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.

This paper presents methods for performing load validation by reconstructing the wind field upstream of a waked wind turbine using nacelle lidar measurements. The load validation methods are simulated using the dynamic wake meandering (DWM) model and aeroelastic simulations, and compared to the performance of the standard IEC recommended DWM method for load validation. The paper is a nice extension of previous work by the authors "Aeroelastic load validation in wake conditions using nacelle mounted lidar measurements," where the authors use wind parameters based on lidar measurements in wake conditions to evaluate the accuracy of load validation as part of a field experiment.

The paper is well written and clearly organized. Furthermore, the topic is relevant given the interest in using nacelle lidars for applications such as power performance and load validation in the wind industry. By investigating the proposed load validation methods in a controlled simulation environment, the authors are able to isolate the impact of the wind field reconstruction methodology, without worrying about aeroelastic model uncertainties. Although there are no major issues with the paper, there are several smaller comments that I believe should be addressed by the authors.

First, more motivation for the proposed lidar-based load validation methods should be presented. For example, if the goal is to achieve load prediction biases that are the same as the baseline method but with lower statistical uncertainty, how will this improve the wind turbine design process? And can you discuss current problems with the IEC-recommended approaches for load validation in wake conditions?

We added two paragraphs describing both the current limitations of the IEC-recommended approaches for load validation in wakes and how nacelle lidar-based procedures can tackle these issues. Further, we discuss the benefits of developing lidar-based power and load validation procedures in general.

Comments:
1. Title: Instead of "using field reconstruction techniques," which is somewhat vague, consider "using wind field reconstruction techniques"

This is now corrected.

2. Pg. 2, ln. 34: "the 10-min statistical properties (mean and variance) of the simulated ambient and operational conditions are set to match the measured ambient wind statistics": This doesn't quite make sense. How can the simulated "operational" conditions be set to match measured "ambient" conditions. Wouldn't you only need to match the ambient conditions?

This has been corrected, it is only the 'ambient' wind conditions that should be matched.

3. Pg. 2, ln 45: ": : :which increases the amount of validation data." This could use a little

more explanation (contrast this to a fixed met tower where only a small sector is valid).

We added more explanations to further clarify this.

4. Pg. 2, ln. 49: "The recent work of Conti et al. (2020) demonstrated that lidar-based load validation procedure in wakes should account for a model of the wake deficit and its dynamics." Since this paper builds on the work of Conti et al. 2020, please discuss this work in a little more detail, especially why it was concluded that lidar-based load validation in wakes should include a wake model.

We have added a paragraph discussing the main findings from the previous work.

5. Pg. 3, ln. 83: "through a field reconstruction technique": "wind field?" Or "wake field?"

'Wake field reconstruction technique' has been added.

6. Pg. 4, ln. 98: "large number of simulations" How many?

We added that we use 18 turbulence field realizations for each 10-min statistic of the inflow wind. Further, we specify that more details on the load validation analysis can be found in Sect. 4.2.

7. Pg. 4, ln. 101: "The mean bias of load predictions... is of the same order of that obtained with the baseline". Please be more specific about how close the lidar-based simulations should be to the baseline. "Of the same order" is a subjective criteria and makes it hard to tell if the new methods are successful.

We replaced "Of the same order" with "equal to".

8. Pg. 4, ln. 111: "the IEC recommends, i.e., the Mann uniform shear spectral tensor model..." This is one model that is recommended. There is also the Kaimal spectral model, etc.

The Kaimal spectral model has been added.

9. Eq. 1: the symbol "i" is used twice, for the spatial location as well as to indicate imaginary numbers. Can you choose unique symbols?

We now use i for index and i for the imaginary numbers.

10. Eq. 2: Should the bold "k" argument on the left hand side be "k_1"?

This has been corrected.

11. Section 3.2: Can you explain more about the tools you are using to implement DWM? In other words, is DWM a software tool that you are using (if so, a reference would be appreciated)? Or is it a model described in the literature that you are implementing yourselves?

We have added more details and rephrased the whole subsection to better describe the assumptions of the DWM model. We also provide a reference of the numerical scheme of the DWM model used to derive the quasi-steady velocity deficit in the paper.

12. Fig. 1: What wind speed is used for the middle plot?

We have removed this plot and replaced with a different figure, where we now define the inflow wind conditions.

13. Pg. 6, ln. 157: "The latter increases the uncertainty of the procedure." It is unclear what "the latter" refers to here.

We have now specified that in the text.

14. Pg. 6, ln. 166: "while the spatial resolution in the longitudinal axis depends on the simulated wind speed." Then what is the temporal resolution of the wind field?

