Authors’ Response to Reviews of

Assessing the Accuracy of a Three-Year High-Resolution Mesoscale Wind Farm Wake Simulation with Lidar and Satellite Radar Data

Alexandros Palatos-Plexidas on behalf of all authors
Wind Energy Science (WES) Journal,

RC: ***Reviewers’ Comment***, AR: Authors’ Response, ☐ Manuscript Text

1. Reviewer #1

RC: ***The paper employs WRF with the Fitch wind-farm parameterization to simulate three years of operation in the Belgian–Dutch offshore cluster. Results are compared with Lidar data at four sites and with SAR data, classified by stability regime to assess performance under different conditions.***

Overall, the paper is clear, well organized, and technically competent. However, the study’s novelty is not well articulated. The paper currently reads as a case-study application of an existing parameterization rather than addressing a clear research gap. The introduction lists several limitations of Fitch’s parameterization and mentions multiple validation studies, yet the manuscript does not explicitly explain what new insight this paper brings. Clarifying the main research question and the key take-home messages would significantly improve the manuscript.

AR: Dear Reviewer,
We sincerely thank you for your time and for the detailed, insightful comments. As stated below, we have addressed each of these comments, and believe this has significantly improved the quality of the paper. All changes and additions in the revised manuscript are highlighted in blue and enclosed in boxes.

We should also highlight that we have corrected the interpolation at hub-height for the BSB-lidar cases, resulting in a more accurate representation of the metrics throughout the full manuscript. More specifically, in the first version of the manuscript, we unintentionally used the lidar-measured heights without adding the BSB platform height of 45 m. This led to an incorrect estimation of both wind speed and wind direction at the intended hub height of 107 m.

1.1. General Comments

RC: ***1. Please clarify the specific research question or gap being addressed. Is the purpose to validate Fitch’s parameterization, or to evaluate WRF’s long-term predictive skill for offshore wakes?***

AR: The purpose of this study is not to validate the Fitch parameterization in isolation, but rather to assess the long-term predictive performance of the full WRF–Fitch modeling system for offshore wake dynamics in an operationally relevant environment. The analyzed Belgian–Dutch cluster is currently one of the largest offshore wind farm clusters, and we consider it crucial to investigate wake dynamics, as according to the 2030 planned commissions, denser wind farm clusters will be deployed in the North Sea Borgers et al. [2025]. We therefore evaluate the model’s ability to reproduce observed wake signatures, stability-dependent wake behavior, blockage, and wind speed deficits using a multi-year dataset. Existing mesoscale studies that use the

Fitch (or any other WFP) typically span short periods (days to months) and use resolutions ≥2km, whereas offshore operational wake behavior requires year-long statistics and finer-scale representation. Even the existence of year-long studies is usually limited in resolution (≥2km). By conducting a three-year 1-km simulation supported by four lidars, including one inside the cluster, and SAR snapshots, we provide a unique long-term validation of offshore wake simulations. The duration (three years) as well as the model resolution (1km) allow us to quantify the wakes in both short-climatological terms and at high mesoscale resolution, as there are not many such studies that exist. The Fitch scheme is an integral component of this setup, but its evaluation is embedded within the broader assessment of the ability of the model to simulate wakes over longer periods and provide better statistics.

To reflect this in the revised manuscript, following modifications have been introduced. The abstract has been changed:

> The rapid expansion of wind farm installations in the North Sea results in an increased need for understanding their influence on the local atmosphere, as well as the interactions between them. Wind farm operation and power production are affected by wakes produced both within the wind farm and by upwind wind turbines. Accurately estimating these impacts requires robust mesoscale atmospheric modeling capable of capturing extended wind speed deficits and their dependence on atmospheric stability. This study presents a three-year-long mesoscale analysis using the Weather Research and Forecasting (WRF) model at a horizontal resolution of 1 kilometer, a comparatively long and high-resolution configuration within the context of offshore wake research. The simulations are evaluated against four lidars located in the Southern Bight of the North Sea in the vicinity of the 3.7 GW Belgian-Dutch offshore wind farm cluster, providing an extensive observational basis that covers upstream, intra-farm, and downstream flow conditions. Coupling the mesoscale atmospheric model with the Fitch wind farm parameterization (WFP) scheme significantly improves simulation accuracy, particularly in regions frequently affected by wake effects. An evaluation across different atmospheric boundary layer stability regimes shows that the model performs best under less extreme stability, while a detailed assessment of upstream, intra-cluster, and downstream wake characteristics highlights the added value of the Fitch WFP scheme for multi-year offshore applications. Finally, synthetic aperture radar images from selected wake events are compared with the model output, demonstrating that the WFP scheme effectively captures the wake structures at the analyzed timestamps. Together, these results provide one comprehensive long-term evaluation of offshore wake behavior in a dense wind farm cluster, helping to address the current gap in mesoscale wake studies.

