, this falls outside of the scope of the present work. Without any data to base it on, any notions about how multiaxial stress might affect the effectiveness of the proposed methodology (e.g. more local behavior or larger dependence on directionality) would be highly speculative. However, all that being said, it would be instructive for the reader to understand the possibility of limitations of the method for applications to analysis based on multiaxial stress calculations. We will therefore make it more clear that the presented results are in principle only valid for approaches based on normal stress calculations and that the effectiveness of the method could be reduced if used with multiaxial stress analysis.

A paragraph discussing this issue and acknowledging the corresponding limitation of the method was added to section 4.1 (page 17, line 11 to 20 in the revised manuscript; page 17, line 32 to page 18, line 6 in the marked up version).

P3. L26-L27. Do the authors mean: the maximum value of the total damage among the eight points after evaluating all possible load cases? If so, make a clarification.

No. This is done per load case. For each load case, the stress is calculated at eight points around the circumference of a given location in the structure (e.g. mudline). The fatigue damage resulting from each stress value is compared and the largest one is selected to represent the fatigue value of this location in the structure, for this specific load case. We can see how the original phrasing of this was a bit unclear.

The description of this procedure in section 2 was updated to make it more clear (page 3, line 20 to 24 in the revised manuscript; page 3, line 29 to 34 in the marked up version).

Regarding Fig. 1-b Does the Normalized fatigue damage correspond to D_k/D_{tot}? If so, add clarification in the figure. How was the proportion of total load cases calculated? How Fig. 1-b would look for the different evaluated points along the tower?

Yes, the partial sums were normalized to the total sum. We agree that making this more explicit also in the figure would be instructive.

The y-axis label in Fig. 1(b) (top of page 4 in the revised manuscript; top of page 5 in the marked up version) was updated to make this more clear.

The proportion of total load cases was calculated as the number k, the number of load cases used to calculate the k-th partial sum, divided by the total number of load cases considered (in this case 3647). This could also have been more clearly stated in the figure/caption.

The x-axis label in Fig. 1(b) (top of page 4 in the revised manuscript; top of page 5 in the marked up version) was updated to make this more clear.

The equivalent curve for the two other points along the tower would be essentially identical and including them in the plot would give little or no additional information. However, this could have been noted in the text so it is more clear that we are not cherry picking the data.

The text in section 2.1 (page 4, line 10 to 11 in the revised manuscript; page 4, line 18 to 19 in the marked up version) now makes this fact explicit.

P4. L11. What does it mean "small" and "intermediate" values of k? How is that scale defined?

This wording is indeed imprecise. What was meant was the values of k for which the curve in the figure is seen to approach very closely to the asymptote. The language here was perhaps too reliant on a qualitative judgment and should have been quantified more clearly, with reference to the figure.

The relevant part of this paragraph was re-written to make the statement more quantitative (page 4, line 7 to 9 in the revise manuscript; page 4, line 15 to 18 in the marked up version).

P6. L18. How many random seeds were used for each load case in this study? What would be the effect of the number of seeds on the final number of load cases to be evaluated?

For the sake of testing this method, only one seed per 10 minute time series (with given wind speed, sea state and direction for the incoming waves) was used. To conform with standards, a minimum of six seeds (or equivalently, one 60 minute time series) per load case should have been used. However, the use of additional seeds would only serve to stabilize the fatigue values per load case against random fluctuations. If there would be noticeable effect on the results, it would hence be a beneficial one. On the other hand, this information could very well be of interest to the reader and should be stated in the text.

A few sentences covering this was added to the relevant paragraph in section 2.4 (page 7, line 13 to 18 in the revised manuscript; page 7, line 10 to 15 in the marked up version).

P6. L28-L31. Could the authors elaborate more about how was the scaling process of the element sizes carried out? Were the element sizes scaled only once or several times until the optimal solution was found?

The elements were scaled just once to obtain each of the models used for testing the method. No further scaling was done. In fact, though each model was meant to "simulate" an optimization process, no actual design optimization was done at any point of this study. Clearly our explanation of the testing setup was not clear enough about what was being done.

The relevant paragraph in section 2.4 was updated to make the description of the testing methodology more clear (page 7, line 24 to 34 in the revised manuscript; page 8, line 1 to 12 in the marked up version).

P7. L3-L6. The statement "From the distribution shown. . ..." is not clear from Fig. 2-a. In this figure, no wind speeds are shown but load cases, which are not clear either. In addition, how can be proved that the load cases with highest normalized fatigue damage are those having the highest probability of occurrence? Is there any reference or way to show this? What does it mean "Normalized fatigue damage"? If you want to show the level of severity, why are you plotting the normalized fatigue damage instead of severity level? I would suggest explaining better this figure both in the figure itself and in the text.

To your first point, the load cases were grouped according to wind speed, with the smallest wind speed to the left and the largest wind speed to the right (as stated in the figure caption). However, this could have been more clear from simply looking at the figure itself. At least, each wind speed bin could have been marked on the x-axis and the extent of each bin could have been delineated more clearly.

Figure 3(a) (previously Fig 2(a)) was updated to include more information about the load cases, both on the x-axis and within the plot itself (top of page 8 in the revised manuscript; top of page 9 in the marked up version).

The rest of your points here follow from an unfortunate error on our part which resulted from not being consistent with our own terminology. What is being plotted here is really the severity, normalized to the largest value. Hopefully this explains the rest as well. For example, the fact that the load cases with the highest severity for a given windspeed are the ones with the highest probability of occurrence is simply a direct observation that we have made when comparing the index of the peaks in each bin with the scatter diagrams used. Displaying this information explicitly in the plot would be very difficult and it was simply meant as a small observation. However, we should have been consistent with the use of the term severity also in the figure labels and captions.

The y-axis label and caption of Fig 3(a) were updated to use severity, consistent with the rest of the paper.

P8. L10. Regarding the statement "However, this turns out to not be the case.", is this statement for this specific case or in general? If it were for this specific case, what would be the consequences on the proposed model in those cases when the sampling sets are much larger than the number of load cases at each location? If it were in general, how can you prove this statement?

In a certain sense, we can of course only verify this statement as far as the loading conditions and support structure models used in the study are concerned. We do not believe there is anything particularly special about the setup in a way that would make it simplify this behavior compared to other setups, but cannot prove this in practice. If this behavior breaks down, such that the load cases selected from each location were generally not the same, then the efficiency of the method would reduce by up to a factor 3 (in this case, more generally a factor equaling the number of locations). In order to not make it seem like we were making very general conclusions from these results, we should probably have written something like "our results indicate that this is not the case." A few words about the consequences if this result breaks down for other setups could also be added in the discussion section.

The relevant sentence in section 3 (page 9, line 16 in the revised manuscript; page 9, line 10 in the marked up version) was changed accordingly.

A paragraph discussing this issue in more detail was added to section 4.1 (page 15, line 3 to page 16, line 3 in the revised manuscript; page 17, line 13 to 21 in the marked up version).

P8. L14-L16. It would be good to show Fig. 2-b for the three evaluated points. That would provide more veracity to the statement given in this paragraph.

Figure 2(b) [now 3(b)] does in fact include all three points. The horizontal axis is the number of load cases taken from each point and the vertical axis is the corresponding total size of the sampling set. In other words, selecting the, e.g., 20 most severe load cases from each point results in a total sampling set size of around 25, and so on. Hopefully this does indeed provide the desired veracity to the statement in the text. However, perhaps this could have been made even more clear in the text and/or the figure caption.

The caption of Fig. 3(b) (previously Fig. 2(b)) (top of page 8 in the revised

manuscript; top of page 9 in the marked up version) and the relevant parts of the text in section 3 (page 9, line 16 to 17 in the revised manuscript; page 9, line 10 to 11 in the marked up version) were both updated to make this issue more clear.

P9. L25-L29. Regarding the statement, "We observe that the method seems to consistently over-predict the fatigue damage. . ." What is the consequence of this? Could there be cases in which the results obtained by the method can lead to under-estimated designs (which are not desirable in any structural design)?

It is indeed possible that the method can in some cases under-predict the fatigue damage of a given design. However, this under-prediction would then only be at the reported error level. From a more practical point of view, there is of course a large difference between an error that leads to under-prediction (unsafe) and an error that leads to over-prediction (safe). Any method like this would have some amount of error, but the difference is that in our case we see that the tendency to over-predict or under-predict the fatigue damage is highly correlated with how the design has been changed compared to the reference design. It would hence be possible for a designer to apply a safety margin in the case where there would be a strong indication that there could be a slight under-prediction of the fatigue damage. We tried to explicate this fact in the discussion section (Page 12, line 3 to 7), but could perhaps have been even more clear in our discussion about these ideas.

The relevant discussion in section 4.1 was expanded and re-written for clarity (page 13, line 30 to page 14, line 3 in the revised manuscript; page 17, line 2 to 6 in the marked up version).

Regarding Figures 3 and 4. If the error can have both negative and positive values, it means that the estimated damage value could be greater than the real damage value. How could be that possible? So, how would you choose the optimal sample set size which makes a balance between the number of loads to evaluate and the final accuracy? Would it be possible to find this value by implementing a simple optimization process? How is the behavior after 180 load cases? It would be good to show more results taking into account that the real number of load cases is larger than 3000. In this way, you could show with more confidence the accuracy of the model.

The estimated fatigue value is simply a scaling of the fatigue value of the initial design. If the scale factor is too large, then the estimated damage value would be larger than the real damage.

In a certain sense, we have left the choice of "optimal" sample set size to the reader. We could have set a target accuracy and simply increased the sample set size until this was reached. However, such a target value would be entirely application specific. Some might be happy with an error of 10%, others might want an error less than 1%. Furthermore, the exact sample set size is likely highly dependent on the support structure design, turbine model and loading scenario. While we expect based on our results that sample set sizes of 30-50 would give errors of only a few percent or less (except for designs that have been altered to a large extent), giving an explicit number would limit the results to only the specific case(s) studied.

We terminated our figures at sample set sizes where the error started to flatten out. We also had in mind that for the method to be considered worthwhile, at least when compared to other methods proposed in the cited previous studies, the number of load cases used needed to be small when compared to the total number of load cases (3647). However, we do see the utility of including larger sizes, at least up to reductions of a factor of 10 (approximately 360 load cases), or even 5 (approximately 720 load cases). This would show more clearly the convergence of the results (indicating that including more load cases gives little additional reduction in error). Perhaps such extended curves would also make it easier for the reader to choose their own "optimal" sample set size.

Two additional figures showing the behavior of the method for larger sets of load cases were added as Fig. 5 (top of page 12 in the revised manuscript; top of page 12 in the marked up version) and Fig. 7 (top of page 15 in the revised manuscript; top of page 15 in the marked up version).

Furthermore, two corresponding paragraphs commenting on these figures were added to section 3.1 (page 10, line 15 to 23 in the revised manuscript; page 10, line 28 to page 11, line 2 in the marked up version) and to section 3.2 (page 11, line 3 to 5 in the revised manuscript; page 13, line 4 to 6 in the marked up version) respectively.

P9. L6-L7. Regarding the statement, "This in turn makes. . ." What would be a possible solution for this?

The most straightforward solution to this issue is to make sure that the changes in the eigenfrequency of the structure are within a certain tolerance. Specifically, that the structure does not approach regions of significant dynamic amplification for certain wind speeds (as indicated by the Campbell diagram of the turbine model). If such a situation is reached, one would either have to "restart" the method by doing a new full fatigue analysis or one would have to apply a safety factor to ensure that the increased uncertainty would not lead to unsafe designs. This was discussed in the discussion section (Page 13, lines 29-33 and Page 14, lines 1-2), but could have been more explicit in its connection with specific results.

The relevant discussion in section 4.4 (page 18, line 32 to page 19, line 9 in the revised manuscript; page 19, line 22 to 33 in the marked up version) was re-written and expanded to make it more clear and to connect it more explicitly with the results.

P10. L7-L12. Not sure how pertinent is this discussion for the purpose of this paper.

This discussion was originally meant to show that the behavior seen for designs MD5, MI5, MD10, MI10 was also seen for designs MR5, MR10 and MRU10. It was further meant to shed some light on why that might be the case, looking into the overall change in mass for these designs. However, during the revision of the manuscript it was discovered that the results shown for the latter designs were misleading in terms of this behavior.

