Review comments

February 1, 2026

1 Review 1

This study aims to provide a quantitative and qualitative comparison of the planned Prinses Elisabeth wind farm cluster across different fidelities of numerical wake models, ranging from mesoscale weather prediction models to fast-running engineering models. However, five of the six sections in the result discussion focus on the mesoscale model, with only the last section discussing the comparison. I believe the authors should make a more detailed analysis comparing the mesoscale model to the engineering model to justify their title. Additionally, the setup and design of the engineering model are not clear to me. Therefore, I don't think the comparison is valid.

We would like to thank the reviewer for the suggestions made which significantly improved the manuscript. Following these suggestions and the suggestions of the other reviewers, we decided to add substantial additional analysis comparing WRF and the engineering wake models. However, this would result in a manuscript that would be too long. As such, and at the request of the Editor and the other reviewers, the manuscript has been split into two manuscripts:

- **Annual wake impacts in and between wind farm clusters - Part I: WRF simulated wake losses for different atmospheric conditions**
- **Annual wake impacts in and between wind farm clusters - Part II: Comparison of WRF and fast-running engineering wake models**

All reviewer comments will be addressed in the appropriate manuscript (Part I or Part II). Responses to the reviewers' comments are provided in bold, and changes made to the manuscript are highlighted in blue bold italics.

Major Comments:

- Section 2.4: More details are needed to describe the setup of the engineering model:

  We agree that further information is necessary. More complete description of the engineering models has been added to Part II manuscript.

  1. How big is the domain? What is the grid resolution? (I assume it is much smaller than the WRF domain, perhaps just the BE-NL wind farm?)

     FLORIS does not have a spatial grid, the wake effects are calculated analytically and sampled at discrete points at the locations of the turbines of interest. Regarding the turbines of interest, the FLORIS simulations we indeed only use the BE-NL wind farm cluster and PE wind farm cluster, this has been clarified in the text.

     Added (Part II line 109-111):
     To reduce computational cost, only the Belgian–Dutch (BE–NL) wind farm cluster and the PE cluster are modeled for the fast-running engineering models used in Part II of the study, while all adjacent wind farms are modeled within WRF.

2. "Driving each model with the resulting homogeneous wind fields" – where does the initial wind field come from?

This has been clarified in the new manuscript.

Added/Modified (Part II line 154-202):

FLORIS simulations are driven by meteorological conditions extracted from the noWF WRF simulations for the entire year 2016 (see Part I). These conditions include wind speed, wind direction, and Turbulence Intensity (TI).

The noWF velocity field is provided to FLORIS using two different inflow setups. The first setup assumes horizontally homogeneous inflow and retrieves the wind speed from a single probe point located at the geographic center of the BE–NL cluster at an altitude of 96 m. In contrast, the heterogeneous inflow configuration accounts for spatial variability in the incoming flow, arising from coastal gradients or large-scale synoptic structures.

In the heterogeneous configuration, FLORIS' built-in implementation described by Farrel et al. (2021) is used. Specifically, the noWF velocity field at 96 m height is extracted at the WRF grid points and fed to FLORIS. For each wind turbine, FLORIS initializes the local inflow wind speed by performing a linear barycentric interpolation of the WRF wind speed field. Wake velocities downstream of each turbine are then computed following the standard FLORIS wake modeling approach, but using this locally interpolated inflow wind speed as the initial condition. No vertical interpolation is applied since both vertical wind shear and veer are omitted in the present study.

Some of the engineering wake model simulations assume a time varying TI estimated based on the noWF WRF data following Larsen (2024). Larsen (2024) proposes to use the Kaimal turbulence spectrum to estimate the wind speed component variances, σ_u, σ_v, and σ_w, at given heights and wind speeds. Empirical ratios between σ_u and other components then allow to transform the TKE extracted from WRF into the along-wind standard deviation. The ambient TI can then be approximated as

$$TI_{amb} = \frac{\sigma_u}{\overline{U}} = \frac{1}{\overline{U}}\sqrt{\frac{2\,TKE}{1+\alpha^{-1}+\beta^{-1}}} \tag{1}$$

with $\overline{U}$, the mean wind speed over the averaging period at the height of interest and α and β, empirical coefficient determined based on Kaimal turbulence spectra.

Four different FLORIS simulation setups are subsequently considered:

- ***Homogeneous FLORIS with fixed TI: A time-varying and spatially homogeneous wind field is assumed in FLORIS, constructed from the wind speed and direction at 96 m altitude at the geographic center of the BE–NL wind farm cluster, taken from the no WF WRF simulation (Part I). Turbulence intensity is fixed in space and time at 6%, a typical value for the offshore North Sea area.***
- ***Homogeneous FLORIS with varying TI: A time-varying and spatially homogeneous wind field and TI is assumed in FLORIS, constructed from the wind speed, wind direction and TI at 96 m altitude at the geographic center of the BE–NL wind farm cluster, taken from the no WF WRF simulation (Part I). This setup aims at evaluating whether incorporating temporal variability in ambient turbulence resulting from changes in different atmospheric conditions improves the representation of wake recovery.***
- ***Heterogeneous FLORIS with fixed TI: A time-varying and spatially heterogeneous wind speed field is assumed in FLORIS, constructed from the wind***

speed at 96 m altitude taken from the no WF WRF simulation (Part I). Wind direction is kept spatially homogeneous and time-varying, constructed from the no WF WRF simulation at 96 m altitude at the geographic center of the BE–NL wind farm cluster (Part I). Turbulence intensity is fixed in space and time at 6%, a typical value for the offshore North Sea area.

  - ***Heterogeneous FLORIS with varying TI:** A time-varying and spatially heterogeneous wind speed field is assumed in FLORIS, constructed from the wind speed at 96 m altitude taken from the no WF WRF simulation (Part I). Wind direction is kept spatially homogeneous and time-varying, constructed from the no WF WRF simulation at 96 m altitude at the geographic center of the BE–NL wind farm cluster (Part I). A time-varying and spatially homogeneous TI is assumed, constructed from TI at 96 m altitude at the geographic center of the BE–NL wind farm cluster, taken from the no WF WRF simulation (Part I).*

  All four setups of the engineering wake models are run in a time-series mode with a 30-min resolution for one full year (2016), where each simulation time step is treated as quasi-steady and independent. Consequently, the model does not account for wake advection or temporal effects in wake propagation contrary to WRF. As noted previously, not all wind farms in the North Sea are modeled, for the engineering models used in this study only those belonging to the BE-NL and PE wind farm cluster are included. This reduced domain of not including all North Sea wind farms allows for the simulation of twelve cases (three layouts (no WF, with PE and without PE similar to WRF) combined with four input setups described above) at a reasonable computational cost, while still capturing the dominant inter-farm wake losses of the PE wind farm cluster on the BE-NL wind farm cluster.

3. "Similar to the WRF case setup, to isolate the effects of the PE wind farm cluster on the existing zone, for every engineering wake model two simulations are performed, i.e., one with and one without the PE." Again, I am curious to know how the wind field is initialized for these simulations.

   Please see answer to the previous remark.

4. Do you run the engineering model for one year using the wind field from WRF? How do you update the incoming wind field in the engineering model using WRF data?

   Please see answer to previous remark.

- Section 3.6: I believe a more thorough analysis can be done to understand the difference in wake loss between the WRF model and the engineering model. For instance, examining the spatiotemporal differences in wind speed and TKE over the wind farm and its wake region, as well as the wake extent. Differences in the vertical profiles of these variables could provide another perspective. The authors should make a more sincere effort to justify the goal of their paper.

  We agree there is an opportunity for more thorough comparisons between the engineering and WRF models. The new manuscript (Part II) focuses specifically on the comparison between WRF and commonly used engineering wake models. Part II quantifies the spread between model approaches (WRF and engineering wake models) by total power production and internal and external wake energy losses comparisons over a meteorologically representative year. The analysis also identifies the atmospheric drivers of these differences, including wind speed and atmospheric stability, and highlights conditions under which discrepancies between modelling approaches are amplified (summer). Now Part II provides a systematic assessment of the differences between engineering wake model estimates and WRF estimates.

Minor Comment:

- The definition of internal wake loss (Equation 2) is technically incorrect. While the equation is correct, the definition is flawed due to the design of the experiment. The internal wake loss for a particular wind farm should be the difference between the simulation with and without that particular wind farm only. However, based on how the simulation is set up in this paper, all internal wake losses are contaminated by the external wake loss from other wind farms. Therefore, Table 7 is incorrect.

Thank you for pointing this out, modifications to the text have been made to make this clear.
Added (Part I line 162-169):

Internal wake energy losses, as defined here, represent the wake energy losses experienced by the existing wind farms in the absence of the PE wind farm cluster. It is important to note that this definition differs from the strict definition of internal wake energy losses, which include only wake effects generated by turbines within the same wind farm. Due to the simulation setup employed in this study, this distinction cannot be made. Consequently, the internal wake energy losses presented here include inter farm wake effects among the existing wind farms, but exclude wake effects originating from the new PE wind farm cluster. Note that these internal wake effects virtually exclusively originate from wind turbines within the dense Belgian-Dutch cluster, as other wind farms included in the without PE simulation are at least 58 km away.

2 Review 2

This paper presents a comparison of engineering wake models (available in FLORIS) to the mesoscale WRF simulator for comparing farm-to-farm wakes in the North Sea. Overall, the paper is well-written and its scope is well-defined. The results may help to explain the strengths and weaknesses of different engineering wake models, although some clarity is needed to understand the modeling procedure. I have also left some open questions that the authors may consider addressing (but these should not be treated as necessary). I will add a disclaimer here that I am a developer of the FLORIS engineering wake modeling package that is used throughout this work.

We would like to thank the reviewer for the suggestions made which significantly improved the manuscript. Following these suggestions and the suggestions of the other reviewers, we decided to add substantial additional analysis comparing WRF and the engineering wake models. However, this would result in a manuscript that would be too long. As such, and at the request of the Editor and the other reviewers, the manuscript has been split into two manuscripts:

- **Annual wake impacts in and between wind farm clusters - Part I: WRF simulated wake losses for different atmospheric conditions**
- **Annual wake impacts in and between wind farm clusters - Part II: Comparison of WRF and fast-running engineering wake models**

All reviewer comments will be addressed in the appropriate manuscript (Part I or Part II). Responses to the reviewers' comments are provided in bold, and changes made to the manuscript are highlighted in blue bold italics.

Comments on modeling procedure.

1. The validation procedure for WRF seems fairly limited. As I understood it, there are a handful of locations used for validation, but even the highest point is still less than half of expected turbine hub heights as it is only for a few locations. Moreover, the WRF simulations used for calibration contained no wind farms, since none were built yet; but this does beg the question of how useful such validation is when the rest of the study is comparing wakes. The authors do, to some extent, point out the WRF simulations are not to be treated as "truth"; however, I think the limitations of validation process should be articulated more clearly.

