

The manuscript “Bayesian uncertainty quantification of engineering models for wind-farm atmosphere interaction” by Aerts et. al. presents a Bayesian UQ framework for calibrating and comparing wind farm flow models. The Bayesian formulation represents both model error and measurement error and is demonstrated for two engineering models given LES data of a wind farm situated in a conventionally neutral boundary layer. It is shown that, for this dataset, the Bayesian framework quantifies the reduced model error in the higher fidelity model when compared to the lower fidelity model, as well as the uncertainty in each model parameter. Overall the paper is well written, and the examples and figures are effective at demonstrating the utility of the Bayesian UQ framework. High-level comments and line-specific comments are enumerated below.

1. Please ensure that all symbols and variables are defined in the text throughout the manuscript. I found myself inferring the meaning of several statistical quantities, which may be difficult for the general wind energy audience. A nomenclature section at the beginning of the manuscript would greatly help to enhance the readability of the paper.
2. The authors highlight the importance of assessing model uncertainty to provide a robust evaluation of AEP predictions from low-fidelity engineering models, particularly for applications such as wind farm flow control and layout optimization. In wind farm flow control, AEP is often estimated using engineering models (such as the standard wake model considered in this study) across a range of control parameters, with typical AEP gains over baseline controls of around 1% or even less observed. Given the methodologies and results presented in this manuscript, can the authors provide any insight into whether such small predicted power gains are at all meaningful in light of the large uncertainty and model error observed with the low-fidelity engineering model, such as those shown in Figure 6? Is this error something that can be improved with more data, or are the deficiencies in the model simply too large for low-fidelity engineering models to make meaningful wind farm planning decisions? While rigorously addressing these questions may extend beyond the scope of this manuscript, commenting on them in the conclusion or suggesting them as directions for future work could provide a valuable perspective for the wind energy industry.
3. L44: Define the first instance of “UQ”, and please ensure all acronyms are defined the first time they are used.
4. L185: What is the justification of using the 0.1 value? Can this be clarified in the text?
5. L320: In standard wake models such as an empirically tuned Gaussian model like the one described in Section 3.1.1, it is common to introduce a multiplication factor to scale the inflow wind speed or power coefficient of the turbine to account for the turbine induction field, blockage effects, and other discrepancies with BEM theory. Is such a scaling factor included here? Would that improve the performance of the standard wake model in Section 4, particularly the large modeling error observed in the first row of turbines in Figure 6?

6. L327: This sentence may give the false impression that the wake expansion coefficients are fixed in this study to values of  $k_a = 0.3837$ ,  $k_b = 0.003678$ . Could this be reworded to clearly state that these are values found in a previous study, and that the wake expansion coefficients will be included as uncertain parameters for the Bayesian UQ analysis?
7. L334: Can additional information be added to this sentence to briefly describe the wake merging method used in this study, and importantly if any additional uncertain model parameters go into this wake-merging model?
8. L349: Consider rewording or splitting up the last sentence of this paragraph — “whereas” may not be the appropriate conjunction here. Are there calibrated parameters associated with the pressure feedback?
9. L359: A primary result of the paper is comparing model performance of wind farm blockage effects. In this sense, it seems counter intuitive to justify ignoring dispersive stresses because they are primarily limited to the farm entrance region. How important is this choice and should that be elaborated on in the text? Is the decision to ignore these stresses represented by the model error term?
10. L393: Please check the symbol for the potential temperature.
11. L405: Please clarify the wording in the last two sentences of this paragraph. Is the intention to state that: (1) for the same number of cores, the standard wake model is ten times faster than the APM model, or (2) because the algorithm can be run in parallel, the total simulation time is effectively the same for both models?
12. L410: Consider adding a description for  $v_e$  and  $v_b$  the first time they are used in the text as the wake model and atmospheric perturbation model parameters, respectively.
13. L414: Intuitively, it would seem the predicted power of the upstream turbines is not sensitive to changes in the wake expansion rate parameters. This sentence seems to suggest the opposite is true. Please double check this sentence. If it is correct, can the authors elaborate on why that is the case?
14. Section 4.1: This section is a nice inclusion of the paper, and provides a clear demonstration of the importance of representing model error in the UQ framework. However it feels out of place in the manuscript and disrupts the discussion between Section 3.3 to 4.2. Would it make sense to put this discussion in Section 2 to motivate representing model error in the Bayesian formulation, or perhaps as an appendix? Please note, this is just a suggestion, and the authors should use their discretion. Either way, nicely done with this example.
15. Figure 8: The posterior distribution suggests negative values of  $k_b$  may be optimal. Would it make sense to allow negative values of  $k_b$ , as long as the total wake expansion rate,  $k_w$ , remains positive? The same considerations may be relevant for Figure 9.
16. L563: “Layouts”

17. Inconsistent usage of “wind farm” versus “wind-farm” throughout the manuscript, including the title.