Responses to Reviewers' Comments for Manuscript 10.5194/wes-2026-19

Validation of RANS-calibrated engineering models and ANN-based surrogate for wind farm flow simulation and layout optimization

Addressed Comments for Publication to

WES Wind Energy Science

by

Jens Peter Schøler, Ernestas Simutis, M. Paul van der Laan, Julian Quick, and Pierre-Elouan Réthoré

Document Overview

This document contains the authors' responses to the reviewers' comments for manuscript 10.5194/wes-2026-19, "Validation of RANS-calibrated engineering models and ANN-based surrogate for wind farm flow simulation and layout optimization". The document is structured as follows:

- **Section 1:** Responses to Reviewer 1's comments
- **Section 2:** Responses to Reviewer 2's comments
- **Community comments** Responses to community comments

General Comments. The manuscript explores the possibility of using Artificial Neural Networks (ANNs) as a surrogate for wake models within the framework of wind-farm layout optimization. The presented approach is tested against engineering wake models from the literature, showing improved results for most of the simulated cases. I find the paper interesting and relevant, and overall, I support publication after minor revisions. My comments below are intended to clarify several technical points and strengthen the framing and validation of the work.

Response: Thank you for the assessment, we agree with it and have sought to improve the manuscript based on your review.

Comment 1

Line 32: The current wording is somewhat misleading. Although engineering wake models are generally of lower fidelity than CFD approaches, this does not necessarily imply lower accuracy. In certain scenarios, engineering models can achieve comparable accuracy, whereas in others they may perform worse. It would therefore be more appropriate to distinguish between fidelity and accuracy, as the latter is case-dependent.

Response:

We agree that fidelity and accuracy are distinct concepts, and have revised L32 accordingly. This excerpt also includes updates to other comments, this will be mentioned with the relevant comments.

> *In the literature, engineering models and CFD methods are often referred to as low- and high-fidelity models, respectively. Engineering models are much faster: $\sim 10^4\times$ faster than Reynolds-Averaged Navier–Stokes (RANS) [1] and $\sim 10^7\times$ faster than Large Eddy Simulation (LES) [2], making CFD too slow for WFLO, where the flow must be evaluated many times during the optimization process. In general, high-fidelity models are considered more accurate. However, under certain circumstances, engineering models can outperform high-fidelity models, e.g., within limited domains, such as the far wake, when operating near a calibration point, or when considering integrated quantities, such as power. Though design decisions and the calibration domain constrain their upper limits of accuracy. Wake models can be split into two categories: analytically solved models, which typically assume a self-similar far-wake profile. Alternatively, there is a group of low-fidelity models that require solving, but do not rely on self-similarity profiles [e.g., 3, 4, 5]. The class of self-similar, analytically solved wake models dominates WFLO practice and is the focus of the present study. In WFLO, a minimum inter-turbine separation is typically imposed to limit fatigue loads; because this constraint keeps downstream turbines out of the near wake, engineering wake models applied in WFLO*

> *have typically been developed and calibrated with the far wake as the design target. Several analytically solved engineering models exist and can loosely be categorized into families: Gaussian formulations [6, 7], a super-Gaussian variation that bridge top-hat near-wake and Gaussian far-wake behavior [8], and the Turbulence Optimized Park (TurbOPark) models [9, 10]. Special formulations have also been developed to account for the wake of turbines operating with a yaw offset, particularly useful for wake-steering scenarios [11, 12]. The breadth of existing models is wide and the models mentioned here are an important subset, but many more exist for different applications. At the farm scale, overlapping wakes have further motivated momentum-conserving superposition schemes [13, 14] beyond the classical linear and quadratic sums [2]. In the present study, recalibration has been performed on two engineering model configurations in* `PyWake` *[15].*

Comment 2

Line 38: The statement appears overly restrictive. Not all engineering wake models rely on the self-similarity assumption (e.g. Klemmer & Howland, 2025).

Response:

To make the statement less restrictive a distinction between analytically solved engineering models assuming self-similarity and those solving simplified physics has been introduced. Examples of the latter has been included. See reply to comment 1 for the excerpt.

Comment 3

The introduction does not sufficiently discuss the breadth of existing engineering wake models and seems to dismiss them prematurely in favor of data-driven approaches. A more balanced overview of current modeling strategies would improve the context and motivation for the proposed framework.

Response:

We have extended the scope of the engineering models discussed significantly and as mentioned in the reply to comment 2 made a distinction between model types. See comment 1 for the excerpt

Comment 4

Table 1: If all wind directions are simulated, please clarify the rationale behind selecting non- uniform values of MN, MW, MS, and ME.

Response:

The inflow is always from west to east (in the grid coordinate system). The different inflow directions are obtained by rotating the farm. This conserves computational resources as upstream direction does not require as many cells to resolve as the downstream direction. This was only mentioned in passing and has now been introduced with the introduction of the mesh and described in greater detail.

Comment 5

Line 205: is it ‘downstream’ of the source turbine or should it be ‘upstream’?

Response:

The downstream instance has been corrected to upstream.

Comment 6

Table 4: Is there a specific reason why turbulence superposition is set to linear here, whereas Table 2 uses Quad. Max. Sum?

Response:

The utilized turbulence model for the SG model was the CrespoHernandez model in PyWake this model defaults to the Quad. Max. Sum., hence it was adopted here. A sentence about using PyWake as default has been added in connection with community comment 4.

