Reviewer Responses

Reviewer #1

Dear authors,

thank you for taking my comments into account. I think the article is improved, but also in the revised article there are a few things that need to be adressed:
The first two references are both called OLAF documentation, and somehow the links do not work for me.
AUTHOR RESPONSE: One of those references was supposed to be “AeroDyn documentation”, this has been fixed. The site was accidently taken down temporarily, which is likely when you tried to access it, but this has been fixed. It also seems like the template has a hard time when citing references that do not have an author, e.g. the OLAF documentation shows up as OLA. We will raise this issue with the journal on how this can be dealt with when the time comes, as we would prefer not to hard code it.

The reference to Meyer Forsting just shows up as 'Forsting'.
AUTHOR RESPONSE: This seems to be an issue with the template. It will not show both names unless they are entered as one word or connected with a dash. We have hard coded both names to appear in the text, and will raise the issue of how it appears in the Reference section with the journal when the time comes.

Figure 2: If I understand correctly, then the unit on the y-axis could be changed to 'CPU hours' for clarity.
AUTHOR RESPONSE: Yes, this has been corrected.

In my opinion there still needs to be some more information on how the SOWFA results were obtained, such as: discretization of the blades, information on the size and resolution of the CFD mesh, ... My apologies if this information is given in a reference that I overlooked.
AUTHOR RESPONSE: Thank you for identifying the missing information. Some of this information is given in another section, and we have added it to the SOWFA section for completeness. The following paragraph has been added to the manuscript:
“
The results presented use 50 actuator points along the blades and epsilon/D=.05, following the work from Churchfield et al. 2017. The grid resolution of the background mesh was .091D with 2 levels of local refinement to achieve a grid size of .023D in the turbine location.”

In two places of the article a 12x error reduction is mentioned, even though the error reduction now with the updated OLAF results is rather in the order of 40%.
AUTHOR RESPONSE: Thank you for pointing this out, the whole article has been checked to make sure all references to the results are correct.

Also the percentage differences for OLAF seem not to be updated in the conclusions. Please check the whole article for descriptions of results that need to be updated.

AUTHOR RESPONSE: Thank you for pointing this out, the whole article has been checked to make sure all references to the results are correct.

It needs to be described how the outliers are defined. For example in Figure 6. Why are the close to 60% difference at min15deg not an outlier, but the 40% difference at zero degree are marked as outlier? I have the same question for many other outliers in the box and whisker plots.
AUTHOR RESPONSE: The process of computing all components of the box and whisker plots is further explained at the beginning of the results section. We have also added a reference to the python function used to compute these values and make the plots.

"These ranges are computed for each box, not across all data in a given figure. Thus, a value that might be an outlier for one set of results may not be for another set, even if the boxes are shown on the same plot."

Related to the outliers: often, there are both a point and an 'x' at the same percentage level but at a different location along the x-axis. Is this the same data point? Then I think only the 'x' should be show or the point and the x should be at the same location.
AUTHOR RESPONSE: This is happening because each individual data point that is used to create the plots is shown. This is not only to be transparent about the number of data points that are used to create each box-and-whisker, but also so that readers can get a sense for the actual values of each QoI result. Thus, there are points throughout the box, whiskers, and at the location of the outliers. This point has been emphasized in the text to avoid confusion.

There are no references to the sheared and turbulent inflow cases in the conclusions. What are the conclusions from these studies?
AUTHOR RESPONSE: Thank you for pointing this out, conclusions about the sheared and turbulent inflow results have been added to this section.

"When comparing the time-averaged OLAF and BEM results, changes to shear exponent and turbulence intensity did not have a substantial impact on the relative accuracy of the models, especially compared to the impact of yaw misalignment. An exception to this is the impact of the shear exponent on the azimuthal trends. In particular, varying the shear exponent has minimal impact on the OLAF results, whereas BEM showed more substantial changes, especially in the regions where the turbine blades are not in the tower shadow."

In Section 4.0.1 you write 'OLAF results show no dependence on yaw misalignment'. The OLAF results change with yaw misalignment. Do you mean that the trends are the same as observed for SOWFA, and thus the percent difference is not so dependent on yaw misalignment?
AUTHOR RESPONSE: Yes, this is what we meant. The text has been changed to better explain this.
"The relative error between OLAF and SOWFA results show minimal dependence on yaw misalignment […]"

I think the output time step of 1 second in the turbulent comparisons needs to be mentioned in the section about turbulent comparisons. I quickly checked how much it would change the standard deviations if the time sampling is changed from 35 Hz to 1 Hz for a case from the Danaero measurements. The standard deviations on the normal forces were reduced by 1-2% in this case due to the low sampling rate. I assume the output time step for BEM and OLAF was the same?

