Authors' response to Referee #2:

We would like to thank Gerard Schepers for reviewing this manuscript, the valuable feedback and the very constructive comments. At this stage of the review process, we respond to the comments and propose improvements for the final manuscript. The referee's original comments are printed in **bold** followed by the corresponding answers. Passages from the manuscript are printed in *italic writing*, in which proposed additions are indicated in blue and deleted parts in ~~red~~.
Thank you very much for your efforts,

Franz Mühle and Jan Bartl on behalf of all authors

Comment (1)

On some places you quantify differences between calculations and measurements but on other places you use very subjective assessments with terms like a good, poor or fair agreement. I realize very well that this is difficult to prevent but be aware that another person may come up with a completely different assessment. For example: You write on page 4, line 1 "It can be seen that the drag coefficient CD is *slightly* different". I would write that the differences are huge. . . .

Thank you for pointing this out. We agree, that a qualitative comparison of results always is very subjective. For this reason, we included quantitative comparisons methods, i.e. the statistical error measured such as the normalized mean square error (NMSE) and the correlation coefficient (r). Thus, the reader can decided him/herself, whether the differences are large or small.
As for the mentioned example, we agree that the differences are rather large than small and therefore suggest the following modification:

p.4, l.1 ff:
It can be seen that the drag coefficient C_D is ~~slightly~~ very different and lift coefficient C_L is diverging significantly from an angle of attack α of approximately 4° between the experimental and XFoil data. This difference is very distinct for high angles of attack that may occur close to stall.

Comment (2)

It is good to compare CP, CT and CMyaw but please be aware that a comparison on basis of these global integrated properties has little meaning. The only conclusion you may draw from a comparison of integrated loads is that a bad agreement means that there is something wrong. A good agreement doesn't say much because we very often see 'compensating errors'. An overprediction at the root may be compensated by an underprediction

at the tip or vice versa. Some examples can be found in my PhD thesis. For yawed conditions with a delicate balance between root and tip vortex effects (Rahimi et al., 2018) I expect this to be even more the case. In that sense I think that a comparison of e.g. local yawing moments from the different calculational methods would be extremely useful to better assess the aerodynamic modelling of the different partners (I realize this is a lot of work and it is not mandatory for me, but I strongly recommend this for a follow-up study).

Thank you for this very good comment. It is absolutely true, that a comparison of these integral parameters CP, CT and CMyaw only allows for a very rough comparison. It is not sure at all, that the airflow over the blade is well-modeled and that different errors might compensate each other. However, we want to point out again, that the main focus of this study is comparing different modeling methods for the wake flow under complex turbine operating conditions. The CP, CT and CMyaw should rather be regarded as input variables. If these are modeled incorrectly, we expect to also see that in the wake flow.
For a more detailed comparison of the actual flow around the blades, a significantly improved experimental setup would have been necessary. This was done in the MexNext project and model turbines as the Berlin Research Turbine would allow for this kind of comparison. As our wind tunnel is limited in size, also our models are limited and consequently the equipment we can allow to be rotated in the blades.

Comment (3)

Does any of the lifting line methods, which use airfoil data, consider dynamic stall effects?

This is another very good comment. The answer is no, unfortunately not, as we consider dynamic stall effects to influence the blade aerodynamics to some degree. We consider this topic as something to look deeper into, given the simulation data available. We have updated the description of the computational methods, now including information that they are not using a dynamic stall model. (For the suggested updated descriptions see "Answers to Reviewer 1" document.)

Comment (4)

On page 6, line 27 you write that the thrust is measured at the tower foot. I do note that the tower is included in the simulation (which is good, so you make a fair comparison) but to my point of view it still obscures the comparison a bit. Some of the differences in CT might come from the tower which are not so relevant for the wake properties since the wake is measured several ROTOR diameters (and very many TOWER diameters) behind the first turbine. In this respect: 1) Do you have "rotor off measurements", i.e. measurements of the tower alone, 2) how large is the tower drag compared to the overall thrust 3) has there been any treatment of the

tower to prevent vortex shedding?