For the ANN, the framework was lifted from the the source material [16]. A sentence was added to acknowledge the configurations was adopted from Pish et al. [16].

Comment 7

Tables 2 and 4 (Rotor averaging): It would be valuable to include, perhaps in an appendix, a sensitivity analysis demonstrating the impact of rotor averaging. Currently, rotor averaging is set to “center”, yet the centerline wind-speed deficit is not necessarily representative of the rotor- averaged deficit (see Fig. G1 in Ali et al., 2025). Additionally, it would be informative to discuss how a more accurate rotor-averaging approach might influence the recalibrated constants of the engineering wake models.

Response:

This is a valid suggestion. A sensitivity analysis of rotor averaging on both flow field accuracy and recalibrated model parameters would be a valuable contribution; however, it falls outside the scope of the current work. We acknowledge this as a worthwhile direction and have added it as a suggested topic for future work in the conclusions section.

Comment 8

Equation 11: Please clarify how the value 8760 is obtained.

Response:

8760 is the number of hours in a year, this explanation has been added in the manuscript.

Comment 9

The validation of re-calibrated engineering wake models would be stronger if comparisons were made against full RANS/LES simulations or experimental data. Validation solely against RANS- based look-up tables is somewhat limited, as these tables inherit the modelling assumptions and uncertainties of the underlying RANS simulations in addition to errors due to interpolation across the parameters space.

Response:

We want to clarify that, the single-wake look-up table is only used for recalibration (Sect. 2.3). The validation described in Sect. 2.5 is conducted against dedicated full farm-level RANS simulations of 18 distinct layouts covering $n_{\mathrm{wt}} \in \{6, 10, 20, 30, 40, 50\}$ turbines at $s_{\mathrm{wf}} \in \{3D, 5D, 7D\}$, which are independent of the LUT-calibration data. RANS do remain the sole reference throughout, and a study including LES or SCADA validation would indeed strengthen the study.

Concluding Response. We thank the reviewer for a precise and balanced review. We believe the raised points have been addressed in the responses above, and that the manuscript has been improved as a result.

Authors' Response to Reviewer 2

General Comments. 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 windfarm 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. While the study presents interesting findings, addressing these points will enhance the credibility and overall quality of the article,facilitating its rapid publication.

Response:

We thank the reviewer for the thorough and positive assessment and for highlighting the relationship between model fidelity and the realized AEP of optimized layouts, a central theme of the paper. We have revised the manuscript along the lines suggested, broadening the introduction and its literature base, refining the presentation of the affected tables and figures, and adding clarifications throughout, while preserving the scope and framing of the main findings. Our responses follow below.

1. Introduction

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

Comment 1

Broaden the Citation Base: Certain parts of the text (e.g., lines 11-15, 23-27, and 34-40) currently have relatively few 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.

Response:

The flagged lines concerning the engineering-models scope, were substantially expanded in response to Reviewer 1 Comments 1–3; we refer the reviewer to those responses for that material. To broaden the coverage of surrogate wake models covering dynamic effects a new paragraph has been added: *From the related field of dynamic surrogate wake models a couple of interesting works are listed to situate the current work better in a wider context. Dynamic models are typically motivated by a desire to improve wind farm control and mitigate loads. The work of Zhang and Zhao [17] used a physics-informed reconstruction of unsteady wake fields from LiDAR measurements, the predictive and stochastic reduced-order model (PS-ROM) of wake dynamics by Andersen and Leon [18], the PS-ROM derivative work on global POD modes by Moreno et al. [19], and real-time physics-guided frameworks [20]. Most recently, Liu et al. [21] proposed PhyWakeNet, a hybrid physics and data-driven model that captures the time-averaged deficit, wake meandering, and small-scale turbulence of an unsteady wake using separate but integrated model components.*

Comment 2

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 For instance, referencing the reinforcement learning-enhanced genetic algorithm by Dongetal (2026) could help bridge the gap between aerodynamic modeling and layout application.

Response:

To broaden the scope of the optimization space slightly mention of gradient-free methods has now been introduced in the introduction: *WFLO algorithms can be broadly grouped into gradient-based and gradient-free methods. The present work focuses on gradient-based optimization. Gradient-free approaches nonetheless remain an area of interest and continue to advance, ranging from classical evolutionary genetic algorithms [22] and particle swarm optimization [23] to recent hybrids such as the reinforcement learning-enhanced genetic algorithm of Dong et al. [24].*

Comment 3

Include Blockage Effects: Additionally, incorporating more discussion and references regarding blockage effects would help provide a more holistic context of the aeroelastic challenges addressed in large-scale wind farm studies.

Response:

We have added some additional context to the introduction regarding blockage and the importance of it in larger clusters.

While wake effects are the dominant turbine-to-turbine interaction, the induction of the turbines also produces a wind-farm-scale blockage effect that slows the inflow upstream and can result in power losses at the leading turbines [25]. This upstream slowdown has been observed in offshore LiDAR measurements [26], and grows in relevance for the large turbine counts of modern offshore wind farms, where the induction of the individual turbines aggregates into a combined wind-farm-scale blockage effect [9].

2. Methodology

Comment 4

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

Response:

We have rearranged Table 3 to be more standard and fit the two-column format. Table 4 has been split into two separate tables.

Comment 5

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.