Specifically that it did not hold for all locations in the structure. Hence, these lines needed to be changed in order to reflect the overall results more correctly. Rather than remove this part completely, we decided to shorten and change it, since it still reflects back on behavior observed for the other designs and gives some clues as to what causes this behavior. This change does not have a major impact on the overall results of the study, nor does it have a significant impact on later discussions in the paper, but some changes needed to be made in order to correct this mistake in the original manuscript.

The relevant sentences (page 10, line 34 to page 11, line 2 in the revised manuscript; page 12, line 6 to page 13, line 3 in the marked up version) were shortened and changed as referenced above.

Furthermore, some changes were made to the discussion in section 4.1 (page 13, line 31 to page 14, line 4 in the revised manuscript; page 17, line 2 to 7 in the marked up version) to reflect these differences and a few sentences were added to discuss the corrected results (page 14, line 5 to page 15, line 2 in the revised manuscript; page 17, line 9 to 12 in the marked up version).

According to this section 3.3., the level of accuracy of the proposed model could decrease considerably when many points in the structure are analyzed since the sample set size could much higher than the number of load cases at each point (i.e. n>>k). How could this limitation be controlled? This is especially important when the entire structure is analyzed under fatigue.

Here it seems we have not been clear enough in our explanations. The absolute level of accuracy does increase, it is merely that the relative accuracy does not. There are two reasons for this. One is that as ϵ_k and ϵ_k^{new} become very small for larger values of k, the relative accuracy becomes less meaningful (a relative difference of 100% at numerical values of 10^{-8} is hardly a significant error in practice). The second, and perhaps more convincing, reason is that while both parts of the fraction in Eq. 8 do tend toward zero, the denominator will tend to zero faster. This does not mean that accuracy is decreasing for higher values of k, it merely means that there is little gained for the method in general once k increases past a certain point. The convergence towards the exact answer is very slow, as each additional load case gives very little new information. All in all, this could have been more clearly written and explained in the text.

Section 3.3 (page 12, line 6 to page 13, line 15 in the revised manuscript; page 13, line 12 to 31 in the marked up version) was re-written and expanded to make the arguments more clear. Additionally, the plots in Fig. 8 (previously Fig. 5) (top of page 16 in the revised manuscript; top of page 16 in the marked up version), as well as the corresponding caption, were changed to display the results in a way that more clearly illustrate what is going on and make the arguments in section 3.3 more evident.

As for the effect of including more points, we think that having selected three points at large separation in the structure, and having shown from Figure 2(b) [now 3(b)] that the number of additional load cases needed in the sample set is fairly small, that the results obtained would indicate that there should be little additional computational effort or loss of accuracy if the entire structure is analyzed.

Regarding Fig. 5, Add the location of the point along the tower at each plot of the figure. How is the error shown in Figure 5 for a greater number of load cases, e.g. 200, 500, 1000?

The location along the tower of each subplot, though indicated in the main caption, could certainly be included in each sub-plot.

As discussed above, the relative error shown in Figure 5 [now 8] will in fact increase for a larger number of load cases. While some additional information could have been included here, the values of k selected were meant to show the behavior at the values determined most relevant from the previous results. However, to illustrate this point a bit more clearly, we could certainly add a few more rows (for example for the suggested values of k) to the figure.

The already mentioned new version of Fig. 8 (previously Fig. 5) (top of page 16 in the revised manuscript; top of page 16 in the marked up version) also includes the location in each sub-plot and a few more rows for larger values of k.

Regarding section 4.1, I would suggest analyzing the viability and the limitations of the proposed model in a general point of view instead of focusing only in the evaluated optimization methodologies (i.e. MD5, MD10, etc.). The readers might have other optimization methodologies and it would be useful for them to know when they can implement this method.

Referring back to our reply regarding the explanation of the testing setup, we would again stress that the different structural models analyzed (MD5, MD10, etc.) do not constitute different optimization methodologies, but rather different fixed states of the structure that would likely be encountered during many types of optimization contained within the relevant scope of the application (mass/weight optimization of the support structure). Of course we cannot claim to have covered everything and we could perhaps have discussed some possible limitations related to this, but overall we think that focusing precisely on the behavior of these test scenarios represent the best way to draw a more general set of conclusions about how the method might behave in different situations. Some further discussion to underline how these models represent (or do not) relevant scenarios that could motivate the reader to use the method for their own purposes will be added.

A paragraph discussing this issue was added to section 4.2 (page 18, line 1 to 10 in the revised manuscript; page 18, line 23 to 32 in the marked up version).

P12. L10 to P13. L3. Elaborate more on these statements, they are not clear as they are now.

This was meant to also refer back to some of the discussion of the results in the previous section, but it was clearly written either too briefly or not clearly enough.

The relevant paragraph in section 4.2 (page 17, line 23 to 29 in the revised

manuscript; page 18, line 9 to 17 in the marked up version) has been re-written and expanded to make the arguments more clear.

P13. L19-L23. I do see important to consider in future works the uncertainty related to the chosen number of load cases k and, even more, the one related to the final sample set size n.

This is true, though it is hard to do so without performing a rather large and comprehensive comparison of different support structures, different turbine models and different environmental data.

It would be good to add a diagram summarizing the proposed model.

A flow chart or similar that summarizes the steps involved in the method could certainly be added.

A flow chart of the basic steps in the presented method was added as Fig. 2 (top of page 6 in the revised manuscript; top of page 7 in the marked up version).

I did not find any comparison or references to previous works during the discussion.

We did not make any specific references here, since we already discussed some details of previous work in the introduction and hoped that this would still be on the reader's mind. However, we agree that this makes this subsection a bit hard to read in isolation and that making some more specific references would help make the points more clear.

References to previous studies were added in section 4.3 (page 18, line 13 to 14 and page 18, line 19 in the revised manuscript; page 19, line 4 and page 19, line 9 in the marked up version.

P1. L13. Change "a few" for "some"

Noted.

Done (page 1, line 13 in the revised manuscript; page 1, line 13 to 14 in the marked up version).

P1. L15-L18. The two first sentences (i.e. "A central practical. . .." and "In order to assess..") could be rewritten in a shorter and clearer way.

Agreed.

The sentences have been re-written (page 1, line 16 to 18 in the revised manuscript; page 1, line 16 to 21 in the marked up version).

P1. L21. Commission

Noted.

Done. Since this was a reference to an entry in the reference list (page 21,

line 14 in the revised manuscript; page 22, line 16 in the marked up version), the change has been marked there.

P3. L13-L17. I would suggest deleting this paragraph. This information is not necessary.

Such a paragraph is standard practice in many cases, but it is not essential. We take this feedback under advisement.

The paragraph was deleted (page 3, line 16 to 20 in the marked up version).

P3. L19. It is not clear what the authors mean in the first sentence. Rewrite it. P5. L2. ". . ..of some new designs of the same structure, with. . .."

Noted.

The sentence (page 3, line 14 to 15 in the revised manuscript; page 3, line 22 to 24 in the marked up version) was re-written.

The suggestion for a change in the sentence (page 5, line 2 to 3 in the revised manuscript; page 4, line 23 to 24 in the marked up version) was implemented with a slight further modification.

P6. L5-L7. Write the last sentence of this paragraph also in equations. That would make the idea clearer.

Noted.

Equations and some corresponding updates to the surrounding text were added to section 2.3 (page 6, line 4 to 11 in the revised manuscript; page 6, line 13 to 20 in the marked up version).

Regarding Fig. 2. Add a legend defining both the green points and the blue points What is the x-axis scale?

Noted.

A legend was added to Fig. 3(a) (previously Fig. 2(a)) (top of page 8 in the revised manuscript; top of page 9 in the marked up version).

The x-axis has no conventional scale, but is rather a collection of all load case indices (initially not shown because there are 3647 indices). As noted previously, we will make some changes to this figure to make this information more clear.

Changes covering this aspect of the figure have already been noted above.

P7. L3. Change "in the left panel of Fig. 2" for "Fig. 2-a"

Done, with a slight modification (page 8, line 5 in the revised manuscript; page 8, line 16 in the marked up version).

P8. L3. Make clearer which type of design is refereeing in "For each design,. . .".

Done (page 9, line 9 in the revised manuscript; page 9, line 3 in the marked up version).

P8. L4. "Specifically, the performance of the method has been quantified. . ."

Done (page 9, line 11 in the revised manuscript; page 9, line 5 in the marked up version).

Regarding Eq. 7, define the variable in the text.

Done (page 9, line 12 in the revised manuscript; page 9, line 5 in the marked up version).

P8. L11. Change ". . .in the right panel of Fig. 2." for "Fig. 2-b".

Done, with a slight modification (page 9, line 17 in the revised manuscript; page 9, line 11 in the marked up version).

P8. L11-L12. "It is reasonably linear, varying between n=7 for k=7 and n=181 for k=150."

Done (page 9, line 17 to 18 in the revised manuscript; page 9, line 12 in the marked up version).

P9. L18. "The relative errors δ (?) for various sample sizes n (?) is shown. . .."

It is indeed δ and n being referred to.

Done (page 9, line 24 in the revised manuscript; page 10, line 3 in the marked up version).

P9. L18. Are the "designs" in this line related to the "new designs" mentioned in P4. L2.?

Yes. Again, we will make some changes to the description of the testing setup to make this more explicit.

Changes related to this have been noted previously.

Regarding Figures 3 and 4. Add the name of the models (i.e. MD5, MI5, etc.) at each plot of these figures in order that each plot can be understood itself without the need for the reader to read the caption of the entire figure. Is the "Relative error" at the y-axis referred to δ from Eq. 7? If so, add δ in the y-axis, as well as the units.

Noted.

Yes, it is δ and this will be added. However, since this is a dimensionless number, there are no units.

Fig. 4 (top of page 11 in the revised manuscript; top of page 11 in the marked up version) and Fig. 6 (top of page 14 in the revised manuscript; top of page 14 in the marked up version) (previously Fig. 3 and Fig. 4 respectively) have

been updated to include design names on top and δ in the y-axis label.

P9. L3. Define "3P frequency"

Done (page 10, line 9 to 10 in the revised manuscript; page 10, line 22 to 23 in the marked up version).

P9. L9. "The relative errors δ(?)"

Yes.

Done (page 10, line 25 in the revised manuscript; page 11, line 4 in the marked up version).

P10. L9. Double "observe"

This sentence was deleted as part of a previously mentioned change.

P10. L18. Use different notations for the value shown in Eq. 3 and the actual one.

Done (page 12, line 6 in the revised manuscript; page 13, line 12 in the marked up version).

Regarding Eq. 8, define in the text all variables of this equation.

Done (page 13, line 6 to 8 in the revised manuscript; page 13, line 23 to 25 in the marked up version).

P12. L10. Change "given that" for "because".

This sentence was deleted as part of a previously mentioned change.

P13. L3. I would say "With regards to applications to design optimization, this method seems to be very promising"

Done, with minor modifications reflecting also some previous changes (page 17, line 29 in the revised manuscript; page 18, line 17 in the marked up version).

P13. L10. I would eliminate this title and add section 4.3 to section 4.2

Since section 4.2 was expanded and section 4.3 was improved according to previous comments, we think these sections work best separately.

P13. L28-L30. Elaborate more on this idea and change "(e.g.)" for ", e.g.," or ", for example,"

This idea is in fact continued in the next few sentences of the paragraph, but this could have been written in a more clear way.

This paragraph was re-written and extended to make the ideas more clear (page 18, line 32 to page 19, line 9 in the revised manuscript; page 19, line 22

to 33 in the marked up version). The specific change requested was included as part of this.

The text is full of informal language, like the ones shown below. I would suggest using a more formal language (e.g. In other words, shown, etc.).

Generally speaking, this is a matter of taste. However, there are some examples here that probably push into informal language in a way that may make the text less understandable for some. We will make the necessary adjustments to avoid this, noted below. Otherwise, the remaining text has been left unchanged except as part of previously mentioned changes.

Substantial machinery in place (P3. L19.)

This was modified as part of a previous change.

Effectively speaking (P5. L5.)

This was simply deleted (page 4, line 27 in the marked up version).

In plain words (P5. L21.)

Done (page 5, line 22 in the revised manuscript; page 6, line 3 in the marked up version).