   We agree that WRF validation was limited and that the limitations of this validation procedure should be articulated more clearly. We have therefore extended the validation by including offshore LiDAR observations providing vertical wind profiles up to hub-height, allowing evaluation of the simulated wind field at heights relevant for wake impacts. Due to the mismatch between the meteorological year and the wind farm operation (PE especially not build yet), a true validation of wake energy losses is not possible and instead, the validation focuses on establishing confidence in the background wind field and atmospheric variability (in between simulations). Accordingly, WRF is not treated as ground truth but as a physically consistent reference for assessing the spread and sensitivity of engineering wake model predictions.

   - **Added LiDAR in Table 2**
   - **Added (Part I lines 203-211):**
     We first seek to establish confidence in the WRF model representation of the wind field. However, we emphasize that the WRF model is not ground truth. The simulations in this study represent idealized scenarios, as only a few of the Belgian-Dutch wind farms (Thorntonbank I, II and III, Belwind, and Northwind, 182 of 572 turbines) were operational in 2016, while the existing turbines used in the simulations correspond to the wind farm layout as of 2022. The year 2016 was nevertheless selected because it was identified as the most meteorologically representative. As a consequence of this mismatch between the

meteorological year and the wind farm layout, a true validation of wake energy losses is not possible. As such, validation will only become feasible once post-construction field observations for the fully developed wind farm area become available (Palatos-Plexidas et al., 2025). Future research may then also explore the optimal representation of the wind field by selecting the most appropriate model configuration and parameterizations.

- **Added (Part I lines 240-252):**
  The LiDAR at LG was designed specifically for assessing offshore wind resources. It provides vertical wind profiles up to 215 meters, enabling model evaluation at heights comparable to the level of turbine hub heights. Comparing the different WRF simulations with the LiDAR observations (Table 5) reveals a small mean bias across all altitudes and different simulation. The bias for wind speed ranges from 0.04 ms^{-1} at 62 m to −0.10 ms^{-1} at 215 m in the no WRF simulation, from -0.08/-0.07 ms^{-1} at 62 m to -0.22/-0.23 ms^{-1} at 215 m in the without PE simulation and with PE simulation, respectivily. Wind direction biases gradually decrease with height, ranging from values around 5° at 62 m to around 1.5° at 215 m. This suggests improved directional agreement at higher altitudes. The mean wind speed and direction bias of different simulations often fall within the observational uncertainty, accounting for model variability and LiDAR measurement error, especially between without PE simulation and with PE simulation. Larger differences are observed when comparing to the no WF simulation, this could be caused by wakes. To conclude, the background wind field can be assumed to be similar across simulations, and justify studying the wake effects in the suggested simulations. Overall, based on all observation comparisons, a good agreement between modeled and observed values at hub height supports using WRF to evaluate wake impacts of the newly built PE on existing wind farms.
- **Added table (Part I Table 5) here table 1:**

Table 1: Mean wind speed [ms^{-1}] and wind direction [°] bias (simulation minus observations) and the 95% confidence interval, which accounts for both model variability and when available measurement uncertainty, for the no WF WRF, without PE, and with PE simulations at different altitudes of the LG LiDAR location (Table 1.

	Wind speed bias [m s^{-1}]			**Wind direction bias [°]**		
Altitude [m]	**no WF**	**without PE**	**with PE**	**no WF**	**without PE**	**with PE**
62	0.04 ± 0.04	-0.08 ± 0.04	-0.07 ± 0.04	5.77 ± 0.32	5.49 ± 0.33	5.10 ± 0.32
90	0.03 ± 0.04	-0.09 ± 0.04	-0.08 ± 0.04	5.06 ± 0.32	4.81 ± 0.33	4.39 ± 0.32
115	0.01 ± 0.04	-0.12 ± 0.04	-0.11 ± 0.04	4.29 ± 0.32	4.09 ± 0.33	3.78 ± 0.32
140	-0.01 ± 0.04	-0.14 ± 0.04	-0.13 ± 0.04	3.43 ± 0.32	3.45 ± 0.33	3.10 ± 0.32
165	-0.04 ± 0.04	-0.17 ± 0.04	-0.16 ± 0.04	2.63 ± 0.32	2.63 ± 0.33	2.52 ± 0.32
190	-0.07 ± 0.04	-0.19 ± 0.04	-0.19 ± 0.05	1.98 ± 0.32	1.90 ± 0.33	1.90 ± 0.32
215	-0.10 ± 0.05	-0.22 ± 0.05	-0.23 ± 0.05	1.53 ± 0.33	1.60 ± 0.34	1.50 ± 0.33

2. Similarly, I feel the statement that WRF should not be treated as "truth" should appear more strongly earlier in the paper. Although the authors do state that WRF is not the truth, my impression is that this message often gets lost in comparisons (in this work and others), which may mean we select parametrizations for engineering models based on faulty criteria.

   We agree with the reviewer's concern. We have revised both Part I and Part II to state more explicitly, and earlier in the papers, that WRF should not be treated as ground truth. Furthermore, Part II now repeatedly highlights that the focus of the analysis is on understanding the spread among modeling approaches, not on selecting parameterizations based on comparisons with WRF.

   - **Added explicitly (Part I lines 203-211):**
     We first seek to establish confidence in the WRF model representation of the

wind field. However, we emphasize that the WRF model is not ground truth.

- **Added already in the introduction (Part II lines 78-84):**
  Despite these advances, it remains unclear which modeling approach (numerical weather prediction model vs fast-running engineering wake models) provides more reliable long term estimates of wake impacts for large offshore wind farm clusters. Since no validation data are available within the scope of this study to establish a ground truth, the objective is not to identify a single best-performing model. Instead, comparing wake predictions from different model approaches allows the uncertainty associated with farm–farm wake assessments to be quantified. Such a comparison makes it possible to assess the spread in integrated metrics, such as annual energy production and wake-induced losses, and to identify the atmospheric factors, such as wind speed, wind direction and atmospheric stability that drive differences between model predictions.
- **Added (Part II lines 267-273):**
  Based on these findings, TurbOPark with $A = 0.04$ *is used as the representative engineering wake model in subsequent analyses when a single wake model is considered. This choice does not imply that TurbOPark provides the most accurate representation of wake effects, given that no observational validation data are available and WRF cannot be regarded as a ground truth.*
- **Added (Part II lines 435-437):**
  This study investigates how wake losses of offshore wind farm cluster depend on the modelling approach, motivated by the lack of systematic, year-long comparisons between numerical weather prediction models and fast running engineering wake models.

3. Finally, along this vane, the statement that "advanced" features of FLORIS (heterogeneity, shear(?), veer) appears only in the conclusions (lines 428–430), and I think should have been stated in section 2.4. There, homogeneity is mentioned, but not shear and veer, and since some level of shear is commonly the default configuration for FLORIS (and appears in the example inputs of FLORIS v3.4), this should be stated clearly.

   Thank you for pointing this out, the description of how shear and veer are handled is now present in the methods explaining the engineering wake model setups (section 2.2.2).

   Added (Part II lines 163-164):
   No vertical interpolation is applied since both vertical wind shear and veer are omitted in the present study.

4. The authors ran FLORIS for a grid of wind speeds and wind directions to generate a look-up table of results, which were then mapped to the year-long case. This procedure is reasonable, but I did not quite understand what wind conditions are used to sample the look-up table (line 182). The authors say "the time series extracted from the no WF WRF run," but I assume the wind conditions varied across the domain—how did the authors chose a scalar wind speed and direction to sample the gridded FLORIS results? Were these selected from a single point in the simulation domain, or some sort of average? Lines 381 and 382 suggest (if I'm reading them correctly) that background flow heterogeneity may be partially the cause of differences, so the choice of representative background flow condition is important here.

   The methodology for extracting input to the engineering models from WRF was not entirely clear in the previous manuscript. In the revised version, this methodology is described in greater detail. Furthermore, since the manuscript has been split into two parts, we have also introduced additional setups for the engineering wake models.

   Added/Modified (Part II lines 154-202):
   FLORIS simulations are driven by meteorological conditions extracted from the

noWF WRF simulations for the entire year 2016 (see Part I). These conditions include wind speed, wind direction, and Turbulence Intensity (TI).

The noWF velocity field is provided to FLORIS using two different inflow setups. The first setup assumes horizontally homogeneous inflow and retrieves the wind speed from a single probe point located at the geographic center of the BE–NL cluster at an altitude of 96 m. In contrast, the heterogeneous inflow configuration accounts for spatial variability in the incoming flow, arising from coastal gradients or large-scale synoptic structures.

In the heterogeneous configuration, FLORIS' built-in implementation described by Farrel et al. (2021) is used. Specifically, the noWF velocity field at 96 m height is extracted at the WRF grid points and fed to FLORIS. For each wind turbine, FLORIS initializes the local inflow wind speed by performing a linear barycentric interpolation of the WRF wind speed field. Wake velocities downstream of each turbine are then computed following the standard FLORIS wake modeling approach, but using this locally interpolated inflow wind speed as the initial condition. No vertical interpolation is applied since both vertical wind shear and veer are omitted in the present study.

Some of the engineering wake model simulations assume a time varying TI estimated based on the noWF WRF data following Larsen (2024). Larsen (2024) proposes to use the Kaimal turbulence spectrum to estimate the wind speed component variances, σ_u, σ_v, and σ_w, at given heights and wind speeds. Empirical ratios between σ_u and other components then allow to transform the TKE extracted from WRF into the along-wind standard deviation. The ambient TI can then be approximated as

$$TI_{amb} = \frac{\sigma_u}{\overline{U}} = \frac{1}{\overline{U}}\sqrt{\frac{2\ TKE}{1+\alpha^{-1}+\beta^{-1}}} \tag{2}$$

with $\overline{U}$, the mean wind speed over the averaging period at the height of interest and α and β, empirical coefficient determined based on Kaimal turbulence spectra.

Four different FLORIS simulation setups are subsequently considered:

- *Homogeneous FLORIS with fixed TI: A time-varying and spatially homogeneous wind field is assumed in FLORIS, constructed from the wind speed and direction at 96 m altitude at the geographic center of the BE–NL wind farm cluster, taken from the no WF WRF simulation (Part I). Turbulence intensity is fixed in space and time at 6%, a typical value for the offshore North Sea area.*
- *Homogeneous FLORIS with varying TI: A time-varying and spatially homogeneous wind field and TI is assumed in FLORIS, constructed from the wind speed, wind direction and TI at 96 m altitude at the geographic center of the BE–NL wind farm cluster, taken from the no WF WRF simulation (Part I). This setup aims at evaluating whether incorporating temporal variability in ambient turbulence resulting from changes in different atmospheric conditions improves the representation of wake recovery.*
- *Heterogeneous FLORIS with fixed TI: A time-varying and spatially heterogeneous wind speed field is assumed in FLORIS, constructed from the wind speed at 96 m altitude taken from the no WF WRF simulation (Part I). Wind direction is kept spatially homogeneous and time-varying, constructed from the no WF WRF simulation at 96 m altitude at the geographic center of the BE–NL wind farm cluster (Part I). Turbulence intensity is fixed in space and time at 6%, a typical value for the offshore North Sea area.*
- *Heterogeneous FLORIS with varying TI: A time-varying and spatially heterogeneous wind speed field is assumed in FLORIS, constructed from the wind speed*

at 96 m altitude taken from the no WF WRF simulation (Part I). Wind direction is kept spatially homogeneous and time-varying, constructed from the no WF WRF simulation at 96 m altitude at the geographic center of the BE–NL wind farm cluster (Part I). A time-varying and spatially homogeneous TI is assumed, constructed from TI at 96 m altitude at the geographic center of the BE–NL wind farm cluster, taken from the no WF WRF simulation (Part I).

