"Interannual variability in fatigue damage estimation from short-term strain monitoring of offshore wind turbines" (https://doi.org/10.5194/wes-2025-173)

In this work, binning-based approaches to extrapolate strain measurement-based short-term fatigue damage estimations for offshore wind turbines to longer periods of interest (for which only EOC data is available) are evaluated. For this purpose, for the first time, long-term measurements (here: eight years) are used. Furthermore, in contrast to the state of the art, a clear differentiation between year-to-year uncertainty and uncertainty due to limited sampling is done.

As many offshore wind turbines start to reach the end of their design lifetime, investigating the performance and uncertainty of fatigue lifetime estimations based on strain measurements is very valuable. The methods used in this work are not new. However, due to the much longer measurement period, new insights were gained, which are of great interest to the readers of WES and probably also to industry.

The paper is well written and nicely structured. There are only a few aspects that could be clarified or where a more comprehensive discussion would increase the added value of this work.

1) In line 43/44 you discuss that neglected EOCs or non-stationary distributions of covered EOCs are the reasons for the year-to-year variability. Perhaps, you can add that structural changes, modified control algorithms, etc. might also lead to changes over several years.
2) Several times you discuss that previous work does not differentiate between year-to-year uncertainty and limited sampling uncertainty. It is correct that so far this differentiation has not been made explicitly. Especially, results like the one in Fig. 14 in your work have not been presented so far and are very valuable. However, it is not fully correct that previous work has not addressed this topic. For example, Hübler & Rolfes (2022) also applied bootstrapping and a sliding window approach to identify separately both types of uncertainty.
3) Section 2 gives a nice overview of the state of the art and research gaps. However, there is a lot of overlap with Section 1. You might think about writing the introduction more concisely.
4) Typo in Fig. 2: It should be "Compute bins' estimators"
5) Eq. 6: You only apply the bin filling if there is no measurement in the bin. This is state of the art. Nonetheless, bins with only one or two measurements do not yield very robust estimates. Therefore, just out of interest: Have you tested to increase the minimum number of measurements in a bin?
6) Section 3.3.1: In this section, you introduce θ as a "general" damage estimator. However, in the entire work, you only use the mean value. Hence, it would be much easier for the reader to understand if you just stick to μ and remove θ from the paper. I know that θ was originally used in Sadeghi et al. (2024), but for the current work, I do not see the benefit of introducing it in Eq. 12 and then "removing" it again in Eq. 14.
7) Fig. 3: If you have time steps of 90 days, the second year is not actually $M^{(5)}$ but starts about 5 days earlier (360 days vs. 365 days). How exactly do you handle the remaining 5 days of the year?
8) Fig. 4: Are the precise positions of the strain gauges confidential or can they be provided? I assume that the strain gauges are evenly distributed every 60°?
9) Fig. 6: If I am not mistaken, the "Geometrical information" is already required in step "Convert the sensors' 10-minute stress …".
10) L. 344/345: You state a hierarchy for the conditional binning models. Have you checked the relevance of these EOCs or why do you use exactly these EOCs and this order. If I remember correctly, your group has analysed this or something similar before. Perhaps, you can refer to this work. Otherwise, the selection/order remains quite arbitrary. Similar applies to the bin sizes. For the wind speed, you analyse the effect of the bin size. However, how did you choose the bin size for the other EOCs? And why did you only investigate it for the wind speed?

11) Fig. 8: Perhaps you can briefly discuss the outliers for 0D and 1D-Case for year 4, 5, and 6. I assume in year 5 and 6, the missing data during winter is responsible. Similarly, in year 4, a lot of data is missing in summer. Hence, the higher damage during winter is overweighted.
12) Table 1: It seems as if high-dimensional binning approaches start to become slightly biased estimators, i.e., $\bar{e}_{norm}$ is no longer "close" to zero. I assume that is due to empty bin filling. Perhaps, you can discuss this briefly.
13) Fig. 9: It is very interesting to see that measurement periods longer than one year improve the results, as previous research sometimes revealed no further reduction for periods above 9 months. A reason could be that previous work mainly looked at measurement periods up to 12 months, where an increase from 9 to 12 months did not have an influence. Hence, previous work never considered two winter periods. Probably, you have not analysed this, but (perhaps for future research) it would be interesting to check whether the improvement is actually due to long measurement periods or more data during winter.
14) Fig. 10: Is there a reason, why you use the mean value here? For the other box plots, you used the median. Changing this within the paper might be confusing for the reader.