Journal of Physics: Conference Series

PAPER • **OPEN ACCESS**

Performance comparison of analytical wake models calibrated on a large offshore wind cluster

To cite this article: Diederik Van Binsbergen *et al* 2024 *J. Phys.: Conf. Ser.* **2767** 092059

Journal of Physics: Conference Series **2767** (2024) 092059 doi:10.1088/1742-6596/2767/9/092059

Performance comparison of analytical wake models calibrated on a large offshore wind cluster

Diederik van Binsbergen [1,2], Pieter-Jan Daems [2], Timothy Verstraeten [2], Amir Nejad[1] and Jan Helsen[2]

[1] Department of Marine Technology, Norwegian University of Science and Technology (NTNU), Otto Nielsens veg 10, 7052 Trondheim, Norway

[2] Department of Mechanical Engineering, Vrije Universiteit Brussel, Pleinlaan 2, Brussels, 1050, Belgium

E-mail: `dirk.w.van.binsbergen@ntnu.no`

Abstract. This study benchmarks the performance of multiple analytical wake models using a multi-level hyperparameter optimization framework for calibrating models with SCADA data in the Belgian-Dutch offshore zone. The calibration targets wind coming the northwest (300 to 330 degrees) with wind speeds ranging from 7 to 9 m/s, a wind direction where wakes are clustered across Belgian wind farms. Six wake models are evaluated, namely those described by Jensen (1983), Bastankhah (2014), Niayifar (2016), Zong (2020), Nygaard (2020), and Pedersen (2022). Relative and accumulated relative error between the calibrated wake models and SCADA data are analyzed both on turbine, farm, and cluster level. Statistical moments of the residual error between each model and SCADA data for individual turbines are presented in boxplots for each wind farm and are analyzed using kernel density estimates. Additionally, the calibrated tuning parameters are used to calculate wake losses, demonstrating a strong overlap across models. The analysis reveals that the top-hat TurbOpark model shows the best performance, followed by its Gaussian variant. The Gaussian models by Niayifar (2016) and Zong (2020), as well as the top-hat model by Jensen (1983) all show relatively good performance, while the Gaussian model by Bastankhah (2014) has the worst performance after calibration. While some models perform better than other models, all models show similar trends post-calibration, indicating that with proper calibration, any model can be viable, given its inherent limitations are recognized and managed.

1. Introduction

With the growing size of offshore wind turbines and wind farms (WFs), understanding the impact of wind turbine wakes on the expected yield and the levelized cost of energy has become crucial. The propagation of the turbine wake results in power losses for downwind turbines and needs to be considered in the yield assessment, controller design, and for layout optimization. The first model for analyzing turbine wakes was introduced by Jensen in 1983 [1], characterized as a top-hat model. This implies a uniform wind speed across the wake from the center to its boundary. Additionally, this model assumes linear wake expansion. The uniform wind speed profile across the rotor plane makes these types of models unrealistically sensitive for small changes in wind direction. Furthermore, top-hat models underestimate the wake effect at the center of the wake, while overestimating it at the edge of the wake. To address these shortcomings, a new model employing a self-similar Gaussian distribution for the wake was introduced [2], aligning with the fundamental nature of turbulent plumes and wakes [3]. Subsequent models have built upon these foundations, integrating elements from the Jensen model or the self-similar Gaussian model [4, 5, 6, 7, 8, 9]. These newer analytical models, informed by a better understanding of wind turbine wakes in different scenarios, have continued to evolve, enhancing the prediction accuracy of wind turbine wakes using analytical models.

An improved understanding of wind turbines operating in larger farms and the increasing size of turbines has led to significant advancements in analytical wake modeling. For instance, it was discovered that earlier models tend to underestimate wake losses in far-wake regions, as indicated by Ørsted in 2019 [10], often overestimating the resulting energy yield. This insight prompted the development of new models, such as the TurbOPark / TurbOGauss [7, 9] models and the Empiricial Gaussian model [11]. In parallel, progress in wake superposition techniques led to the development of the Cumulative Wake model [8] and the model by Zong and Porté-Agel [6].

Analytical wake models typically include tuning parameters that require calibration. Tuning these can be complicated, since the characteristics of wakes observed in the field often vary from those in wind tunnel tests and are complex to measure.

Additionally, accurately quantifying model precision using SCADA data is challenging. The difficulty arises from several factors, such as the stochastic nature of wind, sensor uncertainties, natural fluctuations like diurnal cycles and annual cycles, atmospheric stability, varying site-specific characteristics, and the fundamental assumptions that form the basis of most analytical models, such as assuming a homogeneous and time-invariant flow field. Given these varied sources of uncertainty, which contribute to the residual error between analytical models and SCADA data, it is challenging to assess the accuracy of a specific analytical wake models solely on SCADA data. This therefore highlights the importance of conducting comparative studies between different wake models to fully understand their effectiveness.

2. Objectives

The main objective of this study is to conduct a comparative analysis of analytical wake models, each previously calibrated on data from an offshore wind cluster with a capacity over 2.2 GW of active power at rated capacity. The objectives are further divided into the following key areas:

(i) Quantifying model performance by analyzing the residual error for the calibrated wake models on turbine, farm, and cluster level.
(ii) Quantifying uncertainty through examination of statistical moments.
(iii) Comparing predicted wake losses at both cluster and farm levels.
(iv) Evaluating top-hat and Gaussian profile models.

