Reply to Reviewers:

We would like to thank the reviewers for their comprehensive review and insightful comments, as well as to the associate editor for highlighting important concerns regarding our manuscript. In response, we have addressed each point raised and we believe the paper is now fundamentally better than before. Below, we provide the detailed responses to each comment, and we hope that our revisions addressed all the concerns which the reviewers and the associate editor had.

Associate editor:

Associate editor: *Thank you for your revisions to the previous version of the manuscript. As the reviewers indicate, many of the initial concerns have been addressed. However, the following concerns should still be addressed.*

1. **From Reviewer 1***: regarding the possibility of blockage – reasoning on excluding the blockage from consideration should be incorporated into the text, perhaps using the response to reviewers, 'we did not observe any indications of spatially varying wind directions attributable to blockage for this farm.'*
   Author response: We have now incorporated a portion of text mentioning the reason why a blockage model is not incorporated within the case study wind farm.

2. **From Reviewer 1***: regarding binning, please also incorporate the reasoning for avoiding binning into the manuscript text.*
   Author response: We have now incorporated the reasoning for avoiding binning into the manuscript text.

3. **From Reviewer 1***: a more thorough justification is required of the choice to not compare results against an uncalibrated model.*
   Author response: We have decided to incorporate a comparison of results against the uncalibrated model. This comparison can be observed in Figure 15, with some additional text explaining the error metric. The error metric is based on the metric applied in Nygaard N. G. et al., (2022). We updated Figure 9 and Figure 10 to comply with the same metric, as described using Equations 3 and 4

4. **From Reviewer 2***: a more thorough discussion is required of contributing factors (and resulting limitations of this study) that cause differences between the relationship of wake growth rate and turbulence intensity seen here and the rest of the literature. (See Reviewer 2's extensive first concern.)*
   Author response: We have now incorporated this into our paper, discussing the factors contributing to the observed differences between the wake growth rate compared to existing literature. Here we mainly focus on the turbulence intensity and the scale difference of our cost-function metric when compared to the rest of the literature. Furthermore, we acknowledge the limitation of assuming a constant turbulence intensity in our conclusion section, where we also highlight the potential impact of site-specific characteristics on the calibration outcomes. Consequently, we explicitly advice against the direct application of our calibrated tuning parameters to different sites without further recalibration.

5. **From Reviewer 2***: A thorough discussion of the filtering criteria used to justify the approach.*
   Author response: We identified that this additional filtering criteria was necessary due to either above-rated curtailment, reaching values close to rated power production, or minor underpredictions of power prediction at rated capacity. It became clear that a one-size-fits-all approach might not be suitable for every wind farm, therefore we tried to explore alternative methods. Consequently, we have refined the power-curve filtering process to more precise exclude data points related to curtailment near the rated power. We use the method proposed by Doekemeijer, Simley and Fleming (2022) with stricter margins. This can be observed in Figure 3.

Additionally, we have considered the remaining reviewers' comments as follows:

Reviewer 1:

1. **Reviewer (first review)***: Line 145 – 152: This paragraph just states the blockage topic. But not how it plays into the challenges of calibration through SCADA data.*
   Author: Blockage, when overlooked, can introduce additional complexities in SCADA data interpretation, especially in large wind farms. The blockage effect could adjust the observed wind speed and wind direction. For the specific wind farm in question, which is part of a large cluster, modeling the blockage would present significant challenges. Fortunately, we did not observe any indications of spatially varying wind directions attributable to blockage for this farm. However, considering the larger picture including neighbouring wind farms, blockage cannot be ignored. Therefore, we felt it was necessary to address this in our study.
   Reviewer: *Again, blockage could also occur from a single wind farm. Therefore, the argument "we did not observe any indications of spatially varying wind directions attributable to blockage for this farm" should also be taken into the text.*
   Author response: We have now described this within the text (see author response 1 to the associate editor).