The turbulence fields used in aeroelastic simulations (and in the DWM model) are basically a vector field, where each point in the field represents the local speed of the flow. In the generation of these fields we use the Taylor's assumption of frozen turbulence. Therefore, the large turbulence structures does not really change with time but are simply transported with the mean wind speed of the ambient wind field. As we run simulations with different ambient wind speeds, but the dimension of the turbulence box is fixed in the longitudinal axis to 8192 'points', the spatial resolution is function of $dx = (U_{amb}T_{sim})/8192$, where T_{sim} is the simulation time in seconds (e.g., 600 s for a 10-min simulation).

15. Pg. 7, ln. 171: "continuous-wake" $->$ "continuous-wave"

This has been corrected.

16. Pg. 7, ln. 181: Can you provide a reference for the 4-beam Leosphere lidar?

This sentence has been removed together with the whole paragraph about the various type of nacelle lidars in the literature. For reference, the 4-beam Leosphere is advertised on the Leosphere's website.

17. Pg. 8, ln. 185: "7-beam lidar can potentially increase the accuracy of reconstructed wind fields." Increase the accuracy compared to what?

This sentence has been removed together with the whole paragraph about the various type of nacelle lidars in the literature.

18. Pg. 8, ln. 193-194: There is also a 4-beam Windar CW lidar, and the grid configuration pattern is based on the SWE pulsed lidar. Can you explain why you classified these scan patterns as pulsed and CW, respectively? Furthermore, since you are only modeling a single measurement range, it is unclear how you model CW and pulsed lidars any differently in you simulations. Can you explain this further? Lastly, you are giving up additional measurement points (and therefore potentially wind field reconstruction accuracy) by only using a single range for the pulsed lidars. Why didn't you use multiple range gates?

We have removed the paragraph describing the currently available nacelle lidars. The previous classification between CW and PL lidar was only made to reference the existing type of nacelle-lidars. Still, it did not influence the simulation results, as we mainly simulate the probe volume effects by a pre-defined weighting function. The reason for using a single range is conditional on the fact that we use DWM model-based fields as target fields. Indeed, the DWM model predicts quasi-steady wake deficits, which are computed according to a specified downstream distance. These deficits are meandered transversely, advected in stream-wise direction with the mean wind speed using Taylor's assumption, and superimposed on random turbulence field realizations (we have now described that in detail in Sect. 3.2). As the DWM model does not simulate turbulence evolution, we cannot simulate multiple range gates. This analysis would be suitable using an LES-based wake field. Another aspect to consider when using multiple ranges is that the wake recovers and expands with farther downstream distances; therefore, the wake field characteristics observed further upstream of the rotor may be considerably different from those approaching the turbine rotor.

19. Pg. 8, ln. 199: "A preview distance of 0.7 D is assumed." In addition to the lidar measurement accuracy arguments, there seems to be an interesting dilemma when measuring the wake deficits upstream of a turbine. On one hand, I imagine you would want to measure close to the turbine to capture the true wake velocity deficit at the rotor plane. On the other hand, measuring too close will introduce induction zone effects. Can you discuss how you approached this issue?

We did not investigate this dilemma in detail in this work mainly because we use the DWM model-based fields as the target. As the DWM model does not include turbulence evolution and induction effects, we cannot investigate in detail what is an optimized preview distance for characterizing the inflow wind. In the work of [1], it is shown that an optimum preview distance for free-stream conditions varies between 0.4 – 1.3 D according to specific lidar pattern and the specific wind field characteristics to be estimated. Here, we adopt a fixed preview distance of 0.7D, as we want to measure close to the rotor for the two reasons described in the text (i.e., reducing errors due to turbulence evolution and considering that lidar's probe volume typical increases for farther distances as for a continuous wave system). We have now added a few lines to discuss this.

20. Pg. 8, ln. 204: "A probe volume with an extension of 30 m in the LOS direction is assumed" Can you provide some references for how you chose 30 m for pulsed and CW lidars? Furthermore, how is the probe volume extension defined? For example, the std. dev. of Gaussian weighting function?

We have added that the probe volume length is here defined as the standard deviation of the Gaussian weighting function, and added references. The probe volume length of 30 m does not identify a specific lidar system, but it is an estimate that is comparable with the current CW lidar technology measuring at distances beyond 120 m [2]. Further, we conduct a sensitivity analysis by varying the probe volume lengths in Sect. 4.3.2, to analyze how these lengths influence the accuracy in power and load predictions.

21. Pg. 9, ln. 218: "obtained by simply scaling an isotropic turbulence field..." Can you clarify if the scaling depends on the radial location from the wake center, as shown in Fig. 1?

We have rephrased this sentence and provided a better description of the DWM model, including the wake-added turbulence formulation.