3. Methodology

In this work, the calibration of the wake models is performed on wind farms located in the Belgian offshore zone, as shown in Figure 1. These farms are denoted from top-left to bottom-right as WF_1 to WF_5. While the Pywake framework is utilized within this study [12], the methods are not restricted to this framework alone [13]. The calibration considers the wakes of the entire Belgian-Dutch offshore zone, depicted in Figure 2.

The Belgian offshore zone, with a high power density of 9.5MW/km^2 and a rated capacity of 2.2GW, serves as an ideal case for cluster-wake studies. In contrast, the Dutch offshore zone has a lower power density of 4.8MW/km^2 and a rated capacity of 1.5GW, totaling a rated capacity of 3.7GW.

The analytical wake models selected for calibration are:

- The Jensen wake model (1983) [1, 14].

- The Gaussian wake model by Bastankhah and Porté-Agel (2014) [2].
- The Gaussian wake model by Niayifar and Porté-Agel (2016 [4].
- The Gaussian wake model by Zong and Porté-Agel (2020) [6].
- The top-hat version of the TurbOPark model by Nygaard et al. (2020) [7].
- The Gaussian version of the TurbOPark model by Pedersen et al. (2022) [9].

These models are further highlighted in Table 1. All models are optimized for a constant turbulence intensity of 0.06 and a shear exponent of 0.12.

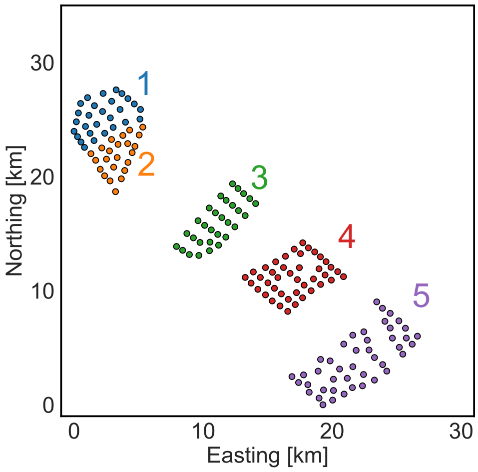

Figure 1. Wind farms within the Belgian offshore zone used for the calibration of wake models.

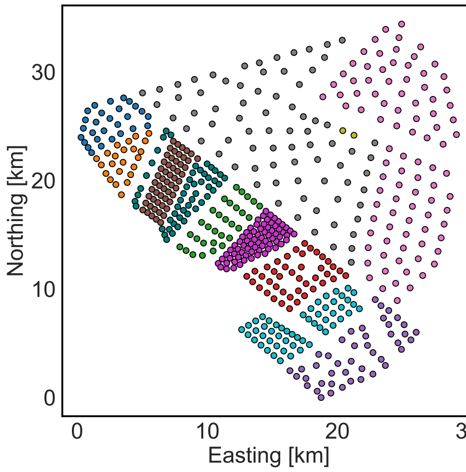

Figure 2. Wind farms within the Belgian-Dutch offshore zone incorporated into the wake calculation framework.

Table 1. Description of the analytical wake models considered and their respective sub models, tuning parameters, and parameter bounds for calibration.

Wake model	Jensen (1983)	Bastankhah (2014)	Niayifar (2016)	Zong (2020)	Nygaard (2020)	Pedersen (2022)
Turbulence model	-	-	Crespo Hernandez (1987)	Crespo Hernandez (1987)	Integrated	Integrated
Superposition model	Squared sum	Squared sum	Linear sum	Weighted sum	Squared sum	Linear sum
Blockage model	-	-	-	-	-	-
Ground model	-	-	-	-	-	Mirror
Tuning parameters $\mathbf{\Omega_{ref}}$	k = 0.04	k = 0.032	$k_a = 0.38$ $k_b = 0.004$	$k_a = 0.38$ $k_b = 0.004$	A = 0.6	A = 0.04
Parameter bounds $[\mathbf{\Omega_{min}}, \mathbf{\Omega_{max}}]$	$k \in [0.001, 0.2]$	$k \in [0.001, 0.2]$	$k_a \in [0.001, 1.0]$ $k_b \in [0.001, 0.06]$	$k_a \in [0.001, 1.0]$ $k_b \in [0.001, 0.06]$	$A \in [0.001, 1.0]$	$A \in [0.001, 0.6]$

The time-series calibration procedure, as described in van Binsbergen et al. (2023) and van Binsbergen et al. (2024) [13, 15], is performed on all wind farms depicted in Figure 1. The analysis within this paper focuses on wind directions ranging between 300 and 330 degrees and wind speeds between 7.0 and 9.0 m/s. As noted in van Binsbergen et al. (2024) [15], this particular wind direction tends to experience a limited amount of heterogeneous inflow and has demonstrated promising results with the Gaussian TurbOPark model by Pedersen et al. (2022) [9]. The calibration procedure is performed on one year of SCADA (Supervisory Control and Data Acquisition) data.