All four setups of the engineering wake models are run in a time-series mode with a 30-min resolution for one full year (2016), where each simulation time step is treated as quasi-steady and independent. Consequently, the model does not account for wake advection or temporal effects in wake propagation contrary to WRF. As noted previously, not all wind farms in the North Sea are modeled, for the engineering models used in this study only those belonging to the BE-NL and PE wind farm cluster are included. This reduced domain of not including all North Sea wind farms allows for the simulation of twelve cases (three layouts (no WF, with PE and without PE similar to WRF) combined with four input setups described above) at a reasonable computational cost, while still capturing the dominant inter-farm wake losses of the PE wind farm cluster on the BE-NL wind farm cluster.

5. Fig. 10. (c) and (d) seems to imply that the FLORIS results depend on the stability class. However, the authors state that turbulence intensity is fixed at 6% throughout for FLORIS simulations. Is there some other dependency on stability that is used, or is the variation in the FLORIS result due to other correlated effects (different wind speed profiles, for example)?

We agree that this comparison was not sufficiently clear in the original manuscript. In Part II, we now explicitly describe the different setups of the engineering wake models, including the use of turbulence intensity fields obtained from WRF (see also our response to the previous comment). To address this issue directly, we have redone the analysis, and the updated results are presented in Fig. 7 of Part II. These results show that wake models which do not explicitly account for turbulence intensity are invariant with respect to stability class, whereas models such as Cumulative Curl and TurbOPark exhibit a clear dependence on atmospheric stability especially for external wake losses.

Added (Part II lines 412-442 and Fig. 7) here Fig. 1:
Building on the seasonal analysis from section 3.3.4, an explicit investigation on how the internal and external wake energy losses from the different modeling approaches behave across atmospheric stability classes will be discussed next (Fig. 7). How these energy losses depend on atmospheric stability classes for WRF only have been discussed in section 3.3 and 3.4 in Part I. In general Fig. 7 shows that, the differences between the modeling approached is primarily controlled by the wake modeling formulation rather than by the inflow conditions.
For the internal wake energy losses with heterogeneous varying TI inflow conditions (Fig. 7 (a)), the uncertainty ranges associated with the different engineering wake models show partial overlap with the WRF modeling approach, especially during unstable conditions. However, differences between modeling approaches remain during stable conditions. This confirms that there is an increased sensitivity of wake recovery to stability under low turbulence regimes. Under higher turbulence conditions, turbulent mixing accelerates wake recovery, leading to closer agreement between the modeling approaches. The same can be observed from Fig. 7 (b) which isolates the effect of inflow conditions for the TurbOPark model ($A = 0.04$). In both Fig. 7 (a) and Fig. 7 (b), the uncertainty ranges of the engineering wake models and of the different inflow conditions, respectively, largely overlap, making it difficult to identify a single modeling approach or inflow configuration that is consistently closer to WRF for internal wake losses.
For the external wake energy losses when comparing the different modeling approaches under heterogeneous varying TI inflow conditions (Fig. 7 (c)), overlap

between uncertainty ranges is observed across multiple stability classes and not only under unstable conditions which was the case for internal wake energy losses. However, in contrast to the internal wake energy losses, the engineering wake models themselves no longer overlap consistently over all stability classes, indicating a stronger sensitivity of external wake losses to the underlying wake formulation. In particular, models that explicitly account for turbulence dependent wake recovery and expansion, such as TurbOPark and the Cumulative Curl model, show closer agreement with WRF across stability classes.
When isolating the effect of inflow conditions for the TurbOPark model (Fig. 7(d)), the uncertainty ranges associated with the different inflow configurations overlap with each other, similar as for the internal wake energy losses (Fig. 7(b)). Nevertheless, a slightly closer agreement with WRF is found for varying TI inflow conditions compared to fixed TI inflow conditions for stable atmospheric conditions.

Overall, this analysis provides a physical explanation for the seasonal patterns discussed in Section 3.3.3, where larger differences between modeling approaches were observed during summer. The increased occurrence of stable atmospheric conditions in summer enhances the sensitivity of wake recovery to the specific wake formulation, particularly for external wakes. As a result, discrepancies between WRF and engineering wake models, and among the engineering models themselves, become more pronounced under these conditions, whereas during periods dominated by unstable stratification the enhanced turbulent mixing leads to closer agreement across modeling approaches.

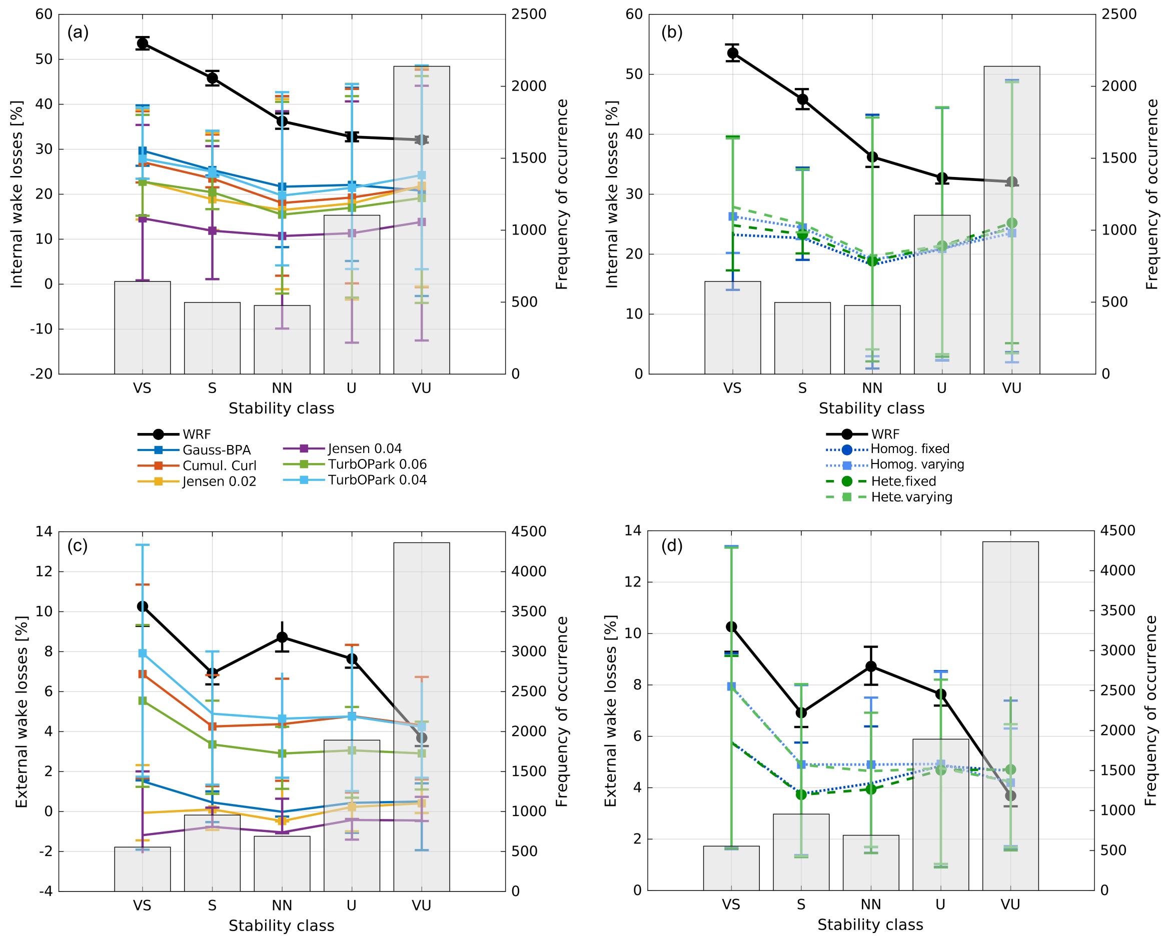

Figure 1: Spatiotemporal conditional averages of wake energy losses by atmospheric stability class (VS: very stable, S: stable, NN: near-neutral, U: unstable, VU: very unstable). (a,b) Internal wake energy losses for wind speeds between 10 and 15 ms^{-1}. (c,d) External wake losses for wind directions from 193 - 317. Lines show engineering wake models for heterogeneous inflow and varying TI and WRF; (a,c) all models, (b,d) TurbOPark ($A = 0.04$) for all inflow setups and WRF. Bars indicate the frequency of occurrence of each stability class. The error bars represent a 95% confidence interval obtained via bootstrapping with 1000 resamples.

Open questions (again, these do not need to be directly addressed in the article, but came to mind as I was reading the text and could perhaps be used to clarify conclusions)

1. To what extent do the authors feel that retuning certain parameters (e.g. the A of the TurboPark model) could have corrected differences between WRF and FLORIS, not only for the single AEP prediction but also for the time-step to time-step comparison? Is it reasonable to assume that engineering wake models will need to be tuned for specific scenarios? This is touched on briefly in lines 93–96 but not really returned to.

   Re-calibrating parameters (e.g., the wake recovery parameter in TurboPark) could potentially reduce the mismatch between WRF and the engineering wake models, both for integrated metrics such as AEP and for time step to time step comparisons. Wake recovery is strongly influenced by site characteristics (e.g., coastal effects and upstream conditions) and atmospheric stability, and regime-dependent tuning (e.g., by wind direction or stability class) could therefore improve agreement. In engineering wake models, a single expansion or recovery parameter aggregates multiple physical processes, making it unlikely that one parameter setting can consistently

capture all conditions. Moreover, calibration is inherently site-specific, which limits its applicability for wind farms that do not exist (e.g., PE) where no operational data are available.

Our goal in comparing WRF and engineering wake models is to elucidate the differences for 'out-of-the-box' and widely-used setups in the models. Future work should investigate the influence of parameter uncertainty, both in WRF and engineering models.

2. The mechanism for further downstream farms might be experiencing "negative" wake losses wasn't clearly explained to me. I understand that the authors are not attempting to investigate this phenomenon closely; however, I didn't really understand the candidate explanation in lines 327–329. I have thought about this (even ignoring higher turbulence) as being explained by a change (decrease) in thrust at the front row turbines (due to the presence of an upstream farm causing lower velocities at the front row turbines), which causes a shallower wake and therefore (in certain cases) more power at the downstream turbines. Is that equivalent to what you are saying?