We have updated the introduction (lines 63-81) as follows:

> The current work presents an assessment of model performance and analysis of wake effects from a three-year-long high-resolution (1 km) WRF simulation of the Southern Bight of the North Sea, with a specific focus on the 3.7 GW Belgian-Dutch wind farm cluster. This cluster consists of several high-density wind farms installed next to each other, making it the largest offshore cluster operational today, hence resulting in a prime test case for studying wind farm wakes. Extending the analysis over a three-year period enables a comprehensive characterization of seasonal patterns and temporal variability. The study is driven by two primary objectives that address notable gaps in current offshore wake research. The purpose of this study is not to validate the Fitch parameterization in isolation, but rather to assess the long-term predictive performance of the full WRF–Fitch modeling system for offshore wake dynamics in an operationally relevant environment. First, we assess the three-year WRF–Fitch simulation performance using lidar measurements positioned both within and around the

Belgian-Dutch cluster, enabling an evaluation of model performance under varying ABL stability regimes. This is particularly relevant given that long-term, wake validations at 1 km (assessed under different stability regimes) resolution remain rare, especially over multi-year periods and in densely built offshore environments. The strategic placement of the lidar probes allows us to assess how the model represents inflow conditions, intra-cluster wake dynamics, and downstream recovery for different wind directions and wind speed regimes, and the availability of four lidars spanning upstream, downstream, and intra-farm positions provides a uniquely comprehensive observational basis for this evaluation. Second, we evaluate the performance of WRF for a set of specific wake events by comparing the simulated wake signatures with satellite-retrieved wind fields. This dual comparison, quantitative through lidar transects and qualitative through two-dimensional SAR wake structures, provides a robust, multi-sensor estimation of wake deficits, blockage effects, and model skill. Together, these objectives provide a comprehensive long-term evaluation of offshore wake behavior in a high-density wind farm cluster.

RC: ***2. Several metrics show that the wind-farm (WF) simulations do not outperform the no-wind-farm (NWF) simulations, which seems inconsistent with some of the conclusions.***

AR: Results and conclusions have been thoroughly updated in the revised manuscript. In summary, the slightly larger negative bias in the WF simulations can be attributed to the fact that WHi, EPL, and LEG are located close to the Belgian–Dutch wind-farm cluster, where wake effects reduce wind speeds downstream. Consequently, the improved biases in the NWF runs likely result from partial error cancellation with the negatively biased ERA5 input, whereas these errors become slightly amplified in the WF simulations (while still remaining noticeably smaller than in ERA5). This behavior is consistent with previous studies showing improved performance of the Fitch scheme in near-wake regions [García-Santiago et al., 2024, Ali et al., 2023, Fischereit et al., 2024]. The validation analysis has been expanded accordingly, now including wind-speed and wind-direction bias metrics per wind-direction sector (polar plots in Figures 7 and 8). Additional relevant discussion is provided in our responses to comments 8, 9, and 11.

Furthermore, the transect analysis in Section 3.2.3, Tables 4 (for SW) and 5 (for NE) wind directions, respectively, have been added to assess bias and cRMSE along upstream, intra-farm, and downstream regions. These results show that the WF simulations reduce the bias in intra-farm and downstream areas, particularly for wind speeds between 5 and 8 m s^{-1} under SW wind direction and between 8 and 11 m s^{-1} when wind direction is NE. Finally, mean absolute error (MAE) between the model and the 10-meter SAR wind speeds across the transects, as well as the absolute error (AE) relative to the lidars, have been added in Section 3.3 (Table 8). These metrics consistently show that the WF runs outperform the NWF simulations for the analyzed events. An updated portion of the conclusions addressing this point is also included in our response to comment 13.

1.2. Specific Comments

RC: ***3. Line 105: Please specify whether one-way or two-way nesting was used. I assume one-way feedback given the use of nested inner domains (d03).***

One-way nesting is used in this study. Section 2.1 has been updated accordingly (lines 96-97):

Furthermore, one-way nesting is applied to prevent any potential noise from the innermost domain from propagating into the parent domains.

RC: ***4. Equations (1-3): Fitch's implementation in WRF applies an energy correction to the momentum and***

TKE tendencies. The equations presented should reflect the actual WRF formulation, not only the original Fitch et al. (2012) expressions.

AR: We thank the reviewer for this comment. The energy correction term ensures that the vertically distributed extraction matches the turbine power curve at hub height. We have included the energy correction terms in Equations 1 and 2, denoted by ec, in Section 2.1. However, we believe that fully detailing this in the manuscript would distract too much attention from the main storyline. Instead, we add a clarifying sentence to the discussion of Equations 1 and 2 in Section 2.1.1 (lines 132-134) as follows:

> The principal equations for describing the influence of the wind farms on the mean flow in a discrete Cartesian computational cell i, j, k are defined as
>
> $$\frac{\partial |\boldsymbol{V}|_{ijk}}{\partial t} = -\frac{N_t^{ij} C_T \left(|\boldsymbol{V}|_{ijk}\right) |\boldsymbol{V}|_{ijk}^2 A_{ijk}}{2(z_{k+1} - z_k)} ec, \tag{1}$$
>
> $$\frac{\partial \mathrm{TKE}_{ijk}}{\partial t} = \frac{N_t^{ij} C_{\mathrm{TKE}} \left(|\boldsymbol{V}|_{ijk}\right) |\boldsymbol{V}|_{ijk}^3 A_{ijk}}{2(z_{k+1} - z_k)} ec, \tag{2}$$
>
> $$\frac{\partial P_{ijk}}{\partial t} = \frac{N_t^{ij} C_P \left(|\boldsymbol{V}|_{ijk}\right) |\boldsymbol{V}|_{ijk}^3 A_{ijk}}{2(z_{k+1} - z_k)}. \tag{3}$$
>
> These equations describe the change in horizontal wind speed $|\boldsymbol{V}|$, the wind power production P, and the turbulence enhancement TKE due to wind turbines, where z_k defines the height at model level k, and A_{ijk} is the cross-sectional rotor area of one wind turbine between the height levels k and $k + 1$. Furthermore, N_t^{ij} is the number of turbines located within cell (i,j). The thrust coefficient C_T and power coefficient C_P depend on the wind turbine type and are obtained from a mix of public and confidential data. The TKE coefficient is defined as $C_{\mathrm{TKE}} = \alpha(C_T - C_P)$, with $\alpha = 1$ in the original Fitch implementation, driven by the idea that momentum lost by the mean flow, which is not converted to electrical power, is transformed into turbulence. However, it was found that this can lead to a significant overestimation of the TKE compared to high-fidelity simulations [Archer et al., 2020]. Although there remains significant debate over the appropriate value of α [Larsén and Fischereit, 2021, García-Santiago et al., 2024], this study uses the current default value in WRF, i.e., $\alpha = 0.25$. Last, ec is the energy correction factor that ensures that the vertically distributed extraction matches the turbine power curve at hub height. The wind speed deficit due to the wind turbine influence is defined as the normalized difference between the unperturbed and the wind-farm-affected wind speed, i.e., $\Delta U_{\mathrm{wake}} = (U_{\mathrm{NWF}} - U_{\mathrm{WF}})/U_{\mathrm{NWF}} \times 100$.

The energy correction factor is defined in the WRF source code as:

$$ec = \begin{cases} \frac{P_{\mathrm{hub}}}{P_{\mathrm{blade}}}, & \text{if } P_{\mathrm{hub}} \neq 0 \text{ and } P_{\mathrm{blade}} \neq 0 \\ 1, & \text{otherwise,} \end{cases} \tag{4}$$

where:

$$P_{\text{hub}} = \frac{1}{2}\rho C_P A_{ijk} V_{\text{hub}}^3, \tag{5}$$

$$P_{\text{blade}} = \sum_{\text{bl}} \frac{1}{2}\rho C_P \left|\mathbf{V}_{ijk}\right|^3 A_{ijk} \Delta X^2. \tag{6}$$

In Equations (5), and (6), $\rho = 1.23$ kg m^{-3} is the air density under standard atmosphere conditions, V_{hub} is the hub height wind speed, and bl is the number of height levels across the area swept by the turbine blades. Turbulence advection is activated in the high-resolution domains.

RC: ***5. Line 182: The statement "The wind speed is calculated as a reflected signal" is unclear. Which wind speed do you mean: surface, 10 m, or another height? Note that the 10 m wind speed derived from SAR is inferred from radar backscatter by assuming neutral stability. So, it is not a measured quantity, but a processed one.***

The overall text in Section 2.3.2, has been updated as follows (lines 191-201):

> The SAR data used in this study were retrieved by the Sentinel-1 European Space Agency mission. More specifically, there are two satellites, Sentinel-1A (2014-present) and Sentinel-1B (2016-2021), equipped with C-band Synthetic Aperture Radar (SAR) instruments operating at 5.405 GHz, and both satellites follow a sun-synchronous polar orbit at an altitude of 693 km [Hasager et al., 2024]. Focusing on the area of interest, including the Belgian-Dutch wind farms cluster in the Southern Bight of the North Sea, as well as additional upstream and downstream areas to study potential wake structures, the analyzed time instances occur approximately at 06:00 UTC every twelve days, and at 17:40 UTC every six days, respectively (Sentinel-1A, 2021-2023). The operation of Sentinel-1B during 2021 provided an additional set of overpasses at around 17:32 UTC every six days. This corresponds to approximately 8 to 10 SAR snapshots per month over the area of interest during the analyzed period. The wind speed at 10-meter height is then estimated by applying the CMOD5N geophysical model function [Hersbach et al., 2007], as explained also in the work of Siedersleben et al. [2020]. This indirect retrieval method assumes that radar backscatter from the sea surface, which reflects centimeter-scale waves generated by instantaneous wind stress, can be converted into wind speed at 10-meter height [Ahsbahs et al., 2017, 2018]. The wind fields are provided at a 500-meter pixel resolution through the Global Wind Atlas Science Portal of the Danish Technical University (DTU) Wind Energy Department (https://science.globalwindatlas.info/#/map/satwinds last access: 30 July 2025).

RC: ***6. For completeness, please include a description of the EMD metric in the Appendix.***

AR: We have addressed this in the revised manuscript (Appendix).

> The Earth Mover's Distance is assessed using the Wasserstein first distance (WD). WD is defined as a distance between two probability measures P and Q on a metric space $(\mathcal{R} \times \mathcal{R})$ by
>
> $$\text{EMD}(P, Q) = \inf_{\gamma \in \Gamma(P,Q)} \int_{\mathcal{R}\times\mathcal{R}} \|x - y\| \, \mathrm{d}\gamma(x, y), \tag{7}$$

where $\Gamma(P, Q)$ denotes the set of all couplings with marginals P and Q.

RC: ***7. Line 256/Eq. A4: Clarify how wind direction is computed. Are wind components rotated from model to Earth coordinates?***

AR: Wind speed (only horizontal components are used in the study both from the lidars and WRF) and wind direction are computed directly from WRF, using Python. The wind components are indeed rotated to Earth coordinates. It has been included in the manuscript, in Section 3.1, lines 255-256:

> All results presented in this study use wind speed and direction rotated to Earth-fixed coordinates.