The optimization consists of three stages, all described in detail in van Binsbergen et al. (2023) [13]. For clarification the cost function and the final optimization stage are shortly described below. The cost function consists of two components, f and g. f penalizes the error on turbine level, while g penalizes the error of the entire cluster, both described by Equations 1 and 2, respectively. Here, N_{WF} indicates the number of wind farms, and N_{WT} the number of wind turbines for a specific farm i. P^{model} is the power derived from the wake model for farm i and turbine j, and P^{scada} is the active power from SCADA data for farm i and turbine j. U, ϕ, and $\mathbf{\Omega}$ are the freestream wind speed, freestream wind direction, and the set of tuning parameters for each wake model, respectively.

$$f\left(U, \phi, \mathbf{\Omega}\right) = \frac{1}{\sum_{i=1}^{N_{WF}} N_{WT,i}} \sum_{i=1}^{N_{WF}} \sum_{j=1}^{N_{WT,i}} \left(P_{i,j}^{model}\left(U, \phi, \mathbf{\Omega}\right) - P_{i,j}^{scada}\right)^2 \tag{1}$$

$$g\left(U, \phi, \mathbf{\Omega}\right) = \left(\sum_{i=1}^{N_{WF}} \sum_{j=1}^{N_{WT,i}} P_{i,j}^{model}\left(U, \phi, \mathbf{\Omega}\right) - \sum_{i=1}^{N_{WF}} \sum_{j=1}^{N_{WT,i}} P_{i,j}^{scada}\right)^2 \tag{2}$$

For each averaged 10-minute timestamp the following calibration procedure is performed:

- Stage 1: Optimization of wind speed.
- Stage 2: Optimization of wind speed and wind direction.
- Stage 3: Optimization of wind speed, wind direction and wake parameters.

For the first and the second stage a Quasi Monte Carlo (QMC) sampler by Bergstra and Bengio [16] is used, exploring the search space of the initially estimated freestream wind speed and wind direction. For the final stage the wind speed, wind direction and the wake parameters are calibrated simultaneously by minimizing the cost-function in Equation 3 using a multivariate Tree-Parzen Estimator (TPE) algorithm by Bergstra et al. [17]. The initial values for wind speed and wind direction of the final calibration stage are based on results from the previous stage, denoted as $\mathcal{U}^*$ and Φ^*, while a and b determine the weight between f and g, taken as $\frac{a}{b} = 4.0$, prioritizing the error on turbine-level. The parameter space, $\tilde{\mathbf{\Omega}}$, is defined by the minimum and maximum values of the parameters, $\mathbf{\Omega_{min}}$ and $\mathbf{\Omega_{max}}$, respectively, from Table 1. While the freestream wind speed and wind direction are not the primary target of the calibration, neglecting them can lead to inaccurate results, since the estimates from SCADA do not always reflect the true freestream wind conditions.

$$\text{minimize} \quad a \cdot f\left(\mathcal{U}, \Phi, \mathbf{\Omega}\right) + b \cdot g\left(\mathcal{U}, \Phi, \mathbf{\Omega}\right) \tag{3}$$
$$\text{subject to} \quad 0.95\mathcal{U}^* \leq \mathcal{U} \leq 1.05\mathcal{U}^* \tag{4}$$
$$\text{and} \quad \Phi^* - 15 \leq \Phi \leq \Phi^* + 15 \tag{5}$$
$$\text{and} \quad \mathbf{\Omega} \in \tilde{\mathbf{\Omega}} \tag{6}$$

For further elaboration on the optimization, van Binsbergen et al. (2023) and van Binsbergen et al. (2024) [13, 15] can be consulted.

To deal with power-curve mismatches, power-curve filtering is performed together with filtering based on operational codes, such as alarm annotations, of the wind turbines. During operational windows accompanied with low active power, the turbine is annotated as inactive, since low active power corresponds with low thrust loads, thereby assuming it has a limited effect on the internal wind farm flow field. The wake simulation is then performed without the turbines with low active power. When the number of inactive turbines reaches over half of the wind farm, the data is not considered for calibration. Data where underperformance and general power-curve mismatches occur above the limit given to low active power is not considered for calibration. Additional details on the filtering procedure can be found in van Binsbergen et al. (2023) [13].

4. Results

The results section is divided into two subsections. The first focuses on uncertainty quantification by analyzing the residual error on cluster, farm, and turbine level and by examining the statistical moments for each wind turbine. This is followed by a comparison of predicted wake losses before and after calibration of the tuning parameters.

4.1. Model and calibration performance

To fully assess the performance of the analytical wake models, both the relative error on cluster level and turbine level needs to be analyzed, since the cost function consists of two components, one penalizing the error of the entire cluster and one penalizing turbine errors. Therefore model performance is evaluated by measuring relative and accumulated relative error between the model predictions and SCADA data at each averaged 10-minute timestamp. This error metric is comparable to the one used in Nygaard et al. (2022) [18], but not identical. The equations defining these metrics, as referenced in Equations 7 and 8, are used to visualize the data in Figure 3. The results are then represented as a distribution for the available data between 300 and 330 degrees and wind speeds between 7.0 and 9.0 m/s using a kernel density estimate. In this context, $P_{rel}^{cluster}$ and P_{rel}^{farm} represent the relative error for the entire cluster and each individual wind farm, respectively.

$$P_{rel}^{cluster} = \frac{\sum_{i=1}^{N_{WF}} \sum_{j=1}^{N_{WT,i}} \left(P_{i,j}^{model} - P_{i,j}^{scada} \right)}{\sum_{i=1}^{N_{WF}} \sum_{j=1}^{N_{WT,i}} P_{i,j}^{scada}} \tag{7}$$

$$P_{rel}^{farm} = \frac{\sum_{j=1}^{N_{WT}} \left(P_{j}^{model} - P_{j}^{scada} \right)}{\sum_{j=1}^{N_{WT}} P_{j}^{scada}} \tag{8}$$

Figure 3 shows the kernel density estimate of the relative error in [%] between the analytical models and SCADA data, while Tables 2 and 3 show the mean and the standard deviation of the relative error in [%] between the analytical models and SCADA data, respectively. For Figure 3 and Table 2, a value below 0 suggest that the model underestimates the energy yield and overestimates the wake losses. Conversely, a value above 0 indicates an overestimation of the yield and an underestimation of the wake losses.