Response:

We thank the reviewer for raising this point and agree that operational layouts are more irregular than the layouts considered in the flow study. The structured design, however, is a deliberate methodological choice rather than an oversight. By varying only the inter-turbine separation and turbine count on a common layout family, the study isolates how model accuracy depends on these two factors. It thereby allows us to conclude at which separations each model performs better or worse. A single irregular layout would yield a verdict on which model performs best, without resolving this distance dependence or generalizing to other irregular layouts. Since an arbitrary layout can be regarded as a collection of interactions at varying separations, the distance-resolved results reported here remain informative when interpreting irregular configurations. We also note that the circular layouts are not equidistant arrays; as they are rotated across the simulated inflow directions, the relative downstream positions of the turbines change substantially, producing a broad range of partial and full wake overlaps rather than a single regular pattern. Finally, the generalisation to irregular, optimisation-generated layouts is in fact assessed in the present study: in the WFLO results (Sect. 2.6, with the resulting layouts shown in Appendix B), each engineering model is applied to the highly irregular layouts produced by the optimiser, and these layouts are validated against RANS-based AEP. This provides a direct test of the models on realistic WFLO layouts at the integrated quantity, AEP, that is decisive for layout selection. Extending this to a dedicated flow-field comparison on irregular layouts would require additional RANS simulations of each such layout, which we consider beyond the scope of the present work.

Comment 6

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.

Response:

We thank the reviewer for the comment, and we would like to clarify a distinction between the inflow wind direction and the turbine yaw alignment, as these are separate quantities. The 45° window refers to the inflow wind direction range used in the flow simulation study (Sect. 2.5), where the wind direction θ is swept from 270° to 315° in 3° increments. It is not a yaw setting. Yaw misalignment is not considered anywhere in this study: all turbines are operated aligned with the local inflow ($\gamma = 0$), as already stated in Sect. 2.4, where excluding yaw misalignment was a conscious decision to limit the number of additional RANS cases required. The choice of inflow-direction window is therefore independent of any yaw angle. Because the flow study layouts are arranged radially (Fig. 3), the relative positioning between turbines repeats across directions, so a single 45° sector is representative of the directional response while keeping the number of RANS cases manageable. We also note that this restricted sector applies only to the flow accuracy study; the layout optimization in Sect. 2.6 uses the full wind rose (Fig. 4b), discretized at 1° during optimization and 5° during RANS validation, so the directional breadth of the wind resource is fully accounted for in the WFLO and AEP results. For these reasons, the current directional setup has been retained.

Comment 7

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.

Response:

We understand this comment to concern the wake superposition model rather than the wake deficit model itself. Two distinct deficit models were considered, a variation of the super-Gaussian model and a variation of the TurbOPark model, but the superposition scheme was held fixed across both: linear superposition was applied to the velocity deficits (Δu), with a quadratic max sum used for the added turbulence intensity (I_a) in the SG configuration. We agree that the superposition method can affect the predicted wake interactions and, in turn, the optimal layout, and that a systematic comparison of superposition schemes would be a valuable extension. This falls outside the scope of the present study, and we note it as a direction for future work.

Comment 8

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

Response:

After re-reading the paragraph introducing the site conditions, we agree it was needlessly confusing. It has been rewritten to make clear that only the wind conditions from the IEA Wind 740-10-MW reference plant are used. An explanation of how the boundary was created has been added. The turbines are now identified as the DTU 10-MW reference wind turbine, with the initial and optimized coordinates for the best optimization runs shown in Appendix B.

We would also like to mention that the full setup is reproducible from the publicly available code.

3. Results and Discussion

Comment 9

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.

Response:

An introduction has been added to the results and discussion section.

Comment 10

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.

Response:

We have added the reference for the super-Gaussian model as this is the reason for said model to exist.

Comment 11

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

Response:

The sentence you have highlighted as missing a reference was highlighted in a community comment as erroneous. We have corrected it in accordance with community comment 16 and references has been reiterated for the models.

Comment 12

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.

Response:

We thank the reviewer for raising this point. It is correct that RANS, being a Reynolds-averaged method, represents the time-averaged flow and does not resolve unsteady wake meandering, unlike LES. We would note, however, that the net effect of meandering on the mean velocity deficit and wake recovery is not absent from the RANS solution; it is represented implicitly through the k-ε-f_P turbulence closure, which is calibrated against measurements. As the reviewer notes, this effect cannot be cleanly separated from the mean flow, but the same holds for the analytically solved engineering wake models considered here, which likewise predict only a time-averaged deficit. Calibration and validation are therefore performed consistently between mean-field representations. We have added a sentence to explain that the effect of meandering on the mean flow is modeled using the closure model in the methodology section on RANS.

Comment 13

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.

Response:

As raised in community comment 16, it was a misunderstanding on our side that TurbOPark was originally intended for inter-farm wakes. We apologize for the confusion, and the relevant passage has been corrected accordingly.

On physical interpretability, we agree that the recalibrated TP parameters have limited physical meaning when read individually. As discussed in our replies to the community comments on this topic, the parameters A, $c_{I,1}$, and $c_{I,2}$ all act on the wake expansion and partly compensate one another, so a marked change such as $c_{I,2}$ from 0.80 to ≈ 3.15 should not be interpreted in isolation. The recalibration targets agreement with the RANS reference rather than a re-derivation of the underlying physics, and we have been careful not to overstate the physical meaning of the individual constants.

In the AEP results, the finding that the recalibrated TP model produces the highest RANS-validated AEP layouts despite its lower flow-field accuracy is a central, deliberately highlighted result. We have discussed it at some length in the results and discussion, and we have expanded on it in response to the reviewers' comments. We believe the matter is now treated in appropriate depth.