RC: ***8. Figure 5: Why does including wind-farm effects increase the bias relative to observations at most sites, given that both WF and NWF simulations share the same background bias from ERA5?***

AR: At WHi, EPL, and LEG, bias is already negative in the ERA5 data, which propagates into the WRF simulations and results in a negative bias in the NWF runs as well. The slightly larger negative bias in the WF simulations arises because WHi, EPL, and LEG are located close to the Belgian–Dutch wind farm cluster, where wake effects generate reduced wind speeds downstream. Consequently, the improved biases in the NWF runs likely stem from error cancellation related to the biased ERA5 input, whereas these errors become slightly amplified (while remaining noticeably lower than the ERA5 errors) in the WF simulations.

This paragraph in Section 3.2.1 (lines 294-301) has been added:

> The bias score (Figure 5A) shows that both ERA5 and the NWF simulation exhibit a consistent negative bias at all sites except BSB, a consequence of the inherent ERA5 underestimation that propagates into WRF. The additional wind speed reduction introduced by the Fitch WFP leads to a slightly more negative bias in the WF run, although WRF generally aligns better with the lidar observations than ERA5. **Part of the improved bias in the NWF runs likely results from an error-cancellation effect with the negatively biased ERA5 input, whereas in the WF simulations, these errors become slightly amplified due to wake-induced deficits, but remain lower than those in ERA5.** At BSB, both ERA5 and the NWF runs show a positive bias, while the WF simulation produces a smaller negative bias, corresponding to an improvement of about 82% over NWF and 73% over ERA5 in the near-wake under the predominant southwesterly flow.

In addition, in the analysis of SW, and NE wind direction in Section 3.2.3, in Figure 12, and Table 4, we observe that in the upstream location (WHi under SW wind conditions), there is always a negative bias, most likely propagated from ERA5 and the overall WRF performance. This under specific conditions, for example, in Table 4, under [3-5) m s^{-1} wind speeds, BSB provides a negligible bias in the NWF simulations that is higher in the WF. We consider that this can happen due to error cancellation. Similarly, under NE wind conditions in Table 5, when the wind speed is low, the bias at BSB is lower in NWF simulations compared to the WF.

In the conclusions, this paragraph (lines 593-605) has been updated as well:

> Our findings demonstrate that the use of the Fitch WFP scheme, in combination with this particular WRF configuration, captures the wake effects and provides an improvement in model performance relative to a simulation without wind farms at the intra-farm region. First, focusing on the time-averaged wind speeds over the full analysis period, the WF simulations yield a substantial improvement in both bias and EMD, with reductions of approximately 82% and 73% compared with the NWF simulations,

respectively, in the intra-cluster (BSB location). **Furthermore, when analyzing the wind speed bias under different wind directions, both NWF and WF configurations reduce the negative bias inherited from the ERA5 data, with the results indicating that the Fitch scheme introduces an additional deficit downstream of the Belgian-Dutch cluster.** However, within the intra-farm area, the Fitch scheme provides a more accurate wind speed distribution across all wind directions, underscoring its importance for representing strongly-waked environments. Beyond the improvements in bias and EMD, the WF simulations perform similarly to the NWF runs in terms of cRMSE and correlation, with only small and marginal differences. Similarly, wind direction metrics are also improved in the WF simulations, as bias is reduced by approximately 51% at BSB, while an analysis of wind direction biases by sector reveals that WF can perform better than ERA5 for wind directions from northeast to south.

RC: ***9. Figure 5 (cRMSE): ERA5 appears to outperform both WRF setups when compared to centralized error. Please discuss why.***

AR: Following the previous comment, Section 3.2.1 has also been updated as follows in lines 316-323:

The larger cRMSE in WRF compared to ERA5 is expected because the coarser ERA5 fields do not resolve kilometer-scale wind speed fluctuations, resulting in a smoother signal and lower instantaneous error. In contrast, the higher spatial and temporal variability represented in WRF leads to phase mismatches between simulated and actual wind fluctuations, producing larger instantaneous discrepancies that manifest as higher cRMSE. The study of [Pronk et al., 2022] shows that ERA5 outperforms WRF simulations in terms of cRMSE and correlation coefficient, both on land and offshore, under different stability conditions. Additionally, a consistent negative bias in wind speed is reported. These findings align with the WRF sensitivity analysis of Olsen et al. [2025] in the North and Baltic Seas, indicating that ERA5 outperforms the various WRF configurations in terms of cRMSE and correlation.

RC: ***10. Model sensitivity: Since the WF simulation rarely performs best (only EMD metric at BSB location), could parameter choices (e.g., TKE coefficient = 1 instead of 0.25) influence the results? A brief sensitivity analysis might help.***

AR: Dear reviewer, thank you for your comment. We have conducted a brief sensitivity analysis on the a values on the Fitch WFP scheme, resulting in no significant differences. More precisely, two additional simulations for a period of about 6 months (mid-January to mid-July 2022, since WHi measurements are available during that period) using $a = 0$, and $a = 1$ have been performed. The main conclusions are described below:

- Bias values across the WF simulations are overall very similar. While some differences are observed at individual locations (e.g., slightly lower bias for $a = 1$ at BSB), no single value of a consistently outperforms the others across all sites.
- cRMSE, EMD, and correlation (r) values are likewise very close between the WF simulations, with only minor, site-specific differences and no systematic advantage for a particular value of a.
- Regarding the wind direction metrics, similar trends to the simulations with $a = 0.25$ are indicated in the rest of the WF simulations, while ERA5 results in better metrics as explained in the response to comment 9.