That is to say (P7. L7.)

Done (page 9, line 4 in the revised manuscript; page 8, line 20 in the marked up version).

this turns out (P8. L10.)

This was modified as part of a previous change.

displayed (P10. L18.)

Done (page 12, line 7 in the revised manuscript; page 13, line 13 in the marked up version).

Put in another way (P13. L21.)

Done (page 18, line 23 in the revised manuscript; page 19, line 13 in the marked up version).

Appendix: Marked up version of revised manuscript

Reducing the number of load cases for fatigue damage assessment of offshore wind turbine support structures by a simple severity-based sampling method

Lars Einar S. Stieng[1] and Michael Muskulus[1]

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

Correspondence: Lars Einar S. Stieng (lars.stieng@ntnu.no)

Abstract. The large amount of computational effort required for a full fatigue assessment of offshore wind turbine support structures under operational conditions can make these analyses prohibitive. Especially for applications like design optimization, where the analysis would have to be repeated for each iteration of the process. To combat this issue, we present a simple procedure for reducing the number of load cases required for an accurate fatigue assessment. After training on one full fatigue analysis of a base design, the method can be applied to establish a deterministic, reduced sampling set to be used for a family of related designs. The method is based on sorting the load cases by their severity, measured as the product of fatigue damage and probability of occurrence, and then calculating the relative error resulting from using only the most severe load cases to estimate the total fatigue damage. By assuming this error to be approximately constant, one can then estimate the fatigue damage of other designs using just these load cases. The method yields a maximum error of about 6% when using around 30 load cases (out of 3647) and, for most cases, errors of less than 1-2% can be expected for sample sizes in the range 15-60. One of the main points in favor of the method is its simplicity when compared to more advanced sampling-based approaches. ~~The method as is~~ Though there are possibilites for further improvements, the presented version of the method can be used without further modifications and is especially useful for design optimization and preliminary design. We end the paper by noting ~~a few~~ some possibilities for future work that extend or improve upon the method.

1 Introduction

~~A central practical obstacle for researchers and designers when it comes to analyzing the performance of offshore wind turbine support structures is the large amount~~ The large number of environmental states that need to be ~~included. In order to assess the structural response to the many loading scenarios resulting from varying wind and wave conditions, simulations must be carried out for a large number of combinations of environmental states, usually called load cases~~considered for design of offshore wind turbine support structures is a significant challenge. A simulation is required for each such state, often referred to as a load case, when analyzing the response of these structures to the offshore environment. Each simulation of this kind, at least when carried out with accurate aero-elastic software, is a non-trivial task in terms of computational effort. Assessing the structural performance in the fatigue limit states for operational conditions alone typically means thousands of load cases when following relevant standards (International Electrotechnical Commission, 2009). Consequently, the computational effort needed in total presents a challenge. The increasing availability of high performance computing clusters in both the industry and at academic institutions has alleviated this issue somewhat for one-time assessments of single designs, but there are other contexts where the problem remains relevant. Design optimization (Muskulus and Schafhirt, 2014; Chew et al., 2016; Oest et al., 2017) in particular is such a case, where having to do repeated structural analyses of evolving designs means that the inclusion of thousands of load cases becomes highly prohibitive. Hence, there is a need for methods that can reduce the computational effort of these analyses, preferably without losing too much accuracy. Motivated by this need, the present study concerns itself with the development of a method that reduces the number of load cases that need to be analyzed down to a more manageable level. Though other loading scenarios are in general relevant, the present work will focus on sets of load cases encompassing the fatigue assessment of operational conditions for the wind turbine.

Several previous studies in the area of simplifying fatigue assessment through load case reduction have been carried out. Zwick and Muskulus (2016) looked at two different methods, piece-wise linear approximation and multi-linear regression, to simplify fatigue analysis for a jacket subject to 21 operational load cases. Using varying wind speeds, with a lumped sea state, the approach aimed to train the methods using fatigue data from several jacket designs and then to use them to predict the fatigue damage of other designs. With this approach, the authors obtained reduced load case sets with sizes of 3-6, with maximum prediction errors for the total fatigue damage of about 6% when using 3 load cases. One limitation with this study was that extensive training of the methods, with substantial computational effort, was required in order to obtain these results. The number of load cases studied was also small compared to the complete set of operational conditions. Häfele et al. (2017) and Häfele et al. (2018) used an approach where reduced load case sets were derived by sampling distributions for the probability of occurrence of the various environmental states, taken from a database of 2048 states. From a hierarchy of load case subsets, the authors estimated the fatigue damage for several different jacket designs. Though the errors were quite high for the smallest subset sizes, this approach demonstrated a clear potential for large reductions in computational effort. Velarde and Bachynski (2017) used a fatigue design parameter in order to select only the most important sea states for detailed fatigue assessment of a monopile.

Multiple studies of load case reduction have also been conducted for floating support structures. Müller et al. (2017) formulated an approach that combined a response surface model with Latin Hypercube Sampling and an artificial neural network. Müller and Cheng (2018) studied an approach making use of Sobol sequences in order to select the optimal load cases to sample. This lead to a more rapid convergence than would have resulted from using just conventional Monte Carlo methods. The approach achieved a maximum error of about 10% in the fatigue estimates when using reduced load case sets of 200-500 out of a total of 5400. Finally, Kim et al. (2018) used an artificial neural network to modify the stress transfer function in order to simplify fatigue assessment in the frequency domain.

While achieving various degrees of success in terms of accuracy and ability to reduce the computational effort, a common trait in most of the cited studies above are that their aims differ slightly from ours. These studies, the one by Zwick and Muskulus (2016) exempted, tried to simplify the fatigue assessments of single designs by making use of methods that were based on considerations of the environmental states alone. Whereas we aim to use also information about the actual fatigue damage for each load case of a base design and then use the combined information to develop a reduced sample set that can be used for designs that have been altered compared to this base design. Since the latter approach is highly relevant for applications like design optimization, we think the present study addresses a gap in the literature.

The method proposed in this study, like in many of the cited studies above, is based on the idea that there is a large amount of information about the total fatigue damage contained in a small subset of the load cases. Furthermore, a fundamental assumption for this method is that the relative fatigue response to each load case remains approximately constant for an extended family of related support structure designs. This makes it possible to train the method on one full fatigue analysis, using the complete set of load cases, and then use the method to propose which load cases should be assessed for future analyses of designs that have been modified. The method itself is based on sorting the load cases by their contribution to the total fatigue damage and then obtaining the partial sum of their contributions, up to a certain, smaller number of load cases. The relative difference between this partial sum and the total fatigue damage is assumed to be constant when the underlying support structure design is modified. From the corresponding partial sum of any new design, multiplied by a scale factor derived from the original relative difference, the total fatigue damage of that design can then be obtained. Hence, using an approach relying simply on sorting and summation, an estimate for the total fatigue damage based on a significantly reduced set of load cases is readily available.

~~The paper is structured as follows: After a quick review of some necessary background, the method itself will be motivated and explained. Then we define the simulation and testing setup used to demonstrate the method. In the next section, we show and explain the results. In the second-to-last section, we further discuss the results and the method and make some overall remarks about their interpretation and implications. In the last section, the findings are summarized and some overall conclusions and thoughts about continued work are presented.~~

2 Background and methodology

Even ~~with the restriction to operational~~ when restricting the area of study to operational loading conditions and fatigue analysis for the support structure, there is a substantial ~~machinery in place~~amount of work that has to be carried out in order to verify that the structure satisfies design requirements. Keeping in accordance with the standards means covering a lot of different environmental conditions (International Electrotechnical Commission, 2009) and following a specific procedure for calculating the fatigue damage (Det Norske Veritas, 2016). Every realization of wind and wave conditions corresponds to a single load case E, which has a probability of occurrence $P(E)$. After a time domain analysis of the support structure, subject to the loading conditions encoded by E, the time series of normal stress is estimated in eight different points along the circumference of each relevant location in the structure. The fatigue damage ~~for each load case, $D(E)$,~~ can be found from the stress by performing rainflow counting (Rychlik, 1987), applying SN-curves (DNV GL, 2016) for each stress range identified and then accumulating the damage by the Palmgren-Miner rule. The maximum ~~damage value~~ fatigue damage value found among the eight points along the circumference of a given location in the structure is chosen to represent ~~each~~ the fatigue damage per load case, $D(E)$, of that specific location. The total fatigue damage from all load cases, D_{tot}, during a lifetime T_{lt}, at a specific location in the structure, is then given by:

$$D_{\text{tot}} = T_{\text{lt}} \cdot \sum_{E} P(E)D(E) \tag{1}$$

A central fact to note here is that the contribution of each load case E to the total fatigue damage is determined by the product of the individual fatigue damage and the probability of occurrence. So the most *severe* load cases in the sense of having the largest contribution to the sum are in fact those where there is a balance between these two factors. Very small damage and high probability, or vice versa, tend to give smaller contributions. Whereas load cases incurring intermediate fatigue damage while also having reasonably high probability of occurrence, tend to be the most severe. This will be important below in determining which load cases get sampled. Normally, a safety factor would be applied to Eq. (1). However, since this only changes the result by a fixed constant, it has been neglected here. By the same reasoning, the lifetime scale factor T_{lt} will also be neglected from now on.

2.1 Sampling based on the k most severe load cases

From Eq. (1), we can define the k-th partial sums of the fatigue damage as:

$$D_k = \sum_{i=1}^{k} P(E_i)D(E_i) \tag{2}$$

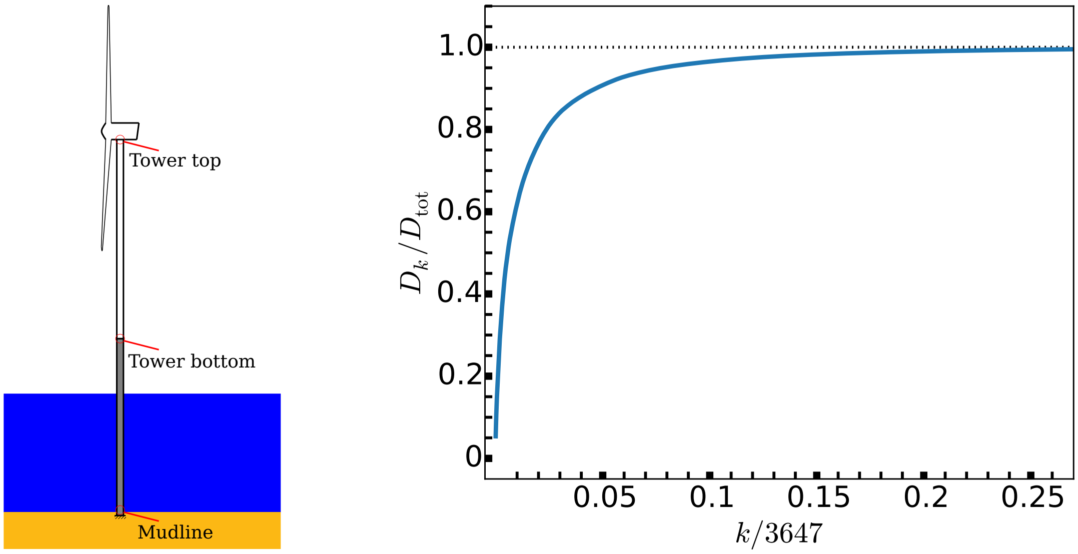
If we now let the set $\{E_i\}$ of load cases be sorted in descending order based on the size of the corresponding product of probability of occurrence and fatigue damage (from now on called *severity*), then from experience D_k should ~~get fairly~~ start to get close to D_{tot} ~~for small to intermediate~~ after values of k corresponding to only a few percent of the total number of load cases. In fact, plotting these partial sums as a function of k gives a curve like the one shown in Fig. 1(b), from which the previous statement can be confirmed. This curve was calculated using data from the tower bottom, but the corresponding curves at other locations in the structure show exactly the same behavior. Furthermore, we may define the relative difference between the sorted k-th partial sum and the actual total fatigue damage as:

$$\epsilon_k = 1 - \frac{D_{\text{tot}}}{D_k} \tag{3}$$

As our fundamental approximation, we may assume that ϵ_k is constant when the underlying support structure design is modified. That is, suppose we want to estimate the total fatigue damage $D_{\text{tot}}^{\text{new}}$ of some new ~~design~~designs of the same basic support structure, with corresponding k-th partial sums D_k^{new}. If we assume that $\epsilon_k = \epsilon_k^{\text{new}}$, then we can obtain an estimate for the new total damage as:

$$\hat{D}_{\text{tot}}^{\text{new}} = D_k^{\text{new}} - D_k^{\text{new}} \cdot \epsilon_k \tag{4}$$

The intuitive interpretation here is essentially that the new total damage is the k-th partial sum plus (~~effectively speaking,~~ since ϵ_k is always negative) an error term that should make up the difference. Some further clarity can be obtained by simplifying



(a) (b)

Figure 1. Illustration of the model used in this study (a) and a plot of the curve formed by the ~~fatigue damage~~ partial sums~~from~~, D_k, as a function of the number of load cases used, k, after sorting the load cases ~~that have been sorted~~ by their severity (b).

the above:

$$\begin{aligned}\hat{D}_{\text{tot}}^{\text{new}} &= D_k^{\text{new}} \cdot (1-\epsilon_k)\\ &= D_k^{\text{new}} \cdot \frac{D_{\text{tot}}}{D_k}\\ \hat{D}_{\text{tot}}^{\text{new}} &= D_{\text{tot}} \cdot \frac{D_k^{\text{new}}}{D_k}\end{aligned} \tag{5}$$

Hence, in practice, the estimate for the new total fatigue damage is the old total fatigue damage times the ratio of the new k-th partial sum to the old k-th partial sum.

