Anonymous Reviewer 2:

Dear Anonymous Reviewer,

Thank you for your feedback. It helped us to improve the manuscript and strengthen our findings. You emphasized some central points that were overlooked in our initial manuscript.

All reviewer comments appear in grey below, while authors' responses appear in blue text. Line numbers referenced in the authors' responses refer to the revised document. Figures included in the manuscript are labeled in italic and using numbers (e.g. *Figure 7*), while figures that only appear in the response to reviewer comments are labeled in smaller font and using roman numerals (e.g. Figure iv).

General Comments

The article 'Quantifying wind plant blockage under stable atmospheric conditions' by Gomez et al. draws conclusions about the magnitude and the sources of an observed velocity reduction upwind (blockage) of an idealized wind farm in two LES wind farm simulations. The main message of the article is that the blockage is higher in a stronger stratified atmospheric boundary layer and that the reason for this is the lack of vertical turbulent momentum transport. The authors further compare different virtual measurement setups to measure the wind speed upwind of the farm and analyse how a signal of blockage can be recognized in the production data. I see this work in general as an interesting addition to the current scientific discussion, but I see a couple of points that need to be addressed before I can recommend the full publication of the work.

Major Comments

1. Introduction: L. 26 - 35 - Definition of blockage, numbers

I have an issue with the introduction of blockage in wind farms. I would say it is not proven that the upstream wind speed necessarily decreases more when turbines are combined in a wind farm. Furthermore, I don't know any credible scientific publication that can relate the observed overpredictions of energy production of wind farms to blockage. The announcement of Orsted does not serve as a credible and sufficient reference. The numbers related to wake deficits (10 %) and blockage (1%) are not explained what they relate to. (wind speed? energy production?) If no reference is found here, I would suggest to rather talk about different orders of magnitude.

We appreciate your comment and include more references to complement our statements. Regarding the difference in blockage for wind plants and single turbines, we included references as follows:

L26-L27: "*And, when multiple turbines are combined into an array, the upstream wind speed decrease is larger than that of a turbine in isolation (Ebenhoch et al., 2017; Bleeg et al., 2018).*"

For the numbers related to wake deficits and blockage, we now include references to lidar observations of wakes and blockage that support these orders of magnitude for the wind speed deficit:

L31-32: "*Though wind speed deficits from wakes are large (~10%) (e.g. Lee and Lundquist, 2017), wind plant blockage produces wind speed deficits of~1% over a wide area upstream (Schneemann et al., 2021), making it much more difficult to quantify*."

Finally, we modify the statement that refers to Ørsted's announcement to <u>suggest</u> rather than <u>state</u> that there are overpredictions:

L29-30: "However, upstream wind plant blockage is usually neglected, possibly resulting in lower-than-forecasted energy predictions and financial losses for wind plant operators (Ørsted, 2019)."

2. Duration of averaging and influence on results and conclusions

The averaging period of 45 min appears quite small. If no longer averaging is possible, the limits of this restriction should be discussed throughout the paper. The paper draws conclusion on the significance of the results based also on the number of samples. For higher turbulence as in the lower stratification case, the significance is automatically lower. Thus, any conclusions about the significance of velocity differences, e.g. Figure 5, should point out that the significance criteria is strongly dependent on the length of measurement period. All conclusions about the significance of the results (wind speed deficit, power measurements) should relate to the selected sampling frequency (0.1 Hz) and the measurement period (45 min). A proper way to make the results between the two boundary layers more comparable is to scale the period of measurement to the turbulence level of the flow.

Thank you for raising this question that we had not yet considered this in our analysis. We now include this restriction in our manuscript as follows:

L177-180: "*Note that we use the same averaging time for two boundary layers that contain different turbulence levels. As such, our analysis automatically reduces the statistical significance of the U12-C0.3 simulations. Turbulence levels in the U12-C0.3 simulation are higher than in the U12-C0.5 simulations, resulting in higher correlation between the data and thus smaller independent sample size."*

3. Discussion of measurement strategies (chapter 5)

I have a hard time grasping the meaning of and the approach in this chapter. I understand the conclusion is that more measurement points (and thus more samples) reduce the uncertainty, which is I would say common sense. So, in this case it would be more interesting to look at the combination of different sampling frequencies and locations. Also, I don't understand why uncertainty is not displayed to evaluate the measurement setups, but rather a bias. Furthermore, I don't think the averaging setups are even supposed to result into the same free stream velocity, as it can be clearly seen in Figures 3 and 9 that the flow is highly inhomogeneous in x. In consequence my suggestion would be to either remove the chapter or put a lot more effort in working out the implications of the different measurement setups.

