Response to referee report 2 (Responses shown with yellow highlight)

Dear authors,

First of all, thank you for your responses to my comments. You have taken my points seriously and answered my questions with care. The figures have significantly improved, the document is easier to read, and notes and explanations have been added contextually. Also, not all comments led to a change in the manuscript, and I fully understand that.

Though, one important concern of me remains with the current publication. The conclusion of this article invokes the idea that axial induction control provided an estimated gain in power production of 1.7% to 2.4%. I am not saying that I disagree with these numbers, but at the same time I am not yet fully convinced after reading the paper. As mentioned before and as you have discussed in your manuscript, there may be other factors at play. I believe you can investigate this more, which would bring more confidence to these numbers. When I look at figure 15, I see that for basically every bin, the mean turbulence intensity for ON is higher than OFF. A higher turbulence intensity may give a higher power production in these situations. This argument is supported by looking at bins with directions 185 to 205 degrees, where axial induction control is not really applied and yet gains are large. Similarly, when looking at the left subplots of Figure 15, the largest gain is at the bin with WS = 10.5 m/s, which also has the largest difference in TI between ON/OFF. At other bins with almost equal TI between ON/OFF, we also see barely any gain in power production.

Now, when we look at the right subplots of Figure 15, things get better and I think this does indicate a gain due to axial induction control, specifically for bins 235 deg and 255 deg, which have almost no difference in TI. I believe it is fair to say that these bins indicate a gain, but I think it is not fair to say that *every* bin indicates a gain due to axial induction control. I think one must investigate the effect of the turbulence intensity (among others) first, before being able to conclude this. Similarly, if you use all data to calculate mean values like on page 21, then you will be comparing ON data with an effectively higher mean TI with OFF data with an effectively lower mean TI. There may indeed be a 2.42% gain between ON and OFF, but perhaps this is 2% due to TI and only 0.42% due to axial induction control. I would suggest the authors to be very careful making such statements. I anticipate this to be a high-impact paper and these numbers are likely to be used as reference values for the potential of axial induction control. An additional comment has been added to emphasise this. I think it's already stated clearly enough for the right-hand plot. A similar comment has been added to the conclusions.

Now, the additional figures, especially Figure 18, adds significant value to the manuscript. From this figure, I conclude that the mean TI between ON and OFF is not very large. This is a very interesting observation and supports the conclusion that the authors already make.

To strengthen the conclusions from the authors, I would suggest diving deeper into Figures 16 and 20. LongSim was used to (re)simulate the measurement points and also reported significant gains. Using LongSim, you could figure out where these large gains come from. If in the “ON” dataset you simulate them with the baseline (OFF) controller, do you still see such large gains? If so, then the gains are not due to axial induction control, but due to other effects. Similarly, what if you simulate the datapoints in LongSim with a fixed value for TI – that could provide insight into the effect of the turbulence intensity on the power gain. By doing these manual simulations, you will gain much insight into where the gains are really coming from, and perhaps why the values are so high. Since such large gains are also seen in LongSim, it must be more than statistical uncertainty. Actually I did some investigations along these lines before submitting the last version, but it was a bit inconclusive, and not useful enough to be worth making the paper even longer.

A second idea to give more insight into the effect of TI is by redistributing the bins in Figure 15 so make sure their average TI is identical. For example, one could make duplicates of low-TI “ON” measurements or make duplicates of high-TI “OFF” measurements to bring the average TI to the same value. Now, with the datasets already being so sparse, one may introduce additional bias, so perhaps the former suggestion is better. Figuring this out will also prevent vague statement such as on page 20 “it is possible that […] might account for [..] some […] power increase.”

Smaller comments:

- Abstract: “show a positive increase in energy production resulting from induction control” is a strong statement. If anything, perhaps rephrase it to something like “the experimental data suggests that induction control leads to both gains and losses in power production, with the gains outnumbering the losses.” Some additional words have been added at line 15 about there not being enough data
- Figure 14: for consistency, it would make sense to put the plot title on the ylabel instead. y-axis label has been addedSame goes for Figure 21. Also, for Figure 21, perhaps change the xlabel to “Time [Hours]” Done
- I am not sure if Table 3 adds significant value I am surprised at the comment, as the table was added to provide more transparency, as a way to deal with previous comments from this reviewer, to help readers draw their own conclusions.
- Figure 23: the authors state that the agreement is very good. Though, I believe at this scale it’s hard to draw this conclusion. The power production may be off by 30% at any point in time, especially at such low absolute power values. A note has been added around line 500 about expected discrepancies at higher frequencies, etc. which we believe answers this point.

