Discussion of:

Data-driven probabilistic surrogate model for floating wind turbine lifetime damage equivalent load prediction

06 March 2025

The manuscript presents an application of a probabilistic surrogate modelling technique with Mixture Density Networks for prediction of floating wind turbine site-specific fatigue damage accumulation. The paper is well written and addresses relevant scientific topics such as probabilistic surrogate modelling and floating wind turbine design assessment. That being said, the scientific novelty does not become evident from the current content of the manuscript. I suggest that a significantly revised paper clearly establishes where the novelty is and focuses the narrative on the novel aspects. Please find some more elaboration in the comments below:

General comments

1) The manuscript discusses quite similar topics and uses similar methodologies as another recent manuscript by the same main author (https://doi.org/10.5194/wes-9-1885-2024, also cited in this paper). Please discuss what is the distinct methodological novelty of the present paper (see also my next comment).
2) As with most surrogate modelling approaches, this one is specific to the model which has been used to run the simulations. This limits the direct applicability of the trained model to just the turbine configuration in question. As a result, the primary scientific contribution of a surrogate modelling paper is normally in the methodology (or some specific findings from the results) rather than the end product. I recommend that the authors clarify what is the methodological contribution in this paper, or highlight some important findings that warrant the publication.
3) Bayesian Neural Networks (BNNs) are another approach to train a heteroscedastic model without the need of making repetitions. The authors may want to mention this and cite e.g. Hlaing et al. (https://doi.org/10.1177/14759217231186048)
4) I am missing a discussion section, which may include thoughts on the limitations of the current study.

Specific comments

5) References style: many references seem to introduce repetitions, such as e.g., " Zhu et al. (Zhu and Sudret, 2020) on line 65. If the authors use the \citep command to refer to a paper, they don't need to repeat the author names in the text as they come automatically from the LaTeX command.
6) Simulation time of 600s seems quite short for floating wind with low-frequency response. This may affect especially the estimation of higher-order moments of the response and may be important for this study which explicitly considers higher-order statistics.
7) Section 3.2.1: The authors suggest the R-squared between the mean and the standard deviation predictions of the distribution as a goodness-of-fit metric. This limits the representativeness of the

comparison as it doesn’t allow comparing higher distribution moments. Also, the R-squared is not sensitive to bias. The other metric proposed by the authors, the Wasserstein distance, is not limited in this way. Is the R-squared then redundant? Results shown in Table 9 may hint at that, since it is only the dw2 that flags the tower bottom FA model as having worse performance than the other three channels.

8) Page 23, line 428: the authors state “Including these additional uncertainties in the feature set would likely increase the variance of the final load estimates.”. I agree with this, but I think some nuance needs to be added. The uncertainties that are propagated through the fatigue model will not necessarily increase the variance of the short-term outputs, they may introduce bias in the long-term mean which will manifest as an uncertainty in the long-term (aggregated) statistics of the outputs. For example, assuming higher annual mean wind speed will lead to a bias in the mean estimate of total accumulated fatigue damage.