Results indicate that, across the entire cluster, the Gaussian wake models by Niayifar [4] and Zong [6], along with the top-hat TurbOPark model [7], exhibit lower mean and standard deviations. On the other hand, the Jensen model [1], Gaussian wake model by Bastankhah [2], and the Gaussian TurbOPark model [9] display relatively larger means and standard deviations. However, it is important to note that the overall performance of the model is not fully represented by the absolute error for the entire cluster or wind farm, since the cost-function used in the calibration also considers turbine level errors.

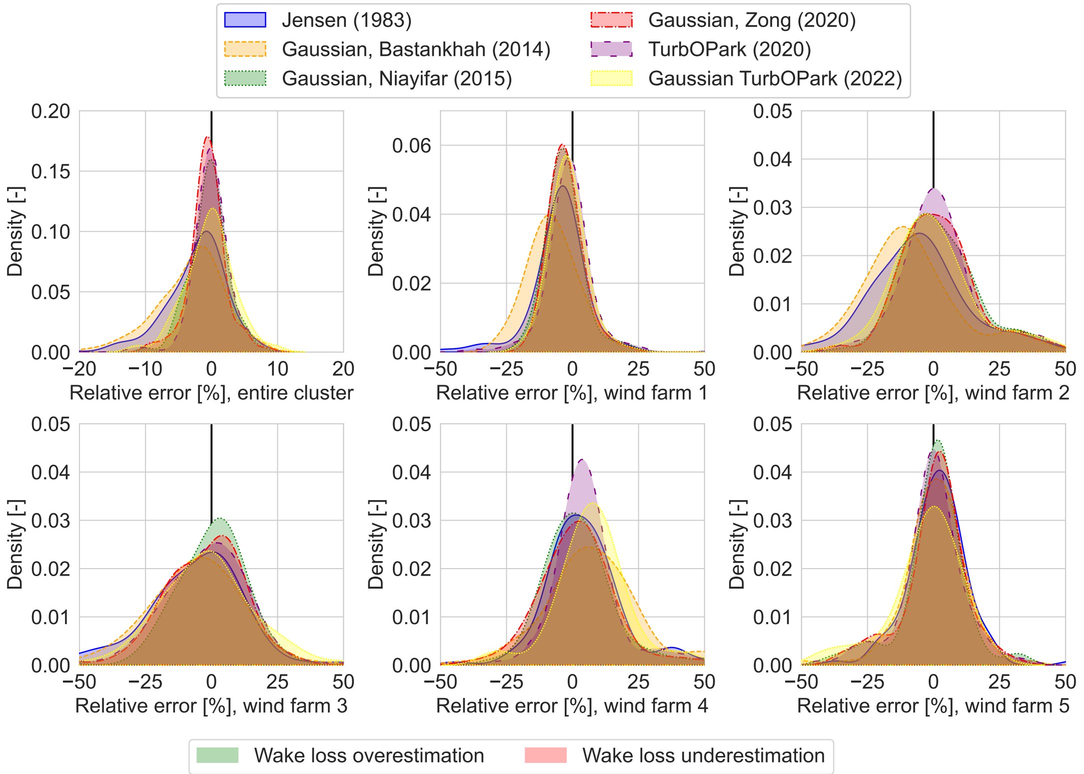

Figure 3. Relative error between calibrated analytical wake models and SCADA data.

Table 2. Mean relative error between calibrated analytical wake models and SCADA data.

Wake model	Jensen (1983)	Bastankhah (2014)	Niayifar (2015)	Zong (2020)	Nygaard (2020)	Pedersen (2022)
Cluster	-3	-3	0	-1	0	0
WF$_1$	-4	-6	-2	-1	0	-1
WF$_2$	-5	-8	3	2	2	0
WF$_3$	-7	-9	0	-2	-2	-2
WF$_4$	5	7	1	0	6	8
WF$_5$	2	1	0	0	-1	-4

To perform a more detailed analysis of the wake model performance, the error between individual turbines is analyzed using Equations 9 and 10. Here, $P_{acc,wt}^{cluster}$ and $P_{acc,wt}^{farm}$ represent the accumulated absolute error for each turbine within the entire cluster and the accumulated absolute error for each turbine per wind farm, respectively.

$$P_{acc,wt}^{cluster} = \frac{\sum_{i=1}^{N_{WF}} \sum_{j=1}^{N_{WT,i}} |P_{i,j}^{model} - P_{i,j}^{scada}|}{\sum_{i=1}^{N_{WF}} \sum_{j=1}^{N_{WT,i}} P_{i,j}^{scada}} \tag{9}$$

Table 3. Standard deviation of relative error between calibrated analytical wake models and SCADA data.