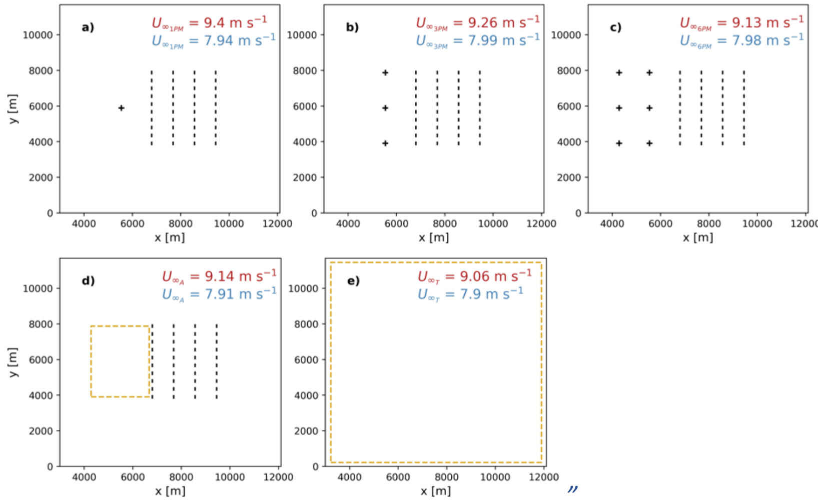
Thank you for this constructive comment. This chapter is intended to point to the difficulties of measuring blockage in experimental setups. The different methodologies we employ to define the freestream velocity are not supposed to result in the same freestream velocity because of the cross-stream and streamwise fluctuations in the flow. However, they point to the difficulties of capturing the blockage effect using point measurements. We changed the nomenclature throughout the section, added an uncertainty analysis and highlighted our findings in a clearer way.

We modified how we reference each measurement methodology as follows:

L256-263: *"After removing the background flow from the velocity field, we calculate the freestream velocity upstream of the wind plant in multiple ways. We test five different approaches as shown in Fig. 10: 1) time-averaged, hub-height wind speed measured at one point 10D upstream of the wind plant $\left(\overline{U}_{\infty_{1PM}}\right)$, such as would be available from a single profiling lidar or meteorological tower; 2) time-averaged, hub-height wind speed measured at three points 10D upstream of the wind plant $\left(\overline{U}_{\infty_{3PM}}\right)$; 3) time-averaged, hub-height wind speed measured at six points 10D and 20D upstream of the wind plant $\left(\overline{U}_{\infty_{6PM}}\right)$; 4) time- and spatially averaged hub-height wind speed measured over the area extending 1D to 20D upstream of the wind plant $\left(\overline{U}_{\infty_{A}}\right)$, such as would* *be available from a scanning lidar; and 5) time- and spatially averaged hub-height wind speed measured over the whole turbulent domain of the no-turbine simulations (referred to as "True freestream" $\overline{U}_{\infty_T}$).”*

Furthermore, the convention for $\overline{U}_{\infty_i}$ can be directly related to Figure 10 and Figure 12.

“Figure 10: Schematic showing the relative location of the wind plant and the sampling locations for defining the freestream velocity of the flow. The freestream velocity in (a), (b), and (c) is calculated using one-, three-, and six-point measurements (PM), respectively. In (d) and (e) the freestream velocity is calculated from areal measurements enclosed by the dashed yellow line. The solid vertical lines represent the individual wind turbines. As such panel (e) represents the simulation with no turbines in the domain. The black crosses represent the locations for sampling the freestream velocity using point measurements. Freestream velocities in each panel are color-coded for each stability condition: red (blue) text represents the U12-C0.3 (U12-C0.5) case.