Comment 14

In Figure 5, the RMSE error colorbar 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.

Response:

We have updated the figure to utilize 3 decimal places and sought to more evenly distribute the ticks on the colorbar.

Comment 15

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.

Response:

The apparent absence of a trend for the TP and ANN models is a consequence of the linear axis rather than a difference in behavior: all models follow the same qualitative trend of increasing error with n_{wt}, as already noted in the manuscript, where the growth is attributed to the modeling error of superposing wakes becoming increasingly problematic as the number of turbines increases. For the TP and ANN models this growth is small in absolute terms, so on a linear RMSE_{Ω} axis, whose range is set by the much larger errors of the SG and originally calibrated models, the lower-magnitude curves are compressed toward the bottom and their growth is difficult to discern. We originally considered a logarithmic ordinate for this reason. However, we chose the linear scale because it preserves the substantial differences in absolute bias between the models, which a logarithmic axis tends to obscure. To make the shared trend explicit without sacrificing this information, we have added a semi-logarithmic version of Fig. 6 in a new appendix, together with a pointer to it in the text. We are of course happy to instead replace Fig. 6 with the semi-logarithmic version should the reviewer prefer a single figure.

Comment 16

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.

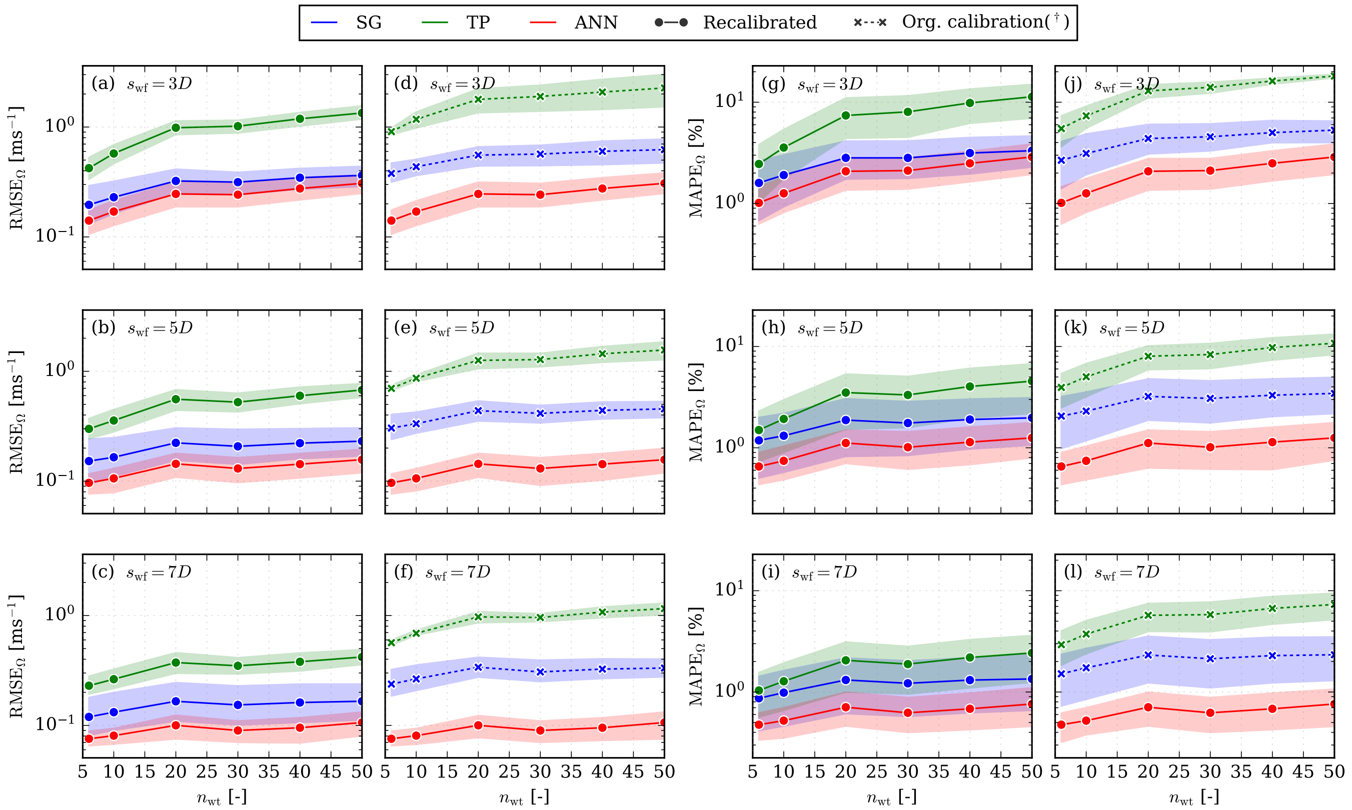

Figure 1: Semi-logarithmic version of Fig. 6, showing that the RMSE_{Ω} of all models, including the TP and ANN models, increases with n_{wt}.

Response:

This appears to be an excerpt from the manuscript without a connected comment, it is assumed that something is missing and the comment is ignored for now.

Comment 17

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.

Response:

We thank the reviewer for engaging with this finding. We believe, however, that the comment characterises our conclusion more strongly than the manuscript states, and we would like to clarify three points.