Wake model	Jensen (1983)	Bastankhah (2014)	Niayifar (2015)	Zong (2020)	Nygaard (2020)	Pedersen (2022)
Cluster	5	5	3	3	4	4
WF$_1$	13	13	11	11	12	11
WF$_2$	18	18	15	15	14	16
WF$_3$	19	27	17	15	17	19
WF$_4$	15	17	17	21	12	14
WF$_5$	11	11	12	12	12	16

$$P_{acc,wt}^{farm} = \frac{\sum_{j=1}^{N_{WT}} |P_j^{model} - P_j^{scada}|}{\sum_{j=1}^{N_{WT}} P_j^{scada}} \tag{10}$$

Figure 4 shows the kernel density estimate of the accumulated error for individual turbines in [%], while Table 4 presents the median values of the accumulated errors for individual turbines. The Gaussian wake model by Bastankhah [2] shows the highest accumulated error for each WF and the entire cluster, followed by the Jensen model [1]. It can be observed that the top-hat TurbOPark model [7] shows the best performance regarding this metric, followed by the Gaussian TurbOPark model [18].

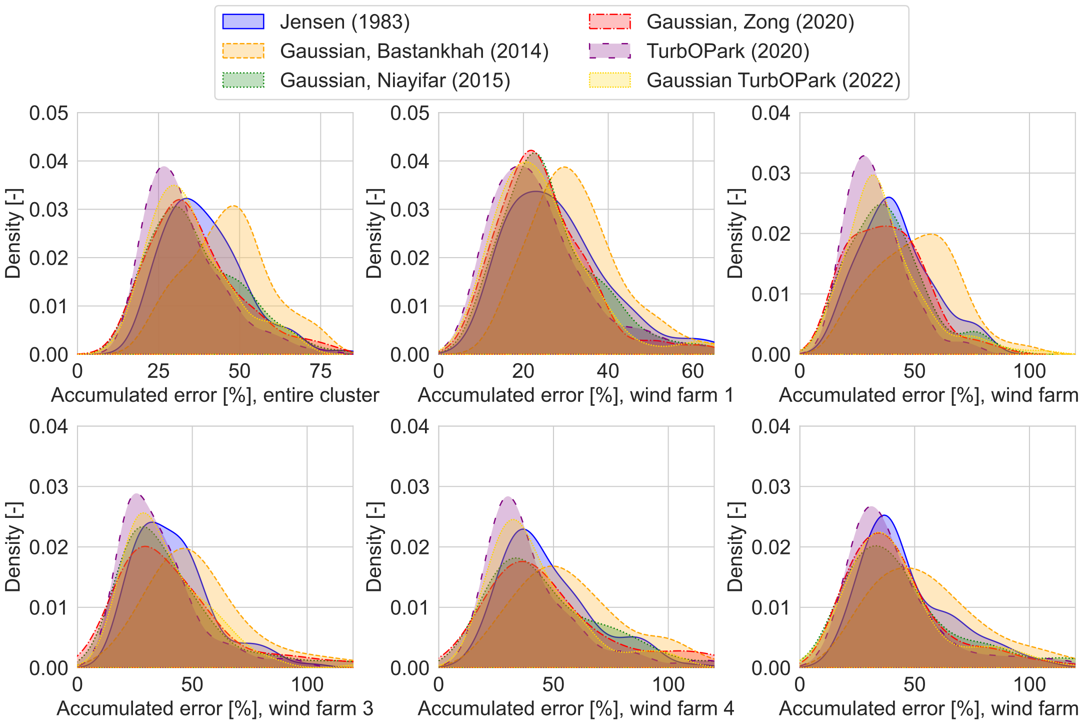

Figure 4. Accumulated relative error between calibrated analytical wake models and SCADA data.

Table 4. Median accumulated relative error between calibrated analytical wake models and SCADA data.

Wake model	Jensen (1983)	Bastankhah (2014)	Niayifar (2015)	Zong (2020)	Nygaard (2020)	Pedersen (2022)
Cluster	37	47	33	33	30	33
WF$_1$	26	30	24	23	22	23
WF$_2$	41	51	36	37	31	34
WF$_3$	40	48	34	35	32	34
WF$_4$	43	55	41	41	32	36
WF$_5$	40	52	38	37	34	38

Figures 5, 6, 7, and 8 support the previously made observations. Here, for each wind turbine the moments of the normalized error between calibrated wake model and SCADA data ($P_{rel}^{wt} = \frac{P^{model} - P^{scada}}{P^{scada}}$) are calculated based on the time series data. The distribution of the moments for each wind farm are then depicted in boxplots. Again, a mean lower than 0 indicates that the model underestimates the energy yield and overestimates the wake losses. Conversely, a mean higher than 0 indicates an overestimation of the yield and an underestimation of the wake losses. Figure 5 shows that the Bastankhah model [2] exhibits the most consistent deviation from a zero mean. Additionally, a trend is noticeable with the Niayifar and Zong models [4, 6], where, further downstream, the spread in mean deviation increases. Moreover, the Jensen, top-hat TurbOPark, and Gaussian TurbOPark models [1, 7, 9] show better consistency with fewer fluctuations in their mean values further downstream.