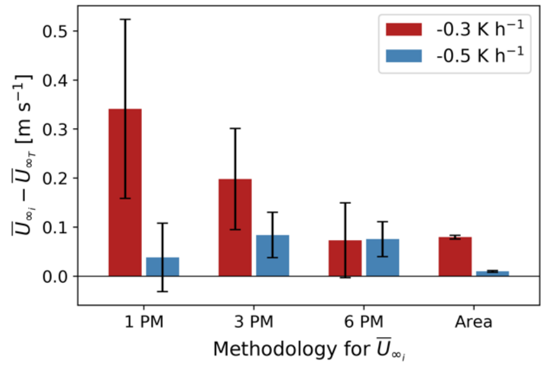


”

For the uncertainty analysis, we include the new Figure 11 and Table 3. The manuscript is modified as follows:

L269-274: “*Depending on how the freestream velocity is defined, cross-stream flow inhomogeneities produce variations in the freestream velocity comparable in magnitude to the blockage effect (Figure 11). For the U12-C0.3 case, the difference between the freestream velocity estimated with a single point measurement and sampling the area upstream of the wind plant is 0.3 m s^{-1}, which is the same order of magnitude as the blockage effect we are trying to measure ($\widehat{\overline{U}}_{def}$~1%). For the U12-C0.5 case, the difference in the various definitions of the freestream velocity is smaller relative to the blockage effect but still present. The largest difference between freestream velocities is 1%.*”

“Figure 11: Uncertainty in estimating the freestream velocity using the methodologies outlined in Fig. 10. Errorbars represent the 95% confidence intervals for each measurement methodology.



"

"Table 3: Furthest distance upstream that displays a statistically significant velocity deficit for each $\overline{U}_{\infty_i}$. Statistical significance is evaluated using a 95\% confidence level.

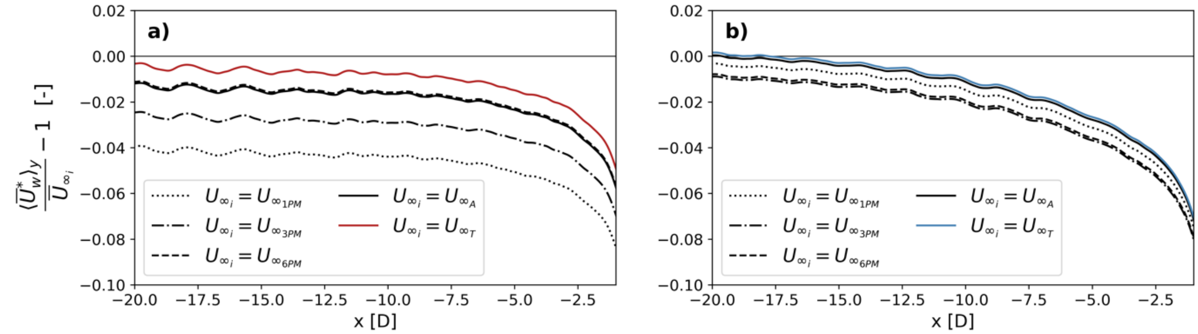
	Induction zone for each $\overline{U}_{\infty_i}$				
Case	$\overline{U}_{\infty_{1PM}}[D]$	$\overline{U}_{\infty_{3PM}}[D]$	$\overline{U}_{\infty_{6PM}}[D]$	$\overline{U}_{\infty_A}[D]$	$\overline{U}_{\infty_T}[D]$
U12-C0.3	*-18.5*	*-18*	*-6.5*	*-6.5*	*-4*
U12-C0.5	*-10*	*-20*	*-20*	*-13.5*	*-13*

"

L286-292: "*We evaluate the statistical significance of the velocity deficit upstream of the wind plant for each $\overline{U}_{\infty_i}$. The furthest distance upstream of the wind plant that displays a statistically significant velocity deficit changes with definition of freestream velocity (Table 3). Generally, the induction zone shrinks as the number of observations used to define the freestream velocity increases. Stronger cross-stream flow inhomogeneities in U12-C0.3 result in larger differences in the induction zone for the different methodologies used to define $\overline{U}_{\infty_i}$ compared to U12-C0.5. Though a smaller density of observations increases the uncertainty in estimating the induction zone of a wind plant, the differences in the induction zone extent are primarily affected by the change in magnitude of the freestream velocity.*"