2.2 Sampling for multiple locations

If we only wanted to know the total fatigue damage at a single location in the structure, Eq. (5) would suffice. However, there is a slight complication when the fatigue damage at multiple locations is needed. While for the most part we expect the order of the severity of the load cases to be about the same at every location, there is no guarantee that it will be *exactly* the same. Hence, using information from just a single location to decide which load cases to sample could lead to significant errors at the other locations. The simplest solution to this is to take the union of the most severe load case sets from each location. Specifically, let V_k^i be the set of the k most severe products $P(E)D(E)$ at location i. We can then define the sampling set, $\tilde{V}_k$, as:

$$\tilde{V}_k = \bigcup_i V_k^i \tag{6}$$

~~In plain words~~Specifically, we combine the k most severe load cases from each location into an expanded set (removing any duplicates), from which we then calculate the partial sums to be used in Eq. (5). It would also be possible to define the sampling set in such a way that it would have an already given size, filling up with as many load cases from the individual location sets as possible. Motivated by, for example, having certain restrictions on how many load cases one can afford to sample given the computational resources and the task at hand. However, this would result in an unbalanced set, biased towards one or more of the locations. Hence, it would be preferable to let the sizes of the individual sets determine the size of the sampling set and then simply choose a value of k such that the resulting sampling set size is acceptable.

2.3 Fatigue damage estimation procedure

By using one full fatigue assessment of a base design, we can then train our method on this data. Sorting the load cases by the severity at each location and then taking the union of the resulting sets, we obtain the sampling set $\tilde{V}_k$ for a given number of k load cases from each location. If we denote the size of the sampling set by n, the n-th partial sums at each location i of the base design, $D_n^{\text{base},i}$, combined with the corresponding total ~~damage~~fatigue damage, $D_{\text{tot}}^{\text{base},i}$, are then used to define ϵ_n^i. The total fatigue damage estimate at location i for any new design, $\hat{D}_{\text{tot}}^{\text{new},i}$, is then obtained by performing simulations and fatigue assessments for the n load cases in the sampling set, estimating the new n-th partial sums ~~and scaling the original~~ as

$$D_n^{\text{new},i} = \sum_{j=1}^{n} \tilde{V}_{k,j}^{\text{new}} \tag{7}$$

where $\tilde{V}_{k,j}^{\text{new}}$ is the set of n severity values obtained for the new design, and finally scaling the base total fatigue damage as prescribed in Eq. (5)~~.~~:

$$\hat{D}_{\text{tot}}^{\text{new},i} = D_{\text{tot}}^{\text{base},i} \cdot \frac{D_n^{\text{new},i}}{D_n^{\text{base},i}} \tag{8}$$

The procedure is summarized in Fig. 2.

2.4 Simulation setup and testing framework

For the simulations used in this study we have used the fully integrated aero-elastic software tool FEDEM Windpower (Fedem Technology, 2016). Our model is comprised of the NREL 5MW turbine (Jonkman et al., 2009) sitting atop the OC3 monopile support structure (Jonkman and Musial, 2010). The structural model was built using three-dimensional Euler-Bernoulli beam elements connected by nodes, one at each end of the elements. At each node there are six degrees of freedom and the internal forces and moments can be automatically estimated and exported for further post-processing. The monopile model was clamped at the seabed. The external wind loads were estimated from turbulent wind field time series given as input. The wave loads

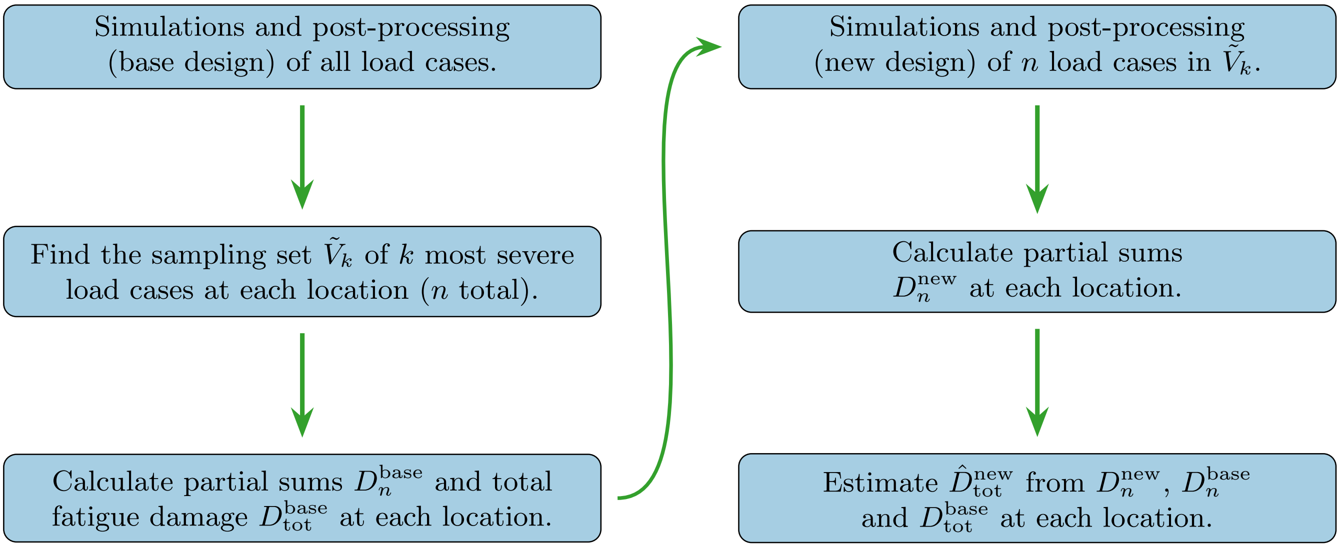


Figure 2. Step by step summary of the estimation method.

were calculated within the software itself by explicit generation of waves from a JONSWAP spectrum according to specified wave parameters and using the Morison equation with drag and added mass coefficients equal to 1.0. Marine growth was included in the model, but current was not. The load cases used in the study have been derived from the Ijmuiden Shallow Water Site wind and wave data reported by Fischer et al. (2010), giving probabilities of occurrence for different wind speeds, sea state parameters and wind and wave misalignment. The selected environmental states represent wind speeds between 4 and 24 m/s with bin sizes of 2 m/s (giving 11 different speeds) with a given turbulence intensity for each wind speed, significant wave height and peak period values depending on wind speed (between 21 and 42 different realizations for each speed) and incoming wave directions varying between 0° and 330° in steps of 30° (giving a total of 12 directions for each sea state and wind speed). 3647 load cases were used in total. One simulation of length 10 minutes (after removing initial transient data) was used for each load case, including different random seeds for each realization of wind and wave input. To make the study more tractable, only one random seed per 10 minute simulation was used, rather than the 6 seeds (or alternately using a single 60 minute simulation) usually required by standards. However, the reason for this requirement is that the fatigue damage per load case becomes more stable, i.e. less subject to statistical fluctuations, when additional seeds (or simulation time) are added. Hence, if the method can be shown to work for fatigue damage values calculated based on a single 10 minute simulation per load case, the method would certainly also work when using 6 random seeds or more. In order to test the method, three different locations along the height of the support structure, thought to be representative of different response behaviors, were selected. These include the tower top, the tower bottom and the mudline. A drawing of the model, which includes identification of the selected locations, is shown in Fig. 1.

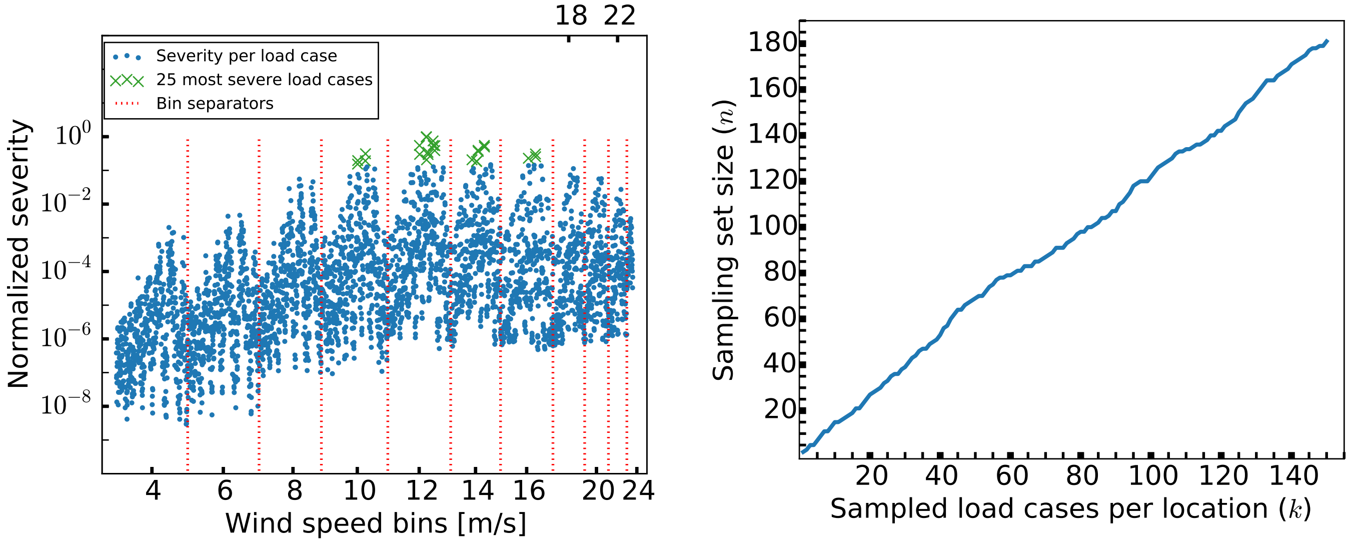
As noted previously, one of the main motivations for this study has been applications to design optimization. Hence, we have found it pertinent to test our method in a setting that would resemble situations likely to be encountered during an optimization loop. Starting with an initial support structure design on which the method is trained, how well would the method perform in predicting the fatigue damage of the modified designs encountered during the optimization? In other words, we want to see how the method performs on designs that correspond to configurations that might represent intermediate steps, or even something close to a solution, of a design optimization problem. This prompts a few different strategies for how to obtain these modified designs. First of all, the type of optimization ~~we are concerned with~~ framework we want to investigate here is mass (or weight) optimization. ~~Essentially, changing~~ In this framework, the diameters and thicknesses of various elements are changed until the design is as light as possible, while satisfying certain constraints on structural performance. ~~The designs chosen are hence ones~~ To see how the method would perform during an optimization procedure of this type, we chose designs where the element ~~sizes have~~ diameters and thicknesses had been scaled either up or down ~~,~~ compared to an original design. To represent different types of scenarios, the scaling was done both systematically across the entire structure and randomly from element to element. For each of these strategies and for two different magnitudes of scaling, the elements of the structure were scaled once according to the given strategy and a new design was thus obtained. In total, seven new designs were generated. Their names (for easy reference later) and quick summaries of how each design was scaled is given in Table 1.

