Author response to reviewer 1

The authors response is shown in red

We thank the reviewer for the valuable comments and suggestions, which we consider very important and help us to sharpen and improve the manuscript. Here our response to each comment.

The manuscript describes the calibration of two of the three model components that constitute the DWM model and the comparison of the full model result to wake flow measurements. Data base for calibration are measurements from a SpinnerLidar in the wake field of one of the V27 wind turbines at the SWIFT facility. Three different scanning strategies are applied to collect different kind of flow field representation. The total amount of collected scans comprises roughly 142 hours that are sorted in three different atmospheric stability classes. By means of Bayesian inference the model parameters of two of the model components of the DWM model, the deficit in the meandering frame of reference and the wake-added turbulence model are calculated. The methodology allows for a calculation of an uncertainty of the parameters established from a probability distribution. The authors show and discuss the results of the parameter fitting also in comparison to other calibration studies of the DWM model. The final evaluation is conducted by comparing the full DWM implementation with the calibrated submodels to a subset of measurements. The manuscript is in general well written and the different steps of the analysis are well documented and explained. The content is relevant to the scientific discussion of using lidar measurements for model validation/calibration, thus I recommend to publish the manuscript after minor revisions.

My main criticism of the manuscript is the lack of discussion of the generalisation of the results. The authors compare their results of the wake deficit modelling with previous studies and claim e.g. that the model parameters in the IEC standard lead to conservative wake predictions. This might be the case in this specific campaign, but I am missing the evidence that the results are fully transferable from the 190 kW V27 to multi-MW wind turbines. This would imply a fully non-dimensionalized problem. It is not the fault of the study that this can not be proved on a single wind turbine, but the generalisation of the results should be handled with more care. I would have expected that discussion at the end of the manuscript.

We now address this point in a new paragraph in the discussion section, where we point out that further work should evaluate whether the obtained results (for a small-sized turbine) are transferable to modern size rotors. As mentioned by the reviewer, this study cannot answer this point explicitly, however, it provides good insights on the quality (and requirements) of the adopted datasets for ensuring reliable and robust calibration of engineering wake models.

This study shows that the model parameters in the IEC standard lead to conservative wake predictions for turbulence levels above 12%. We now specify this result in the abstract, while it was already mentioned in the conclusions. It also discusses why calibrating wake models using power production data (as done in the IEC model) may lead to inaccurate representations of the wake deficit for varying inflow wind conditions. Further, it shows that the IEC-based wake deficit recovery is basically invariant to ambient turbulence raising from 7% to 16% (see Fig. 9). This is not what we observe from measurements, as shown by the lidar data analysis in this study, and also discussed in the recent work of Reinwardt et al. (2020) [1].

Further, the DWM model is an engineering wake model based on simplistic flow modeling assumptions. One of these assumptions is that the calibration parameters (e.g., k_1 and k_2) are universal and independent of the turbine size and inflow wind conditions [2]. However, this assumption might not be valid, as the reviewer mentioned.

Furthermore the final validation of the results compares the calibrated model with the measurements that are used for calibration (even if these are technically not the same measurements as measured in a different period). Thus, it represents rather a verification of the calibration than a model validation.

We called "validation" as we compare model predictions directly with field observations. Also, the validation/comparison of the DWM model-predicted two-dimensional wind speed and turbulence profiles in the FFoR against field observations was not carried out in previous studies.

We rephrased the heading of Sect. 6 from "Validation of the DWM model" to "Validation of the DWM model in the FFoR", in order to specify that we are validating the DWM model-based wake flow field predictions in the FFoR against the lidar data.

Further Comments:

1. Title - As part 2 of this study is only mentioned in the very last sentence of the manuscript and as it includes further measurements, my suggestion is to remove Part I from the title

Part I is now removed as suggested.

2. L10 - We demonstrate that Please edit according to the previous comments on the generalisation of the study

We rephrased the sentence as: "We show that the current DWM-model parameters in the IEC standard lead to conservative wake deficit predictions for ambient turbulence intensities above 12% at the SWiFT site."

3. L42 - load responses mostly Please rephrase.

We rephrased it as: "These three components are presumed to affect wind turbine loading conditions".

4. L46 - still under judgement - still to be assessed.

This has been rephrased as suggested.

5. L71 - and at the DTU ...

We added "the" as suggested.

6. 3.2.1 Atmospheric stability - Can the authors please explain the benefit of sorting the measurements into stability classes for the analysis. as the model has no dependency to atmospheric stability, but just to TI. Why not just sort by TI?

The DWM model's wake deficit formulation in Eq. 1 does not depend on the atmospheric stability but on TI only. However, as the SWiFT campaign provides a comprehensive dataset that allows for characterizing atmospheric stability, we opted for a more detailed analysis of the ambient wind conditions. Such analysis might also be useful for future studies on wake modeling that account for stability effects using the SWiFT dataset.