We do not include the confidence intervals in the velocity deficit because they will add too much clutter to the plot. This would require adding error bars to five different line plots in one same panel. We now make this clear in the caption for this figure in the manuscript:

"*Figure 12: Normalized velocity deficit for the (a) U12-C0.3 and (b) U12-C0.5 case using the various definitions of freestream velocity shown in Fig. 10. Results for the True freestream velocities are color-coded for each stability condition: red (blue) text represents the U12-C0.3 (U12-C0.5) case. Note that the colored lines in (a) and (b) are not the same as the corresponding lines in Figure 6 because here the freestream is single-valued, whereas in Figure 6 the freestream varies in the streamwise direction. Confidence intervals are not shown.*



”

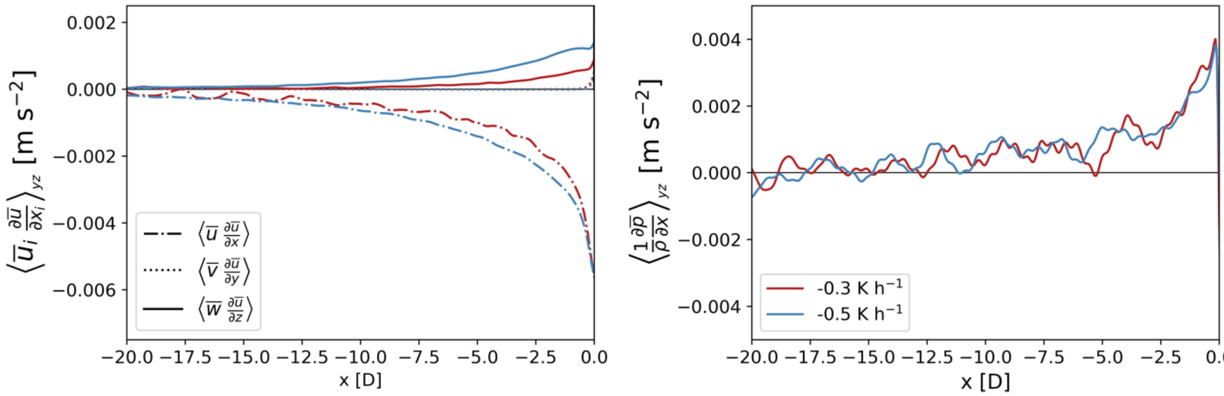
4. Conclusion on difference between the two simulations derived from flux divergence

   I suggest adding the flux divergences from the simulations without any wind farms. As the flow does not appear to be stationary along x, I would assume that there is already divergence even without any wind farm. Like this I am still a bit skeptic to accept the difference in the vertical momentum flux to be the sole reason for the difference in upstream wind speed deficit. Also, what about the mean momentum fluxes?

We appreciate this comment as we agree other terms are also very relevant. We analyze the following terms, and how they differ from the simulations without the GAD: $\overline{u}_i \frac{\partial \overline{u}}{\partial x_i}, \frac{1}{\rho}\frac{\partial \overline{p}}{\partial x}$, and $\frac{\partial \overline{u' u_i'}}{\partial x_i}$.
The induction zone seems to be most impacted by $\overline{w}\frac{\partial \overline{u}}{\partial z}$. The dominant terms in the x-momentum equation are $\overline{u}\frac{\partial \overline{u}}{\partial x}, \overline{w}\frac{\partial \overline{u}}{\partial z}$, and $\frac{1}{\rho}\frac{\partial \overline{p}}{\partial x}$. The pressure gradient term $\left(\frac{1}{\rho}\frac{\partial \overline{p}}{\partial x}\right)$ is nearly identical for both simulations, thus the difference in the wind plant's induction zone is largely determined by the advection of x-momentum by the vertical velocity $\left(\overline{w}\frac{\partial \overline{u}}{\partial z}\right)$. The negative vertical transport of x-momentum across the rotor layer is larger for the -0.5 K h^{-1} simulation compared to the -0.3 K h^{-1} simulation due to stronger vertical shear of the horizontal velocity. This larger negative vertical transport of x-momentum in turn requires a larger positive streamwise transport of x-momentum for the -0.5 K h^{-1} simulation compared to the -0.3 K h^{-1} simulation. We modified the manuscript as follows:

L203-219: "*In evaluating the x-momentum equation, we assume the flow is steady. This is a fair assumption because the cooling rate at the surface, which drives unsteadiness in the flow, results in small changes over the 40-min time averaging period. Furthermore, we neglect the Coriolis force in this analysis because the domain size (12000 m) is small compared to the scales affected by the Earth's rotation L=U/f=*$\mathcal{O}(10^5\ m)$*. In such a way, we consider momentum advection by the mean flow, pressure divergence, and the divergence of turbulent momentum fluxes. The turbulent fluxes are calculated from 20-minute averages of the velocity fields.*

Figure 7: Mean flow momentum advection (a), and pressure gradient (b) terms of the x-momentum equation averaged spatially across the wind plant (y-direction) and the turbine rotor layer (z-direction). The plots show the departure of the terms in the momentum equation from the flow without the GAD.



The induction zone of the wind plant is most affected by the vertical transport of zonal momentum across the rotor layer (Figure 7). The pressure gradient term remains nearly equal for the U12-C0.5 and U12-C0.3 simulations (Figure 7b), given that the drag exerted by the turbines on the flow is very similar for both cases. Conversely, the negative vertical transport of x-momentum across the rotor layer is larger for U12-C0.5 compared to U12-C0.3 due to stronger vertical shear of the horizontal velocity (solid line in Figure 7a). As the flow is forced to move above the wind plant, the vertical momentum transport is balanced by the streamwise momentum advection. The larger vertical momentum loss in U12-C0.5 compared to U12-C0.3 requires additional streamwise advection of x-momentum for the flow to remain steady (Figure 7a), producing more flow deceleration up to 10D upwind of the wind plant. Turbulence divergence plays a minor role in the region upwind of the wind plant (not shown). Though the turbulence divergence terms are larger for U12-C0.3 compared to U12-C0.5, these remain virtually unchanged for the simulation with and without the GAD, suggesting they do not contribute significantly to momentum replenishment upwind of the turbines."

We no longer consider turbulent momentum fluxes to play a major role in the induction zone of the wind plant. The turbulent momentum flux divergence terms still act to replenish momentum in the U12-C0.3 simulation, however, these terms are one order of magnitude smaller than the mean flow momentum transport and pressure divergence terms. Furthermore, we compare these terms for the simulations with and without the GAD (solid and dashed lines in figure below) and there is virtually no difference for both simulations. This implies turbulent momentum fluxes do not influence the induction zone of the wind plant.

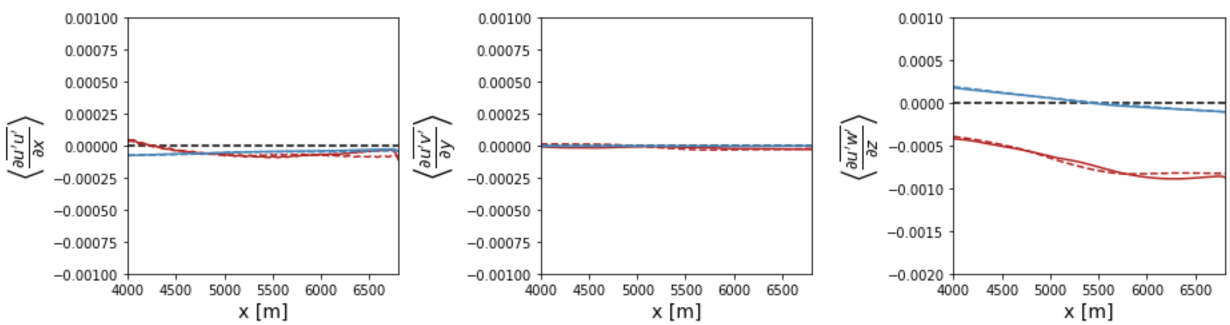


Figure i: Turbulence momentum flux divergence terms for the U12-C0.3 (red) and U12-C0.5 (blue) simulations. The solid and dashed lines represent the fluxes for the simulation with and without the GAD, respectively.