First, we do not conclude that lower-fidelity models perform better in general. Our claim is narrower: that AEP-prediction accuracy is not the sole determinant of the realised AEP of an optimised layout (L450–452, and Conclusion point 3). However, to make sure this is not mistakenly taken to mean "lower-precision models perform better", we have added a sentence stating that it is possible that additional random seeds may enable the ANN to produce the better layout. Overall, the TP result is intended to be reported as an observation within the present study, not as a general rule that simpler models are preferable.

Second, the multimodality mechanism is offered as a hypothesis, not a conclusion. We state that we "hypothesize" the effect (L458–460), and the supporting experiment in Appendix C is described as admittedly limited, showing only a slight indication, with the matter explicitly deferred because "further studies are required to reach a definitive conclusion" (L480–483). We therefore do not present multimodality as an established cause, and the manuscript does not attribute the WFLO outcome to it with the certainty implied in the comment.

Given that the conclusion attributed to us is stronger than the one we draw, and that the relevant limitations are already acknowledged in the text, we have retained the current framing rather than adding further optimisation experiments.

Comment 18

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.

Response:

The cross-comparison in Fig. 8 does not support a systematic link between blockage and a decline in optimization performance. The TP* case is the only blockage configuration for which a marginally better layout exists. By contrast, no better layout is available for the SG* and ANN* cases. This is the reverse of the behavior without blockage, where better layouts existed for the SG model and the ANN, but not for TP. Since all three blockage configurations share the same wind farm flow model (All2All iterative) and optimization algorithm (SGD), a blockage-induced effect would be expected to act on all three models rather than on TP* alone. Given the marginal magnitude and the absence of a systematic pattern, we refrain from attributing this isolated difference to a specific mechanism. We have improved the sentence explaining this, to make it clearer.

5. Conclusions

Comment 19

Please summarize concisely and consolidate the four points into three.

Response:

We have tightened the wording of the concluding findings and consolidated them from four points to three by folding the blockage result into the discussion of WFLO results and optimization behaviour, where it is most relevant. The revised conclusions now cover recalibration, flow simulation performance, and WFLO behaviour including blockage.

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.

Concluding Response. We thank the reviewer for the careful and constructive review, and for recognizing the value of the systematic validation. The revisions broaden the introduction, standardize the affected tables and figures, add an introductory summary to the results, and consolidate the conclusions. We have clarified the role of RANS as the designated reference and the consistent mean-field basis on which the engineering models and the ANN surrogate are calibrated and validated. Where the experimental design was raised, we have made the underlying rationale explicit: the structured flow study is a deliberate choice that isolates how model accuracy depends on turbine count and spacing, and the WFLO result is reported with its stated limitations, with a broader multi-seed characterization and LES or SCADA validation identified as directions for future work. These revisions address the points raised and strengthen the credibility of the study.

Community Comment

General Comments. I read your preprint on validation of RANS-calibrated engineering models and ANN-based surrogate. Interesting stuff! The paper triggered many thoughts and questions that I list below in chronological order.

Response:

Thank you for the interest, we believe community engagement is key to ensuring that research is both correct and relevant. We have answered your questions below and made changes to the manuscript where appropriate to correct misunderstandings and improve clarity.

Comment 1

P.2 – In my experience a minimum distance constraint is imposed in WFLO not because of model deficiencies in the near wake but to limit fatigue loads

Response:

You are correct the passage has been updated to reflect that. See reply to reviewer 1 comment 1 for the excerpt.

Comment 2

P.6 - Can you explain how the configuration of TurbOPark is altered in PyWake?

Response:

The sentence is not meant to refer to PyWake in general, only this study. We have made that explicit now. Depending on the configuration it is the inclusion of a blockage model and in both cases using a linear superposition model. The configuration is listed in Table. 2.

Comment 3

P.6 – For the linear superposition do you use method A or method C as described in Zong and Porté-Agel, A momentum-conserving wake superposition method for wind farm power prediction (2020)?

Response:

Method C. In both configurations the individual velocity deficits are scaled by the local effective wind speed and then summed linearly.

Comment 4

P.7 – Why is the use of a quadratic superposition of wakes incompatible with the blockage model? Is this due to how the All2AllIter. Method works in PyWake? I do not have any experience with PyWake so forgive me if this is obvious

Response:

This is probably not entirely obvious and we have therefore added an explanation of why these are incompatible. The key problem is that the blockage can be negative and that contribution would not always be presereved through a quadratic superpositon.

Comment 5

P.7 – Table 2 lists a turbulence model and a turbulence superposition method for TP. Is it correct to assume that these are for calculating the turbulence in the wind farm but that these model components do not play a role in the AEP and WFLO results discussed in the paper?

Response:

This holds for the configurations without blockage but not for those with blockage. For TP the standalone turbulence model and the associated I_a superposition are only active in the blockage cases, where they supply the turbulence input to the Self-Similarity Deficit blockage model [27]. Through the blockage deficit they therefore enter the velocity field, and hence the AEP and WFLO results, in those cases. In the cases without blockage no standalone turbulence model is used, since the TurbOPark wake expansion already embeds the Frandsen [28] turbulence formulation (Eq. 2), and the TP deficit depends on the ambient I_0 rather than the in-farm turbulence, so the turbulence field does not feed back into the deficit and plays no role in AEP or WFLO there.