Further, the DWM model formulation of the wake meandering depends on the spectral properties of the ambient turbulence. Indeed, both TI and the turbulence length scale influence the intensity of the wake meandering. Thus, classifying inflow conditions according to atmospheric stability is necessary when studying meandering dynamics [3].

Finally, atmospheric stability and ambient turbulence are highly correlated, as shown in Table 1, and the adopted classification does not influence the resulting calibration parameters.

7. L249 - T is not a variable in equation (5), virtual potential temperature needs to be introduced

The reviewer refers to the following equation:

$$L = -\frac{u_*^3 T}{kg\overline{w'\Theta_v'}}, \tag{1}$$

where T is the mean surface-layer temperature, and Θ_v' is the virtual potential temperature. We added "virtual" in the text that was previously missing. In the manuscript, we use the definition of the Obukhov length of Peña et al., (2010) [4] - Eq.(3).

8. 262 - this is sufficient ... Why is this sufficient?

We deleted "sufficient" and rephrased the sentence as: "The dataset collected during *Strategy II* is reported in Table 2 and is used to characterize wake turbulence and meandering under different stability conditions."

9. 442 - These deviations are mainly due ... Can you please elaborate what you mean here?

We rephrased the sentence as: "The largest deviations between predicted and measured deficits are found at shorter distances (2–3D) and are mainly due to the model inadequacy to simultaneously fit all the experimental measurements and experimental uncertainties."

10. Figure 8 - As far as I understood all 4 distances are used for the parameter calibration. In the near wake there is not only a quantitative but also a qualitative disagreement between modeled and measured profiles as the double-Gaussian structure is not measured by the lidar. Have the authors considered to discard the near-wake measurements so they do not disturb the parameter fitting? Also, the uncertainty of the model predictions lie in, I would say most of the plots, outside the measurements. How is that possible? Might the uncertainty be underestimated by the uncertainty quantification approach?

We discard the lidar data at 2D in the fitting process; this is done to improve the quality of the fitting in the far-wake region. As this was not specified in the manuscript, we now add a sentence that describes the adopted procedure (ln. 411).

The uncertainty of the model predictions was not correctly propagated in Fig. 8. The error originated from an erroneous definition of the cross-correlations matrix of the calibration parameters when propagating uncertainties for plotting purposes. We now re-plot Figs. 8 and 9. We also updated the statistics of the parameter $\sigma_{\epsilon_{def}}$ in Table 4.

11. Figure 9 - This is an interesting figure, but it should be pointed out that the parameters from the other studies were derived from completely different data sets. The different results might also be due to a lack of transferability to larger dimension turbines.

We added a paragraph in the text in Sect. 5.2.1 to better discuss the results.

12. Figure 11 - I am confused by the Figure 11 and the corresponding description in the text. From the text I would expect two versions of the wake-added turbulence in the DWM model, one based on the parameter-fitting of eq. 3 and one based on the parameter fitting of eq. 17.

We now add both wake-added turbulence models from Eqs. 3 and 17.

13. 6.1 Correction for rotor induction effect - It's unfortunate that the measurements had to be conducted in the induction zone of the turbine. The uncertainty introduced by the model correction has to be at least mentioned and preferable quantified.

We mentioned in the text (Ln. 562) that the wind speed is reduced by up to 12% in the induction zone. We now plot the SpinnerLidar-measured statistics of the rotor-effective wind speed (U_{eff}) without accounting for induction effects in Fig. 14 for completeness.

14. Figure 14 - Please also show and discuss the results of DWM**, because the DWM* is as far as I understood not the DWM model as it would be applied based on wind measurements at or upstream of the turbine.

We added the predictions from DWM**.

15. L655 - Our result indicate ... Please also here consider the limits to generalise from these results.

We now address the validity of our results in the discussion session.

References

[1] Inga Reinwardt, Levin Schilling, Peter Dalhoff, Dirk Steudel, and Michael Breuer. Dynamic wake meandering model calibration using nacelle-mounted lidar systems. *Wind Energy Science*, 5(2):775–792, 2020.

[2] Rolf-Erik Keck, Dick Veldkamp, Helge Aagaard Madsen, and Gunner Chr. Larsen. Implementation of a mixing length turbulence formulation into the dynamic wake meandering model. *Journal of Solar Energy Engineering*, 134(2):021012, 2012.

[3] Ewan Machefaux, Gunner C. Larsen, Tilman Koblitz, Niels Troldborg, Mark C. Kelly, Abhijit Chougule, Kurt Schaldemose Hansen, and Javier Sanz Rodrigo. An experimental and numerical study of the atmospheric stability impact on wind turbine wakes. *Wind Energy*, 19(10):1785–1805, 2016.

[4] Alfredo Peña, Sven-Erik Gryning, and Jakob Mann. On the length-scale of the wind profile. *Quarterly Journal of the Royal Meteorological Society*, 136(653):2119–2131, 2010.