We also modified the discussion section accordingly and removed the discussion on turbulent momentum fluxes:

L337-339: "*A highly stratified atmosphere hinders turbulent motions, increasing vertical shear of the horizontal velocity and thus modifying mean momentum advection across the rotor layer.*"

Minor Comments

L 1: It’s not true that only the first row of the plant is influenced

Thank you for pointing out this imprecision, we modified Abstract and Discussion sections accordingly.

L 95: *with a smaller time step*

The manuscript was modified accordingly.

Table 1: should also have the height of the two domains

Thank you for pointing this out. We now include the domain size in the vertical direction as well.

Figure 3: The graph looks like a much longer domain would be necessary to derive at a quasi-stationary region along x. Were any sensitivity studies done for the choice of the simulation domain?

It is true that turbulence is not completely stationary in our domain. However, it remains close to stationary in the region of interest. We modified this section on turbulence as follows:

L147-154: “*We determine whether turbulence has propagated throughout the entire domain using the variance of the vertical velocity, which is calculated using 20 min time windows. Turbulence is close to steady 20 min after initializing the nested domain for U12-C0.3 (Figure 3a), and 30 min after initializing the nested domain for U12-C0.5 (Figure 3b), and results are discarded before these times. Furthermore, turbulence in the surface layer becomes quasi-stationary after x = 4000 m for both simulations (Figure 4), and we also discard results upstream of this location. Although we trigger turbulent motions up to the capping inversion, turbulence in the residual layer decays rapidly throughout the domain and becomes small after x = 4000 m. Note that minimal turbulence persists in the residual layer, as is sometimes observed in regions with flat terrain (Banta et al., 2006; Banta, 2008; Bonin et al., 2020).*

Figure 3: Time evolution of the vertical velocity variance for the U12-C0.3 (a), and U12-C0.5 (b) simulation without the wind turbines. The profiles are averaged over a 0.8 km^2 region centered at the first row of the wind plant (x = 6804 m, y = 5890 m). The perturbations of the vertical velocity are calculated using a 20-min moving average.

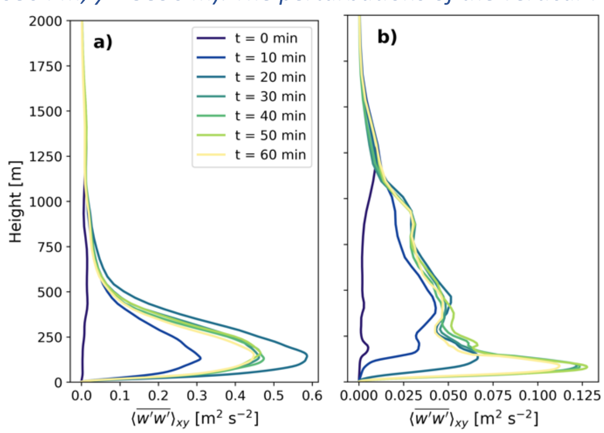
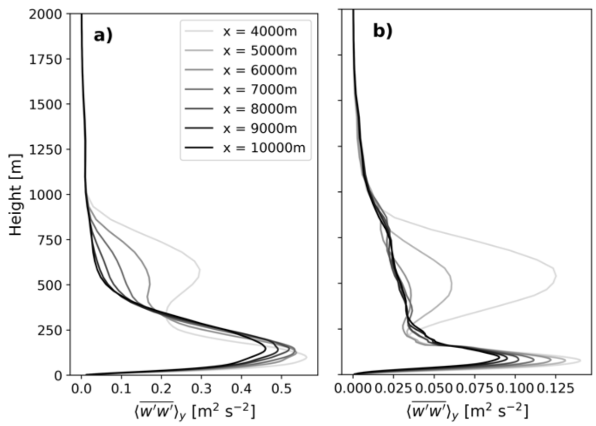


Figure 4: Evolution of the vertical velocity variance averaged in the y-direction across the domain for the U12-C0.3 (a), and U12-C0.5 (b) simulations without the turbines. Vertical profiles are color coded for each x-location in the domain and plotted in 1000 m increments.



