Review WES-2020-88

Dear authors,

I enjoyed reading your article on the field experiment at Sedini investigating axial induction control. I find this work to be extremely important in the further validation of wind farm control technologies. The article contains a lot of information and generally has a good story. Though, I have a number of major comments on the article. I want to stress that these comments are my opinion, and with the sole intention of improving the scientific relevance of the manuscript. These comments are not meant to discourage the authors in their work, and I want to again emphasize that I am convinced of the article's importance. I hope this feedback finds you well.

General comments:

1. The core results of this article, based on the title and my interpretation of the article, would be represented by Figures 15 and 16. These figures show the reported field experiment gains in power production using axial induction control compared to baseline operation. The authors show gains of up to 70% and losses of up to 40% using this algorithm. Based on what the literature has shown before on axial induction control, I find such values difficult to believe; I would expect the gains/losses to be in the order of 5% at the highest. When looking deeper into the data, I have two observations:

    a. The amount of datapoints is very small. Uncertainty bounds would be very useful in showing this data. Moreover, a discussion on how many datapoints are necessary for a reliable estimate in the first place would help the reader a lot.
    b. The data has been binned according to WS and WD, but not TI. This means that one bin can contain measurements with a very low turbulence intensity, and also measurements with a very high turbulence intensity. The underlying assumption of clubbing these data together in a single bin is that TI has no effect on the power production of the array of turbines (on page 20: "the effect of turbulence intensity is not expected to be as important"). I find this assumption to be unacceptable without a proper proof/validation/discussion.

    This assumption may likely explain the large variations in gains/losses seen in Figure 15. For example, if I compare the average of 3 hypothetical 'OFF' measurements with a very low turbulence intensity with 5 hypothetical 'ON' measurements with a very high turbulence intensity, I may see such large gains of tens of percent. However, this does not directly mean that the gains are due to axial induction control, but are also highly probable to be due to external factors (TI, in this example). Another explanation might be turbines turning on and off due to the slightest difference in wind speed, but I find this to be less likely. Actually, I believe Figure 18 supports my reasoning, since gains of similar magnitude are being predicted by LongSim. I strongly doubt whether LongSim (or any engineering wind farm model) in any setting would predict a gain of 70% solely due to axial induction control vs. greedy wind farm control. Rather, Figure 18 argues for the argument that the current data processing methodology does not accurately show the gain due to axial induction control in the field experiment.

    Using LongSim, you could check how much the relative power capture varies with wind speed and with TI, and then choose to bin the data according to the one that varies the most (or both, if necessary). My guess would be that the relative gains do not significantly change with WS. If they do, and you have to bin the data according to both WS and TI, this may mean you are left with very few datapoints. Perhaps you have enough data for a handful of bins. Though, I would argue that it's better to have one or two reliable bins/gains than a large set of relatively unreliable gains.

Furthermore, I am missing a more detailed explanation about the results of the field experiment specifically. I understand that the entire process of synthesizing the controller is extremely important, but the focus seems to be a little bit off in this article. The results of the field test are not described until page 18 of the article. Then, the results are quite surprising and in my eyes deserve a detailed discussion. Finally, based on my reasoning above, I do not agree with the conclusion that axial induction control has been demonstrated/validated to increase the power production by several percent. The authors also state that statistically no quantification can be made, thereby seeming to contradict their own statement on the gains that can be achieved.

2. I believe the manuscript can be reduced in size, additionally improving readability. Namely, the authors tackle many topics and the paper therefore becomes very dense, yet I feel like it misses a clear direction at times.
    a. I believe this article should focus on the results of the axial induction control experiment. Validation of the LongSim model is very important and should, in my eyes, be a separate article.
    b. Sections such as 3.3 can be reduced significantly. The authors explain things in a very general sense, and then explain how they have done it themselves. I think the former is often not necessary (perhaps cite literature) or can be reduced to a minimum.
    c. Three different LUTs are compared in Section 4. The percentual increase in power production is 1.50%, 1.57% or 1.58%, respectively. These values fall within the uncertainty limits, no? To me, this seems like too much detail to show in the article. The authors could simply motivate the underlying ideas behind adding uncertainty/smearing (cite Rott, Quick, Simley), and state that smearing did not lead to losses in LongSim while alleviating the pitch actuation and increasing robustness.

3. An accurate reflection on the current literature on axial induction control is missing in this article. Namely, there are publications in the literature on axial induction control field experiments by ECN.TNO. There are also wind tunnel experiments by several researchers, among which Campagnolo et al. from TU Munich. The authors should explain what is new about this experiment, why it is necessary, place their findings in the appropriate context, and explain why their findings do or do not agree with the literature.