2. **Reviewer (first review):** *Line 171: What is prohibiting this type of analysis for binned observations?*
   Author: Binned analysis assumes balance: It is valid when the magnitude and frequency of overestimation are in balance with the frequency of underestimations. Otherwise, results can be skewed. Additionally, the volume of usable data becomes limited in binned observations, since even the downtime of a single turbine can introduce significant skewing.
   Reviewer: *This reasoning should also be reflected in the paper text.*
   Author response: This has now been added to the paper text (see author response 2 to the associate editor).

3. **Reviewer (first review):** *Figure 4 & 5: Can the authors provide a definition of the displayed metrics? Which quantity was used for normalization?*
   Author: We have now added the Equations to the Figures

Reviewer: *Technical suggestion: Because of the fraction you could put the definitions not in the figure caption but in the paper text where there is more space to introduce them.*
Author response: We have now removed the Equations from the Figure captions, and instead added it to the text, together with an additional explanation of the parameters.

4. **Reviewer (first review):** *Section 4.2: It would be good if the subclusters can at least be described a bit more in their configuration. Furthermore, the discussion should also include the results from a baseline model that is not optimized for comparison.* **Author**: A Figure has been added that shows the coordinates of the wind turbines within the farm.
Reviewer: *An answer to the second part of the comment was given to the other reviewer. The authors have decided against comparing their results to an uncalibrated/different model. In my opinion this is of critical importance in the future to benchmark the proposed method.*
Author response: We have decided to incorporate a comparison of results against the uncalibrated model. This comparison can be observed in Figure 15, with some additional text explaining the error metric (see author response 3 to the associate editor).

Reviewer 2:

1. **Reviewer:** *I was not satisfied with the authors' response to the first general comment of Reviewer 2. The found relationship between the wake growth rate and the turbulence intensity differs from a body of literature that observed a stronger relationship from field experiments (e.g. Trabucchi et al., 2017), wind tunnel observations (Ishihara and Qian, 2018*), and large-eddy simulations (e.g. Niayifar and Porte-Agel, 2016). In my opinion, a discussion of what might contribute to the differences and mentioning possible limitations of the present study is required in the manuscript given the differences.*

   Reviewer: *The authors replied to the general comment no. 1 of Reviewer 2 that their 'decision to focus on the velocity deficit model and to exclude others, like the wake-added turbulence model, comes from the interdependencies these models show. Including both submodels could lead to a multimodal solution space, where certain parameters in one submodel influencing adjustments in another." From this, I understand that the optimized parameters might change if those submodels are changed. Therefore, the found wake growth rates and the conclusion that 'comparing the optimized parameters to the baseline reveals that the baseline parameters underestimate the wake effects, which subsequently leads to an overestimation of the expected yield' should be softened by mentioning the limitation in the paper explicitly in my opinion.*

   Reviewer: *However, there are also possible reasons that can be mentioned why it might be different from the studies cited in the first paragraph above. They only used isolated wind turbines to determine the relationship between turbulence intensity and wake growth rate. It might be that the relationship is naturally different inside a wind farm, because the turbulence of the interior wind farm flow changes not only in its intensity, but also in scales. Some papers showed that the thrust coefficient of the wind turbine affects the wake recovery as well.*

Author response: We acknowledge the limitation arising from our decision to analyze the wind farm at a single TI of 0.06. The choice to go for a TI of 0.06 was based on the work by Doekemeijer B. M. (2022), but we recognize that this assumption will make the results only representative for a given TI of 0.06, meaning that changing the TI in the model will require additional calibration. We have incorporated this to both the optimization results section and the conclusion section.

Moreover, we are aware of the differences between the above-mentioned papers and our acquired results. Our analysis focuses on the collective behavior of the wind farm, where the mentioned studies focus on a single turbine wake, where the focus is on a given number of rotor diameters behind the turbine. Tuning on such scale can make the tuning parameters account for flow physics, which are inherently different to those compared to the single wake case.

