

We first like to thank the reviewer for the comprehensive, fair and constructive comments. Please find our answers to the comments below. We marked the reviewers comments blue and our answers black.

Main concerns:

1. The authors have a tendency to state physical explanations of observed flow behaviour as a fact rather than as a hypothesis or backed-up by specific analysis. While the explanation is often plausible, the authors should more clearly indicate when they are presenting a hypothesis or whether they actually have evidence supporting their claim.

Checking the mentioned statements revealed that some of them are not correct and have to be revised.

For example

- Line 166: “Due to the self-reinforcing behaviour of this process, the crosswise variations can build up quickly without y-shift, even if the wind farms have a distance of 15 km to the recycling plane.” It is unclear to me whether you actually saw this in preliminary calculations or if this is a hypothesis?

We added a sentence on that:

"Test simulations without y-shift showed that, due to the self-reinforcing behavior of this process, the crosswise variations of wind speed can build up to +- 2%."

- Line 295: “This BL height dependency occurs because a thicker BL contains more kinetic energy that can be transported down to the wind turbine level by turbulent vertical mixing than a shallow BL.” Do you have evidence supporting this hypothesis? If this is something that will be discussed later on, please mention so explicitly.

Yes, the vertical kinetic energy flux is greater for the thicker BL, as can be seen in Fig. 10.

Rephrase: "This BL height dependency occurs because the turbulent vertical kinetic energy flux is greater for the case with the thicker BL, cf. Fig. 10 c and d."

- Line 296: “The wake length and speed deficit of small wind farms (e.g. N-1 and N-2) is relatively unaffected by the BL height because the wind farm induced internal boundary layer does not reach the top of the BL.” Did you investigate the internal

boundary layer development? If not, how can you know that this is the reason? It would be interesting to add IBL development to figure 7 (and extend the analysis to small wind farms as well, see below) to support this hypothesis.

We now included IBL-lines in Figure 7 and added a new figure with vertical cross sections through a small wind farm ( $y = 50$  km), that shows the IBL behaviour (Figures shown later under reviewers point 2). We rephrased line 296 ff like this:

"The wind speed deficit and the wake length of small wind farms (e.g. N-1, N-2 and N-3) is relatively unaffected by the BL height because the wind farm induced internal BL does not reach the inversion layer (NBL cases) or only reaches it several 10 km downstream of the wind farm trailing edge (CBL cases), cf. Fig. 7.5\_new. Consequently, the BL height only affects further wind speed recovery (e.g. to 9.5 m/s) in the far wake of small wind farms. For example, the wind speed recovery from 9 to 9.5 m/s in the wake of N-2 takes longer for the case CBL-700-7D (~40 km) than for the case CBL-1400-7D (~15 km), cf. Fig 4 c and d."

- Line 314: "The case SBL-300-7D covers several flow features that cannot be seen in the other cases." Are these flow features not visible because they do not occur or because they are smaller? Did you verify quantitatively whether there is any flow deceleration in front of the wind farm in neutral or unstable conditions?

They are smaller. We rephrased this part:

"The case SBL-300-7D covers several flow features that are not as significant in the other cases."

In line 320 we added information about the flow blockage in the other cases:

"At a distance of 2.5 D upstream of the first wind turbine row of the wind farms in Zone 3 the wind speed is reduced by approximately 10% relative to the inflow wind speed. For all other cases the speed reduction is approximately 2%."

We carefully checked the Froude number (according to the definition in Wu and Porté-Agel, 2017) for all cases and found that this parameter is not suitable for predicting flow blockage (at least for the 5 cases that we investigated). The flow is subcritical (and thus indicating flow blockage, according to Wu and Porté-Agel) only in the case CBL-1400-7D, but in this case the flow blockage is very small. Additionally we mixed up subcritical with supercritical, probably due to a mistake in Table 2 of Wu and Porté-Agel (2017) (they mixed up the Froude numbers). So we rephrased the text from line 320 to 323 as:

"Wu and Porté-Agel (2017) reported 11 % speed reduction 2.5 D upstream of the first turbine row of a 20 km long wind farm in a CNBL with a FA stratification of  $\Gamma = +5 \text{ K km}^{-1}$ . However, for  $\Gamma = +1 \text{ K km}^{-1}$  they reported a speed reduction of only 1.2 %, because the flow is supercritical (Froude number  $F_r < 1$ ). Using the same definition<sup>2</sup> as in Wu and Porté-Agel (2017), the Froude number in

the case SBL-300-7D is  $Fr = 1.47$ , indicating a supercritical flow. This should, according to the reasoning of Wu and Porté-Agel, result in a weak flow blockage, which does not correspond to the significant flow blockage observed in the case SBL-300-7D. The only case that is subcritical (and should thus show significant flow blockage) is CBL-1400-7D ( $Fr=0.81$ ), but in this case the flow blockage is only very weak. Hence, for the cases that are investigated in this study, the Froude number, as defined by Wu and Porté-Agel, is not an appropriate parameter for predicting flow blockage."

- Line 418: "Because the mean wind speed inside the BL decreases in the streamwise direction, the IL must be displaced upwards in order to maintain a constant mass flux inside the BL." Note that the mass flux can also be conserved by means of an acceleration above the wind farm (below the IL) or by airflow to the sides. In your results, the IL is displaced upward, but that does not necessarily mean that this will always be the case (e.g., a stronger capping inversion might lead to flow acceleration or flow diversion to the sides). Please rephrase.

Yes, that is not precise enough, right. But if we write 'mean wind speed inside the BL', than that already includes the effect of acceleration above wind farm (below IL). We rephrased this to:

"The IL displacement is a result of the reduced wind speed in the bulk of the BL: To obtain a divergence-free flow inside the BL, the wind speed reduction (streamwise convergence) is compensated by vertical divergence (IL displacement) and crosswise divergence (flow around the wind farms)."

- Line 532: "The wind farm efficiencies for the NBL-case are greater because the inflow wind speed in the bulk of the BL is higher for the NBL-case than for the CBL-case." Higher bulk wind speed causing higher efficiencies seems a very plausible explanation, but I'm not sure you can deduce this conclusion from your results with 100% certainty. Maybe rephrase?

Yes, this is a hypothesis and should not be written as a statement. Rephrase (from line 531):

"A comparison between the cases NBL-700-7D and CBL-700-7D shows that greater wind farm efficiencies are obtained for the NBL, although better efficiencies are expected for the CBL due to the better vertical mixing. Comparing the wind speed profiles of this cases (cf. Fig. 3) shows that the inflow wind speed in the bulk of the BL is higher for the NBL than for the CBL, which is probably the reason for the higher wind farm efficiencies."

- Line 596: "As stated earlier, the power output of infinitely large wind farms is

determined by the energy input of the pressure gradient. Hence, the power output of infinitely large wind farms does not depend on the BL height, at least for this idealized setups with a stationary CBL inflow.” Did you actually run simulations of infinitely large wind farms with stationary CBL inflow to confirm this hypothesis?

No, we only expect this behaviour. Rephrase:

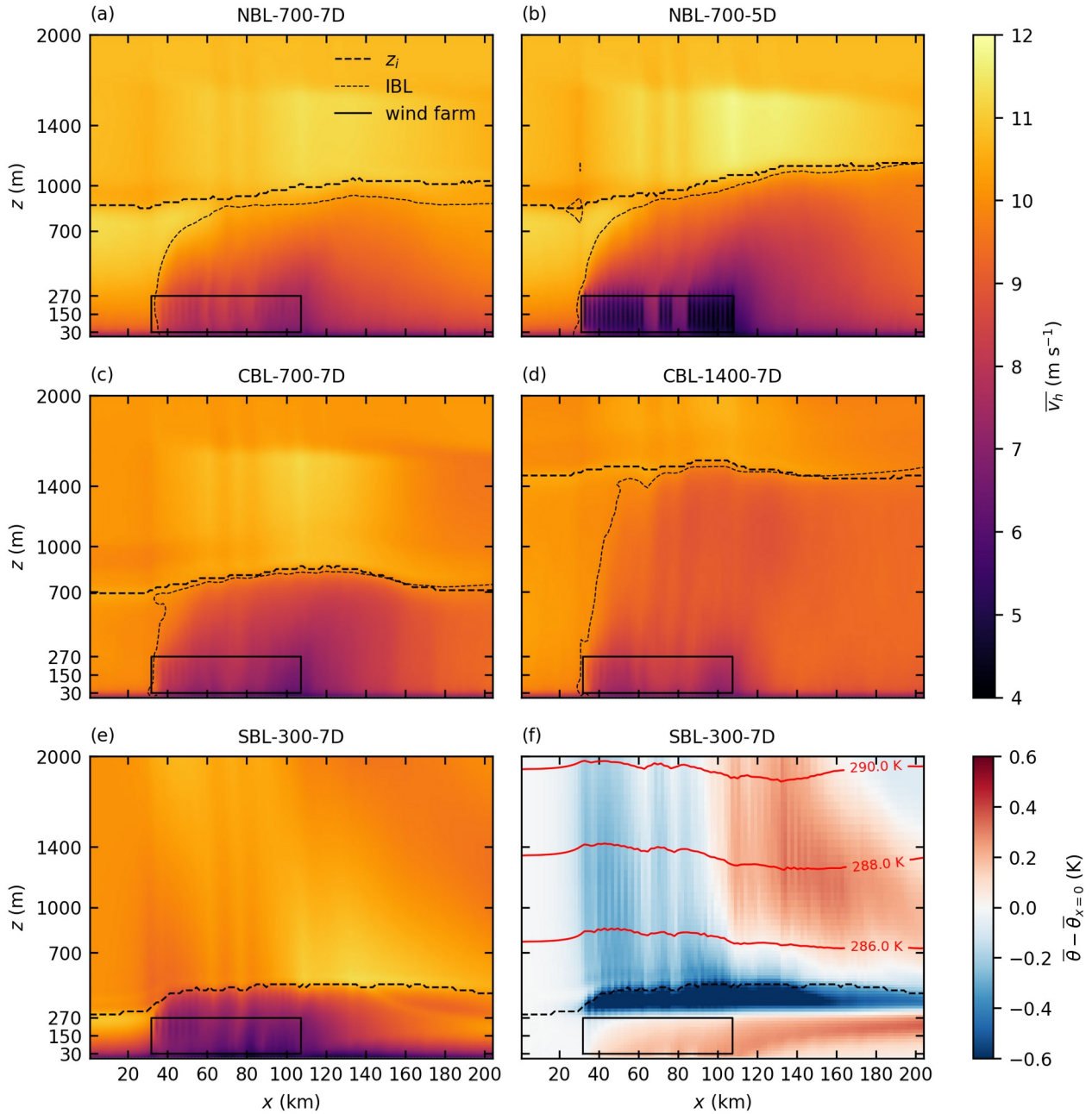
"As stated earlier, the power output of infinitely large wind farms is determined by the energy input of the pressure gradient, which does not depend on the BL height. Hence, the power output of infinitely large wind farms is expected not to depend on the BL height, at least for this idealized setups with a stationary CBL inflow."