3 Results

As an initial point of entry, we may ask which of the load cases are in fact the most severe for the base design and hence which ones will be sampled by the method. From the distribution shown in ~~the left panel of~~ Fig. 3(a), it is clear that the most severe load cases are clustered among just a few wind speeds. In particular, these speeds are (in order of which speed has the highest number of severe load cases) 12 m/s, 14 m/s, 16 m/s and 10 m/s. Though less clear from the plot, these load cases otherwise represent the wind and wave misalignment angles and sea state parameter values having the highest probability of occurrence. ~~That is to say~~In other words, while the severity of the wind speeds is a result of a balance between incurred fatigue damage and probability of occurrence (at the particular site used in this study, 6 m/s has the highest probability of occurrence among the wind speeds), the severity of particular wind and wave misalignment angles and sea state parameter values within a given

Table 1. The modified designs used in this study, with names and how they have been modified (scaled).

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



(a) (b)

Figure 3. Normalized ~~fatigue damage~~ severity per load case ~~(from low wind speeds to high wind speeds when going left to right)~~ at tower bottom, with load cases separated into different wind speed bins, with the 25 most severe load cases specially marked (a) and the size of the total sampling set as a function of the number of load cases ~~used per~~ included from each of the three locations (b).

wind speed bin is dominated by the probability of occurrence. The analysis here is based on data taken from the tower bottom, but completely analogous conclusions can be drawn from the two other locations.

For each support structure design listed in Table 1 and the unaltered base design, a full fatigue analysis was performed (that is, not just for the load cases selected by the method) in order to be able to quantify the performance of the method. Specifically, the performance of the method ~~as~~ has been quantified in a way similar to Eq. (3), using now the relative difference, δ, of the estimate and the true value for the total fatigue damage of each design:

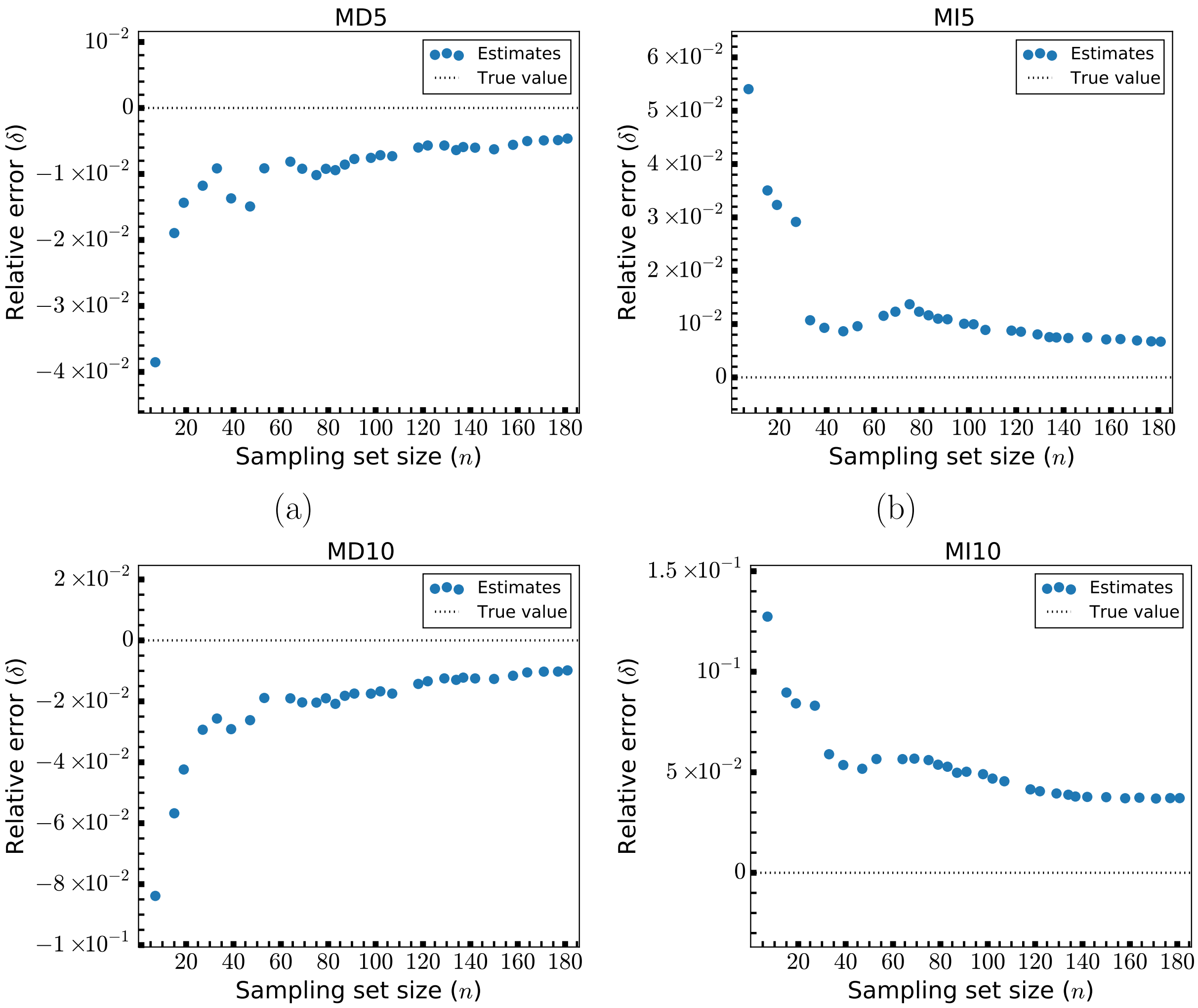
$$\delta = 1 - \frac{\hat{D}^{\text{new}}_{\text{tot}}}{D^{\text{new}}_{\text{tot}}} \tag{9}$$

One concern might be that there are large differences in the order of the severity for the load cases in each of the three locations. This would in principle lead to sampling sets that are very large compared to the number of load cases selected per location. However, ~~this turns out to not be~~ our results indicate that this is not the case. A plot of the size of the total sampling set as a function of the number of load cases selected ~~per~~ from each location is shown in ~~the right panel of~~ Fig. 3(b). It is reasonably linear, varying between ~~7~~ $n = 7$ for $k = 5$ and ~~181~~ $n = 181$ for $k = 150$. Hence, as an approximation, k can be said to be fairly close to the actual number of sampled load cases, at least for smaller sample sizes. Finally, for the sake of not showing data that yield little additional insight, the results ~~displayed~~ shown below are in each case taken from a single location only. Specifically, for each design, the location with the maximum error was chosen to represent the behavior of all three locations within a given plot. In practice, the chosen location is usually either the tower bottom or the mudline, since the behavior at tower top seems generally more favorable.

3.1 Uniformly scaled designs

In Fig. 4, the relative errors~~for various sample sizes~~, δ, for various sampling set sizes, n, is shown for the four uniformly scaled designs (MD5, MI5, MD10, MI10). Except for in the case of MI10, the estimates fairly quickly converge to a level of roughly 2% error or less. For MD5 and MI5 this level of accuracy requires 20-30 samples (a reduction in the load case set by more than a factor 100), whereas for MD10 it takes about 50 samples to reach this level (though at 30 samples the error is no more than 3%). For MI10, the convergence is slower and the error is generally a bit higher. In this case, the error level is at around 6% or less after 30 samples, goes below 5% at around 100 samples and then slowly tends toward 4% or less for the larger samples sizes. The maximum error encountered is at about 13% for MI10 and is otherwise less than than 10% for the other designs. In other words, for the the first three designs, errors of about 4-8.5% are attainable using only 7 load cases. We observe that the method seems to consistently over-predict the fatigue damage (giving negative errors, see Eq. (9)) when the design has been consistently scaled down and under-predicts (giving positive errors) when the design has been scaled up. Inspecting Eq. (5) we may surmise that this means that for down-scaled designs the proportion of the fatigue damage in the k-th partial sum has increased compared to the base design, whereas for the up-scaled designs this proportion has decreased. The overall convergence is not quite smooth, presenting some occasional jumps in the estimation error. These jumps are ultimately quite small (usually at no more than a single percentage point) and are likely signs of small instabilities in the method for reduced sample sizes. In these cases, the sudden inclusion of certain additional load cases (with the effect of either improving or decreasing performance) can have a visible effect on the overall estimate. As for why MI10 seems to under-perform when compared to the others, this is likely because the changes to the global eigenfrequency induced by scaling all elements by 10% can lead to dynamic amplification for lower wind speed load cases when the frequency increases (corresponding to the structure being scaled up). In this particular situation, there is a significant shift towards the 3P frequency of the turbine, defined as the rotation speed-dependent frequency at which any of the three blades passes by a fixed point, as seen from the Campbell diagram of the NREL 5MW turbine (Jonkman and Jonkman, 2016). The result is a significant increase in the severity of lower wind speed load cases, which means that the error in including only the most severe load cases in the fatigue estimation changes more drastically for this design. This in turn makes the method less accurate than for the other designs, where the changes in fatigue damage are more uniformly distributed among the load cases.

In Fig. 5 the same results are shown for some selected larger sets of load cases, including also some smaller sets of load cases for reference. The accuracy of the method keeps increasing as the number of load cases used increases, but the gain in accuracy for each additional load case (as indicated by the slope of the curve) decreases drastically after a certain point. After 183 load cases, corresponding to a reduction of the number of load cases by a factor 20 and corresponding to the smallest error shown in Fig. 4, the benefit is fairly minor. While a further reduction in error by one order of magnitude can be achieved by the use of 730 or 911 load cases, the error at 183 load cases is already small enough that the cost in additional computational effort is likely prohibitive. The exception is again for model MI10, where the convergence is much slower and the errors generally



(a) (b)

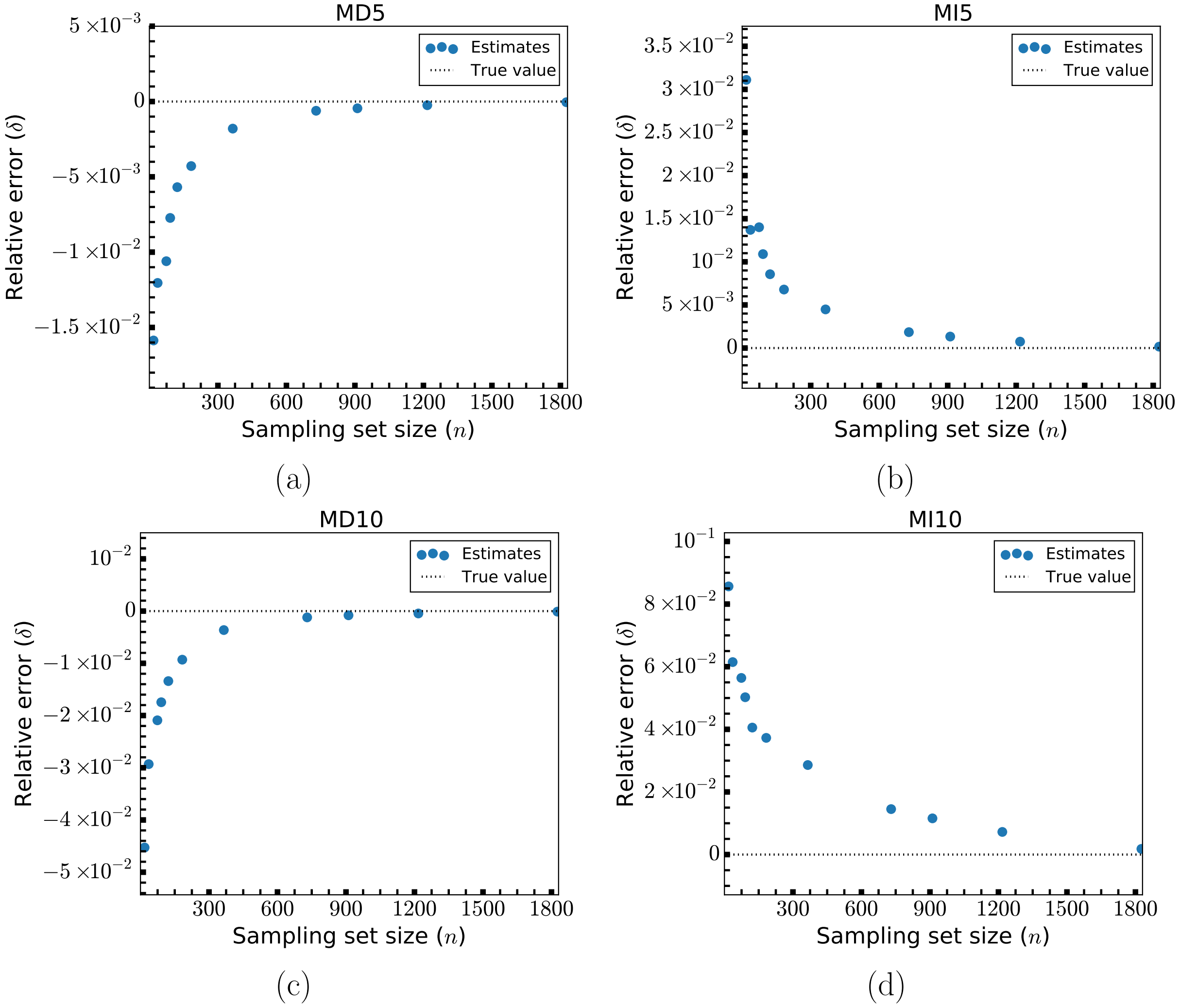
(c) (d)

Figure 4. Relative errors, δ, of fatigue estimates for models MD5 (a), MI5 (b), MD10 (c) and MI10 (d).

higher. Going from 183 load cases to 730 (a reduction factor of 5 compared to the full set of load cases) takes the error from around 4% to around 1.5%.