   In the revised manuscript we calculated the energy losses based on power averaged and not on efficiency averages. Additionally we added uncertainty intervals for the AEP losses and the energy losses. For the AEP the negative losses can probably be attributed to numerical uncertainties, for the external wake losses the negative values fluctuate around 0 and thus is not a significant result. The differences between the two methodologies AEP and wake losses is also explained (see below). The negative wake losses for the BE-NL wind farm cluster for some wind directions is probably linked to lateral speedups. In Part II of the paper the negative losses are linked to the heterogeneous setups for engineering wake models not including explicitly variations in TI. But as mentioned by the reviewer, this would need further research that is not in the scope of these manuscripts.

   - **Clarified (Part I lines 273-278):**
     It should also be noted that very small increases (0.11-0.15%) in AEP are seen for three wind farms after the construction of the PE wind farm. Given the low magnitude and the fact that these occur at only three locations, it is difficult to draw any strict conclusions regarding actual increases in AEP. These increases in AEP may be attributable to numerical variability, possibly related to coastal effects or model variability in WRF, and further research would be needed to determine whether they have physical (flow speedup) or economic significance.
   - **Added (Part I lines 312-319):**
     Comparing the external wake energy losses (Table 7) with the AEP losses (Table 6) shows that both tables exhibit similar spatial trends. The BE–NL wind farm cluster, Galloper & Greater Gabbard, London Array, Thanet, and Luchterduinen show the largest impacts, while more distant farms show much smaller effects. Differences between the two tables arise because the metrics are defined using different normalizations. The AEP losses are computed as relative changes between the without PE and with PE simulations and normalized by AEP without PE, whereas the external wake losses are defined as ratios of temporally averaged power normalized by the noWF simulation (Eqs. 4–6). This difference in normalization, together with the implicit weighting of high-production periods in the AEP calculation, explains why some farms show statistically significant AEP impacts while their external wake losses are not statistically significant.
   - **Clarified (Part I lines 363-379):**
     Interestingly, negative external wake energy losses (-1%), i.e., positive interference from the construction of the new PE wind farm cluster, are observed for winds coming from the north to northeast (0-60) or south (180-210) (Fig. 4 (b)). The negative external wake energy losses for wind directions from the south can be explained by a speedup at the lateral locations of the PE wind

farm clusters (Fig. 8). It appears that the PE wind farm cluster act as a windbreak, causing global blockage and accelerating the flow around sides of the wind farm. A comparable lateral acceleration caused by global blockage is reported for certain North Sea wind farm clusters (Borgers et al., 2025). While similar effects have been studied for windbreaks, where they mainly cause vertical speed-ups (Liu and Stevens, 2021), lateral flow accelerations influencing farm-to-farm wakes have only recently been examined in more detail (Hasagar et al., 2015; Nygaard and Hansen, 2016; Borgers et al., 2025). Nygaard and Hansen (2016) observed this farm effect based on measured wind turbine data from two offshore wind farms. This speedup effect in the lateral direction has been observed multiple times for individual wind turbines (Ammara et al., 2022; Araya et al., 2014; Puccioni et al., 2023). In the coming years, it would be beneficial to enhance our understanding of this acceleration phenomenon around wind farms, through idealized simulations, however, this is outside the scope of this work. Understanding these speedups is important, as they may have a significant impact on external wake energy losses, necessitating consideration in the planning of high-density offshore regions. Why negative external wake energy losses are observed for wind directions from the north to northeast is less clear. As mentioned previously for the AEP losses, these may result from numerical variability or other uncertainties in the simulation. The underlying mechanisms of these minor effects, which might be physical or not, should be investigated in the future.

- **Added figure (Part I Fig 8), here Fig 2.:**

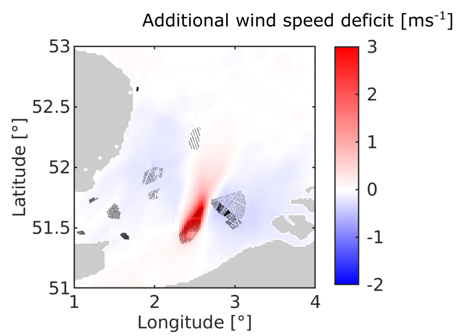

Figure 2: External BE-NL wind farm wake deficit for wind directions between 180 and 210, the sector showing negative external wake losses. This map shows the additional wind speed deficit after the construction of the PE wind farm for timesteps with negative external wake energy losses over the farm.

3 Review 3

This article provides a comparison of WRF vs engineering wake models for comparing internal and external wake losses in the North Sea, focusing in particular on the impact from a new cluster, Princess Elisabeth. As other reviewers have provided feedback on the engineering model comparison, I focus here only on the WRF analysis. Overall, it is an interesting study and many insights are provided. However, I am concerned about some of the results which I find difficult to understand. Most importantly, the large amount of negative wake losses presented throughout the plots are difficult to understand. Although flow acceleration around wind farms is indeed a phenomenon, the magnitude and breadth of these negative wake losses is concerning. I think there may be an issue in the calculation that should be investigated. I also feel the paper could use a tighter focus. There is a significant section devoted to WRF performance at different sites and atmospheric conditions, and another section devoted to comparison against engineering models, but only at the Belgian-Dutch cluster. In my opinion, this a natural split and could motivate two papers, each of which targeted at answering a specific problem. I would recommend revisiting the analysis and consider splitting the paper before resubmission.

We would like to thank the reviewer for the suggestions made which significantly improved the manuscript. Following these suggestions and the suggestions of the other reviewers, we decided to add substantial additional analysis comparing WRF and the engineering wake models. However, this would result in a manuscript that would be too long. As such, and at the request of the Editor and the other reviewers, the manuscript has been split into two manuscripts:

- **Annual wake impacts in and between wind farm clusters - Part I: WRF simulated wake losses for different atmospheric conditions**
- **Annual wake impacts in and between wind farm clusters - Part II: Comparison of WRF and fast-running engineering wake models**

All reviewer comments will be addressed in the appropriate manuscript (Part I or Part II). Responses to the reviewers' comments are provided in bold, and changes made to the manuscript are highlighted in blue bold italics.

Regarding the calculation of wake losses, the original manuscript averaged wind farm efficiency. Following the reviewer's suggestion, we revised it to first average the power output before normalization. In all the results the 95% confidence interval has been added. Combining these two we feel we are more confident about the results. Some of the negative losses do now overlap with 0, which suggests that there is no significant negative loss. However, in some cases we believe that a speedup around the wind farm is causing the negative losses. This has been carefully addressed in the revised manuscript. We also suggest that this need further investigation in the future, but this is outside the scope of this manuscript. This new method is now more clearly explained in the manuscript. Modified (Part I lines 176-204):

The absolute internal wake energy loss is defined in Eq. (3) as the power reduction of the reference power (no WF) associated with the existing wind farms only (without PE). The reference power $\overline{P}_{\mathrm{noWF}}$ is obtained by sampling the hub-height wind speed from the no WF WRF simulation at each turbine location and converting it to power using the turbine power curves. The total (internal + external) absolute wake energy loss is defined in Eq. (4) as the difference between the reference power and the mean power obtained in the presence of both the existing wind farms and the PE cluster (with PE). The absolute external wake energy loss induced by the PE wind farm cluster is defined in Eq. (5) as the difference between the total and internal losses.

$$\Delta\overline{P}_{\mathrm{internal}} = \overline{P}_{\mathrm{noWF}} - \overline{P}_{\mathrm{withoutPE}}, \tag{3}$$

$$\Delta\overline{P}_{\mathrm{internal+external}} = \overline{P}_{\mathrm{noWF}} - \overline{P}_{\mathrm{withPE}}. \tag{4}$$

$$\Delta\overline{P}_{\text{external}} = \Delta\overline{P}_{\text{internal+external}} - \Delta\overline{P}_{\text{internal}} = \overline{P}_{\text{withoutPE}} - \overline{P}_{\text{withPE}}. \quad (5)$$

From the absolute wake energy losses defined in Eqs. (3)–(5), the corresponding percentage wake energy losses are obtained by normalization with the reference power $\overline{P}_{\text{noWF}}$. The percentage internal wake energy loss is defined in Eq. (6) as the relative power reduction associated with the existing wind farms only (without PE), normalized by the reference power. The total (internal + external) percentage wake energy loss is defined in Eq.(7) as the relative power reduction in the presence of both the existing wind farms and the PE cluster (with PE), normalized by the reference power. The percentage external wake energy loss induced by the PE wind farm cluster is defined in Eq. (8) as the difference between the total and internal percentage losses.

$$\text{Loss}_{\text{internal}} = 100 \times \frac{\Delta\overline{P}_{\text{internal}}}{\overline{P}_{\text{noWF}}}, \quad (6)$$

$$\text{Loss}_{\text{internal+external}} = 100 \times \frac{\Delta\overline{P}_{\text{internal+external}}}{\overline{P}_{\text{noWF}}}, \quad (7)$$

$$\text{Loss}_{\text{external}} = 100 \times \frac{\Delta\overline{P}_{\text{external}}}{\overline{P}_{\text{noWF}}}, \quad (8)$$

All (absolute or percentage) wake energy losses are computed from the averaged power values defined above, rather than from instantaneous (absolute or percentage) losses at individual time steps. As described earlier, these losses (internal, external, and total), will be analyzed using temporal, spatiotemporal, and spatiotemporal conditional averages to further elucidate the drivers of the interactions between turbines and farms. Throughout this manuscript, the term wake energy losses refers to the percentage definitions given in Eqs. (6)–(8), unless explicitly stated as absolute wake energy losses, in which case the definitions of Eqs.(3)–(5) are used.

In addition to wake energy loss metrics, annual energy production (AEP) is computed by integrating the power of the turbines power over the full simulation period. AEP is calculated separately for the without PE and with PE simulations and AEP losses are reported as a relative change between these two cases. In contrast to the wake energly loss definitions above, AEP changes are normalized by the baseline production of the existing wind farms, without PE, rather than by the noWF reference.

Comments:

- Line 21: small point, but it is convention to express ranges from lowest to smallest number

  Not there in new manuscripts

- Line 23: Unclear what "qualitatively" refers to here. Aren't spatial distributions of wakes or waked wind speeds still fundamentally quantitative? Perhaps a better term could be used here.

  True, changed throughout the whole manuscript into spatiotemporal average and spatial patterns (latter is now in Part II).

- Line 45: "limiting their application to specific atmospheric conditions". Not true anymore. Companies like Whiffle are running GPU-powered LES models that can run full year simulations at costs often deemed reasonable by stakeholders. Today, computational limits are less prohibitive.

  We agree that this is becoming possible these days. However it is still not common practice.

  Modified to (Part II lines 51-54):

  Historically, numerical weather prediction models have not been considered the models of choice for evaluating yearly power production or wake loss estimates from wind farms due to their relatively high computational cost. However, recent advancements in computational resources have allowed to apply these models to longer time periods (Akhtar et al., 2021; Borgers et al., 2024).