Overall, this assessment indicates that the sensitivity of the wind speed and wind direction metrics analyzed in this study to the choice of a is limited, and that differences between configurations are small. It is also worth noting that the sensitivity assessment was conducted over a selected six-month period and is therefore intended to provide guidance on parameter choice rather than a comprehensive evaluation across the full 2021–2023 analysis period. We also note that the value of $a = 0.25$ was originally introduced to improve the representation of both wind speed and turbulent kinetic energy (TKE) in the Fitch WFP scheme (e.g., Archer et al. [2020], García-Santiago et al. [2024]), which is not (TKE) evaluated in the present study but remains an important consideration. The value $a = 0.25$ was therefore retained as a reasonable and balanced choice

RC: ***11. Figure 6: The same issue applies to wind-direction metrics.***

AR: Wind direction metrics are better in the WF simulations in comparison to the NWF runs. However, compared to ERA5, there is indeed a worse matching for the same reasons explained in the previous comment, as well as in the studies of Pronk et al. [2022], and Olsen et al. [2025].

In addition, in the analysis of wind direction under different wind directions, in Figure 8 (in the revised manuscript), it appears that ERA5 does not always perform better when wind direction is from 0 to 180^o (clockwise) (e.g., at WHi, EPL, and LEG WF is better in these cases), but it is slightly more accurate when wind goes from 180 to 360^o (clockwise). This is also mentioned in the response to comment 8 (lines 336-340):

> Similar to the wind speed bias per wind direction, Figure 8 presents the wind direction bias. Although the mean bias in Figure 6A suggests a closer agreement between the lidars and ERA5, WRF provides a better representation of the wind directions when the wind originates from the north to south-southeast(0 to 180^o). Conversely, for westerly wind directions between 180 and 360^o ERA5 performs better. At BSB, the mean bias is nearly identical between the WF runs and ERA5, but ERA5 shows a slightly closer agreement with the lidar measurements across all wind directions.

RC: ***12. Figure 15: How do you reconcile the comparison with SAR, which assumes neutral stability, while WRF simulates varying stratification? The WRF–Lidar comparison may be more representative, and the SAR transects can be more of a qualitative reference.***

AR: Cases A and C, where WRF indicates stable-neutral and unstable-neutral conditions, respectively, align well with the neutrality assumption in SAR imagery. In contrast, cases B and D, characterized by very unstable stratification at the analyzed locations, introduce uncertainty both in the SAR neutrality assumption and in the stability classification derived from L in WRF. More specifically, Section 3.3. has been updated accordingly: (with bold in lines 544-549 is the part that addresses this specific comment)

> Figure 16 presents the comparison of SAR and WRF wind fields at 10-meter height. The left column shows the SAR-derived wind speeds, denoted by the subscript 1. The middle column depicts the WF wind speeds, indicated by a subscript 2, and in the right column, the reference NWF wind speeds are illustrated, and denoted with a subscript 3, as a reference case. In case A, higher wind speeds are also related to stable-neutral conditions and a shallow boundary layer (see Table 7). The wake structure of the Belgian-Dutch wind farm cluster is well captured by the WF simulation. However, a slight deviation in wind direction is observed in WRF, exhibiting a positive bias toward more westerly flow. Similar bias has also been highlighted in the wind roses (Figure 4) and in the wind direction validation heatmap in Figure 6A. The prolonged wake extending more than 50km downstream under stable-neutral stratification agrees with the findings of Cañadillas et al. [2020] and Platis et al. [2018], who also report that wake structures are more pronounced and longer during stable conditions. Furthermore, Siedersleben et al. [2020] examined a stable event simulated with the Fitch scheme in WRF (version 3.8.1), comparing the original formulation ($a = 1$) with a variant excluding the added TKE term, and showed that including the added TKE source under stable stratification results in a more realistic

representation of the prolonged wake of a SAR-retrieved 10-meter wind speed pattern. In our study, although we use a finer grid spacing (1km instead of 1.67km) and a reduced TKE coefficient ($a = 0.25$), which limits the influence of the added TKE term, we observe a very good agreement between the SAR image and the WF simulations during stable conditions. Furthermore, Ali et al. [2023] evaluated several WFP schemes to characterize the wake of a specific event using 10 m SAR data. Their results show that, when comparing the Fitch model with $a = 1$ and $a = 0.25$, the latter produces a longer wake downstream of the wind farms, similar to the behavior observed in case A analyzed in the present study.