We note one minor inconsistency: the blockage model is fed by the Crespo and Hernández [29] turbulence model rather than the Frandsen formulation built into the wake deficit. Using the latter for the blockage as well would be more consistent with the TurbOPark Gaussian deficit, and we suggest this as a refinement for future work. We do not expect the effect to be significant, as the wake deficit, the dominant driver of the AEP and WFLO results, is unaffected and only the secondary blockage contribution would change.

Comment 6

P.7 – I find it strange that you treat the C_T limit as a tuneable parameter in TP but not in the SG model. The expression for the near wake streamtube expansion is the same and TP only includes the limit to avoid situations (C_T >= 1) where the expression becomes divergent or complex. We need this limit as we do see a few thrust curves with C_T>1 at low wind speeds. Even if this does not make sense, it is better to have a safeguard in the model to prevent the code from breaking. What do you do in PyWake if this case is encountered?

Response:

The asymmetry reflects how the two deficit models are parameterized in PyWake: we calibrated the parameters each model exposes rather than re-deriving the formulations. The TurbOPark deficit exposes $C_{T,\text{lim}}$. In contrast, the super-Gaussian deficit introduces no such cap, so there is no analogous parameter to tune. The question of how to handle the $C_T \geq 1$ case does not arise here, as the DTU10MW turbine thrust curve we use is not defined for $C_T \geq 1$, and the clipping is built into the curve itself.

It is then true that the super-Gaussian model should also have been allowed to have this parameter tuned to unify the methodologies. We have added a sentence to clarify the mismatch and suggest that future work should be mindful of this difference.

Comment 7

P.7 – What do you mean by the sentence about I_k, I_u and TI in lines 1443-145? What makes you state that engineering wake models and in this case TP assumes I_u = I_k? Turbulent kinetic energy is not a concept we ever use. To me it seems more relevant for RANS simulations

Response:

Indeed, this is meant to clarify the difference between using wake models and RANS. The sentence is targeted at CFD practitioners who try to understand what the TI models in the engineering modeling framework means.

Comment 8

P.11 – Why would the bounds on c_s in SG and c_epsilon in TP be different. The two parameters seem to have same role in the models

Response:

The connection between the two parameters was not apparent to us at the time of recalibration. Had it been they should have been the same.

Comment 9

P.16 – You write true wind speed, but I assume you mean RANS wind speed

Response:

Yes, this was a mix up between the term "true data" and "target wind speed". We have corrected it to target wind speed.

Comment 10

P.16 – I would generally make a distinction between (re)calibrating towards RANS and (re)calibrating towards observations

Response:

We have renamed the subsection to "Recalibration of wake models with RANS data". This should help remind the reader that this is with respect to RANS data. Combined with the updates made for comment 18 it should be sufficient to avoid further misunderstandings.

Comment 11

P.16 – It is nice that the recalibrated values for c_s and c_epsilon are both lower than the starting point. This indicates the connection between the two variables I raised under item 8

Response:

Yes, we are fortunate here that the oversight we made with respect to these two variables was not crucial.

Comment 12

P.16 – Both a lower value of c_epsilon and of C_T_lim makes the initial wake narrower (the latter only for high C_T values above the limit). This tends to compensate the much higher A-value which leads to faster wake spreading

Response:

Thank you for making this observation, we have added the explanation to the manuscript.

Comment 13

P.17 – With the recalibrated values for c_I,1 and C_I,2 the wake-added decays faster with increasing downstream distance than in the original Frandsen turbulence model, therefore if everything else is equal the wake would recover slower. But A also increases. To me this illustrates the difficulty of calibrating a model with many parameters of which several influence the same behaviour in this case wake expansion. We made a conscious choice in TP to keep the original values of the Frandsen turbulence model both because we did not have something better to use in their place and because we wanted to keep the number of free parameters at a minimum. If we were to change c_I,1 and c_I,2 I would prefer to calibrate them by comparing with turbulence data alone.

Response:

You are correct, and the point is well taken: $c_{I,1}$, $c_{I,2}$ and A all act on wake expansion and partly compensate one another, so the recalibrated values are not physically interpretable in isolation. We let these parameters vary because the calibration target is RANS rather than the measurement data the original Frandsen values were derived from. Fixing the parameters would have placed the engineering models at a disadvantage in a study where RANS is the reference. Had the goal been more classical in nature i.e. calibration against measurements or LES, we agree that calibrating $c_{I,1}$ and $c_{I,2}$ against turbulence data alone is preferable.

Comment 14

What would you find I wonder if you only allowed A in TP to be changed?

Response:

Because A is one of several parameters (A, $c_{I,1}$, $c_{I,2}$, c_ϵ, $C_{T,\mathrm{lim}}$) all acting on the wake expansion, we would expect A alone to change differently, since the other terms could no longer offset it. However, a constrained single-parameter recalibration would require a complete rework and is considered outside the scope of the current paper.

Comment 15

As a comment another part of the larger probably comes from changing from quadratic to linear superposition of the wakes. The latter method leads to more intense wakes with the same parameters, and the increased value of A compensates for that

Response:

Thank you that is a good observation we have added it to the manuscript.

