

Reviewer:

Comments to the Author

The paper "Validation of RANS-calibrated engineering models and ANN-based surrogate for wind farm flow simulation and layout optimization" by Jens Peter Schøler, Ernestas Simutis, M. Paul van der Laan, Julian Quick, and Pierre-Elouan Réthoré presents a comparison between an Artificial Neural Network (ANN)-based surrogate trained on Reynolds Averaged Navier-Stokes (RANS) data and two representative engineering wake models based on the TurbOPark and Super-Gaussian formulations. The study also involves recalibrating the engineering models. a systematic flow simulation study across varying turbine counts and spacings, and WFLO benchmarks validated against RANS-based Annual Energy Production (AEP). Results show that the ANN surrogate achieves the lowest RMSE and MAPE across all scenarios in flow estimation, albeit at a higher computational cost. In WFLO, the TurbOPark-based model produced the highest RANS-validated AEP layouts, despite having lower predictive accuracy, suggesting that optimization complexity influences outcomes. Blockage modeling increased computational cost without improving accuracy.

This article is a highly valuable systematic validation paper. The author has conducted a thorough comparison and re-calibration of the commonly used engineering wake models (TurbOPark and Super-Gaussian) and the agent model based on artificial neural networks (ANN) in wind farm layout optimization (WFLO). Through controlled experiments, the paper reaches a striking conclusion: the ANN and Super-Gaussian models, which have the highest accuracy and the smallest error in flow field prediction, do not produce the layout scheme with the maximum annual energy production (AEP) in the final WFLO; conversely, the TurbOPark model, which has relatively lower flow field accuracy, yields the optimal solution. This finding reveals the complex relationship between "model fidelity" and "multi-modality of the optimization objective function". Overall, the research method is rigorous, and RANS data is used as the "ground truth" for closed-loop validation. However, I have some questions regarding the design and editing of the article, which will be beneficial for the credibility and rapid publication of the article. I recommend **accepting** this manuscript for publication after minor revisions addressing the following issues. While the study presents interesting findings, addressing these points will enhance the credibility and overall quality of the article, facilitating its rapid publication.

1. Introduction

The Introduction section could be further strengthened by expanding the breadth and depth of the literature review. To provide a more comprehensive background for your study, please consider the following suggestions:

- **Broaden the Citation Base:** Certain parts of the text (e.g., lines 11-15, 23-27, and 34-40) currently have relatively few references. It is recommended to introduce additional citations to better demonstrate familiarity with recent industry developments, such as the dynamic wake modeling (PhyWakeNet) discussed by Liu et al. (2026), or other similar recent works.
- **Balance the Research Scope:** While the current introduction focuses well on wake model reconstruction, the main body of the paper extends into Wind Farm Layout Optimization (WFLO). It would be beneficial to supplement the review of optimization methodologies to ensure the introduction aligns with your core work. For instance, referencing the reinforcement learning-enhanced genetic algorithm by Dong et al. (2026) could help bridge the gap between aerodynamic modeling and layout application.
- **Include Blockage Effects:** Additionally, incorporating more discussion and references regarding blockage effects would help provide a more holistic context for the aerodynamic challenges addressed in large-scale wind farm studies.

2. Methodology

(1) In the Methodology section, tables 3 and 4 are not standardized. Please split the tables.

(2) Although the author has introduced the research methods and formula descriptions in detail, there are still areas for improvement and refinement in the experimental design. In the 2.5 Wind farm flow study section, although it has a certain systematicity in evaluating the basic scaling laws of the model (number and spacing of wind turbines), there is still a significant deficiency in terms of representativeness for real engineering scenarios. The author evaluates the model by scaling the basic layout {3D, 5D, 7D}. This purely regular and equidistant array can only be achieved in an idealized scenario. During the intelligent layout optimization process (whether using genetic algorithms, particle swarm algorithms, or other heuristic algorithms), the algorithm inevitably generates highly irregular and non-uniform layouts in order to minimize wake overlap. The presence of a large number of eccentric wake interferences and local dense clusters of wind turbines in such irregular layouts means that a model's good performance on regular arrays does not guarantee its robustness in the irregular optimization exploration space. It is suggested that the authors add a preliminary optimized irregular layout in the experimental design to test the model's generalization ability in real WFLO scenarios.