- Line 46: Why can engineering wake models account for more atmospheric conditions but high fidelity microscale models can't? I feel like the opposite is more true. Engineering models have tuning often ignore atmospheric conditions completely, apart from maybe turbulence intensity. I think some clarity on what high fidelity models can do vs engineering models would be useful here.

  Clarified (Part II lines 41-45):

  On the other hand, low-fidelity, fast-running engineering wake models provide a cost-effective alternative for wake studies that can account for multiple atmospheric conditions (Nygaard et al., 2020; Nygaard and Hnsen, 2016; Munters et al., 2022). Yet, they highly simplify the underlying physics and often rely on empirical tuning, limiting their ability to fully capture the actual atmospheric conditions.

- Line 51: Range is more like 500m-3km. Not sure who is confidently using data at 5km resolution. Also, the size of the turbine isn't the relevant comparison here; it's the spacing of the turbines, distance between wind farms, etc. that matters

  Thank you for being more precise,

  Modified (Part I lines 37-40):

  However, it is important to note that typical horizontal resolutions in mesoscale models for wind energy research mostly range from 1 to 3 km, which is larger than the spacing between the wind turbines (Fischereit et al., 2022a). However, horizontal resolutions of 333 m and 5 km are also possible (Prosper et al., 2019).

  Modified (Part II lines 48-51):

  For wind energy purposes, these models have typically spatial horizontal resolutions between 1 and 3 km, which is larger than the typical inter-turbine spacings (Fischereit et l., 2022a). In some occasions, this spatial horizontal resolution can be as small as 333 m or as large as 5 km (Prosper et al., 2019). Wind farm parameterizations are used within these models to represent the effects of wind turbines that occur at the subgrid scales (Veers et al., 2019).

- Line 53: Some brief description of how wind farms are parameterized in mesoscale models would be useful here, with emphasis on the default Fitch parameterization.

  Agree, added/Modified (Part I lines 42-48):

  Here, wind turbines are typically parameterized as momentum sink and turbulence kinetic energy source within the planetary boundary layer. A widely used scheme is the Fitch parameterization (Fitch et al., 2012), which represents the extracted momentum by a single wind turbine based on the turbine thrust curve and the added turbulent kinetic energy is proportional to the difference between the thrust and power coefficients. This allows the model to account for both the wake velocity deficit and the enhanced mixing caused by turbine-induced turbulence. Other wind farm parameterizations focus on further elaborating on this parameterization by tuning model coefficients or incorporating additional physical effects such as wind shear, turbine layout, etc. (Redfern et al., 2019; Pan and Archer, 2018).

- Line 54: Have different wind farm parameterizations been used, or different parameters to the WRF model? Important distinction here.

  We agree, we hope the following modification makes it more clear that the different studies not always used the same parameterization.

  Modified (Part I lines 50-53):

  These investigations, employing different wind farm parameterizations, have included studying the effects of wind veer and shear using the rotor equivalent wind speed model (Redfern et al., 2019), wind-waves and swell (Porchetta et al., 2021), wind farm layout (Pan and Archer, 2018), and model resolution (Pryor et al., 2020), among others on the power production and wakes of wind turbines.

- Line 58: Not sure I agree here. Lately, running 1 year + simulations is the norm and is certainly sufficient to compute annual impacts, especially if a typical meteorological year is chosen. Again, cost is becoming less and less of an issue, and it could be useful to make that point, i.e., we used to run shorter periods due to computational limits, but are now getting longer.

  Indeed we agree, we have shifted the focus of the sentence.

  Modified (Part I lines 54-57):

  Until recently, mesoscale wake studies have predominantly focused on limited time periods that are insufficient to directly assess annual mean wake effects or sub-annual wake variability. These shorter time periods were often chosen because of the high computational cost associated with long-term simulations. However, with increasing computational resources, it is becoming more common to perform simulations over full years or longer. For example, earlier studies approached seasonal...

  Modified to (Part II lines 51-54):

  Historically, numerical weather prediction models have not been considered the models of choice for evaluating yearly power production or wake loss estimates from wind farms due to their relatively high computational cost. However, recent advancements in computational resources have allowed to apply these models to longer time periods (Akhtar et al., 2021; Borgers et al., 2024).

- Line 75: I think a better distinction between mesoscale and engineering models is needed. Fundamentally, mesoscale models make no assumptions about wake behavior. How wakes propagate, interact, dissipate etc. is determined entirely by prevailing atmospheric conditions simulated in the model. Engineering models, on the other hand, fundamentally make explicit assumptions about wake behaviour. It's very important to make this clear.

  Thank you, this has been clarified

  Modified (Part II lines 54-71):

  In numerical weather prediction models, wakes are not explicitly prescribed. Instead, wake dynamics are determined by the atmospheric flow field, where the mean flow is explicitly resolved at the grid scale for the mesoscale model and subgrid-scale turbulence is represented through parameterizations. In these models there are no explicit assumptions about wake propagation, recovery or superposition but the resolution remains insufficient to resolve intra-farm wake interactions down to the turbine scale (< 1 km).
  Fast-running engineering models, on the other hand, have been widely used to estimate the yearly production and wake losses of wind farms, due to their very low computational cost (Pena et al., 2018). The computational cost of most engineering wake models scales linearly with the number of turbines and is therefore generally not a limiting factor when modeling large wind farm clusters. Instead, the main limitation of engineering wake models lies in the assumptions and simplifications used and how these assumptions remain valid for large wind farms and farm–farm wakes. Engineering wake models can be categorized into bottom-up, top-down, and coupled approaches. Bottom-up models simulate wakes at turbine scale and combine individual wakes to represent farm-scale effects (Bastankhah and Porte-Agel, 2014; Katic et al., 1986). These models perform well for intra-farm wakes but often predict a faster recovery of wakes than observed for large wind farms, leading to an underestimation of farm–farm effects (Fischereit et al., 2022b). Top-down models describe the impact of a wind farm on the atmospheric boundary layer as a whole and are better suited to capture the total momentum deficit induced by large wind farms, but generally lack flow heterogeneity dependence (Emeies, 2010, Antonini and Caldeira, 2021). Coupled approaches aim to combine turbine scale wake formulations with boundary layer information, offering improved physical consistency at the cost of increased complexity and calibration requirements (Stevens et al., 2015; Nygaard et al., 2020).

- Line 99: I'd be specific about language. Mesoscale models are not wake models. They parameterize wind farms directly in the simulations, and we can deduce wakes from those simulations. They aren't wake models though, unlike the engineering models.

  We agree that the language was not used consistently in the original submission. The goal of Part II has been articulated more clearly.

  Added/Modified (Part II lines 81–102):

  Despite these advances, it remains unclear which modeling approach (numerical weather prediction model vs fast-running engineering wake models) provides more reliable long term estimates of wake impacts for large offshore wind farm clusters. Since no validation data are available within the scope of this study to establish a ground truth, the objective is not to identify a single best-performing model. Instead, comparing wake predictions from different model approaches allows the uncertainty associated with farm–farm wake assessments to be quantified. Such a comparison makes it possible to assess the spread in integrated metrics, such as annual energy production and wake-induced losses, and to identify the atmospheric factors, such as wind speed, wind direction and atmospheric stability that drive differences between model predictions.

  To address this gap, this study examines the annual mean and sub-annual variability of wake effects induced by a planned wind farm cluster on surrounding operational wind farms. Wake impacts are predicted using both a numerical weather prediction model (Part I) and commonly used fast-running engineering wake models (Part II). While both approaches are suitable for longer-term analyses, a systematic comparison over a full, meteorologically representative year has so far been lacking. This comparison is conducted for the planned Princess Elisabeth (PE) wind farm cluster, which will be constructed in the near future in the Belgian part of the North Sea. The year under investigation is the most meteorologically representative of the past 30 years (see Part I). This allows for a realistic assessment of the impact of wakes on adjacent operational wind farms. This study only includes commonly used wind farm parameterization and engineering wake models using standard calibration parameters, even though some models or calibration parameters may perform better under specific conditions.

  This paper constitutes Part II of a two-part study investigating wake impacts of offshore wind farm clusters. While Part I focuses on the characterization of annual mean and sub-annual wake effects using a numerical weather prediction model, the present paper (Part II) extends this analysis by comparing numerical weather prediction results with commonly used fast-running engineering wake models. By comparing the results of these different model strategies, this study aims to quantify the spread in predicted wake impacts and to identify key drivers of model approach discrepancies, thereby providing insight into the uncertainty associated with long term farm-farm wake assessments.

- Line 130: (placeholder) why are we comparing against obs given PE doesn't exist yet? Why not use deployed lidars in the region?

  Thank you for pointing this out, to validate the simulations at hub height a comparison to LiDAR measurements at the LEG locations has been added. Additionally, the validation of the three simulations (no WF, without PE, with PE) has been added to the manuscript. It can be noted that the validation of the no WF simulation of some stations is modified compared to the original manuscript, this is because a post processing mistake has been found, incorporation of the recovery rate, and inclusion of circularity for wind direction. Additionally, the uncertainty on the mean bias has been included as suggested later. This uncertainty has been calculated as the uncertainty from the model sampling and from the measurements (for wind speed 3.3 % for all measurement locations except 5.6% for Westhinder). N is the sampling size and ρ_1 the effective sample size taking autocorrelation into account.

– **Modified Table 1 (Here Table 2) including LiDAR location, the uncertainty and recovery rate per location.**

Table 2: Observational mast name, location, measurement altitude, uncertainty of the wind speed measurements and recovery rate for 2016.

Observational mast	Longitude	Latitude	Height [m]	Uncertainty [%]	Recovery rate [%]
Westhinder (WH)	2.4378° E	51.3883° N	26	5.6	98.30
Wandelaar US (WA)	3.0458° E	51.3945° N	26	3.3	35.55
Wandelaar vane (WA)	3.0458° E	51.3945° N	26	3.3	97.04
Scheur-Wielingen (SW)	3.2985° E	51.4183° N	25	3.3	96.23
Sandettie-Lightship (SL)	1.800° E	51.102° N	14	3.3	92.5
Lichteiland-Goeree (LG)	3.67° E	51.9258° N	38	3.3	98.29
Lichteiland-Goeree (LG) LiDAR	3.67° E	51.9258° N	62 - 290	3.3	99.46 - 92.63

– **Modified Table 4 (Here Table 3) including the different simulations and uncertainty.**

Table 3: Mean wind speed [ms^{-1}] and wind direction [°] bias (simulation – observations) at LG and the 95% confidence interval, which accounts for both model variability and when available measurement uncertainty, for the no WF WRF, without PE and with PE simulations at the five different observation locations (Table 2).