3.2 Randomly scaled designs

The relative errors, δ, in the estimates for the randomly modified designs (MR5, MR10 and MRU10) are shown in Fig. 6. These all generally show improved performance compared to the uniformly modified designs. Except for the smallest sample estimate for each model, every estimate has an error of less than 2%. For MR5 and MRU10, errors of no more than 1% occur with sample sizes of no more than 35-40 (a reduction of the load case set by a factor of about 100). MR10 crosses this



(a) (b)

(c) (d)

Figure 5. Relative errors, δ, of fatigue estimates for larger sets of load cases; models MD5 (a), MI5 (b), MD10 (c) and MI10 (d).

same error threshold at around 50-60 samples. There is in general a reasonable convergence behavior for all three models. MR10 exhibits marginally higher errors than the two other models. This could be because element scaling of $\pm10\%$ could lead to a higher degree of overall uniform changes than in the other cases. Since each element in the structure has a different size, one would expect a certain bias towards either overall decrease or increase when the scaling is done randomly from a uniform distribution. The larger the scaling, the larger the resulting bias. In fact, inspecting the changes to the overall mass for these models, MR10 has a bias twice as large as MR5. ~~More interestingly, and in agreement with the previous results, we observe observe a correspondence between overall down- or upscaling and the tendency to either~~ However, this bias does not consistently lead to over- or ~~under-predict~~ under-prediction of the fatigue damage ~~. Both MR5 and MRU10 consistently under-predict the fatigue damage and it turns out that these models both have a slight bias towards an overall increase in size.~~

~~The reverse holds for MR10.The overall change in mass for MR10 is actually very close to that of MD5 and indeed the results for these two models are very similar~~like it did for the uniformly scaled models. While the results shown here makes it look like the behavior is the same as before, this is not the case for all the points in the structure.

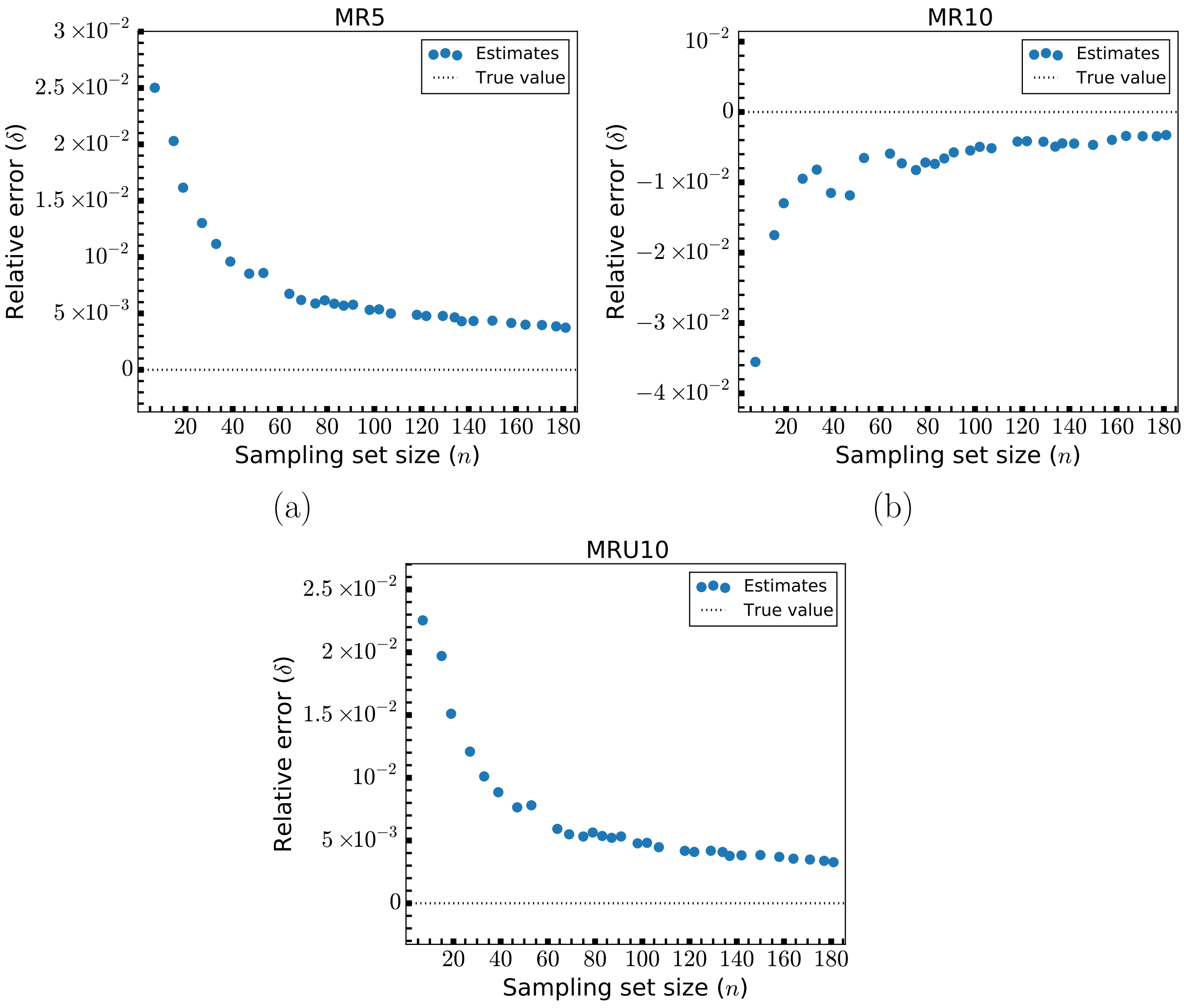
In Fig. 7 results for larger number of load cases for the randomized models are shown. As was similarly noted for Fig. 5, the benefit of increasing the number of load cases beyond what was shown in Fig. 6 is small when considering the speed of convergence of the error and its overall order of magnitude at that point.

3.3 Real behavior of ϵ_k

When we initially defined the method, it was based on a basic assumption: That the relative error when using only the k most severe load cases would remain approximately constant under modification of the support structure design. The results shown so far indicate that this is indeed the case, but this should be verified explicitly in order to have confidence in the theoretical basis of the utilized methodology. To investigate this, we have calculated the absolute value of the relative difference between the value of ϵ_k (as defined in Eq. (3)) for the base design, ϵ_k^{base}, and the actual value of ϵ_k for each modified design, ϵ_k^{new}. This is ~~displayed~~ shown, together with the respective values of ϵ_k as heatmaps in Fig. 8, where the color of each cell indicates the absolute value of the relative difference. There is generally very good agreement between the values of ϵ_k for each design, though again there is some more deviation for design MI10. This is presumably for reasons similar to why the estimation method had larger errors in this case. The attentive reader might notice that unlike above, the relative differences do not decrease for increasing values of k. If anything they seems to either fluctuate or increase. ~~The reason for this is that as the number of load cases sampled increases, the numerical value for ϵ_k decreases~~However, as can be seen by inspecting the numerical values, the absolute differences do decrease. Hence, ~~the absolute differences are certainly decreasing. Though more specifically~~while the relative differences increase, the actual numerical values of these differences become quite small and therefore less relevant in practice. Furthermore, it is not hard to show from Eq. (3) and Eq. (5) that:

$$\left|1-\frac{\epsilon_k^{\text{new}}}{\epsilon_k}\frac{\epsilon_k^{\text{new}}}{\epsilon_k^{\text{base}}}\right| = D_{\text{tot}}{}^{\text{base}} \cdot \frac{\left|1-\frac{D_{\text{tot}}^{\text{new}}}{\hat{D}_{\text{tot}}^{\text{new}}}\right|}{D_{\text{tot}}-D_k}\frac{\left|1-\frac{D_{\text{tot}}^{\text{new}}}{\hat{D}_{\text{tot}}^{\text{new}}}\right|}{D_{\text{tot}}^{\text{base}}-D_k^{\text{base}}} \tag{10}$$

Where D_{tot} refers to the real total fatigue damage, D_k is the k-th partial sum of severity products as defined in Eq. (2), the ˆ refers to an estimate made using the method introduced in this study and the superscripts "base" and "new" refer to the initial design and any modification of this design respectively, as above. Both the numerator and the denominator ~~decrease~~ tend to zero as k ~~increases~~tends to the total number of load cases, so the actual behavior depends on the convergence of the method (controlling the numerator) compared to the proportion of the total fatigue damage in a given partial sum D_k (controlling the denominator). Essentially, one can roughly compare the convergence shown in Fig. ~~4~~ 5 and Fig. ~~6~~ 7 to that shown in Fig. 1~~.~~ ~~Since~~ (b). In other words, since the denominator converges faster than the numerator, the relative difference in Eq. (10) will tend to increase for increasing values of k. A practical consequence of this, which was also noted previously, is that the benefit of increasing k, i.e. including more load cases in the estimate, becomes very small after a certain point. Additionally, since the