4. The article sometimes feels like a technical report more than it feels like a scientific article. Namely, the article contains descriptions of certain procedures and data which need not be mentioned for the understandability and reproducibility of the work (also, results cannot be reproduced anyhow since the data is not publicly available). For example,
    a. the paragraph on page 2 "The original intention […] Kern et al. (2019)." can be removed, in my eyes.
    b. the model acronyms in Figure 3 may be written in a more understandable way

c. Similarly, the manuscript contains statements about in which format the data was handed over from GE to DNV GL, and that it was updated periodically (page 17), which should be omitted in my eyes.
d. Also, the description on when certain turbines were curtailed is irrelevant for the results/conclusions in this manuscript. Figure 12 and the corresponding text can be removed; it suffices to state that curtailed data entries were removed from the dataset.
e. Generally, the text on page 18, section 5.1, can be reduced significantly. The authors can remove the following text snippets without any loss of generality or information: "namely the time stamp […] toggle state.", "as no setpoints … relevant records", "For the sake of … toggle flag changes", and "of the filter flag".
f. Rephrase statements such as "[…]possible to use measured stability as a lookup table input" to "[…] possible to measure stability".

5. I find it hard to read and understand several figures. Some figures miss an informative caption, labels on the axes, and the text in figures vary significantly in size throughout the document. For example,
    a. Fig 3 has grey lines and a black box around it while other figures have black lines
    b. Fig 4: the subplot titles in Fig 4: "A4-33/A4-38". The caption could read something along the lines of "Power production normalized by the power production of WTG 38", and the ylabel could read "Power ratio of WTG 37 [-]". Also the xlabels need not be repeated if the exact same axis/xlabel is used in the subplot directly underneath it. This makes the figure more compact. Generally, it would make sense to put Fig. 4 completely on one page, including caption, to avoid confusion.
    c. Figure 6: "RotorAvDir" instead of "Rotor-averaged wind direction"
    d. Figure 14: "results.mat", also missing units in labels
    e. Figures 19, 20, 21: at this scale, it is not possible for the reader to draw any meaningful conclusions from them. Also, I believe these figures could be removed from the document to improve readability. Instead, I think tables with quantitative values would be much more interesting for validation.
    f. Generally, legends would be appreciated in plots, though I understand that the captions also contain the information.

Minor comments:

- The title may be more informative: what kind of wind farm, what size of turbine array.
- The authors sometimes use vague language in the text. For example, "*convincing* validation", "*some* field tests", "the controller was toggled on and off at *regular* intervals", "the turbine is yawed *a little* out of the wind direction", "and *nearly* the lowest overall", "*almost* as good", "*higher*-frequency turbulence" (what frequencies?), "would be desirable to have *a lot* more datapoints", "the agreement is *very good*", "to reduce *some* individual turbine setpoints", "*excellent* agreement". I would be useful if the authors refrain from such vague statements, and rather use exact terms. For example, "the turbine is yawed, *typically by up to 30 degrees*, out of the wind direction".
- Most of the references are technical documents of the CL-Windcon report. I suggest substituting these references as much as possible with scientific articles, though I understand this is difficult.

Specific remarks:

Some of these comments may have already been addressed earlier, but just for completeness sake, I am putting down all the small things I have noticed here while reading the article:

- Page 1: Abstract: *Horizon 20-20* or *Horizon 2020*?
- Page 1: Introduction: I am missing an inherent motivation of why this work is important. Perhaps the recent "Expert Elicitation on Wind Farm Control" paper by Wingerden et al. can help motivate this work. That paper shows a survey among experts which concludes that validation is currently the most important step before adoption by the industry.
- Page 1: A discussion on the effect of axial induction control on the pitch actuator duty cycle and the structural loads is missing. It would be good to address this, at least briefly.
- Page 2: Section 2: mention that it is an *onshore* wind farm.
- Page 2: Section 2: A wind rose and possibly a flow field from LongSim with wake interactions for WTG 31-38 could be insightful for the reader.
- Page 3: Figure 1: Why is there a blue box around this figure?
- Page 3: The citation "Knudsen et al." should be "van Wingerden et al." or "Doekemeijer et al.", considering TUDelft led this deliverable / report in CL-Windcon.
- Page 3, section 3.1: SCADA data was used for model comparison. Does this dataset contain turbine curtailment/derating? This would important to discuss in your article.
- Page 4: Perhaps add a citation for the bulk Richardson number
- Page 5: Perhaps add a citation for the Obukhov length
- Figure 2: Why does the x axis go until 25, rather than 24? Perhaps just hide 25.
- Figure 3: Missing units next to RMS, and remove the grey border around the figure
- Figure 4: The figure spans multiple pages, so when looking at page 6, it's not clear that the caption belonging to it is on page 8.
- Page 8: Explain the statement "However, for the purposes of the Sedini experiment this would not be possible to arrange".
- Page 9: Section 3.3: the first paragraph is written hypothetically. "in general", "would be", "can provide", "can usually", "can be", "could then". I would suggest the authors to keep a narrow focus on their own work, rather than explaining general methodologies/guidelines. A similar writing manner occurs in Section 3.4: "… in whatever measurements are actually used" and "wind conditions may not be the same".
- Page 10, section 3.4: when talking about including uncertainty in the optimization, perhaps cite Rott et al., Quick et al. and/or Simley et al. that have published on this topic.
- Page 10, "this has the advantage of faster optimisation, but also …". Perhaps it would be good to also discuss the disadvantages of smoothing the setpoints (and in this manner).
- Section 4: Generally, I would omit this section (in line with major comment 2). Sections such as 4.1 are general descriptions of LongSim and it might improve clarity if the authors would cite existing literature instead. If anything, important values can be collected in a table. In line with this, I think figures 5-13 contain too much information/would be better placed in a separate article.
- Figures 5 and 6 appear to be inconsistent. In Fig. 5, Turbine #38 is the title, while in Fig. 6, turbine #38 is part of the label. Also, the ylabel is missing and only units are given in Fig. 6. Generally, I suggest the authors to reconsider whether each figure is essential to the article and whether the style and size are consistent throughout.
- Figure 7 is missing an informative ylabel

- Figure 8: I can only see one line: the final smoothing one. Perhaps the authors can consider removing this figure and just stating this in their text.
- Figure 9 only has units on the ylabels, not the actual variable. The figure also appears before being mentioned in the text, which may be confusing to readers. Moreover, I suggest the authors to consider removing this figure.
- Figure 10: missing units on ylabel. Also, the purpose of this figure is unclear to me
- Figure 11 is missing a ylabel, and I suggest the authors to reconsider the need for this figure.
- Figure 12: may be removed, also not units or description on ylabels
- Figure 13 is missing a ylabel and units, xlabel should be **Time [hours]**, and may be removed
- Page 18: "at 1-minute resolution", this has been mentioned at least twice before in the article. I would suggest the authors to not repeat such information.
- Page 19: "it would clearly be desirable to have a lot more datapoints to give more confidence in the results", based on my reasoning in my major comment (1), I would expect that even with an infinitely large dataset, you will still not get rid of the large variations that you are seeing in the results. I think it would be essential to (at least) bin data according to turbulence intensity as well. (see major comment 1)
- Figure 14 contains no units on xlabel
- Page 20: the authors talk about a "mean" increase. Is this the mean of the mean of all bins? Is this the mean of all datapoints? Is this weighted according to the wind rose? Also, with variations of up to 70% and -40%, I have my doubts about taking a *mean*. The error margins would be very large. Please discuss this in your article.
- Figure 15 and 16 have differently sized text
- Page 22: "the predicted overall increase, 2.38% is very similar to the field test result of 2.42%". Though, from what I can see, the actual values between Fig 15 and 18 are quite significant. The fact that the mean values coincide does not serve as a validation by itself. It seems somewhat coincidence that the mean values are so close together.
- Figure 23: I was surprised the article went back to simulations here in Section 5.3.2 after having already discussed simulations until page 17, and then having discussed the field experiments.
- Figure 19, 20, 21: are these figures essential to the article? For me, I find it hard to derive any information from these figures, since they are so noisy and show data over such large timescales. Nonetheless, if the authors decide to keep these figures, please look at the ylabels and units. Also, in previous figures, units were given with square brackets around them. The font size also changes significantly between these figures.
- Page 25: generally, if a model validation is to be made in this article, I think quantification is essential. Perhaps first explain what is important: is it the absolute power by the turbines? Is it the relative power production compared to T1? Then how can you quantify the error/accuracy of the model? Do you want to validate steady-state or dynamic effects? I generally find that the authors make important steps in comparing their model, but I am missing a more informative discussion on what is necessary, how to quantify it, and the quantitative results.