AUTHOR RESPONSE: Yes, this is correct and a very good point. We have added this detail to the paper. Since the focus is on the relative error and this same sampling rate was used for all simulations, this shouldn't have a significant impact on the conclusions of this section.
"It is important to note that the output frequency for all results was 1 sec, which has an effect on the computed standard deviation when turbulence intensity is included. However, as the main drive of this comparison is the relative error and the same sampling rate was used for all simulations, this is not expected to change the conclusions of this section."

Best regards,
Georg

Reviewer #2

The authors addressed some of my questions/requests in their revised text.
First I would like to commend the authors for their attempt to improve the quality of their original submission. However, I still find that the rebuttal to some of the points I raised in my first review is rather weak. So, I would like to come back to those points and ask for some further explanations/clarifications.

In my opinion the presentation of the results remains the weak point of the paper. The use of bars and whisker plots is not the best way to present results of periodic responses/signals. In particular the way chosen in the paper, by introducing a global error metric that incorporates all signals. Canonical cases are meant to serve the scope of a more detailed comparison/benchmarking. Performing an FFT analysis is not necessary for the yaw cases. A discrete Fourier analysis would be sufficient. However, if this is not available I would at least recommend the authors to enlarge figure 7b (maybe also separate the curves for the different yaw angles). As I explained in my first review the variation of the flapwise moment drives the differences in all other signals. Therefore, presenting tower moments (in particular when using a common structural tool) is pointless.

AUTHOR RESPONSE: Thank you for your considered comments. In agreement with your comment, part of the paper is devoted to looking at specific QoI and understanding the specific impacts of the various models in some example and canonical cases (for instance extreme yaw values). With the many output data and quantities of interest investigated, we believe that there is value in having an overall sense of how a model performs for all QoI, and we think that box and whiskers plots are helpful in achieving this, providing the overall model performances at a quick glance.

You are correct, the differences in tower moments are principally a consequence of differences in aerodynamic loads (such as the flapwise moments). We have chosen to keep them because they are an important and common quantity of interest for wind turbine design. The reader can get an idea of the differences (or errors) expected for this quantity.

Regarding an FFT analysis, the authors agree that it can offer some insight. We would be happy to include FFT analyses, but have chosen not to do so in this revised version because our simulations were run with a low output frequency of 1Hz, which would likely result in FFT plots that are maybe less interesting than if the sampling frequency was higher. Rerunning the simulations with a smaller output time step can be done for BEM and OLAF, but we are in the unfortunate situation where we do not have the staff hours or computational resources to rerun the CFD simulations. We are sorry about that. If the reviewer judges that the FFT are essential to the paper, we will of course reconsider, though our resource constraints will remain.

To address part of your comment, we have implemented your suggestion to enlarge Figure 7b and separate the curves for different yaw angles. While this is not ideal, we hope that you will understand the resource limitations we face and accept this change in lieu of an FFT analysis.

FFT plots provide information on how energy is distributed to the various frequencies (natural or excitation). I still believe that azimuthally averaged plots don't offer much more information than that already included in the plots of the canonical cases paragraph. I still believe that it is worth trying to rerun one case at least (producing the necessary output in terms of output sampling frequency) in order to check the energy content of the OLAF simulation against the BEM one.
AUTHOR RESPONSE: (We have discussed FFT plots in the paragraph above). The authors believe that the azimuthally-averaged plots offer important insight to the results. For instance, the root bending moment obtained with BEM shows larger amplitude than the ones captured by OLAF and CFD, and similar behavior were also observed as part of IEA Task 29. This plot therefore offers a mean of comparison with publications from the IEA Task 29. Additionally, the plots provide insights on where in the blade path differences can arise, which can aid in understanding root differences and potentially help with further modeling improvements.

As a general remark I would like to say that in my humble opinion, the argument "it will be the focus of future work" is only valid when you are asked to perform new work, which requires substantial amount of new developments. In the present case, the simulations and the results are there and what the reviewer is asking for, is a clarification (or a further check) of their validity. I would say that this by no means priority of a future work but of the present one. This is for the benefit of the researchers who submit the work (in order to identify the range of validity of their models) but also for the benefit of the readers who are then aware of the strong and weak points of a method.
AUTHOR RESPONSE: We agree with your comment. We are happy to offer as much clarification as possible, and we have tried our best to address your comment in this revised version. We do wish that we had planned for additional resources, and will take the FFT time step issue into consideration for future projects.