In case B, the wake is weaker, but still detectable both in SAR image and in the WF simulation. In this case, stability conditions are very unstable, and the PBLH is around 700 meters, explaining the attenuated wake pattern. Case C is characterized by higher wind speeds, L varies from -300 to -450 m, and PBLH is approximately 300 m across the three locations, resulting in a more unstable-neutral stratification profile, resulting in the observed downstream wake of about 40 to 50km. Similar wake patterns but slightly shorter, about 25 to 35km have been observed during neutral stratification in the study of Platis et al. [2018]. Last, in case D, very unstable ABL conditions prevail, and the PBLH varies between 250 meters upstream at the WHi location to 358 meters downstream at the EPL area. This is the only SAR image with clear wakes for SW wind conditions. Overall, the patterns of the wakes are accurately represented by the WF simulations, highlighting the importance of using the Fitch WFP scheme in WRF. **Although SAR assumes neutral stability to compute the 10-meter wind speed, there are discrepancies with the WRF model, especially in cases B and D, where very unstable stratification dominates based on the values of Obukhov length. Uncertainty exists regarding the assumption of SAR image neutrality used to calculate wind fields, an issue that has also been reported in previous studies across different applications Hasager et al. [2011], Ahsbahs et al. [2017]. Furthermore, there is uncertainty in the estimation of L within WRF, as it strongly depends on the PBL and surface schemes applied in each configuration.**

In addition, in the revised manuscript, we provide a comparison of WRF against SAR and lidar (when available) for the four analyzed events using mean absolute error (MAE) and absolute error (AE). This has been added in the revised text in Section 2.3.4 (lines 242-245):

Last, for the SAR transects, we compute the mean absolute error, MAE $= \sum_1^n |U_{\text{SAR}} - U_{\text{WRF}}|$, where n is the number of points across the upstream, intra-farm, and downstream regions of each transect (70 points in total after interpolation of the SAR and WRF data). In addition, for the same time instances, we evaluate the point-wise absolute error, AE $= |U_{\text{lidar}} - U_{\text{WRF}}|$, using the lidar observations as reference.

The results are presented in Section 3.3. Table 8 has been added and the text has been revised accordingly in lines 554-585:

Therefore, to further evaluate the performance of the WRF model, hub height-lidar measurements are incorporated when available, depicted as gray scatter points in Figure 17. In addition to the qualitative assessment of the wind speeds in Figure 17, Table 8 presents the MAE across the upstream, intra-farm, and downstream transect regions, as well as the point-wise AE, when lidar measurements are available, and the evaluation methodology is described in section 2.3.4. In cases A, B, and C (NE wind flow), the transect follows the direction EPL (upstream) $\rightarrow$ BSB (intra-farm) $\rightarrow$ WHi (downstream), whereas case D corresponds to a SW flow event, where the sequence reverses.

Case A shows good agreement in the representation of the transect wake. Specifically, while the WF simulation tends to underestimate wind speeds upstream and within the Borssele zone, it accurately captures the wind speed deficit in the Belgian wind farms and downstream of the wind farm cluster. This is also reflected in Table 8, where only in the upstream region WF generates a slightly higher MAE = 0.49 m s^{-1}, compared to MAE=0.37 m s^{-1} in the NWF simulation. However, the improvement between the WF against the NWF simulations in MAE in the intra-farm region and downstream is 81% and 85%, respectively. Similarly, in Case B, both the intra-farm and downstream wind speed patterns are well represented by the WF simulation. Although the overall trend is accurately reproduced, the WF simulation tends to overestimate the wind speeds upstream and downstream of the wind farms. Additionally, a flow acceleration is observed as the wind exits the Belgian-Dutch cluster. This pattern is consistent with the broader flow conditions downstream of the Belgian–Dutch cluster, since a similar acceleration is evident in the SAR observations and in both the WF and NWF simulations. Both the lidar measurements and the WF simulation effectively capture the intra-farm wind speed deficit and the downstream flow acceleration. Furthermore, in Table 8, MAE and AE are better in the WF than NWF simulation across the full transect regions.

Case C also features NE incoming wind conditions. Similar to the previous cases, the WF simulation successfully captures the wind speed deficit trends both within the intra-cluster zone and in the downstream wake recovery region. Both the NWF and WF simulations show noticeable discrepancies upstream at EPL compared to lidar observations with an AE reaching above 3 m s^{-1}. This is significantly improved both at BSB and downstream at WHi, where WF simulations outperform the NWF runs both in terms of MAE across the transect regions and AE against the lidar measurements. During this event, the Obukhov length at EPL is relatively high and negative, indicating unstable to near-neutral conditions. This may be attributed to less negligible shear, which can contribute to increased deviations from the lidar measurements. In the NWF simulation, an unrealistic and nearly linearly increasing pattern in wind speed of approximately 0.5 m s^{-1} between EPL and WHi is observed. Case D is the only selected timestamp with a SW wind direction. Although the SAR data exhibit considerable noise, particularly within the intra-farm region, the WF simulation is still able to reproduce the wind speed deficit and the downstream wake-recovery trend. In this case, only EPL lidar measurements are available, which show good agreement with the WF simulation, with a slightly improved AE compared to the NWF run. Furthermore, the MAE provides a more accurate matching with the 10-meter SAR wind speeds across the three transect regions in the WF simulations.

RC: ***13. Line 441 and Conclusions: The claim that "the model is systematically validated" and that "WF simulations improve the general accuracy" is not fully supported by the presented results. Please moderate these statements.***

AR: Statements like the aforementioned one have been moderated. Conclusions have been updated overall, based on updated results and the comments of all three reviewers. For instance, lines 593-605 have been changed accordingly (see response to comment 8 as well).