(3) In terms of wind direction selection, the experimental design only chose a 45° wind direction sector, while the yaw angle of the wind turbine is generally set at 60°. Therefore, it is suggested that the authors set a larger interval angle.

(4) The study only considered the impact of a single wake model on the optimal layout of wind farms. It is necessary to consider whether the wake superposition model has an impact on the optimal layout of wind farms.

(5) Line 282, please provide the parameter characteristics of the IEA740-10-MW.

3. Results and Discussion

(1) In the Results and Discussion section, please summarize the content described in this section, including the research content and expected research results of the paper, and place it at the beginning of this section.

(2) Lines 334-336, "This exception was surprising, but as the remaining metrics improved, it is expected that this was due to a large error near the rotor." Please add a reference to support this statement.

(4) Lines 339-341, "The TP model exhibits the greatest overall improvement, which aligns with expectations: it was initially intended." Please add a reference.

(5) I noticed that the ANN proxy model in the paper was completely trained based on RANS data. You calibrated the engineering models (SG and TP) based on real measurement data using the same RANS single wake data. This has a certain logical consistency in the experimental design. However, RANS is a steady-state approximation and cannot capture the meandering and transient turbulent structures of the wake like large eddy simulation (LES). Forcing the calibration of empirical parameters originally fitted based on real meteorological towers or SCADA data with RANS may cause larger flow field prediction errors, leading to a greater deviation from real wind farm planning and design. It is suggested that the authors discuss this in the paper.

(6) TurbOPark was originally designed to address the wake attenuation between super-large wind turbine clusters. The author used a single-machine RANS wake database containing various CT and I0 to recalibrate it. To adapt to the single-machine wake, the parameters of the TP model changed dramatically. Does it still have a certain physical interpretability? At the same time, the TP model actually achieved the highest AEP. The conclusion drawn is worthy of in-depth discussion.

(7) In Figure 5, the RMSE error color bar retains four decimal places, which is incorrect. It is recommended to retain three decimal places. The labels should be evenly spaced, rather than these unevenly spaced numbers.

(8) In Figure 6, the author conducted wake simulation analysis experiments on wind farms

with different spacings and different numbers of wind turbines from different perspectives, explaining the sources of errors and trend changes. However, the RMSE of the TP model and the ANN model did not show a significant increasing trend with the increase in the number of wind turbines, which is completely different from the SG model. Please explain the reason in the text.

(8) Lines 418-420, The ANN was an exception, where the blockage case was cheaper; this could be because the IPOPT algorithm, which succeeded in the case without blockage more aggressively explores the loss landscape than the SGD, which is known to be resistant to local minima. (9) The TP wake model with the lowest accuracy in evaluating AEP surprisingly achieved the best results in WFLO, while the higher-precision high-fidelity model ANN performed poorly. The authors attributed this to the stronger "multimodality" of the objective function space caused by the high-fidelity model, which led the gradient-based optimization algorithm to get trapped in local optima. WFLO is a typical highly non-convex problem, and gradient-based optimization algorithms are highly dependent on the initial layout (initial random seed). The authors' conclusion that "lower-precision models perform better" based on "limited random seeds" is extremely unrigorous. Please add experiments to support the core conclusion.

(10) In lines 434-436, "In cases with blockage, however, there is a single better layout for the TP model: the one generated by the TP model without blockage." After introducing the blockage effect in the TP model, the layout found by the optimizer was worse than that found by the TP model without blockage. Please discuss why the introduction of the blockage effect leads to a decline in optimization performance.

5. Conclusions

Please summarize concisely and consolidate the four points into three. This article has a strong logical experimental design. The authors are requested to carefully revise and supplement the experimental design to achieve high credibility. I believe that after revision, it is worthy of publication in Wind Energy Science.