	Wind speed bias [m s^{-1}]			Wind direction bias [°]		
Observational mast	**no WF**	**without PE**	**with PE**	**no WF**	**without PE**	**with PE**
Westhinder (WH)	-0.33 ± 0.04	-0.37 ± 0.04	-1.14 ± 0.04	13.71 ± 0.53	13.09 ± 0.53	12.96 ± 0.54
Wandelaar (WA) US	0.14 ± 0.06	-0.02 ± 0.07	-0.05 ± 0.07	-9.85 ± 1.26	-10.42 ± 1.30	-11.15 ± 1.30
Wandelaar (WA) vane	0.29 ± 0.04	0.19 ± 0.04	0.13 ± 0.04	11.57 ± 0.55	10.50 ± 0.56	10.53 ± 0.56
Scheur–Weilingen (SW)	0.66 ± 0.03	0.60 ± 0.04	0.58 ± 0.03	9.24 ± 0.56	8.42 ± 0.57	8.40 ± 0.56
Sandettie–Lightship (SL)	-0.95 ± 0.27	-0.94 ± 0.27	-0.94 ± 0.27	9.14 ± 4.05	9.33 ± 4.07	9.43 ± 4.11
Lichteiland–Goeree (LG)	-0.06 ± 0.04	-0.17 ± 0.04	-0.16 ± 0.04	2.10 ± 0.53	1.57 ± 0.54	1.49 ± 0.53

– **Added Table 5 (Here Table 4)including LiDAR mean bias of wind speed and wind direction for the different simulations including uncertainty.**

Table 4: Mean wind speed [ms^{-1}] and wind direction [°] bias (simulation minus observations) and the 95% confidence interval, which accounts for both model variability and when available measurement uncertainty, for the no WF WRF, without PE, and with PE simulations at different altitudes of the LG LiDAR location (Table 2).

	Wind speed bias [m s^{-1}]			Wind direction bias [°]		
Altitude [m]	**no WF**	**without PE**	**with PE**	**no WF**	**without PE**	**with PE**
62	0.04 ± 0.04	-0.08 ± 0.04	-0.07 ± 0.04	5.77 ± 0.32	5.49 ± 0.33	5.10 ± 0.32
90	0.03 ± 0.04	-0.09 ± 0.04	-0.08 ± 0.04	5.06 ± 0.32	4.81 ± 0.33	4.39 ± 0.32
115	0.01 ± 0.04	-0.12 ± 0.04	-0.11 ± 0.04	4.29 ± 0.32	4.09 ± 0.33	3.78 ± 0.32
140	-0.01 ± 0.04	-0.14 ± 0.04	-0.13 ± 0.04	3.43 ± 0.32	3.45 ± 0.33	3.10 ± 0.32
165	-0.04 ± 0.04	-0.17 ± 0.04	-0.16 ± 0.04	2.63 ± 0.32	2.63 ± 0.33	2.52 ± 0.32
190	-0.07 ± 0.04	-0.19 ± 0.04	-0.19 ± 0.05	1.98 ± 0.32	1.90 ± 0.33	1.90 ± 0.32
215	-0.10 ± 0.05	-0.22 ± 0.05	-0.23 ± 0.05	1.53 ± 0.33	1.60 ± 0.34	1.50 ± 0.33

– **Rewritten section 3.1 WRF validation discussing these new tables. (Part I lines 209–259):**

We first seek to establish confidence in the WRF model representation of the wind field. The simulations in this study represent idealized scenarios, as only a few of the Belgian-Dutch wind farms (Thorntonbank I, II and III, Belwind, and Northwind, 182 of 572 turbines) were operational in 2016, while the existing turbines used in the simulations correspond to the wind farm layout as of 2022. The year 2016 was nevertheless selected because it was identified as the most meteorologically representative. As a consequence of this mismatch between the meteorological year and the wind farm layout, a true validation of

wake energy losses is not possible. As such, validation will only become feasible once post-construction field observations for the fully developed wind farm area become available (Palatos-Plexidas et al., 2025). Future research may then also explore the optimal representation of the wind field by selecting the most appropriate model configuration and parameterizations.

Although optimizing the model setup through an extensive sensitivity study is beyond the scope of this work and has already been addressed elsewhere (Hahmann et al., 2024; Siedersleben et al., 2020; Tomaszewski and Lundquist, 2022; Li et al., 2021; Vermuri et al., 2022), it remains essential to assess the performance of the simulated flow fields relevant to this study. For this purpose, wind speed and wind direction from the no WF, without PE, and with PE simulations are compared against observations at five different locations (Table 1). The mean bias (simulated minus observed) is reported in Tables 4 and 5 for both wind speed and wind direction.

The validation of the ambient wind estimation largely focuses on the no WF case because the inclusion of wind farms in the simulation causes wakes at the observational mast locations, thus deviating from the observed values, where only a few turbines generated wakes in 2016. This is especially visible at the location of WH while the location in SL is less affected by wakes (smaller differences between the different simulations no PE, without PE and with PE). For the latter, this is the case since the main wind direction in this area of the North Sea is dominated by southwesterly winds (Ivanova et al., 2023).

The bias in the simulation without wind farms (no WF) is relatively small, ranging from -0.95 to 0.66 ms^{-1} for wind speed and $-9.85\ ^{\circ}$ to 13.71° for wind direction. This falls within the range reported in previous studies on offshore wind fields (Rosecrans et al., 2024; Fischereit et al., 2022b, Borgers et al., 2024; Li et al., 2021). Examining specific observation locations, namely WH, SL, and LG, the wind speed predicted by WRF tends to underestimate observed values offshore, consistent with findings by Li et al. (2021), Rosecrans et al., (2024) and Borgers et al. (2024). In contrast, at WA and SW, closer to the coast, the WRF model tends to overestimate wind speed (Hahmann et al., 2014), suggesting a potential coastal effect not well captured by WRF. Due to the relatively coarse coastal resolution, WRF sees a longer fetch which results in higher wind speeds in the model (Hahmann et al., 2014). No clear trend is evident in the no WF biases concerning the altitude of the measurement masts considered, except that the highest observation point exhibits the lowest bias, similar to Klemmer et al. (2024), but opposite to Munoz-Esparza et al. (2012) and Sward et al. (2023).

When comparing all three simulations (no WF, without PE, and with PE), the strongest wake related deviations appear at WH, WA, and SW, where the inclusion of wind farms leads to a significant change in bias compared to the no WF case. At LG, which is located downwind, some differences between the three simulations can also be linked to wake effects. At WH the differences between the simulation no WF and without PE are within the uncertainty range (95 % confidence interval which accounts for both model variability and, when available, measurement uncertainty, see Appendix Section ??), while the wake effect of the new PE is clearly present in the wind speed bias with PE. At SL, where wakes are not expected to play a large role, the biases are nearly identical across the three simulations. This indicates that the background flow is consistently reproduced in all simulations, providing confidence in the subsequent wake energy loss analysis.

The LiDAR at LG was designed specifically for assessing offshore wind resources. It provides vertical wind profiles up to 215 meters, enabling model evaluation at heights comparable to the level of turbine hub heights. Comparing the different WRF simulations with the LiDAR observations (Table 5) reveals a small mean bias across all altitudes and different simulation. The bias for wind *speed ranges from 0.04 ms^{-1} at 62 m to –0.10 ms^{-1} at 215 m in the no WRF simulation, from -0.08/-0.07 ms^{-1} at 62 m to -0.22/-0.23 ms^{-1} at 215 m in the without PE simulation and with PE simulation, respectively. Wind direction biases gradually decrease with height, ranging from values around 5° at 62 m to around 1.5° at 215 m. This suggests improved directional agreement at higher altitudes. The mean wind speed and direction bias of different simulations often fall within the observational uncertainty, accounting for model variability and LiDAR measurement error, especially between without PE simulation and with PE simulation. Larger differences are observed when comparing to the no WF simulation, this could be caused by wakes. To conclude, the background wind field can be assumed to be similar across simulations, and justify studying the wake effects in the suggested simulations. Overall, based on all observation comparisons, a good agreement between modeled and observed values at hub height supports using WRF to evaluate wake impacts of the newly built PE on existing wind farms.*

- **Added appendix C**

  The uncertainty calculation accounts for both the variability of the bias time series and the uncertainty associated with the measurements . For wind speed, a relative measurement uncertainty (U_m) of 3.3% is assumed for all stations, except Westhinder where a value of 5.6% is used. For wind direction, no measurement uncertainty is available and thus only the model variability is considered.

  The effective number of independent samples (N_{eff}) is estimated to account for temporal autocorrelation (10 minute output) in the bias time series as shown in Eq. C1.

$$N_{\text{eff}} = \frac{N}{1 + 2\rho_1} \tag{9}$$

  Here, N is the total number of samples and ρ_1 is the lag-1 autocorrelation coefficient of the bias. The total uncertainty on the mean bias (U_{tot}) is calculated as the quadratic sum of the uncertainty due to model variability and the uncertainty associated with the measurements, shown in Eq. C2.

$$U_{\text{tot}} = \sqrt{\left(\frac{\sigma_{\text{bias}}}{\sqrt{N_{\text{eff}}}}\right)^2 + \left(\frac{U_m\,\bar{X}_{\text{obs}}}{\sqrt{N_{\text{eff}}}}\right)^2} \tag{10}$$

  Here, σ_{bias} is the standard deviation of the bias time series, and $\bar{X}_{obs}$ is the mean measured value at the observation mast. The first term in Eq. 10 represents the standard error of the mean bias due to model variability, while the second term represents the measurement uncertainty. When the measurement uncertainty is not available (e.g., for wind direction), only the first term in Equation 10 is used.

- Line 135: What value was used for the TKE generation parameter?

  Thank you, this has been clarified (Part I lines 140-141):

  When wind turbines are included, the parameterization of Fitch et al. (2012) is used with a turbulent kinetic energy generation parameter of 0.25 (Archer et al., 2020).

- Line 180: where does the wind rose come from? What wind data are you using to inform the engineering wake models, and how does this data compare to the WRF simulations?

  The methodology for extracting input to the engineering models from WRF was not entirely clear in the previous manuscript. In the revised version, this methodology is described in greater detail. Furthermore, since the manuscript has been split

into two parts, we have also introduced additional setups for the engineering wake models.

Added/Modified (Part II lines 154-202):

FLORIS simulations are driven by meteorological conditions extracted from the noWF WRF simulations for the entire year 2016 (see Part I). These conditions include wind speed, wind direction, and Turbulence Intensity (TI).

The noWF velocity field is provided to FLORIS using two different inflow setups. The first setup assumes horizontally homogeneous inflow and retrieves the wind speed from a single probe point located at the geographic center of the BE–NL cluster at an altitude of 96 m. In contrast, the heterogeneous inflow configuration accounts for spatial variability in the incoming flow, arising from coastal gradients or large-scale synoptic structures.

In the heterogeneous configuration, FLORIS' built-in implementation described by Farrel et al. (2021) is used. Specifically, the noWF velocity field at 96 m height is extracted at the WRF grid points and fed to FLORIS. For each wind turbine, FLORIS initializes the local inflow wind speed by performing a linear barycentric interpolation of the WRF wind speed field. Wake velocities downstream of each turbine are then computed following the standard FLORIS wake modeling approach, but using this locally interpolated inflow wind speed as the initial condition. No vertical interpolation is applied since both vertical wind shear and veer are omitted in the present study.