22. Pg. 9, ln. 221: How might the ambient wind conditions be measured in practice?

This is explained just a few lines below; see ln. 230. Ideally, from a met mast installed at the site or a nacelle lidar measuring the inflow wind.

23. Pg. 10, ln. 226: What do you mean by 'The u-velocity fluctuations are recovered from the target wake fields?'

We have rephrased to: 'Only the u-velocity fluctuations are reconstructed from the *target* wake fields.'

24. Pg. 11, ln. 256: "By denoting: : : as the constrained turbulence field that incorporates lidar measurements: : :" It seems that in Eq. 9, $u'_{CS,B,i}$ represents the turbulent fluctuations with the mean ambient wind profile removed. Do you first remove the mean ambient wind speeds from the lidar measurements before they are used to generate the constrained turbulence field?

That's correct. We have now described this step in the procedure.

25. Eq. 10: I'm confused about how $K_{def,lidar}$ is defined. From Fig. 1, K_{def} is presented as a scaling factor applied to the ambient wind field (= 1, when wake losses are not present). But here, it appears to be defined as the normalized deficit (= 0, when wake losses are not present). Can you clarify this and make sure the definitions of K_{def} are consistent?

That's correct, we now define K_{def} as the normalized deficit (= 0, when wake losses are not present) and keep this definition consistently.

26. Eq. 10: Since the left hand side of this equation is being fit to the Gaussian function, they are not actually "equal." It would make more sense to present this equation as a minimization objective function (e.g., based on the difference between the measured deficit and the Gaussian model) Also, should $U_{amb(z)}$ have the mean operator applied to it, like in Eq. 9?

We have corrected the equation accordingly, and provided a minimization objective function instead.

27. Eqs. 10 and 11: The explanation of Eq. 11 is confusing. In your final method are you using the Gaussian fit from Eq. 10 as part of Eq. 11, or does Eq. 11 entirely replace Eq. 10? It would help to present both equations as minimization problems, so it's easy to see where the lidar measurements are being used and what exactly is being fit to the Gaussian profile.

We have corrected Equations 10 and 11 and provided a better description of the procedure.

28. Section 4.1: In addition to the analyses presented, a nice way to quantify the accuracy of the reconstructed wind fields could be to compare the RMSE of the rotor average wind speed u_{eff} as well as the best-fit linear horizontal and vertical shear coefficient time series between the target and reconstructed wind fields. These variables should play a large role in determining the turbine loads.

We agree that it could also be an option. However, the analysis presented in Sect. 4.1 should provide sufficient information for evaluating how different lidar scanning configurations, complemented with the proposed wake field reconstruction techniques, perform. Further, we assess three wake field-related indicators (U_{eff}, ρ_E^2, σ_u^2) in the load validation analysis in Sect. 4.2, which should explain to a large extent, the observed deviations in power and load predictions.

29. Pg. 12, ln. 308: "run at the downstream distance of 5 D" Please be more specific. The turbine of interest is located 5 D downstream of the upstream turbine?

This has been corrected.

30. Fig. 7: On the left plot showing U_{eff}/U_{amb}, can you explain why the ratio converges to 0.93 at high wind speeds? As wind speed increases, the turbine thrust should keep decreasing causing wake losses to continue to decrease, so I would expect the ratio to approach 1.

It does not converge to 1 because although the trust coefficient decreases for higher wind speeds, the ambient turbulence is relatively low, and therefore the wake field does not fully recover at a distance of 5D, which is the one analyzed in this study. The ratio U_{eff}/U_{amb} will converge to 1 for higher ambient turbulence or farther downstream distances due to the increased turbulence mixing. We have now described that in the paper.

31. Pg. 18, ln. 440: "In addition, improved estimates of both ρ_E^2 and σ_u^2 are seen in Fig. 10c,d." There seem to be improvements at low wind speeds, but slightly worse performance at high wind speeds. Can you comment on this in the paper?

We have now discussed this result in the paper.

32. Figs. 10 and 11: I would suggest full captions.

The full captions have been added.

33. Pg. 20, ln. 463: "focus the analysis on the SL, Grid, and Grid* configurations" These might be the most promising scan patterns, but also not the most likely, given currently available commercial lidar technology. It would be interesting to analyze the time series for one of the commercially-available lidar scenarios as well.

We opted to show only the most promising results, as the currently available commercial lidar technology (i.e., the 4P, 7P, and the Cone patterns) will introduce significant biases in the power and load predictions, as one can see from results in Figs. 8, 9, 10, and 11. This figure intends to show that provided a sufficient number of wind measurements taken upwind of the rotor, both the *CS*- and *WDS*-approach can reconstruct power and load time series that are highly correlated with the *target* observations. These results explain why we obtain lower statistical uncertainty X_R values.