”

We discuss the repercussion of this non-stationarity in section 5:

L241– L246: *“Our simulations also display some variability in the streamwise "background" flow (Figure 9b). The streamwise variability in the hub-height horizontal velocity results from turbulence development throughout the domain. As turbulent motions develop throughout the domain, higher momentum is transported downwards across the rotor layer, increasing horizontal velocity at hub height. This downward transport of momentum and turbulence kinetic energy is sometimes observed in stable boundary layers with low-level jets (Karipot et al., 2008; Banta et al., 2002; Mahrt and Vickers, 2002; Conangla and Cuxart, 2006; Wang et al., 2007).”*

We also analyzed in more detail the relationship between turbulence and horizontal velocity. We perform a principal component analysis on the time-averaged horizontal velocity, and TKE across the domain for the simulations without the GAD parameterization. We evaluate the spatial pattern of each variable in the *x*- and *z*-directions for the U12-C0.3 simulation, which displays the largest streamwise variations in horizontal velocity. The dominant spatial pattern for the evolution of TKE in the streamwise direction matches a dominant spatial pattern for the evolution of the horizontal velocity. The correlation in the empirical orthogonal functions of hub-height horizontal wind speed and TKE along the x-direction is 0.75 for the U12-C0.3 simulation. The empirical orthogonal functions for TKE and horizontal wind speeds at $y = 5200\ m$ that explain 24% and 18% of variance, respectively, for the U12-C0.3 simulation are shown below.

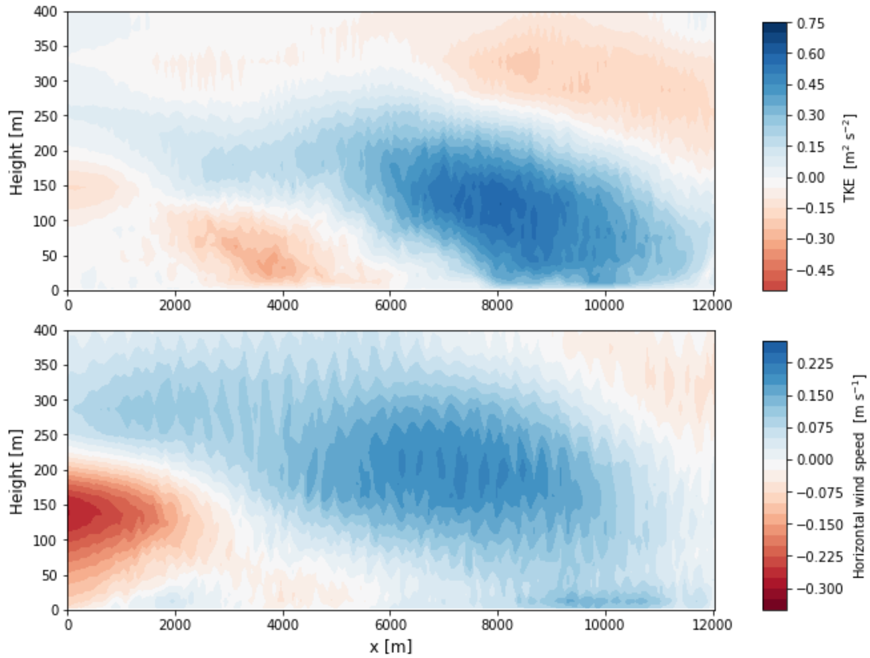


Figure ii: Empirical orthogonal functions for turbulence kinetic energy (top) and horizontal wind speed (bottom) at the center of the domain in the y-direction.