Comment 16

P.18 – It is a misunderstanding that TP was initially intended for farm-to-farm interactions and therefore not expected to perform optimally in the intra-farm context without recalibration! The discovery that other engineering wake models underestimated the wake loss between neighbouring wind farms was part of what led us to formulate TP, but the model was never intended primarily as a farm-to-farm model- Ørsted has always used TP as the wake model since its introduction. Our calculations do not distinguish between turbines in the target wind farm and in neighbour wind farms. They are modelled in one and only one calculation using TP, meaning no separate calculation for farm-to-farm interactions. Obviously, the intra-farm wake loss exceeds the farm-to-farm loss, so the model we use (TP) should be most suitable for this situation. TP was calibrated on 19 offshore wind farms under situations that included both inter- farm and inter-farm wakes, but the latter constitutes the vast majority of flow situations

Response:

Thank you for that correction, we have corrected the paragraph to focus on TurbOParks calibration with farm data instead of farm-to-farm wakes and contrasted it to the single wake calibration of the super-Gaussian model.

Comment 17

P.18 – On the other hand, I am not surprised the TP does not perform well when compared with RANS in a single wake scenario. The model was never calibrated to perform well in this situation but instead tuned to perform as well as possible averaged across a large portfolio in all situations. This includes wake superpositions and cover all stability conditions, not just neutral conditions

Response:

Neither are we, hence why we recalibrated the model.

Comment 18

P.18 – Instead of uncalibrated you should consider saying original calibration and similarly stress that the recalibration is with RANS simulation data

Response:

We agree that it is confusing to use "uncalibrated" as it misrepresents the work of others. We have adopted the distinction between "original calibration" and "recalibration". On the point about mentioning that recalibration is towards RANS data, we beleive that is a core idea in the work and is not necesarry to mention everytime the models are mentioned. However, we have made sure to mention it in the section title as per comment 10, to remind the reader or make it noticeable for people skimming the article.

Comment 19

P.22 – Your conclusion about AEP accuracy not being the only important factor in WFLO is interesting. Maybe it points to a difference between AEP accuracy and getting the wake expansion right. The former requires not only the wind to power conversion to be accurate but also the size of the centreline deficit. TP may get the centreline deficit wrong compared with neutral stability RANS, but it could get the wake expansion mostly right. It also has the Gaussian overlap, which is important in partial wake situations, whereas the SG model uses the wind speed in the centre of the rotor (as I read table 2)

Response:

The evidence we are seeing does not support that considering the results in the flow study part of the paper. Here the TP model does not perform well overall indicating that the flow field is not well captured, it is this observation that there is a discrepreancy between the result of the flow study and the WFLO study that leads us to the conclusion about accuracy in flow and ability to interact well with WFLO.

Comment 20

P.31 – Why is there a tendency for some of the layouts to have turbines organised on lines in the interior of the domain? This seems counter-intuitive to me

Response:

The first thing to note is that the optimizer packs most turbines along the domain boundary, which is expected here since the objective accounts only for wake losses and not for cabling, foundations, or installation costs that would otherwise penalize the perimeter. As for the interior structure, we can speculate, but one plausible explanation is that, with only AEP driving the objective and the minimum-separation constraint active, the remaining turbines settle into streamwise rows roughly aligned with the dominant wind directions, so that each turbine tends to sit in the gap between the wakes shed by the row upstream rather than directly behind a neighbour, which would reduce the wake overlap seen across the wind rose.

Concluding Response. We thank the commenter for the detailed and knowledgeable engagement, which was especially valuable given the firsthand familiarity with the TurbOPark model.

References

[1] M. P. van der Laan et al., “An improved k-ε model applied to a wind turbine wake in atmospheric turbulence,” *Wind Energy*, vol. 18, pp. 889–907, 5 May 2015. DOI: `10.1002/we.1736`.

[2] F. Porté-Agel, M. Bastankhah, and S. Shamsoddin, “Wind-turbine and wind-farm flows: A review,” *Boundary-Layer Meteorology*, vol. 174, pp. 1–59, 1 Jan. 2020. DOI: `10.1007/s10546-019-00473-0`.

[3] J. Ainslie, “Calculating the flowfield in the wake of wind turbines,” *Journal of Wind Engineering and Industrial Aerodynamics*, vol. 27, pp. 213–224, 1-3 Jan. 1988. DOI: `10.1016/0167-6105(88)90037-2`.

[4] S. ; Ott, J. ; Berg, and M. Nielsen, “Linearised cfd models for wakes. danmarks tekniske universitet, risø nationallaboratoriet for baeredygtig energi,” Risø National Laboratory, Tech. Rep., 2011.

[5] K. S. Klemmer and M. F. Howland, “Wake turbulence modeling in stratified atmospheric flows using a novel k-ℓ model,” *Journal of Renewable and Sustainable Energy*, vol. 17, 3 May 2025. DOI: `10.1063/5.0249278`.

[6] M. Bastankhah and F. Porté-Agel, “A new analytical model for wind-turbine wakes,” *Renewable Energy*, vol. 70, pp. 116–123, 2014. DOI: `10.1016/j.renene.2014.01.002`.

[7] A. Niayifar and F. Porté-Agel, “Analytical modeling of wind farms: A new approach for power prediction,” *Energies*, vol. 9, 9 2016. DOI: `10.3390/en9090741`.

[8] F. Blondel and M. Cathelain, “An alternative form of the super-gaussian wind turbine wake model,” *Wind Energy Science*, vol. 5, pp. 1225–1236, 3 Sep. 2020. DOI: `10.5194/wes-5-1225-2020`.

[9] N. G. Nygaard, S. T. Steen, L. Poulsen, and J. G. Pedersen, “Modelling cluster wakes and wind farm blockage,” in *Journal of Physics: Conference Series*, vol. 1618, IOP Publishing Ltd, Sep. 2020. DOI: `10.1088/1742-6596/1618/6/062072`.