34. Pg. 21, ln. 469: "It should be noted that the structural resonance occurring at low wind speeds, which excites the tower can potentially affect the correlation results." Can you discuss why this resonance appears? Could it be removed by improving the controller tuning?

It appears because of the structural design of the DTU 10 MW, which is a reference (theoretical) turbine model. At low wind speeds (thus low RPM), the 3P rotational frequency (0.3–0.48 Hz) excites the eigenfrequency of the tower (≈ 0.25 Hz). Considering that the wake induces unbalanced load distribution on the rotor, which in turn amplifies the rotor harmonics (1P, 2P, and 3P), this results in structural resonance. Besides that, we also observe that the bending moment of the tower bottom for large turbines is highly driven by the 3P frequency, as also shown in Fig. 13 (where the imprint of the turbulence wind is almost non-existence). Some internal work at DTU has been conducted to reduce the resonance, and the controller utilized in this work should be optimized to reduce resonance effects, which are still present and amplified under wake conditions. Future studies that evaluate these lidar-based reconstruction approaches can be conducted with different wind turbine designs that do not experience these resonances.

35. Pg. 22, ln. 481: Usually magnitude-squared coherence is written as $\gamma^2 = abs(S_x, y)^2/(S_x * S_y)$. Therefore, I would expect your definition to be $\gamma = abs(S_x, y)/sqrt(S_x * S_y)$. Is this correct?

This has been corrected as $\gamma^2 = |S(\tilde{y}, \hat{y})(f)|^2/(S(\tilde{y})(f)S(\hat{y})(f))$

36. Fig. 14: On the left plot, why is the baseline coherence so high at low frequencies (above the noise floor)?

This follows as the MxBR signal (blade root flapwise bending moment) is driven by both the wake meandering frequency and most importantly by the 1P rotational frequency (both are relatively low frequency signals as shown in Fig. 14). As the target and baseline simulations operates at similar 10-min average RPM (the ambient wind speed is the same), the 1P peak does not vary significantly between the Power Spectral Density of the target and the baseline MxBR loads. So we can see a non-zero coherence at low frequency, which is still lower than 0.3.

37. Pg. 25, ln. 538: As mentioned earlier, the "need for reducing the statistical load prediction uncertainty" in wake conditions could be motivated more clearly in the paper. More discussion or references talking about the need for improved methods would strengthen the message of the paper.

We have now addressed this point in the introduction, and delete this paragraph in the discussion section.

38. Pg. 26, ln. 562: You say that the lidar-based predicted load statistics are comparable to the results from the baseline DWM method (Δ_R between 0.97 - 1.01). However, from Figs. 9 and 11, it seems more accurate to say that Δ_R is between 0.92/0.94 and 1.01. Is 0.94 still an acceptable difference?

We have removed this paragraph from the discussions, as we discuss those results in the appropriate section (see Sect. 4.2). But, it is correct that Δ_R is between 0.92/0.94 and 1.01, depending on the load component, probe volume size and the adopted wake field reconstruction techniques.

39. Pg. 27, ln. 610: Similarly, the range of Δ_R with the lidar-based method is more like 0.94-1.01 instead of 0.97-1.01. When saying that this is comparable with the baseline method, please be more specific about what "comparable" means.

We have rephrased the conclusions accordingly.

40. Pg. 27, ln. 615: In addition to these lidar parameters, the load prediction accuracy is sensitive to the turbulence intensity as well.

This has been added.

41. Pg. 28, ln. 628: "largest energy content at higher frequencies ($>$ 3P, 0.3 Hz)." From the plots, the largest energy content is at very low frequencies and right at 3P, but $>$ 3P does not contain as much energy content.

This has been corrected to "largest energy content at higher frequencies up to 3P (0.3 Hz)."

References

[1] Eric Simley, Holger Fürst, Florian Haizmann, and David Schlipf. Optimizing lidars for wind turbine control applications-results from the iea wind task 32 workshop. *Remote Sensing*, 10(6):863, 2018.

[2] Alfredo Peña, Charlotte Bay Hasager, Merete Badger, Rebecca Jane Barthelmie, Ferhat Bingöl, Jean-Pierre Cariou, Stefan Emeis, Sten Tronæs Frandsen, Michael Harris, Ioanna Karagali, Søren Ejling Larsen, Jakob Mann, Torben Mikkelsen, Mark Pitter, Sara Pryor, Ameya Sathe, David Schlipf, Chris Slinger, and Rozenn Wagner. Remote sensing for wind energy, 2015.