Some of the engineering wake model simulations assume a time varying TI estimated based on the noWF WRF data following Larsen (2024). Larsen (2024) proposes to use the Kaimal turbulence spectrum to estimate the wind speed component variances, σ_u, σ_v, and σ_w, at given heights and wind speeds. Empirical ratios between σ_u and other components then allow to transform the TKE extracted from WRF into the along-wind standard deviation. The ambient TI can then be approximated as

$$TI_{amb} = \frac{\sigma_u}{\overline{U}} = \frac{1}{\overline{U}} \sqrt{\frac{2\, TKE}{1 + \alpha^{-1} + \beta^{-1}}} \quad (11)$$

with $\overline{U}$, the mean wind speed over the averaging period at the height of interest and α and β, empirical coefficient determined based on Kaimal turbulence spectra.

Four different FLORIS simulation setups are subsequently considered:

- ***Homogeneous FLORIS with fixed TI: A time-varying and spatially homogeneous wind field is assumed in FLORIS, constructed from the wind speed and direction at 96 m altitude at the geographic center of the BE–NL wind farm cluster, taken from the no WF WRF simulation (Part I). Turbulence intensity is fixed in space and time at 6%, a typical value for the offshore North Sea area.***
- ***Homogeneous FLORIS with varying TI: A time-varying and spatially homogeneous wind field and TI is assumed in FLORIS, constructed from the wind speed, wind direction and TI at 96 m altitude at the geographic center of the BE–NL wind farm cluster, taken from the no WF WRF simulation (Part I). This setup aims at evaluating whether incorporating temporal variability in ambient turbulence resulting from changes in different atmospheric conditions improves the representation of wake recovery.***
- ***Heterogeneous FLORIS with fixed TI: A time-varying and spatially heterogeneous wind speed field is assumed in FLORIS, constructed from the wind speed at 96 m altitude taken from the no WF WRF simulation (Part I). Wind direction is kept spatially homogeneous and time-varying, constructed from the no WF WRF simulation at 96 m altitude at the geographic center of the BE–NL***

wind farm cluster (Part I). Turbulence intensity is fixed in space and time at 6%, a typical value for the offshore North Sea area.

  - ***Heterogeneous FLORIS with varying TI: A time-varying and spatially heterogeneous wind speed field is assumed in FLORIS, constructed from the wind speed at 96 m altitude taken from the no WF WRF simulation (Part I). Wind direction is kept spatially homogeneous and time-varying, constructed from the no WF WRF simulation at 96 m altitude at the geographic center of the BE–NL wind farm cluster (Part I). A time-varying and spatially homogeneous TI is assumed, constructed from TI at 96 m altitude at the geographic center of the BE–NL wind farm cluster, taken from the no WF WRF simulation (Part I).***

  All four setups of the engineering wake models are run in a time-series mode with a 30-min resolution for one full year (2016), where each simulation time step is treated as quasi-steady and independent. Consequently, the model does not account for wake advection or temporal effects in wake propagation contrary to WRF. As noted previously, not all wind farms in the North Sea are modeled, for the engineering models used in this study only those belonging to the BE-NL and PE wind farm cluster are included. This reduced domain of not including all North Sea wind farms allows for the simulation of twelve cases (three layouts (no WF, with PE and without PE similar to WRF) combined with four input setups described above) at a reasonable computational cost, while still capturing the dominant inter-farm wake losses of the PE wind farm cluster on the BE-NL wind farm cluster.

- Table 5: The varying degrees of bias are concerning, as is the 1.22 m/s bias at SW. What QC was performed on the measurement data, and over what period were the data analyzed? Some info on data quality, recovery rate, etc. would be helpful. Also, in the end, validation against measurements under 40m nearly as important as validation at hub height, again pointing at the importance of using lidar data instead

  See previous comment on Line 130: WRF validation

- Figure 3: Struggling a bit to interpret here. I don't see any real impact apart from the main downwind cluster. Also why does site B have losses on the north end but nowhere else, and which also doesn't align with the wind speed deficit map? Further, the wind speed deficit map is spotty and difficult to explain (why the northern deficit?). This is likely due to numerical anomalies between the different simulations, and I recommend limiting impacts to 0.05 m/s or so, or some value that eliminates these artifacts. Finally, why limit wind deficits to max of 0.2? We're missing a lot of structure in that big blob in the middle.

  We agree that Figure 3 can be confusing and it looks that site B only has losses in the north, however this is only because of limits of the colorbar. As you can see in Figure 3, when the colorbar is changed you can see that there is AEP loss in all neighboring wind farms of PE, corresponding to the values shown in Table 6 of Part I. In Part I the colorbar of the AEP loss is set to max 10 to show the effects in the closest neighbouring BE-NL cluster.

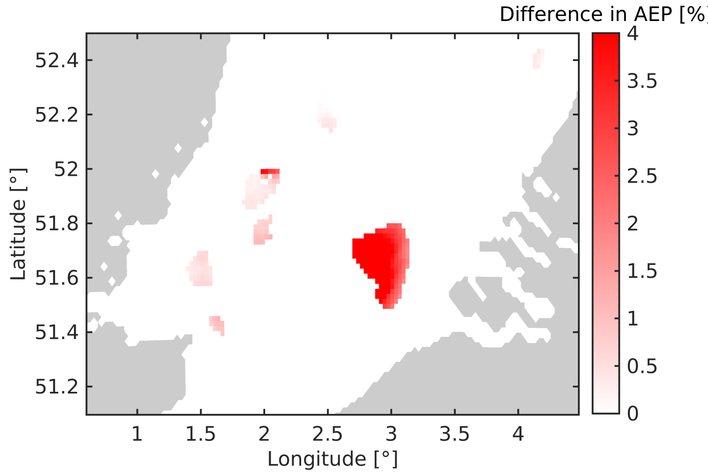

Figure 3: Loss in AEP [%] caused by the PE wind farm

We follow the reviewers advice on limiting the impacts of 0.05 m/s to remove noise and extend the wind speed deficit to 0.8 m/s to get additional structure on the additional wind speed deficit.

- Table 6: Do you discuss the negative values anywhere in the paper?

  In the revised manuscript we calculated the energy losses based on power averaged and not on efficiency averages as suggested by the reviewer. Additionally we added uncertainty intervals for the AEP losses and the energy losses. For the AEP the negative losses can probably be attributed to numerical uncertainties, for the external wake losses the negative values fluctuate around 0 and thus is not a significant result. The differences between the two methodologies AEP and wake losses is also explained (see below). The negative wake losses for the BE-NL wind farm cluster for some wind directions is probably linked to lateral speedups. In Part II of the paper the negative losses are linked to the heterogeneous setups for engineering wake models not including explicitly variations in TI. But as mentioned by another the reviewer, this would need further research that is not in the scope of these manuscripts.

  - **Clarified (Part I lines 273-278):**
    It should also be noted that very small increases (0.11-0.15%) in AEP are seen for three wind farms after the construction of the PE wind farm. Given the low magnitude and the fact that these occur at only three locations, it is difficult to draw any strict conclusions regarding actual increases in AEP. These increases in AEP may be attributable to numerical variability, possibly related to coastal effects or model variability in WRF, and further research would be needed to determine whether they have physical (flow speedup) or economic significance.
  - **Added (Part I lines 312-319):**
    Comparing the external wake energy losses (Table 7) with the AEP losses (Table 6) shows that both tables exhibit similar spatial trends. The BE–NL wind farm cluster, Galloper & Greater Gabbard, London Array, Thanet, and Luchterduinen show the largest impacts, while more distant farms show much smaller effects. Differences between the two tables arise because the metrics are defined using different normalizations. The AEP losses are computed as relative changes between the without PE and with PE simulations and normalized by AEP without PE, whereas the external wake losses are defined as ratios of temporally averaged power normalized by the noWF simulation (Eqs. 4–6). This difference in normalization, together with the implicit weighting of high-production periods in the AEP calculation, explains why some farms show

statistically significant AEP impacts while their external wake losses are not statistically significant.

– **Clarified (Part I lines 363-379):**
Interestingly, negative external wake energy losses (-1%), i.e., positive interference from the construction of the new PE wind farm cluster, are observed for winds coming from the north to northeast (0-60) or south (180-210) (Fig. 4 (b)). The negative external wake energy losses for wind directions from the south can be explained by a speedup at the lateral locations of the PE wind farm clusters (Fig. 8). It appears that the PE wind farm cluster act as a windbreak, causing global blockage and accelerating the flow around sides of the wind farm. A comparable lateral acceleration caused by global blockage is reported for certain North Sea wind farm clusters (Borgers et al., 2025). While similar effects have been studied for windbreaks, where they mainly cause vertical speed-ups (Liu and Stevens, 2021), lateral flow accelerations influencing farm-to-farm wakes have only recently been examined in more detail (Hasagar et al., 2015; Nygaard and Hansen, 2016; Borgers et al., 2025). Nygaard and Hansen (2016) observed this farm effect based on measured wind turbine data from two offshore wind farms. This speedup effect in the lateral direction has been observed multiple times for individual wind turbines (Ammara et al., 2022; Araya et al., 2014; Puccioni et al., 2023). In the coming years, it would be beneficial to enhance our understanding of this acceleration phenomenon around wind farms, through idealized simulations, however, this is outside the scope of this work. Understanding these speedups is important, as they may have a significant impact on external wake energy losses, necessitating consideration in the planning of high-density offshore regions. Why negative external wake energy losses are observed for wind directions from the north to northeast is less clear. As mentioned previously for the AEP losses, these may result from numerical variability or other uncertainties in the simulation. The underlying mechanisms of these minor effects, which might be physical or not, should be investigated in the future.

– **Added figure (Part I Fig 8), here Fig 2.:**

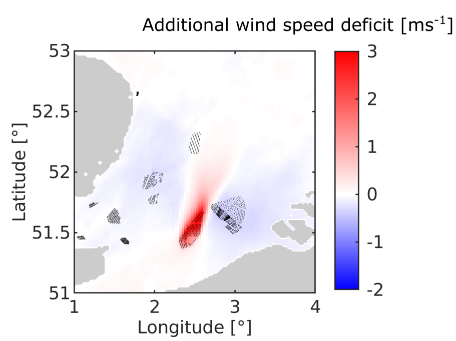

Figure 4: External BE-NL wind farm wake deficit for wind directions between 180 and 210, the sector showing negative external wake losses. This map shows the additional wind speed deficit after the construction of the PE wind farm for timesteps with negative external wake energy losses over the farm.

- Line 274: The trend should be very statistically significant. We should see considerably higher wake losses in stable conditions. And how are you quantifying statistical significance? There should be a much stronger signal here. Part of the problem may be the Obukhov length, which is surface specific. Suggest using bulk Richardson number, or vertical temperature gradient, as alternates to compare.

  The result in the initial submission sought to provide a general summary of key impacts, rather than parsing the individual effect of each variable, such as stability.