[10] N. G. Nygaard, J. G. Pedersen, S. D. Hansen, and P. Krastins, “A turbulence optimized park model,” Ørsted, Tech. Rep., 2022.

[11] M. Bastankhah and F. Porté-Agel, “Experimental and theoretical study of wind turbine wakes in yawed conditions,” *Journal of Fluid Mechanics*, vol. 806, pp. 506–541, Nov. 2016. DOI: `10.1017/jfm.2016.595`.

[12] L. A. Martínez-Tossas, J. Annoni, P. A. Fleming, and M. J. Churchfield, “The aerodynamics of the curled wake: A simplified model in view of flow control,” *Wind Energy Science*, vol. 4, pp. 127–138, 1 Mar. 2019. DOI: `10.5194/wes-4-127-2019`.

[13] H. Zong and F. Porté-Agel, “A momentum-conserving wake superposition method for wind farm power prediction,” *Journal of Fluid Mechanics*, vol. 889, A8, Apr. 2020. DOI: `10.1017/jfm.2020.77`.

[14] L. Lanzilao and J. Meyers, “A new wake-merging method for wind-farm power prediction in the presence of heterogeneous background velocity fields,” *Wind Energy*, vol. 25, pp. 237–259, 2 Feb. 2022. DOI: `10.1002/we.2669`.

[15] M. M. Pedersen et al., *Pywake 2.5.0: An open-source wind farm simulation tool*, Feb. 2023.

[16] F. Pish, T. Göçmen, and M. P. V. D. Laan, “Generalization of single wake surrogates for multiple and farm-farm wake analysis,” *Journal of Physics: Conference Series*, vol. 2767, p. 092 058, 9 Jun. 2024. DOI: `10.1088/1742-6596/2767/9/092058`.

[17] J. Zhang and X. Zhao, "Spatiotemporal wind field prediction based on physics-informed deep learning and lidar measurements," *Applied Energy*, vol. 288, p. 116 641, Apr. 2021. DOI: `10.1016/j.apenergy.2021.116641`.

[18] S. J. Andersen and J. P. M. Leon, "Predictive and stochastic reduced-order modeling of wind turbine wake dynamics," *Wind Energy Science*, vol. 7, pp. 2117–2133, 5 Oct. 2022. DOI: `10.5194/wes-7-2117-2022`.

[19] J. F. C. Moreno, J. P. M. León, and S. J. Andersen, "Convergence and efficiency of global bases using proper orthogonal decomposition for capturing wind turbine wake aerodynamics," *Wind Energy Science*, vol. 10, pp. 597–611, 3 Mar. 2025. DOI: `10.5194/wes-10-597-2025`.

[20] B. Li, M. Ge, X. Li, and Y. Liu, "A physics-guided machine learning framework for real-time dynamic wake prediction of wind turbines," *Physics of Fluids*, vol. 36, 3 Mar. 2024. DOI: `10.1063/5.0194764`.

[21] X. Liu, Z. Li, and X. Yang, "Phywakenet: A dynamic wake model accounting for aerodynamic force oscillations," *Wind Energy Science*, vol. 11, pp. 771–793, 3 Mar. 2026. DOI: `10.5194/wes-11-771-2026`.

[22] G. Mosetti, C. Poloni, and B. Diviacco, "Optimization of wind turbine positioning in large windfarms by means of a genetic algorithm," *Journal of Wind Engineering and Industrial Aerodynamics*, 1994. DOI: `Mosetti1994`.

[23] S. Pookpunt and W. Ongsakul, "Optimal placement of wind turbines within wind farm using binary particle swarm optimization with time-varying acceleration coefficients," *Renewable Energy*, vol. 55, pp. 266–276, Jul. 2013. DOI: `10.1016/j.renene.2012.12.005`.

[24] G. Dong, J. Qin, C. Wu, C. Xu, and X. Yang, "Reinforcement learning-enhanced genetic algorithm for wind farm layout optimization," *Renewable Energy*, vol. 259, p. 125 093, Mar. 2026. DOI: `10.1016/j.renene.2025.125093`.

[25] J. Bleeg, M. Purcell, R. Ruisi, and E. Traiger, "Wind farm blockage and the consequences of neglecting its impact on energy production," *Energies*, vol. 11, p. 1609, 6 Jun. 2018. DOI: `10.3390/en11061609`.

[26] J. Schneemann, F. Theuer, A. Rott, M. Dörenkämper, and M. Kühn, "Offshore wind farm global blockage measured with scanning lidar," *Wind Energy Science*, vol. 6, pp. 521–538, 2 Apr. 2021. DOI: `10.5194/wes-6-521-2021`.

[27] A. R. M. Forsting, G. P. N. Diaz, A. Segalini, S. J. Andersen, and S. Ivanell, "On the accuracy of predicting wind-farm blockage," *Renewable Energy*, vol. 214, pp. 114–129, Sep. 2023. DOI: `10.1016/j.renene.2023.05.129`.

[28] S. Frandsen, "Turbulence and turbulence-generated structural loading in wind turbine clusters - risoe-r no. 1188(en)," Forskningscenter Risoe, Tech. Rep., 2005.

[29] A. Crespo and J. Hernández, "Turbulence characteristics in wind-turbine wakes," *Journal of Wind Engineering and Industrial Aerodynamics*, vol. 61, pp. 71–85, 1996.