**Review for WES-2023-98**

**General Comment:**

The study proposes a novel method of site-specific wake model calibration using SCADA data. This is a relevant area of scientific research. SCADA data is becoming more abundant and low-fidelity wake models are essential for wind farm site-assessment and control. The study first gives an overview of different wake models and the challenges in calibration. Then, the study describes the data filtering of the investigated offshore wind farm. Finally, the optimization framework and results are presented.

The presented idea for calibration is worth exploring, furthermore, the study highlights calibration problems due to the influence of neighboring farms. A problem that will become more important in the future. However, the study contains some shortcomings. Firstly, the novel calibration method itself is not the center of the paper. Instead, a lot of room is given for explanation of different wake models (some specific to the FLORIS framework) and filtering of SCADA data. The wake models have been described in many other papers. Furthermore, probably due to non-disclosure issues, the study does not describe the investigated farm, no layout is provided. This makes it much harder to visualize it for the reader.

The result section does not thoroughly analyze the results of the method. The method does not only contain the wake model calibration, but also the wind direction and wind speed adjustment which influence the results heavily. Furthermore, no comparison to a baseline model is provided.

The authors should rethink what message they want to convey with this study. *Some* possible: Is it about the optimization method? Then it should be less FLORIS-centric and give not so much room for the applied wake model. The framework should be general. Instead describe the optimization, the hyperparameter tuning and results more in detail. Is it about the uncertainty of the parameters of the Gaussian wake model? The parameters should in principle be applicable to a wide range of conditions and not change every 10 minutes. Then the authors have to describe what we learn from this volatility. Is it about the whole toolchain with filtering etc.? Then, it should be better highlighted what could be an application for the community or industry.

**Specific comments:**

Section 1.1 This section gives an overview about different wake models in FLORIS. Can the authors motivate why this section is necessary for the paper? In the end, the developed methodology should be applicable for any wake model.

line 54-55: “...there is no velocity profile...” is not correct, the wake has a top-hat velocity profile

line 55: What do the authors mean with linear wake decay? The velocity deficit does not decay linearly in the Jensen model.

line 58: “overestimates the *wake* at the edge of the wake” probably “velocity deficit” is meant here

line 60: “Gauss-legacy” is just a name of an implementation inside the FLORIS V2 package. Non-FLORIS users are not familiar what “legacy” means. On top of that, in the FLORIS V3 version (which is cited) the model is called “Gauss”

line 60. “...widely used in industry...” can you provide a reference for this statement, e.g. a survey?

line 66: “...it has become clear that traditional wake models often underestimate wake losses...“ can you provide a reference for this statement?

line 68-75: Why are several wake models described? If the point is that other wake models have additional parametrizations, it would be enough to give just one example

line 76: The word “scaling” is a bit confusing and unnecessary to refer to the wake model parameters

line 79: Here it sounds like if it is a *must* to calibrate the model with CFD results and it is always done like that. However, it could be also a wind tunnel or LIDAR measurement in the first place.

line 99: What do the authors mean with “...different results were obtained.”

line 109-113: It seems that in these sentences the authors actually mean the nacelle anemometer instead of wind vane? The wind vane is used to measure the wind direction. What does a bias of 5° mean for a wind speed measurement?

line 113: The wind speed also has a non-linear relation to the active power of the wind farm. What is mean with the “… calibration of the wind direction ...”

line 122 It sounds like *spatial* correction factors were applied on the wake deflection parameters which is not the case in that study.

line 136: What does “homogenoeus” mean here?

line 145-152: This paragraph just states the blockage topic. But not how it plays into the challenges of calibration through SCADA data

line 155: Can the authors motivate the use of the GCH model? As mentioned in sec. 1.1, the additional modifications are only for yaw effects, which are neglected in this study.

line 160: Can you justify not accounting for transients? Is the farm small enough or can you show that these effects are not present?

line 160-164: In the previous section (line 139ff) the authors state that stability, which is correlated to TI, is important for calibration. Can you justify neglecting it? Would it not increase the uncertainty in the calibration parameter?

line 164: “wake blockage” is not a common term. Do you mean “farm blockage”? The farm also has a blockage effect by itself.

line 171: What do the authors mean with “...essential to not skew the calibration results...”.

line 171: What is prohibiting this type of analysis for binned observations?

line 190: The later part of the study suffers from the fact that the reader has no image of the farm layout. It would be helpful if the authors state the number of turbines and the type of layout (irregular/regular), average spacing.

Figure 2: The 3D depiction is a bit misleading as it could be a pie-chart that reports percentages. The figure should be clearly identifiable as a compass rose e.g. also with labels in degree

Figure 4 & Figure 5: Can the authors provide a definition of the displayed metrics? Which quantity was used for normalization?

line 245: What do the authors mean with "...above rated SCADA data..." ?

Figure 8: Normally TI is defined as std(ws)/ws so figure 6 and 8 are not very different. It is just that 6 is additionally normalized by the wind speed. Why is it important to show also figure 8?

line 265: Typo "… , The ..."

Table 1: Secondary steering as model is not necessary as no yaw is considered in this work

line 295ff.: Can the authors elaborate on the point that the sensitivity increases for specific directions? Are there higher or lower wake losses? The lack of a provided layout makes it difficult here.

Equation 8: Can the authors motivate the use of this error type? It seems that this is an error on farm level. Positive and negative turbine errors can cancel out. To improve the wake calibration, shouldn't the error be calculated on turbine level?

line 323: Can the authors state the found weighting between a and b?

line 331: The allowed range seems quite high. As the authors stated themselves, the wind speed has the largest impact on the farm power. How often did the optimization go to this extreme value? How can the estimated wind speed from the SCADA data be so far off?

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.

Figure 15 & 16: The figures have similar captions. It should be clear that they represent different subclusters.

line 395ff. It is not clear whether some directions were excluded from the optimization. In sec. 2.3 the authors state that directions 250-50° were excluded, yet the figure 19 (and also 15 and 16) display all directions.

Code availability: This is just the code of the open-source packages used. Not the own code developed for this study