

Dear authors,

Thank you for your contribution that covers a really interesting topic.

The paper addresses the important subject of fatigue load assessment for floating offshore wind turbines. It proposes a probabilistic, sampling based approach instead of the state-of-the-art bin based approaches. For this sampling based approach, an advanced quasi-random sampling is proposed that is supposed to lead to better results than plain Monte Carlo sampling.

Starting from set of more than 10000 time domain simulations with different environmental conditions, sampled from three correlated statistical distributions, subsets with smaller numbers of simulations are selected, and the resulting uncertainty is analysed. It is concluded that 100 to 200 samples are sufficient to reach satisfying accuracy.

Although the topic is really interesting and the paper is mostly well written, some modifications or clarifications are needed. And an assessment of the performance of the current approach should be added (see remark 23).

- 1) P.2, l.19: Ocean currents are not really relevant for fatigue, as they induce nearly no dynamic loads. Hence, leave it out, if you do not have a reference that proves the relevance of currents.
- 2) P.2, l.21: Although, there is no sensitivity analysis for the considered structure, studies for other structures could be a good starting point. For example, you chose your three environmental conditions, based on some knowledge (“most relevant parameters” p.2, l.30). Hence, it would be good to cite a sensitivity analysis e.g. [1].
- 3) P.2, l.29: You correctly state that a high number of simulations is needed due to the “curse of dimensionality”. However, here, you do not say anything about the additional number of random seeds that are needed (later on, you use 3 seeds). However, the stochastic nature of wind and waves increases the number of simulations. This effect can be significant and should be mentioned here.
- 4) P.3, l.28-34: You say that there are two approaches (improved sampling and surrogate modelling), and “some experience with both approaches has already been established in the past” (p.3, l 28). However, the literature mentioned in the following focuses nearly completely on surrogate models, while your work uses improved sampling. There should be some references to work concerning improved sampling as well, even if the references might not cover exactly the problem of fatigue of floating wind turbines (e.g. [2])
- 5) P.7, l.1: Some more information on the simulation setup would be nice, e.g. turbulent wind model, wave spectrum etc.
- 6) P.10: Figure 6 (left) and Figure 5 seems to be not consistent by cutting the wind speed distribution at cut-in wind speed.
- 7) P.10, l.3: You are using “only” 3 seeds. This is not a lot, if you compare it, for example, with the results in [3, 4]. However, you can argue that you are, firstly, using 1h simulations, and secondly, the use of a probabilistic, sampling approaches inherently includes different seeds (for sampling points with similar parameters). However, this should be briefly discussed here. Otherwise, the reader could claim that 3 seeds are not sufficient.
- 8) P.11: Eq. 9: You introduce this quantity as “your” DEL. However, it might be really confusing for the reader that this is NOT a DEL, but an DEL^m . This also changes the unit to e.g. $(kNm)^4$. Either rename it, or redefine it and stick to “real” DELs. This is really important, as you switch between

real DELs and DEL^m throughout the paper. This is confusing and also introduces the problem of different uncertainties in Section 4.2.