Our findings demonstrate that the use of the Fitch WFP scheme, in combination with this particular WRF configuration, captures the wake effects and provides an improvement in model performance relative to a simulation without wind farms at the intra-farm region. First, focusing on the time-averaged wind speeds over the full analysis period, the WF simulations yield a substantial improvement in both bias and EMD, with reductions of approximately 82% and 73% compared with the NWF simulations, respectively, in the intra-cluster (BSB location). Furthermore, when analyzing the wind

> speed bias under different wind directions, both NWF and WF configurations reduce the negative bias inherited from the ERA5 data, with the results indicating that the Fitch scheme introduces an additional deficit downstream of the Belgian-Dutch cluster. However, within the intra-farm area, the Fitch scheme provides a more accurate wind speed distribution across all wind directions, underscoring its importance for representing strongly-waked environments. Beyond the improvements in bias and EMD, the WF simulations perform similarly to the NWF runs in terms of cRMSE and correlation, with only small and marginal differences. Similarly, wind direction metrics are also improved in the WF simulations, as bias is reduced by approximately 51% at BSB, while an analysis of wind direction biases by sector reveals that WF can perform better than ERA5 for wind directions from northeast to south.

In lines 611 to 617, there is an updated version in the conclusions:

> Extending the performance metrics to wake characteristics under different wind speed magnitudes, and focusing on SW and NE wind directions, the analysis revealed several consistent patterns. First, under SW wind direction, the WF simulations accurately capture the wake structure when wind speeds range between 5 to 8 m s^{-1}, providing an almost negligible bias at BSB and downstream at EPL. Under NE wind direction, the optimum performance of WF simulations occurs at wind speeds between 8 and 11 m s^{-1}. Although the model tends to overestimate intra-farm wind speed deficits at BSB under higher wind speed conditions, it shows comparable improvements in long-distance downwind wakes at EPL (for SW winds) and WHi (for NE winds), where smaller bias and cRMSE values are obtained.

And in lines 623 to 634:

> Furthermore, a comparison with SAR images at specific events shows that WF performs well in representing the wake structure generated by the Belgian-Dutch cluster. Both in the intra-farm and the wake recovery regions, WF provides a realistic wind speed deficit. To assess the performance of the model during these events, MAE and AE are computed. More specifically, MAE is estimated across the upstream, intra-farm, and downstream regions, while AE is calculated at each lidar location when measurements are available. Building on the previously discussed improvements in the qualitative wake-pattern representation, the error metrics further confirm that the WF simulations provide a better alignment with observations than the NWF runs across all locations, with the most pronounced improvements occurring in the intra-farm and downstream regions.
>
> This study can accurately evaluate the model performance by utilizing upwind, intra-farm, and downwind lidar measurements as well as SAR images at specific events. Systematic biases in both wind speed and direction are evident, particularly in the wind speed metrics. However, it should be highlighted that the model is driven by ERA5 hourly data, which consistently introduce biases, affecting the boundary conditions in WRF and propagating wind speed underestimations to the analyzed locations.

RC: ***14. The reported 36 % bias reduction appears inconsistent with Fig. 5a, where WF performs worse at most sites. Please clarify how this number was derived.***

AR: That was a typo: we were referring only to the BSB location. However, the entire paragraph in the conclusions has been changed. The detailed paragraph is attached as a response to the previous comment, 13, lines 593-605 in the revised manuscript.

RC: ***15. The suggestion that future work should "benchmark different parameterization schemes" is valid but not novel as this topic has been addressed many times in the literature, particularly Fitch's parameterization and the Explicit Wind-farm Parameterization (EWP).***

AR: It's a general recommendation for WRF parameterization schemes, not limited to WFP schemes but also including PBL schemes, as they play a key role in turbulence production in WRF and are important for assessing model performance. In this study, we used the default WRF version (v4.5.2), which supports only the Fitch scheme for WFP, and this can only be combined with the MYNN PBL scheme. The part below has been added to the revised text, in the conclusions, in lines 635-639:

> The scope of this study is limited to using the Fitch WFP scheme combined with the MYNN 2.5 PBL scheme. Additional multi-year simulations using different WFP schemes, like the explicit wind farm (EWP) parametrization scheme [Volker et al., 2015, García-Santiago et al., 2024] applied with different PBL schemes or the induction correction to the original Fitch scheme [Vollmer et al., 2024], could also provide insights into the year-long effects of large wind-farm clusters.

RC: ***16. Line 6: Spell out "1 km" as "one kilometer".***

AR: We have corrected this in the revised text.

RC: ***17. Line 118: "...is the number of turbines..."***

AR: This has been changed in the revised version.

References

T. Ahsbahs, M. Badger, P. Volker, K. S. Hansen, and C. B. Hasager. Applications of satellite winds for the offshore wind farm site anholt. *Wind Energy Science*, 3(2):573–588, 2018. . URL `https://wes.copernicus.org/articles/3/573/2018/`.

Tobias Ahsbahs, Merete Badger, Ioanna Karagali, and Xiaoli Guo Larsén. Validation of sentinel-1a sar coastal wind speeds against scanning lidar. *Remote Sensing*, 9(6), 2017. ISSN 2072-4292. . URL `https://www.mdpi.com/2072-4292/9/6/552`.

Karim Ali, David M Schultz, Alistair Revell, Timothy Stallard, and Pablo Ouro. Assessment of five wind-farm parameterizations in the weather research and forecasting model: a case study of wind farms in the north sea. *Monthly Weather Review*, 151(9):2333–2359, 2023. URL `https://doi.org/10.1175/MWR-D-23-0006.1`.