Nevertheless, assuming a TI equal to 0.06 is a conservative estimate for AEP calculations compared to other available options, such as calculating TI from SCADA, met mast measurements, or by using IEC standards as reference, which could lead to larger deviations from the reference value determined by Niayifar and Porte Agel (2016). Our focus is on quantitatively optimizing the performance of the entire wind farm, rather than qualitatively analyzing the wake effect of an isolated wind turbine. The effect of turbulence is something we are willing to study in the future.

In conclusion, we recognize the limitations associated with our chosen TI and the observed differences between our findings and the reference literature and we have now addressed this within our paper, both in the optimization results section and the conclusion section.

Reviewer: *Regarding the implementation of the model that was optimized here, the reply also did not address my initial concern. It is true that nacelle mounted anemometers are not perfect, which can be seen from a comparison to upstream looking lidars/upstream located met towers that have been published for some field campaigns. However, the power curve provided by the manufacturer is of unknown quality (or at least it is not provided in the manuscript) and the power curve might have been created for an undisturbed inflow (e.g. an inflow that can be described with a log- or power-law). It remains the question if and how it can be applied to a waked or partially waked wind turbine in the interior of a wind farm. The reply of the authors did not rule out that the differences in wake growth rates found here and those in other literature might be partly explained by such issues. Therefore, I am missing here as well that the author mention such possible limitations of their findings.*

Author response: We have refined the power curve filtering process, as visualized in Figure 3, by narrowing the filtering boundaries. While I do agree that the power-curve might have been developed assuming an undisturbed inflow, it is still unclear which approach more accurately estimates wake losses. For instance, in scenarios where a wind turbine experiences a small partial wake effect on either its left or right side, the impact on the nacelle-mounted anemometer is expected to be minimal, yet a reduction in power production is expected. In our opinion, a comprehensive analysis of such cases

would necessitate a parallel study with the cost-function based on wind speed rather than power production. Nevertheless, we will indeed acknowledge this limitation, along with the previously mentioned ones, in our paper.

2. **Reviewer:** *Section 2.4 / line 252 – 259: The filtering criterion is still mysterious to me and the clarification by the authors have not improved this. The authors say that they sometimes observe discrepancies between the active power and the power set point at wind speeds above rated wind speeds. Addressing this issue with threshold for the ratio of the wind speed variance is still not intuitively clear to me. I suspect that the criterion might work for the authors special circumstances at their experimental site, but I doubt whether it is a universal approach to detect this issue with the turbine operation (I wanted to test this on a set of SCADA data to verify but I did not find the time in the end). Therefore, I believe the authors should add the following information to the paper:*
   (1) A summary of the impact of this filter criterion (How many data points does it remove from the total? Can it be shown that that it is successful at removing what the authors describe?).
   (2) It should be clearly stated whether the filter criterion is based purely on a correlation or whether there is also a physical causation to support it (and in case of the former that it might not be a procedure that can be universally recommended to identify this state of turbine operation).

Author response:

We identified that this additional filtering criteria was necessary due to either above-rated curtailment, reaching values close to rated power production, or minor underpredictions of power prediction at rated capacity. It became clear that a one-size-fits-all approach might not be suitable for every wind farm, therefore we tried to explore alternative methods. Consequently, we have refined the power-curve filtering process to more precise exclude data points related to curtailment near the rated power. This can be observed in Figure 3.

References:

Doekemeijer, B. M., Simley, E., Fleming, P.: Comparison of the Gaussian Wind Farm Model with Historical Data of Three Offshore Wind Farms, Energies, 15, 1964, https://doi.org/10.3390/en15061964, 2022

Nygaard, N. G., Steen, S. T., Poulsen, L., and Pedersen, J. G.: Modelling cluster wakes and wind farm blockage, Journal of Physics: Conference Series, 1618, 062-072, https://doi.org/10.1088/1742-6596/2265/2/022008, 2020