Finally some additional comments stemming from the revised material:
In 2.3.1 nothing is said about the "frozen wake". Is this always 1D long as stated in section 2.1? Does the length of this wake portion affects the results at all?
AUTHOR RESPONSE: We have indeed kept a frozen wake length of approximately one diameter for all the simulations. The frozen far wake only contributes to a small portion of the induction at the rotor (most comes from the free near wake and free far wake). The frozen far wake affects the convection and rollup of the "free" (far and near) wake but mostly at the boundary between the free far wake and frozen far wake. The frozen far wake is mostly used to avoid the rollup of the wake onto itself at the end of the wake (the wake truncation being a wrong boundary condition, not satisfying Helmholtz laws). Overall, the length of the frozen far wake has a negligible effect on the rotor loads as long as the free near wake and free far wake are long enough, which was the case for our simulations.

What is a bit unclear in figure 2 is the value of the other two parameters when changing the specific one presented in the plot. For example when changing near wake extent, what is the wake discretization in degrees and how long is the far wake.

AUTHOR RESPONSE: This is an excellent point. We have added a line to Table 2 and the use of these "Nominal" values has been clarified in the text.

Moreover, some additional less or more important, specific comments (on the revised material/text) can be found in the accompanying pdf file (performed again directly on the body of the paper).

dynamic stall is the correct. Are you sure about this statement? Aoa excursions due to turbulence are already quite big even in less flexible blades. Therefore accurate unsteady aerodynamics and dynamic stall models were indispensable before blades became flexible.
AUTHOR RESPONSE: This is correct, dynamic stall is not relevant only for flexible blades. We have modified the text to address your comment.
"The unsteady motion of the blade resulting from the increased flexibility of modern blade designs lead to large angle of attack fluctuations, meaning an accurate and robust dynamics stall models is important."

What do you mean by "filtered". Do you imply Reynolds Averaged?
AUTHOR RESPONSE: The term "filtered Navier-Stokes equations" is typically used to describe the spatially filtered set of equations which are the ones used in large-eddy simulations. The Reynolds-averaged NS equations are averaged in time, as opposed to space. We are solving the large-eddy simulation equations when using SOWFA. We have added "spatially filtered" to clarify this.

Do you need 10min to reach periodic conditions in uniform inflow cases? I doubt! What is the point of that?
AUTHOR RESPONSE: It's true that we might not have had to run the simulations for this long for the uniform inflow cases. However, we did for the turbulent inflow cases and when running many simulations, it is best to keep inputs as consistent as possible. 10-minute simulations are the standard length of time to run such simulations, and the authors did not see a problem with simply keeping that standard, as it did not add much computational time and resulted in an easier setup across all the simulations.

It is a bit confusing that near wake extent study is distinguished by wake discretization study herein, while in the next paragraph near wake extent study is considered as part of the wake discretization study.
AUTHOR RESPONSE: The next section is titled "Wake Discretization and Extent" (changed from Wake Discretization and Length) which the authors feel is an appropriate title to cover the included studies. We have, however, renamed the overall section to be "Wake Parameter Specification Study Results", and have also updated this wording in Tables 2 and 4.

in degrees?
AUTHOR RESPONSE: This wording has been changed.

Please rephrase. Repetition of "length"
AUTHOR RESPONSE: This has been done.

Why do you call this straight line filaments? It is a straight line between the marker points but the filaments still follows a spiral. The same way it is done in the near part. I don't see why do you make such a distinction. The far wake is the one consisting of the tip filaments only isn't it.
AUTHOR RESPONSE: The wording here has been changed to better reflect how these calculations are performed. Here, we have changed "straight line filaments" to "tip- and root-vortex filament"

is that correct?
AUTHOR RESPONSE: No, thank you for pointing this out, the reference has been fixed.

See previous review. Units are missing.
AUTHOR RESPONSE: The units have been added as well as a statement in the figure caption. This has also been done for Figure 4.

What's the diffrerence between core spread value in SOWFA and in OLAF. The reported values are of a different order of magnitude.
AUTHOR RESPONSE: Yes, the different values of core-spreading provide different results. This is an active area of research, and we are trying to understand this effect. We note that the tools/methods in this study have been used with some of the recommended values in the literature at the time. And as we have learned, the core spread value is a key component of some of the differences in this study. We have added the following sentence to the conclusion:
"We note that the core radius of the vortex methods and actuator line model are an important parameter that can affect the results. This study used the recommended parameters from the literature, although more research is needed to understand the effect of core radius on blade aeorelastic response."