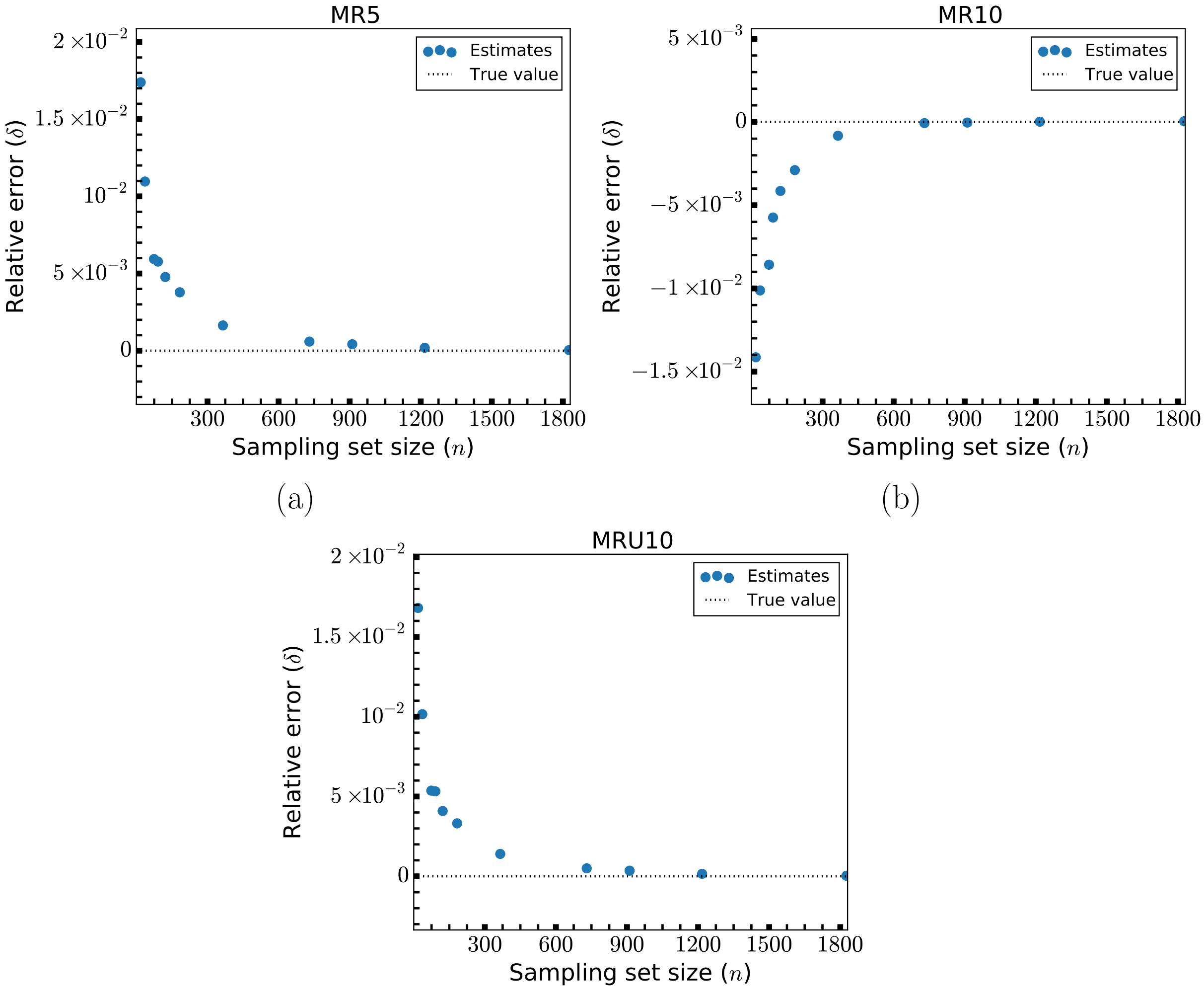
(a)

(b)

(c)

Figure 6. Relative errors, δ, of fatigue estimates for models MR5 (a), MR10 (b) and MRU10 (d).

convergence of the fatigue estimates for MI10 was particularly slow, the behavior seen in the heatmaps for this design at both tower bottom and mudline (a significant increase for increasing k) seems reasonable.



(a) (b)

(c)

Figure 7. Relative errors, δ, of fatigue estimates for larger sets of load cases; models MR5 (a), MR10 (b) and MRU10 (d).

4 Further discussion

4.1 Viability of the method

As seen above, the proposed method is able to predict the total fatigue damage of the modified designs with a high degree of accuracy. With the exception of design MI10, all estimates eventually converge towards errors of 2% or less (in some cases much less) and with drastic reductions in the load case set (factors of 50-200 in most cases). Even for the case of MI10, where the error is about 4-6% for all but the smallest sample sizes, this result is quite convincing in terms of the level of accuracy that can be expected for such an approach given the extent of the modifications to the structural models. In fact, higher accuracy

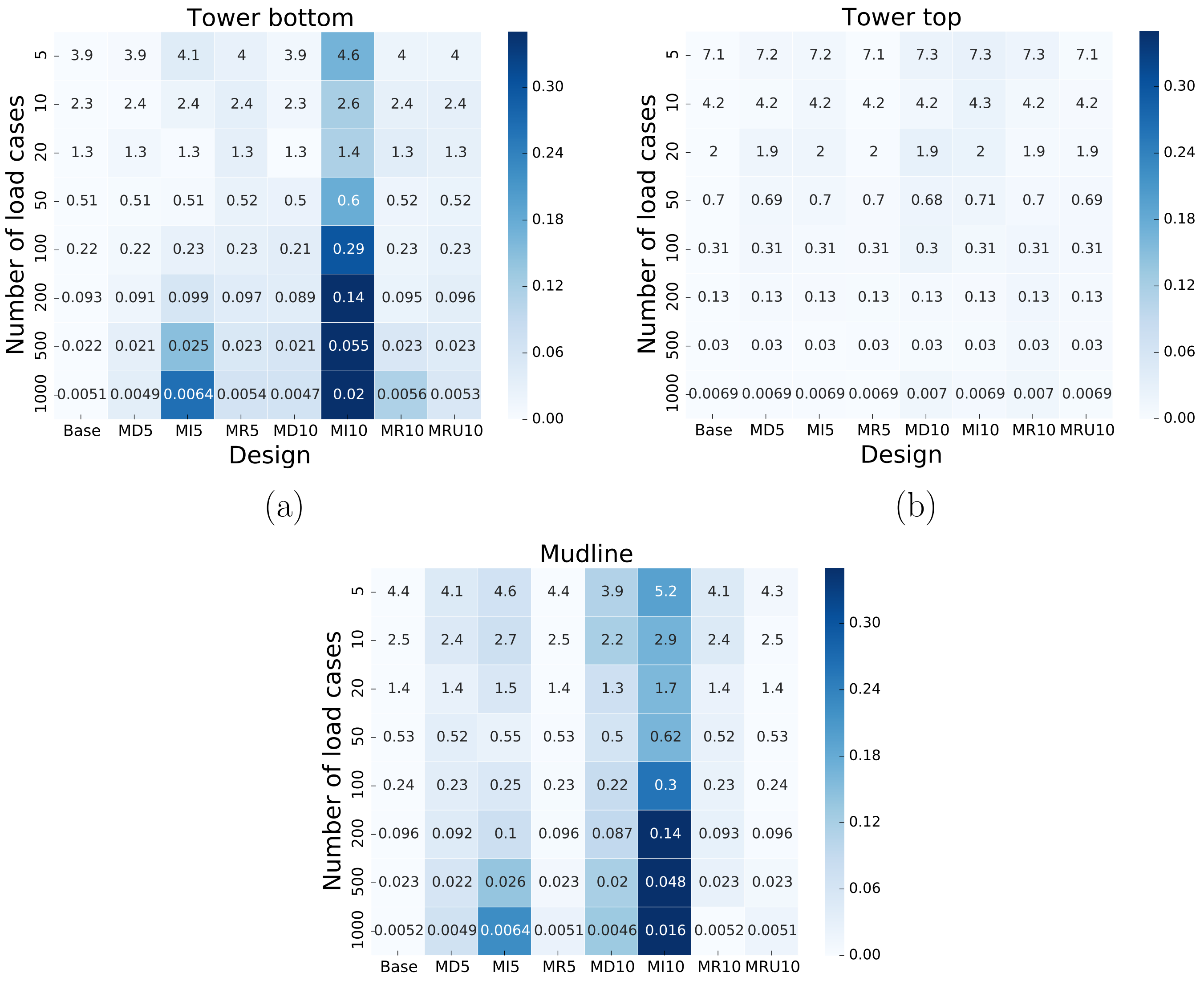


Figure 8. ~~Relative differences between the value~~ Values of ~~ϵ_k~~ ϵ_k for the base design and each additional design, for selected values of k. The colors of each cell are set by the magnitude of the relative difference between the base design value and the value in the given cell. At tower bottom (a), tower top (b) and mudline (c).

than that reported for design MI10 might not even be required. A 5% error in the prediction of total fatigue damage represents a change in the lifetime of a support structure by 1 year if the real expected lifetime is 20 years. This is certainly within the range of other types of errors one might expect in terms of uncertainties in the modeling or the environmental conditions, both of which are usually accounted for by multiplying the total fatigue damage by partial safety factors of 2-3. In such a framework, errors on the order of 10% might even be acceptable, in which case a very large load case reduction is possible for all models. Additionally, there seems to be a clear connection between ~~overall systematic~~ consistent changes to the size (mass) of the ~~structure~~ structural elements and whether the estimates for the fatigue over- or under-predict the true value. In fact, the two properties are directly correlated. Though in practice the consequence of under-predicted fatigue damage is much more severe than over-predicted fatigue damage, the fact that the correlation is as visible as indicated in Fig. 4 and Fig. 5 means that it is possible to correct for this effect. Keeping track of ~~such overall changes then~~ systematic changes in the structure hence makes it possible to account for the errors in the estimates in correspondingly systematic ways. For instance, if the estimate is known to be an over-prediction, then it may be deemed ”safe” in a conservative sense and in the opposite case one might want to add in a small safety factor. However, the results for the randomly modified designs indicate that overall changes to the structural mass is not enough to account for this behavior. Correspondence between the overall changes and the changes to the elements where the fatigue damage is calculated is also important. Hence, some care must be taken when attempting to correct for consistent over- or under-prediction of the fatigue damage.

One of the reasons the method is as efficient as it is when analyzing more than one location in the structure, is the behavior seen in Fig. 3(b): That the number of load cases in the sampling set does not increase significantly when considering all three locations. Since the three locations chosen are so far away from each other, located at each end of the support structure and around the middle of the structure respectively, we do not expect that the addition of even more locations should make a significant difference. However, we do note that in the worst case scenario where there is no overlap between the set of the k most severe load cases at each of the l locations with that of any of the other locations, the size of the sampling set would be $k \cdot l$, drastically reducing the efficiency of the method. Though our results indicate that anything close to this behavior is unlikely, at least for any monopile support structure, some attention should be paid to ensure similar performance if applying the method to other types of support structures.

The methodology has been shown to be quite effective for a range of different support structure designs, but there is one limitation that should be noted: The presented results were all obtained while using the same turbine model. The turbine model will have a very important impact on global dynamics, e.g. 1P and 3P frequencies and the total system mass, damping and stiffness, and it is hence likely that changing turbines would induce changes in the fatigue distribution that could be challenging for the method to handle. For example if severe resonance effects are encountered. On the other hand, a support structure design is usually constructed with a specific turbine model in mind and the results have shown that the method can handle significant changes to global dynamics to a certain extent (as seen for model MI10 in Fig. 4). Furthermore, severe resonance is hardly desirable in any case and such designs would likely be ruled out by other means. Hence, while we recommend that the method be trained for use with only a single turbine model at a time, as long as the impact on global dynamics is not too significant, the method could still be viable for related turbine models within somewhat relaxed error criteria.

Another possible limitation, at least for some applications of the method, is that the results here have been derived using only normal stress. On the one hand, this is standard practice in the industry and therefore also for many research applications. Furthermore, the methodology has for some time been seen to give fairly accurate (often conservative) fatigue estimates for applications in the oil and gas industry (see e.g. Lotsberg (2016)). On the other hand, there are certainly cases where multiaxial stress is important to consider. However, the procedure required to account for this is quite involved. Calculating multiaxial stress requires the use of shell elements rather than beam elements. This makes the modeling and time-domain analysis much more complicated than what has been done in this study. Additionally, the estimation of fatigue damage from multiaxial stress is also more complicated and less standardized, in particular the cycle counting (see e.g. Stephens (2001)). Hence, we consider the effect of multiaxial stress to be outside the scope of this study and would therefore advise caution when using the method for such applications.

4.2 Applications to design optimization and preliminary design

One of the most discernible outcomes of the testing framework is the indication that the method works best for designs that have been randomly modified~~. This is understandable given that the method is based on an assumption about how~~, as seen when comparing Fig. 6 with Fig. 4. As noted previously, and especially evident for design MI10, systematic changes in the structure will to a larger extent cause changes in global dynamics that decrease the performance of the method. Specifically, the method relies on proportional changes in the fatigue damage ~~changes that works best when there are no major changes in structural behavior~~across all load cases. As we have seen, this property is sensitive to, e.g., changes in eigenfrequency. Random changes ~~will to a larger extent leave the eigenfrequencies ,~~ to the structure have a much smaller impact on the eigenfrequencies and other global phenomena that are expected to skew the fatigue damage distribution across all load cases~~, unaltered. With regards to applications to~~. Since random changes more closely resemble the configurations most relevant for design optimization, ~~this is very promising~~ the method seems very promising for this application. While it can occur that large systematic changes result from an optimization loop, e.g. if the original structure is significantly over- or under-designed with respect to fatigue resistance, most of the computational work in most cases will occur in stages where the overall changes to the structure are small. One can certainly also envision applications of this method to preliminary design, where perhaps a larger extent of the work is in rough scaling of the design. In most cases, the errors reported here are small enough also for these design situations. Even the larger errors reported (in the case of 10% up-scaling) might be acceptable in the early phases of design.