Cristina L Archer, Sicheng Wu, Yulong Ma, and Pedro A Jiménez. Two corrections for turbulent kinetic energy generated by wind farms in the wrf model. *Monthly Weather Review*, 148(12):4823–4835, 2020. URL `https://doi.org/10.1175/MWR-D-20-0097.1`.

R Borgers, J Meyers, and N P M van Lipzig. Energy production and inter-farm wake losses in future north sea wind farms. *Environmental Research Letters*, 20(7):074036, jun 2025. . URL `https://doi.org/10.1088/1748-9326/add8a2`.

Beatriz Cañadillas, Richard Foreman, Volker Barth, Simon Siedersleben, Astrid Lampert, Andreas Platis, Bughsin Djath, Johannes Schulz-Stellenfleth, Jens Bange, Stefan Emeis, and Thomas Neumann. Offshore wind farm wake recovery: Airborne measurements and its representation in engineering models. *Wind Energy*, 23(5):1249–1265, 2020. . URL `https://onlinelibrary.wiley.com/doi/abs/10.1002/we.2484`.

J. Fischereit, H. Vedel, X. G. Larsén, N. E. Theeuwes, G. Giebel, and E. Kaas. Modelling wind farm effects in harmonie–arome (cycle 43.2.2) – part 1: Implementation and evaluation. *Geoscientific Model Development*, 17(7):2855–2875, 2024. . URL https://gmd.copernicus.org/articles/17/2855/2024/.

O. García-Santiago, A. N. Hahmann, J. Badger, and A. Peña. Evaluation of wind farm parameterizations in the wrf model under different atmospheric stability conditions with high-resolution wake simulations. *Wind Energy Science*, 9(4):963–979, 2024. . URL https://wes.copernicus.org/articles/9/963/2024/.

Charlotte B. Hasager, Merete Badger, Alfredo Peña, Xiaoli G. Larsén, and Ferhat Bingöl. Sar-based wind resource statistics in the baltic sea. *Remote Sensing*, 3(1):117–144, 2011. ISSN 2072-4292. . URL https://www.mdpi.com/2072-4292/3/1/117.

Charlotte Bay Hasager, James Imber, Jana Fischereit, Aito Fujita, Krystallia Dimitriadou, and Merete Badger. Wind speed-up in wind farm wakes quantified from satellite sar and mesoscale modeling. *Wind Energy*, 27 (11):1369–1387, 2024. URL https://doi.org/10.1002/we.2943.

H. Hersbach, A. Stoffelen, and S. de Haan. An improved c-band scatterometer ocean geophysical model function: Cmod5. *Journal of Geophysical Research: Oceans*, 112(C3), 2007. . URL https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/2006JC003743.

X. G. Larsén and J. Fischereit. A case study of wind farm effects using two wake parameterizations in the weather research and forecasting (wrf) model (v3.7.1) in the presence of low-level jets. *Geoscientific Model Development*, 14(6):3141–3158, 2021. . URL https://gmd.copernicus.org/articles/14/3141/2021/.

B. T. E. Olsen, A. N. Hahmann, N. G. Alonso-de Linaje, M. Žagar, and M. Dörenkämper. Low-level jets in the north and baltic seas: mesoscale model sensitivity and climatology using wrf v4.2.1. *Geoscientific Model Development*, 18(14):4499–4533, 2025. . URL https://gmd.copernicus.org/articles/18/4499/2025/.

Andreas Platis, Simon K. Siedersleben, Jens Bange, Astrid Lampert, Konrad Bärfuss, Rudolf Hankers, Beatriz Cañadillas, Richard Foreman, Johannes Schulz-Stellenfleth, Bughsin Djath, Thomas Neumann, and Stefan Emeis. First in situ evidence of wakes in the far field behind offshore wind farms. *Scientific Reports*, 8(1): 2163, 2018. ISSN 2045-2322. . URL https://doi.org/10.1038/s41598-018-20389-y.

V. Pronk, N. Bodini, M. Optis, J. K. Lundquist, P. Moriarty, C. Draxl, A. Purkayastha, and E. Young. Can reanalysis products outperform mesoscale numerical weather prediction models in modeling the wind resource in simple terrain? *Wind Energy Science*, 7(2):487–504, 2022. . URL https://wes.copernicus.org/articles/7/487/2022/.

S. K. Siedersleben, A. Platis, J. K. Lundquist, B. Djath, A. Lampert, K. Bärfuss, B. Cañadillas, J. Schulz-Stellenfleth, J. Bange, T. Neumann, and S. Emeis. Turbulent kinetic energy over large offshore wind farms observed and simulated by the mesoscale model wrf (3.8.1). *Geoscientific Model Development*, 13(1): 249–268, 2020. . URL https://gmd.copernicus.org/articles/13/249/2020/.

P. J. H. Volker, J. Badger, A. N. Hahmann, and S. Ott. The explicit wake parametrisation v1.0: a wind farm parametrisation in the mesoscale model wrf. *Geoscientific Model Development*, 8(11):3715–3731, 2015. . URL https://gmd.copernicus.org/articles/8/3715/2015/.

L. Vollmer, B. A. M. Sengers, and M. Dörenkämper. Brief communication: A simple axial induction modification to the weather research and forecasting fitch wind farm parameterization. *Wind Energy Science*, 9(8):1689–1693, 2024. . URL https://wes.copernicus.org/articles/9/1689/2024/.