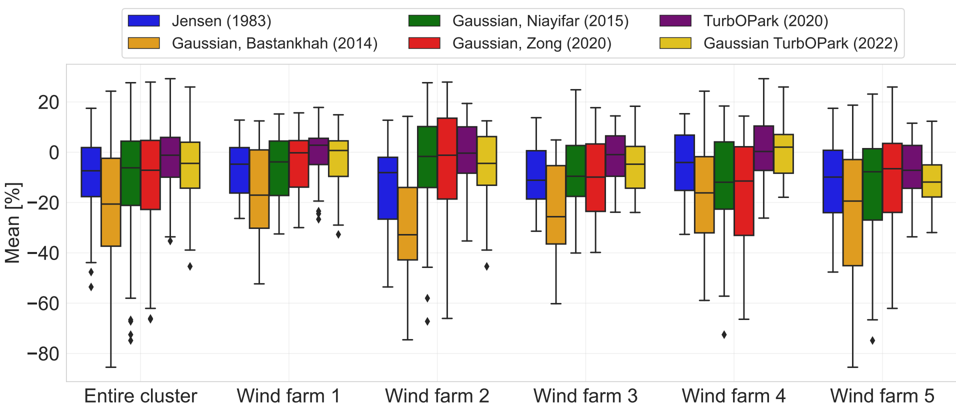

Figure 5. Time-series mean of the normalized error between model and SCADA for each wind turbine within the entire cluster and each wind farm.

In Figure 6, the analysis of the standard deviation between calibrated wake model and measurements shows that the TurbOPark model [7] outperforms others, followed by the Gaussian TurbOPark [9], the Jensen [1], Niayifar [4], and Zong [6] wake models. Consistently, the Bastankhah model [2] shows the highest standard deviations. Additionally, an increase in standard deviation further downstream is observed.

The asymmetry in optimization is quantified by analyzing the skewness for each wind turbine, as shown in Figure 7. For the upstream wind farm, a positive skew indicates a frequent underestimation of wake effects, possibly due to heterogeneous inflow or blockage, which are both not considered within the framework. Further downstream the models tend to skew negatively,

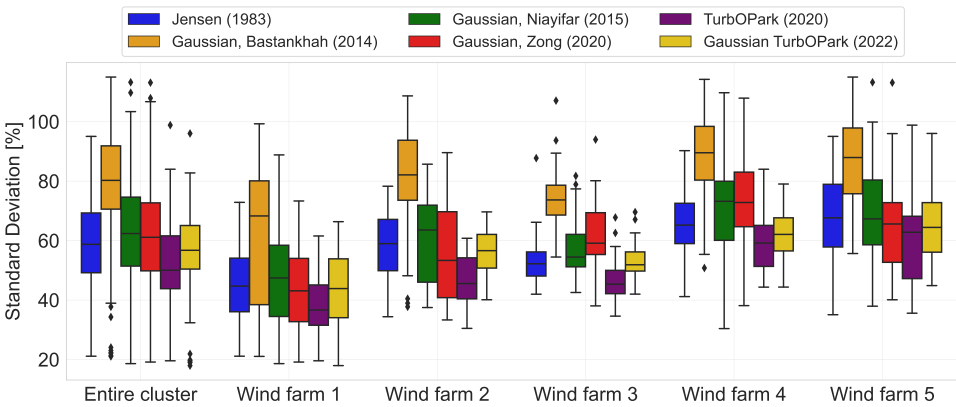

Figure 6. Time-series standard deviation of the normalized error between model and SCADA for each wind turbine within the entire cluster and each wind farm.

suggesting an overestimation of the wake effect.

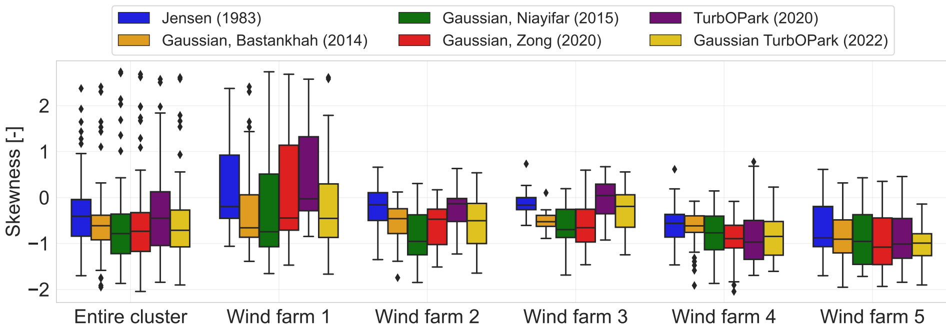

Figure 7. Time-series skewness of the normalized error between model and SCADA for each wind turbine within the entire cluster and each wind farm.

Kurtosis results are presented in Figure 8. A higher kurtosis corresponds to heavier distributed tails, but it is important to interpret kurtosis together with variance, since a low variance combined with high kurtosis implies that extremes are not as far from the mean as they would be as with high variance. Wind farm 1, with the lowest standard deviation, shows the highest kurtosis. Additionally, a trend can be observed where wake models with a higher standard deviation exhibit lower kurtosis. This trend cannot be seen between the wind farms, where the kurtosis slowly increases form wind farm 2 to wind farm 5.

When assessing key indicators for model performance in this deep-array scenario, namely by analyzing statistical moments and the relative and accumulated relative error, the top-hat TurbOPark model [7] shows the best performance, followed by the Gaussian TurbOPark model [9], especially further downstream. Interestingly, the performance of the Jensen model [1] does not decrease further downstream, unlike that from the Niayifar and Zong models [4, 6], which initially perform relatively well but show decreased performance further downstream. Across all scenarios, the Bastankhah model [2] consistently shows the worst performance.