In other words, the wind direction was not controlled, and therefore the effect of stability is mixed with the inherent interactions between stability, wind speed, wind direction, etc. To isolate the impact of stability, we have identified a prevailing wind condition important for external and internal wake losses, with a particular focus on a narrow range of wind directions and wind speed respectively. Within this additional condition bin, we investigate the impact of stability, which shows the anticipated trend that stable stratification results in stronger wake effects.

Additionally, we calculated the bulk Richardson number at the same location of where we extracted L, and this between 10 m and 100 m altitude. The classification for stability based on Ri is shown in Table 5. On the full year time series there is a Pearson correlation of 72% between L and Ri_b. Together with the fact the L method has been used in previous literature (Porchetta et al., 2019; Nybo et al., 2020; Platis et al., 2022; Howland et al., 2022), we believe that the method described in the manuscript is valid.

Table 5: Atmospheric stability classification based on the bulk Richardson number (Ri_b).

Stability class	**Range of Ri_b**
Very unstable	$Ri_b < -0.023$
Unstable	$-0.023 \leq Ri_b < -0.0036$
Neutral	$-0.0036 \leq Ri_b < 0.0072$
Stable	$0.0072 \leq Ri_b < 0.084$
Very stable	$Ri_b \geq 0.084$

– **Added spatiotemporal conditional average (Part I lines 169-175):**
Three different averaging procedures are employed, where N_s is the number of turbines, N_t the number of time steps, and $P_i(t)$ the instantaneous power output of turbine i at time t:
  * ***Temporal average: $\overline{P}_i = \frac{1}{N_t}\sum_{t=1}^{N_t} P_i(t)$, i.e., the average over time for turbine(s) containing grid cells.***
  * ***Spatiotemporal average: $\overline{P} = \frac{1}{N_s N_t}\sum_{i=1}^{N_s}\sum_{t=1}^{N_t} P_i(t)$, i.e., averaging across all turbine(s) containing grid cells and time steps simultaneously.***
  * ***Spatiotemporal conditional average: $\overline{P}_{\text{spatio-bin}} = \frac{1}{N_{\text{bin}}}\sum_{i,t\in\text{bin}} P_i(t)$, considering only specific conditions such as wind speed, wind direction, atmospheric stability, or a combination of the previous.***

– **Modified (Part I lines 341-355 added Fig. 6) Here Fig. 5 :**
Figure 4 (c) shows external wake energy losses as a function of atmospheric stability. In this figure, the highest losses are present for near-neutral conditions. This arises because the figure does not control for wind conditions, and aggregating over all atmospheric situations allows competing effects to cancel or amplify, which obscures the underlying stability dependence. To elucidate how the wake effects depend on stability more clearly, we define a narrow wind direction bin of 193 to 317. This wind direction bin only includes wind directions for which the BE–NL wind farm cluster is in the wake of the PE wind farm cluster. These results are shown in Fig. 6 and show decreasing external wake energy losses from a very stable to a very unstable atmosphere. The lowest external wake energy losses are present for the most frequently occurring very unstable atmospheric conditions. It has been reported in earlier studies that, for the North Sea, (very) unstable atmospheric conditions are commonly seen (Porchetta et al., 2019; Munoz-Esparza et al. 2012; Nybo et al., 2020). Also at the location of the planned Vineyard wind farm at the East Coast of the US, unstable conditions are present 53.6% of the time (Rosecrans et al., 2024). Stable conditions in the North Sea (BE–NL wind farm cluster) are often associated with southwesterly wind directions, when warmer air is transported over the colder sea, especially in summer. However, this pattern is site-dependent,

as stability is strongly influenced by the temperature contrast with the nearest coastline. For other wind directions, particularly those bringing colder air over a relatively warmer sea, especially in winter, (very) unstable conditions tend to dominate (Nybo et al., 2020; Cheynet et al., 2018).

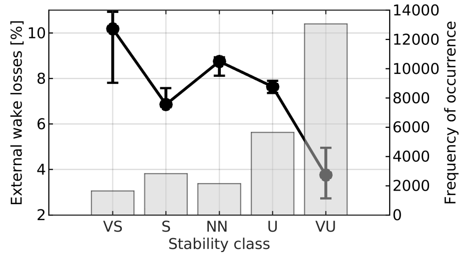

Figure 5: (Line) Spatiotemporal conditional average of external wake energy losses as a function of atmospheric stability classes (VS: Very Stable, S: Stable, NN: Near Neutral, U: Unstable, VU: Very Unstable) only including wind directions from 193 - 317, (bars) frequency of occurrence for the different atmospheric stability classes. The error bars represent a 95% confidence interval obtained via bootstrapping with 1000 resamples.

– **Modified (Part I lines 392-400 added Fig. 10) Here Fig 6.:**
Figure 9 (c) shows the internal wake energy losses as a function of atmospheric stability. The internal wake energy losses generally decrease from stable to unstable conditions, with an apparent increase for very unstable conditions. Similar to the external wake energy losses, this may seem counterintuitive, but arises because this figure does not control for wind speed or wind direction and instead aggregates all atmospheric conditions. After additionally conditioning on wind speeds between 10 and 15 ms^{-1}, a clear pattern of decreasing internal wake energy losses with increasing atmospheric instability emerges (Fig. 10), as expected. Since very unstable conditions occur most frequently, a large number of low wind speed events are included, which increases the relative percentage losses and therefore results in higher apparent internal wake energy losses. This is consistent with the findings of Xia et al. (2025), who showed that wind speed deficits do not translate linearly into power or energy losses.

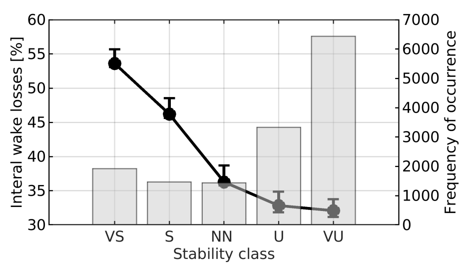

Figure 6: (Line) Spatiotemporal conditional average of internal wake energy losses as a function of atmospheric stability classes (VS: Very Stable, S: Stable, NN: Near Neutral, U: Unstable, VU: Very Unstable) only including wind speeds between 10 and 15 ms^{-1}, (bars) frequency of occurrence for the different atmospheric stability classes. The error bars represent a 95% confidence interval obtained via bootstrapping with 1000 resamples.

- Figure 4: Why are there external wake losses in sectors other than SW, and what do you make of the negative losses?

  See reply to remark about Table 6

- Line 283: Not necessarily related to thrust. At high winds, you're well above rated capacity, and even with wakes you can stay above rated, hence no losses. It's more a question of wind speed magnitude. At low wind speeds, your denominator is low, hence percentage based losses are high. It'd be useful to see a side by side of percentage and power/energy based losses

We thank the reviewer for this helpful comment. We have added a side by side comparison of percentage and absolute external wake energy losses and revised the text accordingly.

– **Added (Part I lines 356-362 and Fig. 7) here Fig. 7**:
  External wake energy losses are expressed as percentages (Eq. 8). At lower wind speeds, turbines produce little power, so even small external wake energy losses appear as relatively large percentage losses (Fig. 4 (d)), even though the absolute wake effect may not be stronger. At the same time, turbine thrust coefficients are higher at low wind speeds, leading to longer wake recovery distances. In contrast, when examining absolute losses (Fig. 7), the largest deficits occur for wind speeds between 10 and 15 ms^{-1} (mostly Region II). In this range, turbines operate close to rated power while still generating significant thrust, so wakes cause large power losses. At higher wind speeds, the thrust coefficients decrease and turbines typically remain at rated power even in the presence of wakes.

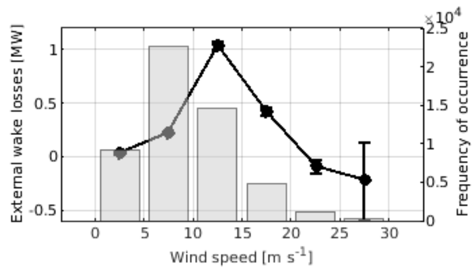

Figure 7: (Line) Spatiotemporal conditional average of absolute external wake energy losses as a function of wind speed, (bars) frequency of occurrence for the different wind speed bins. The error bars represent a 95% confidence interval obtained via bootstrapping with 1000 resamples.

– **Added (Part I lines 401-406 and Fig. 11) here Fig. 8**:
  For the internal wake energy losses in function of velocity we observe the same behavior as for the external wake losses. When expressed as percentages, they are larger at lower wind speeds because the denominator is small, so even modest absolute deficits translate into large relative losses (Fig. 9 (d)). In contrast, when considering absolute losses, the largest deficits occur between 10 and 15 ms^{-1}, when turbines operate close to rated power with high thrust coefficients and wakes cause substantial power losses (Fig. 11). At higher wind speeds, turbines remain at rated power, which limits the relative impact of internal wake losses.

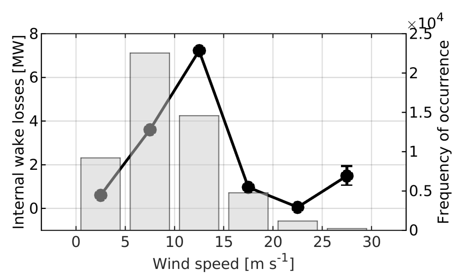

Figure 8: (Line) Spatiotemporal conditional average of absolute internal wake energy losses as a function of wind speed, (bars) frequency of occurrence for the different wind speed bins. The error bars represent a 95% confidence interval obtained via bootstrapping with 1000 resamples.

- Figure 6: I don't understand these negative wake losses in multiple wind directions. May I ask how you're calculating the average? Are you calculating a percentage loss for each timestep, then averaging those percentages by sector? Or are you (and more correctly) averaging/summing power for each timestep for each sector, and only calculating percentage loss based on those final averages? The two approaches will yield very different results and the latter is more defensible.

We agree. We have modified the calculation procedure as suggested by the reviewer, see introduction on response to reviewer, and we do now average the power production and then compute efficiency, rather than computing efficiency at each time step and then averaging after. This ensures a more clear and accurate representation of the energy effects of wakes (i.e. the energy ratio instead of the power ratio). Additionally to make sure this is clear we work with taking the average at the end, we talk about wake energy losses consistently in the manuscript.

- Table 7: I don't understand how KF and GS can have negative external wake losses arising from the Prinses Amalia deployment. They're too far away. I think there is an issue with your calculations.

  We now included the 95% confidence interval calculated by bootstrapping with a resampling of 1000. Here we see that the confidence interval overlaps with zero, which indicates that there is no significant impact of PE on KF and GS and should thus not be considered a physical effect. The same applies to other distant farms, including Scroby Sands, East Anglia ONE, Kentish Flats, Egmond aan Zee, Gunfleet Sands and Prinses Amalia, whereas only the closely related farms such as the BE–NL cluster, Galloper & Greater Gabbard, London Array, Thanet, and Luchterduinen show a statistically significant effect.

  Also see reply to remark about Table 6