This is again a very good comment. 1) Yes, we have previously made "rotor-off" measurements of the tower-thrust only and 2) the tower thrust $C_{T,Tower} \approx 0.10$ at $\lambda = 6.0$, which corresponds to a little less than 10% of the total thrust. 3) On the upstream turbine LARS1, there has not been made any treatment suppressing vortex shedding. This might for practical applications be beneficial, but would add increased complexity to the test cases.
In contrast to earlier Blind test experiments, where we have tried to compare the rotor thrust only, we now decided to include the tower in the comparison. Earlier, we had to subtract the tower thrust from the measured total thrust, which itself included some uncertainties, as the major part of the tower is in the wake of the rotor. This solution seemed to be the fairest, for comparison purposes.

Comment (5)

I am extremely happy that you include measurement uncertainties in the results. Still I am a little bit surprised to see that the uncertainties are independent of the conditions. I would expect a dependency?

The measurement uncertainties are indeed different for the different test cases and operating conditions. This can be observed in Figures 6 and 10, where CP, CT and CMyaw are compared. The uncertainty values presented in Section 3.2 indicate the maximum calculated uncertainties measured for the different variables.

Comment (6)

Are there any calculations which include the tunnel geometry. How large are tunnel effects? They might be more important than the measurement uncertainties in particular for yawed conditions.

This is a very good thought. Indeed, all numerical simulations include the wind tunnel geometry as it certainly affects the performance and wake development (specifically wake deflection in yawed conditions). It would have been very interesting to quantify the effects of tunnel blockage on turbine performance and wake development, by running a simulation with and without the wind tunnel boundaries. This is, however, considered to be a specific study on its own. A study of the effects of different wind tunnel blockage ratios on the wake development and expansion of a non-yawed turbine has previously been performed by Sarlak et al. (2016), using the geometry of our NTNU model turbine. A follow up study on the effects of yaw would be very interesting.

Comment (7)

What about the turbine quality: Have the blade geometries been scanned

(small differences in blade geometry may lead to huge differences in airfoil polars at these low Reynolds numbers) and how accurate are the pitch angles? Are the pitch angles of all blades similar, and are the blade geometries similar? If not the aerodynamic unbalance may obscure many of the results. How accurate is the rotational speed?

Thank you for this very good comment focusing on the accuracy of the experiment. No, the blade geometries have not been scanned by a 3D scanner. We agree, that small geometry inaccuracies (which are definitely present) might influence the wake results more than expected. We have, however, used optical laser methods as well as phase-locked photographic methods to ensure that the pitch settings of our three blades are deviating as little as possible (as much as these methods allow). Furthermore, the pitch settings have been adjusted such that the vortex shedding of the three blades in the very near wake is equidistant. The distance between the three vortices shed were measured by the means of phase average hot wire measurements (Eriksen and Krogstad, 2017).
The rotational speed is constantly measured by an optical sensor in the nacelle. The variations in rpm are observed to be below 1.0%.

References

[1] Rahimi, H., Martinez Garcia, A., Stoevesandt, B., Peinke, J., and Schepers, G.: An engineering model for wind turbines under yawed conditions derived from high fidelity models:. Wind Energy, doi: 10.1002/we.2182., 2018.

[2] Sarlak, H., Nishino, T., Martinez-Tossas, L.A., Meneveau, C., and Sørensen, J.N.: Assessment of blockage effects on the wake characteristics and power of wind turbines, Renewable Energy, 93, 340–352, doi: 10.1016/j.renene.2016.01.101, 2016.

[3] Eriksen, P.E. and Krogstad, P.-Å: Development of coherent motion in the wake of a model wind turbine, Renewable Energy, 108, 449–460, doi: 10.1016/j.renene.2017.02.031, 2017.