Given the similarities between both TurbO models, apart from the radial wind speed profile,

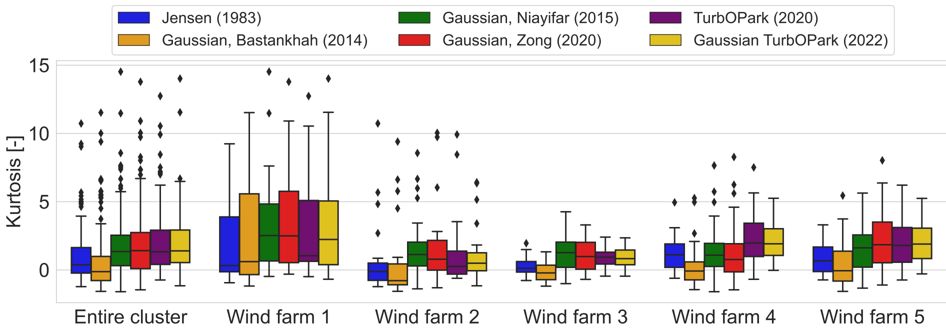

Figure 8. Time-series kurtosis of the normalized error between model and SCADA for each wind turbine within the entire cluster and each wind farm.

one hypothesis for the better performance of the TurbOPark model [7] is its overestimation of the wake deficit at its edges. The wake model assumes no turbine yaw misalignment, which can be present to some extent and can steer the wake away from a perpendicular direction to the wind. The absence of modeling yaw misalignment can lead to an underestimation of the wake effects when partial wake overlap of downwind turbines occurs. The top-hat models characteristic of overestimating the wake at its edges can indirectly mitigate this issue, compensating for the potential underestimation by Gaussian models. Furthermore, wake meandering can be considered as cause for the better performance of the TurbOPark model [7]. Braunbehrens and Segalini (2019) [19] suggest addressing wake meandering by widening the modeled wake, a phenomenon a top-hat model may indirectly account for more effectively than the Gaussian wake models.

4.2. Wake loss prediction

The calibrated tuning parameters, calculated using the optimization framework, are now used to compare wake losses before and after calibration for each analytical wake model. The calibrated values are determined as the median from the bin corresponding wind directions between 300 and 330 degrees and wind speeds from 7 to 9 m/s. Subsequently, wake losses at a wind speed of 8 m/s and a wind directions ranging from 300 to 330 degrees are calculated for the entire cluster and each individual wind farm.

The comparative results, shown in Figure 9, illustrate how the calibration framework affects expected wake losses. It reduces the predicted wake losses for the Gaussian TurbOPark model [9] and increases the wake losses for the Jensen [1], Bastankhah [2], and TurbOPark [7] models. Interestingly, for the entire cluster, the Niayifar [4] and Zong [6] models showed little change with calibration, however, a subtle shift is observed where the upstream losses are slightly more underestimated, while the downstream losses are more overestimated with the calibrated tuning parameters. The performance mismatch between the Bastankhah model (2014) and other models is more pronounced, likely due to the limitations in accurately accounting for wake losses at varied turbine spacings. This is apparent in wind farm 5, where the calculated wake losses by the Bastankhah model [2] deviate from those of other models for all wind directions. Similar discrepancies are observed in wind farms 3 and 4, particularly with wind coming from the Dutch zone, characterized by larger turbine spacings in comparison to the Belgian zone. A similar, but less evident, pattern is seen for the Jensen model [1].

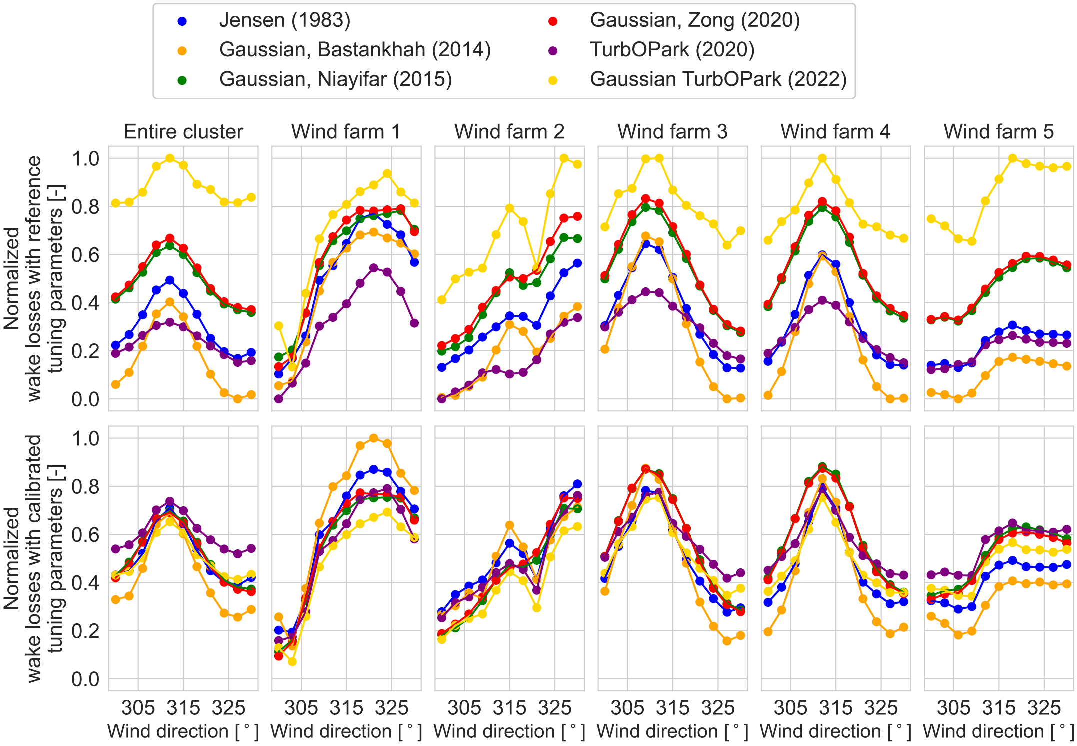

Figure 9. Comparison of normalized wake effects: On the top are the predicted normalized wake losses using reference tuning parameters, and below are the normalized wake losses using calibrated tuning parameters. As the basis for normalization within each column the lowest and highest wake loss values from either reference or calibrated parameters are used.