The various design configurations that were used to test the method were chosen in an attempt to cover as many scenarios of interest as possible. However, not all types of scenarios could be accommodated and hence there are some configurations about which we cannot make strong conclusions. The most obvious of these is the fact that we have scaled both diameters and thicknesses by the same factor. Even for the randomized designs, the scale factor was only randomly sampled on an element-wise basis. A situation where either diameters are increased and thicknesses decreased or vice versa, could easily occur in practice during design optimization. On the other hand, based on our result, it seems that the most significant factor in determining the effectiveness of the method is whether or not there are global changes in eigenfrequency. Hence, though we are unable to explicitly confirm this based on our results, we expect that even in configurations like that described above (or other potential untested ones), the method should be viable under the same criteria: As long as there are no global changes that induce non-proportional changes in fatigue damage for only a certain subset of load cases, the method performs well.

4.3 Comparison with previous work

Comparing the approach taken in this study with most previous work on load case reduction, certainly the studies cited in the introduction of this paper, one of the main advantages is the simplicity of the method. Because most of the other studies (e.g. Häfele et al. (2018) and Müller and Cheng (2018)) have slightly different aims, i.e. reducing the number of load cases for single design situations, it is not necessarily sensible to compare directly the achieved accuracy for a given amount of load case reduction (though if one were to do so, it would be a reasonably favorable comparison). Something similar could be argued in terms of the methodology, that such a simple approach is only possible in the current setting, but we would still stress the overall simplicity as a major reason why this method would be useful. Especially the avoidance of more advanced statistical and computational procedures (like in Müller and Cheng (2018) and Kim et al. (2018)), will likely make this approach more appealing for industrial applications. There is also little reliance on software, requiring only the ability to sort the fatigue data and then create sampling sets where duplicate load cases have been removed. Furthermore, we note that since the method is completely deterministic (as opposed to many sampling-based approaches), there is little or no uncertainty in the results reported here. ~~Put another way~~In other words, while the specific results (say whether k samples gives an error of exactly $x\%$) are tied to specific background details of the study (the models used, the load case data, etc.), if the method gives a certain accuracy for a certain set of data, it will always give this accuracy for that data.

4.4 Possible continuations

The simplicity of the method might also suggest the possibility of improvements, at least in some of the scenarios shown. While some attempts at applying sequence acceleration techniques were made, with little or no positive effects (hence why this was not shown), it is certainly possible that such approaches, or similar ideas, might decrease the error of the estimates or at least decrease the number of samples needed to reach a certain level. We additionally note that further ideas for how to apply the method for specific applications could also be developed. For example, since systematic design modifications of a certain size can impact the accuracy of the method, as seen especially for design MI10 in Fig. 4, it would be possible to apply ~~it~~ the method in an adaptive way for~~(~~, e.g.~~) optimization. On the one hand,~~ , optimization. One could argue that such adaptive strategies are not necessary, since it is often possible to avoid such ~~situations~~ inaccuracies by enforcing eigenfrequency constraints. However, if it is known a priori that certain changes in the eigenfrequencies can decrease the performance of the method due to dynamic amplification for some wind speeds, then one ~~can~~ possible adaptive strategy would be to implement a check for this situation which when triggered ~~has~~ would have an effect on how the method ~~is utilized. While several other solutions are possible, one could, when such large changes are detected,~~ was utilized. When such large global changes to the structure would be detected, one could for example either increase the number of samples used or perhaps require a new full analysis to update the data used to train the method. In other words, such an adaptive strategy would define a kind of "safe" region of design configurations in which the method could be applied very accurately (somewhat analogous to trust region methods in mathematical optimization, see e.g. Nocedal and Wright (2006)) and would change the way the method was applied whenever the design was no longer in this region. One can also envision other types of applications, where something other than (or at least not exclusively) the design is modified. For example probabilistic design/reliability analysis, where the statistical behavior under the variation of a set of input parameters is investigated. While this would have to be verified in a separate, future study, one can envision the method being employed in a similar fashion as here: Training the method on a base parameter configuration and then reducing the number of load cases needed for fatigue assessment when the parameters are allowed to vary.

One limitation of the results obtained in this study is the fact that only operational loading conditions (power production) were analyzed. Since many other conditions are relevant for design, it would pertinent to ask whether the method could be extended to these cases as well. Based on the results obtained here, it seems clear that the effectiveness of the method in these other scenarios would depend on whether or not the fatigue damage also in these cases changes proportionally when the design is modified. If this property still holds, then most likely the error level when using only the k most severe load cases would still be approximately invariant and the method should work fairly well. If this property does not hold, the accuracy of the method could be significantly reduced. Investigating the performance of the method for other types of load cases would be an interesting continuation of the present study.

5 Conclusions and outlook

In this study we have presented a simple approach for reducing the number of load cases required for accurate fatigue assessment of an offshore wind turbine support structure under operational conditions. By making a simple assumption about the relative error incurred by only using the most severe load cases in the total fatigue sum, specifically that this error remains approximately constant as the design is modified, we are able to make accurate predictions for the fatigue damage of a set of seven modified designs. One key part of the method is that the ordering of the severity of each load case is slightly different from location to location. Hence, we have used the union of the reduced sets at each location to form a total sampling set that is used in the method. While slightly increasing the number of samples needed, this has a significant impact on the overall performance in terms of balancing the accuracy at each location in the structure. The overall results of the method are very promising, achieving errors of a few percent or less for sample sizes of 15-60, depending on how the designs have been modified. Only in one case, where the increased dimensions of the design caused significant changes in the eigenfrequency and subsequent dynamic amplification for some wind speeds, were the errors a bit higher. Though still in this case less than 6% for comparable sample sizes. Considering that even a sample size of 100 means a reduction of the load case set (initially numbering 3647) of about a factor of 36, the method generally allows for very large savings in computational effort for fatigue assessment. The method is particularly effective for designs where modifications have been made randomly from element to element, achieving errors of less than 1% for reasonably small sample sizes. This in particular, though also the overall performance, makes the method useful for applications to design optimization. The fact that the method seems to consistently under- or over-predict the fatigue damage based on whether the design has been ~~scaled more up than down or vice versa~~ consistently scaled up or down even makes it possible in some situations to further correct the estimates in order to ensure that the method is always conservative.

One clear advantage compared to ~~state of the art~~ state-of-the-art approaches for load case reduction, aside from the overall accuracy, is the simplicity of the method. Whereas the most common approaches rely on various types of sampling techniques that require some amount of statistical and computational complexity, our approach relies entirely on sorting, the union of small sets (combining and then discarding duplicates) and basic arithmetic. Aside from the overall attractiveness of such simplicity, this makes the method more useful for applications in industry where complex methodologies can lead to unacceptable bottlenecks in the work flow. The simplicity of the method presented in this study (on both a conceptual and implementation level) could also be attractive for other scientists, who may not be as comfortable with advanced sampling methods.

While the method as is can readily be applied in many settings, some future developments can be envisioned. For example, one could study possibilities for improving the convergence of the estimates or investigate specific ways of applying the method to design optimization that adapts to regimes where the estimates are expected to lose accuracy. A future study might also look into whether, or to what extent, the method could be extended for use within a probabilistic design or reliability framework. In practice, this would mean seeing whether the fundamental assumption of the method, the invariance of the relative fatigue estimation error when sampling only the most severe load cases, also holds when parameters other than those related to the structural dimensions are altered. Finally, the performance of the method for other support structure types (jackets, floating support structures, etc.), other turbine models and other loading scenarios (other than power production) are all open questions for future work.

Code and data availability. The data used for plotting the figures, and corresponding python scripts to make the plots, are available as supplementary material. The raw fatigue data is available upon request. The underlying raw simulation data is too large to distribute.

Competing interests. The authors declare that they have no competing interests.

Acknowledgements. 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”.

References

Chew, K.-H., Tai, K., Ng, E. Y. K., and Muskulus, M.: Analytical gradient-based optimization of offshore wind turbine substructures under fatigue and extreme loads, Marine Structures, 47, 23–41, 2016.

Det Norske Veritas: ~~Fatigue~~ Design of Offshore ~~Steel~~ Wind Turbine Structures, ~~Recommended Practice~~Offshore Standard, ~~DNV-RP-C203, 2005.~~ DNV-OS-J101, 2016.

~~Det Norske Veritas~~DNV GL: Fatigue Design of Offshore ~~Wind Turbine~~ Steel Structures, ~~Offshore Standard~~Recommended Practice, ~~DNV-OS-J101, 2014.~~ DNV-RP-C203, 2016.

Fedem Technology: Fedem User's Guide, release 7.2.1. Fedem Technology AS: Trondheim, Norway., 2016.

Fischer, T., de Vries, W., and Schmidt, B.: Upwind Design Basis, Technical Report, Endowed Chair of Wind Energy (SWE) at the Institute of Aircraft Design, Stuttgart University, 2010.

Häfele, J., Hübler, C., Gebhardt, C. G., and 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, pp. 266–273, International Society of Offshore and Polar Engineers, San Francisco, CA, USA, 2017.

Häfele, J., Hübler, C., Gebhardt, C. G., and Rolfes, R.: A comprehensive fatigue load set reduction study for offshore wind turbines with jacket substructures, Renewable Energy, 118, 99–112, 2018.

International Electrotechnical ~~Commision~~Commission: Wind turbines - Part 3: Design requirements for offshore wind turbines, International Standard, IEC 61400-3, 2009.

Jonkman, J. and Jonkman, B.: FAST modularization framework for wind turbine simulation: full-system linearization, Journal of Physics: Conference Series, 753, 082 010, 2016.

Jonkman, J. and Musial, W.: Offshore Code Comparison Collaboration (OC3) for IEA Task 23 Offshore Wind Technology and Deployment, Technical Report. NREL/TP-5000-48191, 2010.

Jonkman, J., Butterfield, S., Musial, W., and Scott, G.: Definition of a 5-MW Reference Wind Turbine for Offshore System Development, technical Report. NREL/TP-500-38060, 2009.

~~Jonkman, J. M. and Jonkman, B. J.: FASTmodularization framework for wind turbine simulation: full-system linearization, Journal of Physics: Conference Series, 753, 082 010, 2016.~~

Kim, H.-J., Jang, B.-S., Park, C.-K., and Bae, Y. H.: Fatigue analysis of floating wind turbine support structure applying modified stress transfer function by artificial neural network, Ocean Engineering, 149, 113–126, 2018.

Lotsberg, I.: Fatigue design of marine structures, Cambridge University Press, New York, 2016.

Müller, K. and Cheng, P. W.: Application of a Monte Carlo procedure for probabilistic fatigue design of floating offshore wind turbines, Wind Energy Science, 3, 149–162, 2018.

Müller, K., Dazer, M., and Cheng, P. W.: Damage Assessment of Floating Offshore Wind Turbines Using Response Surface Modeling, Energy Procedia, 137, 119–133, 2017.

Muskulus, M. and Schafhirt, S.: Design optimization of wind turbine support structures - a review, Journal of Ocean and Wind Energy, 1, 12–22, 2014.

Nocedal, J. and Wright, S. J.: Numerical Optimization, Springer, New York, Second edn., 2006.

Oest, J., Sørensen, R., Overgaard, L. C. T., and Lund, E.: Structural optimization with fatigue and ultimate limit constraints of jacket structures for large offshore wind turbines, Structural and Multidisciplinary Optimization, 55, 779–793, 2017.

Rychlik, I.: A new definition of the rainflow cycle counting method, International Journal of Fatigue, 9, 119–121, 1987.

Stephens, R. I.: Metal Fatigue in Engineering, 2nd edn., 2001.

Velarde, J. and Bachynski, E.: Design and fatigue ana lysis of monopile foundations to support the DTU 10 MW offshore wind turbine, Energy Procedia, 137, 3–13, 2017.

Zwick, D. and Muskulus, M.: Simplified fatigue load assessment in offshore wind turbine structural analysis, Wind Energy, 19, 265–278, 2016.