- 9) P.12: Figure 7: Units of the DELs are missing.
- 10) P.12, l.5: How many bootstrap samples do you use?
- 11) P. 13, l. 1: You state that you reach a fast convergence, as a 10% error margin is obtained by using 140 samples. However, later on you discuss correctly that this error is valid for DELs and not for damages or lifetimes. In the end, the designer is mainly interested in damages of lifetimes. Your 10% error margin leads to 40-50% errors in the damage for $m=4$, or even errors far above 100% for $m=10$. You briefly discuss this topic later on, but it has to be mentioned here, as it somehow relativises the results.
- 12) P.14: Figure 10: Are the units correct? If you are showing “your” DEL (Eq. 9), then you have other units. And add units to all subplots in Figure 10.
- 13) P.14, l.14: You state that there is no significant impact of the wave period. Later on, you discuss the effect of the wave period. This contradicts each other. Hence, reformulate the statement.
- 14) P.15, l.19: Sample based approaches consider this effect by definition. What about state-of-the-art bin based approaches? It should be possible to include these effects by introducing wave period bins. However, it is obvious that sampling based approaches are beneficial here.
- 15) P. 15: Figure 11: Check units of “your” DELs again.
- 16) P. 16: Figure 12: You cut off a region of high wave heights and low wave periods. This is correct. But still, it seems as if there are still quite high waves for small periods (e.g. 6.5m with 4s). These are quite special conditions that you normally do not see (e.g. FINO data [5], or standards [6] $11.1\sqrt{H_s/g} = 9.0s$). Are these conditions actually present at the considered site (in the raw data), or is it only due to the defined correlation?
- 17) P. 17, l.8: “over-weighting” sounds as if this weighting would be wrong. Perhaps “high” and “low weighting” is better.
- 18) P.17: Eq. 10: Again, this is not really a DEL, but DEL^m . Either rename it, or redefine it.
- 19) P.18, l.9: “Figure 14 shows the accumulation of normalized DELs from equation (10) towards the full sum of DELs as described in equation (9)”. You should probably say that you are starting with the sample with the lowest damage. (see remark 21 as well)
- 20) P.18, l.18: You say that a large DEL (actually DEL^m) can change the sum of DELs (DEL^m) by 20%, but the lifetime DELs (real DEL) only by 5%. However, if we are going back to damages or lifetimes, we have again 20%. Hence, the 20% are the more important number (or the 45% in line 20, compared to 10%). This should be clarified.
- 21) P.19: Figure 14 (left): The figure is not directly self-explaining. I would add the highest DEL first (and not the lowest) so that you have a large increase for a small number of DEL indices. Furthermore, it might be helpful to rename the horizontal axis, e.g. “data proportion” or “sample proportion”.
- 22) P.19, l. 15: 100-200 samples are only valid for these three environmental conditions. This should be mentioned here.
- 23) In general: A comparison of this approach with others is missing. If you want to demonstrate the performance of this approach, you should either compare it to the state-of-the-art binning method or to plain MCS. For example, you could add another figure, comparable to Figure 8, showing the convergence of MCS and of your low-discrepancy sequence approach for samples up to 200. Such a comparison would really make clear the advantages of your work.

Editorial remarks:

- 1) P.2, l.2: The sentence is incomplete.
- 2) P.5, l.9: $n=N$? Be consistent with the nomenclature
- 3) P.6, l.12: Space after “2017)”
- 4) P.7: The text in Figure 3 is barely readable. Improve quality of this figure.
- 5) P.10, l.15: Remove the second space after “from the”.
- 6) P.10, l.16: Don’t use the $2e6$ notation.
- 7) P.12: Figure 7: Be consistent with your notation in the whole paper. Either use 1×10^4 or $1 \cdot 10^4$, not both.
- 8) P. 16: Figure 12: Please improve readability.

References

- [1] Hübler, C., Gebhardt, C. G., & Rolfes, R. (2017). Hierarchical four-step global sensitivity analysis of offshore wind turbines based on aeroelastic time domain simulations. *Renewable Energy*, 111, 878-891.
- [2] Choe, Y., Byon, E., & Chen, N. (2015). Importance Sampling for Reliability Evaluation With Stochastic Simulation Models. *Technometrics*, 57 (3), 351–361.
- [3] Zwick, D., & Muskulus, M. (2015). The simulation error caused by input loading variability in offshore wind turbine structural analysis. *Wind energy*, 18(8), 1421-1432.
- [4] Häfele, J., Hübler, C., Gebhardt, C. G., & Rolfes, R. (2017). A comprehensive fatigue load set reduction study for offshore wind turbines with jacket substructures. *Renewable Energy*, 118, 99-112.
- [5] Scott, G. (2007). Applications and Analysis of Offshore Wind and Wave Measurements. In 52nd IEA Topical Expert Meeting Wind and Wave Measurements at Offshore Locations, Berlin. https://www.ieawind.org/task_11/TopicalExpert/52_Wind_Wave.pdf
- [6] Guideline for the Certification of Offshore Wind Turbines. Germanischer Lloyd Rules and Guidelines, IV—Industrial Services, Part, 2005