5. Conclusion

A multi-level hyperparameter optimization framework is utilized to calibrate analytical wake models using SCADA data from five wind farms in the Belgian-Dutch offshore zone. The calibration is performed on wind coming from north-west (300 - 330 degrees) and wind speeds ranging from 7 to 9 m/s, a direction identified as promising for calibration due to its limited inflow heterogeneity and large amount of clustered wind farms [15]. A constant turbulence intensity of 0.06 and a shear exponent of 0.12 are assumed, without considering blockage models. Six models are compared in this study, namely the analytical wake models of Jensen (1983), Bastankhah (2014), Niayifar (2016), Zong (2020), Nygaard (2020) and Pedersen (2022). The relative and accumulated relative error between the measurement data and the model predictions are analyzed both per wind farm and for the entire cluster, as well among individual turbines. Statistical moments for each wind turbine are calculated and presented as boxplots per wind farm.

In general, the TurbOPark model (2020) shows the best overall performance, followed by the Gaussian TurbOPark model (2022), Niayifar (2016) model, and Zong (2020) model. Where the Niayifar (2016) and Zong (2020) models perform relatively well close to the upstream turbines, further downstream their performance decreases. This is opposite to the top-hat and Gaussian TurbOPark models (2020,2022) and the Jensen model (1983). The Bastankhah (2014) model performs poorest of all models.

Using calibrated tuning parameters, new wake losses are calculated for the specified wind conditions. The performance shortfall of the Bastankhah (2014) model appears to be primarily linked to variable and larger turbine spacings, particularly noticeable in Wind Farm 5 and under conditions of more northerly directed winds, where larger turbine spacings are present. Conversely, the TurbOPark models shows their effectiveness in these scenarios, underscoring its development purpose.

It's important to note that despite varying levels of accuracy, all models show similar trends in wake losses after calibration. This suggests that with proper calibration, any model can be effective, provided its fundamental limitations are understood and accounted for.

Acknowledgments
The authors would like to acknowledge the support of VLAIO in the context of the Blauwe Cluster project Cloud4Wake. The authors would also like to acknowledge the Energy Transition Funds for their support of the Poseidon project. This research was supported by funding from the Flemish Government under the "Onderzoeksprogramma Artificiële Intelligentie (AI) Vlaanderen" programme.

References

[1] Jensen N 1983 A note on wind generator interaction
[2] Bastankhah M and Porté-Agel F 2014 *Renewable Energy* **70** 116–123 ISSN 0960-1481
[3] Pope S B 2000 *Turbulent Flows* (Cambridge University Press) ISBN 9780511840531
[4] Niayifar A and Porté-Agel F 2016 *Energies* **9** 741 ISSN 1996-1073
[5] Ishihara T and Qian G W 2018 *Journal of Wind Engineering and Industrial Aerodynamics* **177** 275–292 ISSN 0167-6105
[6] Zong H and Porté-Agel F 2020 *Journal of Fluid Mechanics* **889**
[7] Nygaard N G, Steen S T, Poulsen L and Pedersen J G 2020 *Journal of Physics: Conference Series* **1618** 062072
[8] Bastankhah M, Welch B L, Martínez-Tossas L A, King J and Fleming P 2021 *Journal of Fluid Mechanics* **911** ISSN 1469-7645
[9] J G Pedersen E Svensson L P and Nygaard N G 2022 *Journal of Physics: Conference Series* **2265** 022063
[10] Ørsted 2019 Ørsted presents update on its long-term financial targets
[11] NREL 2023 Floris. version 3.5 URL `https://github.com/NREL/floris`
[12] Pedersen M M 2023 URL `https://gitlab.windenergy.dtu.dk/TOPFARM/PyWake`
[13] van Binsbergen D, Daems P J, Verstraeten T, Nejad A and Helsen J 2023 *Wind Energ. Sci. Discuss. [preprint]*
[14] Katić I, Højstrup J and Jensen N 1987 A simple model for cluster efficiency
[15] van Binsbergen D, Daems P J, Verstraeten T, Nejad A and Helsen J 2024 Scalable SCADA-based Calibration for Analytical Wake Models Across an Offshore Cluster *Journal of Physics: Conference Series, under publication*
[16] Bergstra J and Bengio Y 2012 *Journal of Machine Learning Research* **13** 281–305 URL `http://jmlr.org/papers/v13/bergstra12a.html`
[17] Bergstra J, Yamins D and Cox D D 2012 Making a science of model search
[18] Nygaard N G, Poulsen L, Svensson E and Pedersen J G 2022 *Journal of Physics: Conference Series* **2265** 022008
[19] Braunbehrens R and Segalini A 2019 *Journal of Wind Engineering and Industrial Aerodynamics* **193** 103954 ISSN 0167-6105