Response to referee report 4 (Responses shown with yellow highlight)

The article is interesting, well written, and deals with the field experimentation of a cooperative wind farm control strategy that aims at increasing the overall power of a wind farm through an appropriate reduction of the power produced by the upstream machines, whose effect is to reduce their axial induction and thus to mitigate the speed deficit in their wakes. Precedent experiments conducted in wind tunnel and simulated environment (properly cited by the authors) have shown that this method seems to be not very effective. As well as the previous experiments conducted in real life have shown extremely small gains in the order of the measurement uncertainty, leading to inconclusive results.

This paper has the merit to describe in very detailed way the experimental setup and the methodology used for the synthesis and the implementation of the controller, as well as the simulation model used to the purpose. For this reason, I think that it is worthy to be published in the journal.

As stated by the authors, the data obtained through the experimentation are not sufficient to derive statistically robust results. Moreover, the used simulation model shows a partial agreement with the experimental data, especially regarding the predictions for the machines located more downstream in the array. The differences between the model-predicted and the experimental overall-cluster power (not shown in the paper, even if they should be) are probably superior to the power gains measured experimentally, a fact that suggests that the model cannot be used to validate the obtained results.

I believe, therefore, that the results obtained do not allow to conclude on the effectiveness or not of the tested method in terms of boosting the wind farm power output. This aspect should be emphasized in the conclusions, A further comment has been added to the conclusions. which should also include what should be, according to the authors, the actions required in order to reach a satisfactory conclusion on the effectiveness or not of the proposed method. A further sentence added to the conclusions for this.

I also report in the following some other suggestions for improving the manuscript.

- Page 1, line 21-22. The control concepts the sentence refers to are not introduced before. This has been corrected
- Section 3.2 what exactly the control set-point is? Is it the power reduction, expressed in percent of the available one? This is described, in as much detail as permitted by the manufacturer, on page 13 (line 267) I think it is important that the authors clarified this aspect. Moreover, I think it is important to show here some resulting LUT, and quickly comment how close the computed optimal set-points are with respect to those adopted by other authors whose findings have been cited in the introduction. This is not possible as the setpoints are completely dependent on layout and spacing, so there are no comparable examples in the literature.
- Page 18-Line 359: to allow the reader understanding why what is shown in Fig 14 is a "small set-point", it must be clarified before what the set-point is. The definition of the setpoint is dealt with as far as possible as explained in reply to the second bullet point. The reference to small values of the setpoint has been removed as it does not add anything significant.
- Page 26, Line 669-470. I would not claim that 20 degrees of difference in the wind direction is a slight difference. The wake-to-turbines interaction with 225 degree wind direction is totally different from the one with 245 wind direction. the word 'slightly' has been deleted
- Figure 20. It would be very beneficial to put aside of Figure 20 a similar figure that shows the delta between measured and simulated “power ratio ON/OFF”. There is the space for it, and it would allow the reader to quickly get how good the predictions of power gains are with respect to the measured data. I don't think this would add much. The reader can get the full picture by visually comparing figures 20 and 16. A plot of the differences is more noisy and not so easy to interpret.
- Page 28- Figure22, WT13 data. I would not claim that the agreement is that good as it was for WT 38-37 and, partially, also for WT36. In many instants, indeed, the predicted power is more than double the SCADA data. This probably means that the simulation model is overestimating the wake recovery as we look further downstream in the WTs array. An additional comment about this has been added around line 500.
- Fig. 22: it would be very interesting to show here also the comparison between the measured and simulated overall cluster power. The control set-point are indeed derived with the goal of maximizing the overall cluster power, and LongSim is used as simulation model. it would be therefore very interesting to check how good the model is in capturing this quantity. This extra plot has been included.
- Fig. 22 and 23: I personally don't think that visually comparing time series is the best way to judge the agreement between numerical and experimental data. I would have instead plotted the numerical data w.r.t. the experimental data, also including the correlation factor and the related RMS. Because the unmeasured high frequencies are synthesised, small mismatches in time would lead to large scatter in such plots which does not reflect on model accuracy. The time series plots allow one to see such effects more clearly.
- For what shown in the paper, I would not be confident claiming that LongSim provided an excellent agreement with the experimental data for all WTs. Fig 22, for example, clearly shows that WT13 power is quite often overestimated. I have already commented on this.