The empirical orthogonal functions demonstrate a downward vertical transport of turbulence kinetic energy and momentum, which result in an increase in hub-height wind speeds. It is not uncommon for stable boundary layers over land to display a downward transport of momentum in the presence of a low-level jet (Karipot et al., 2008; Banta et al., 2002; Mahrt and Vickers, 2002; Conangla and Cuxart, 2006; Wang et al., 2007), as indeed occurs in both of our experimental setups.

Figure 5: Why are the lines not converging to zero at 20 D?

We appreciate this comment and we added clarification in the manuscript. Our simulations show there is indeed an effect out to 20D, however, the effect is not statistically significant (95% confidence level). We define the induction zone using a statistical analysis on the velocity fields. Therefore, although other studies state the wind plant influences the velocity field upstream all the way to the domain inflow boundary (e.g. Wu and Porté-Agel, 2017; Allaerts and Meyers, 2018) based on the mean velocity, we do not consider the induction zone extends that far upstream because our statistical analysis suggests this small deviation is not statistically significant.

The figure below shows the confidence intervals on the normalized velocity deficit from the strong stably stratified simulation. This shows the mean velocity deficit is slightly less than zero at 20D, but the confidence intervals are not below zero. Therefore, the induction zone as we define it does not extend beyond 20D. Confidence intervals are not included in Figure 6 in the manuscript because we want to aggregate the results from the different cases (i.e. different stability cases, and single turbine and wind plant) into one same plot, while showing the statistical significance of our results.

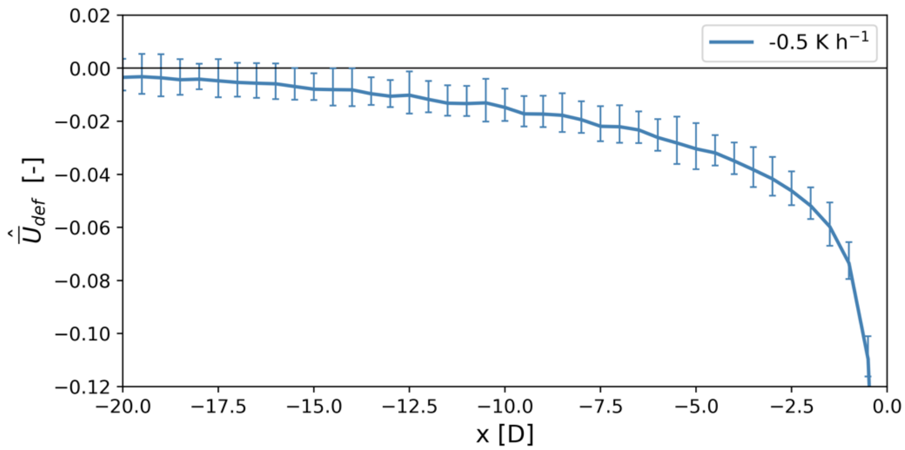


Figure iii: Normalized velocity deficit for the U12-C0.5 case. The error bars represent the 95% confidence intervals.

The manuscript is modified as follows:

L168-169: "*We define the induction zone with the statistical significance of the velocity deficit upstream of the first row of turbines in the plant*."

L186-188: "*The wind plant modifies the flow in a statistically significant manner up to 15D upstream for U12-C0.5. Conversely, there is not enough statistical evidence that the induction zone extends further than 2D upstream for U12-C0.3.*"

L346-347: "*We do not find statistical evidence of a far-reaching induction zone for the U12-C0.3 case. For this weaker stable layer, the velocity deficit is only statistically significant up to 2.5D upstream of the wind plant.*"

Figure 8b: From Figure 8a I would assume that the difference between the inversion height upstream and downstream should be a lot higher in the strongly stratified case than displayed here.

Thank you for catching this. We were plotting the potential temperature profiles at the wrong location downstream. We modified the figure to include the correct locations downstream. Please see revised Fig. 8.

L 289: See comment for L1

Thank you for pointing out this imprecision, we modified Abstract and Discussion sections accordingly.

Discussion & Conclusions: For me the chapter is too long and hard to read. I suggest restructuring the chapter. The part of the non-existent gravity waves for example could be written much shorter and more concise.

Thank you for this comment, we modified the discussion section to make it more concise and clearer.