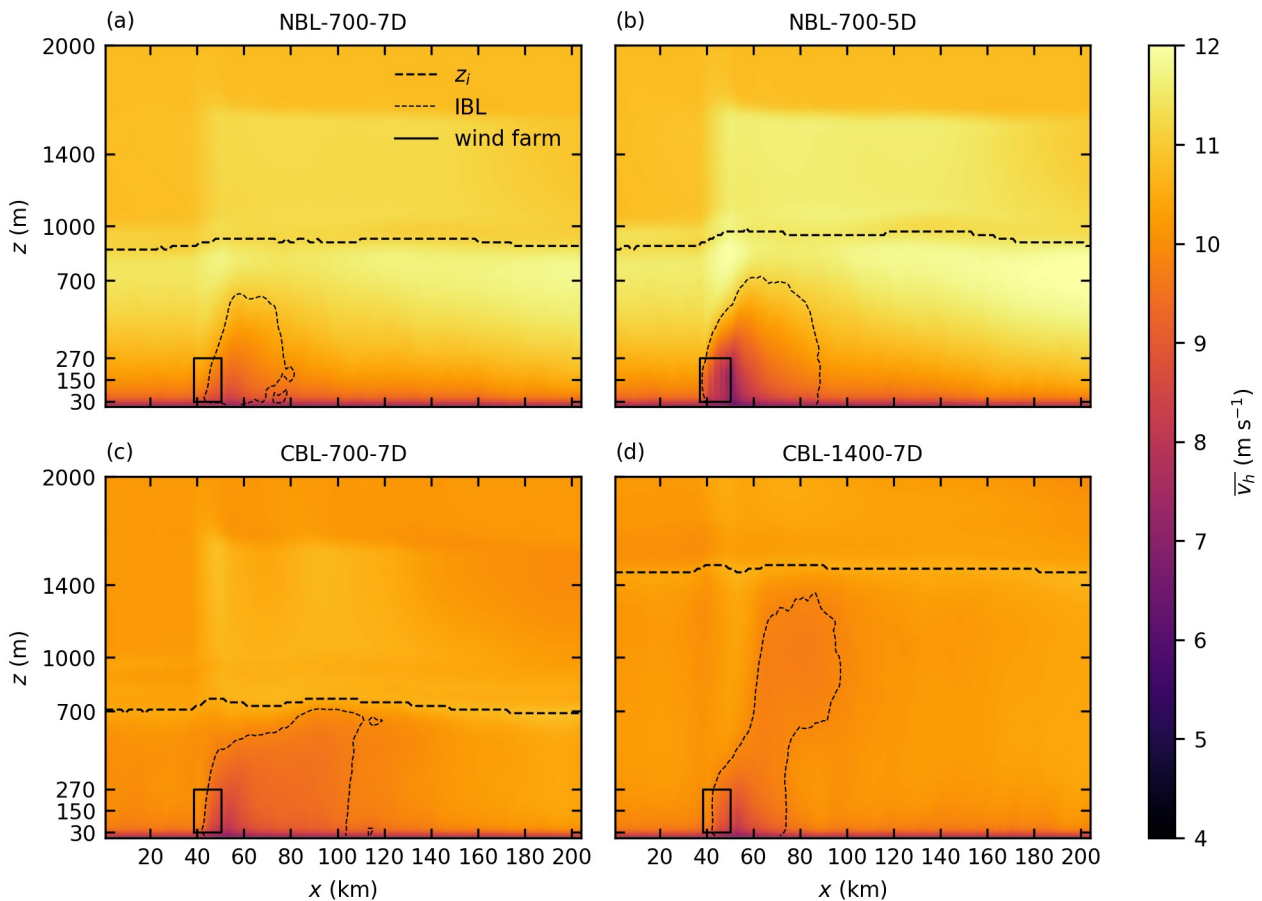
2. The authors made an effort to cover an enormous area in their simulations, but the bulk of the analysis is based on the large wind farm cluster in zone 3 (e.g., many analyses look at cross-sections at  $y=120$  km). I think it would be useful to include more analyses of the smaller wind farms to be able to contrast the behaviour of “small” and “very large” wind farms, as the difference in flow behaviour between small and very large wind farms appears to be one of the main messages of the paper. For example, figures 5, 7, 8 and 10 focus solely on the very large wind farm cluster and there is no counterpart analysis for one of the small wind farms. Moreover, it is interesting to see in fig 10 that the vertical kinetic energy flux continues to decrease over the wind farm and does not reach a plateau, unlike what was found by Allaerts and Meyers (2017, J. Fluid Mech.). I wonder whether this is because you are looking at much longer wind farms, and therefore a similar analysis for a smaller wind farm would be useful. As a matter of fact, line 615 states that “These results show that the power output and the wake of very large wind farms behave very differently compared to small wind farms.” At least in terms of the power output, you haven’t shown this in the paper even though the data is available.

This is a legitimate objection and we agree that it is a good idea to show more results for the small wind farms. However, the paper is already very long and includes many large figures. Showing figures 5, 7, 8 and 10 additionally for small wind farms and discussing these figures would lead to at least 50 pages in total, which is too much in our opinion. As a compromise, we decided to add a figure like Fig. 7 (see below), that shows the BL development for a small wind farm ( $N=2$ ,  $y=50$  km). We think that this additional material on small wind farms has to be sufficient, also because the main focus of this paper is on the effects of very large wind farms. The investigation of

differences between small and large wind farms will definitely be part of a follow up study that we are currently working on.

We added a paragraph which discusses the new figure in the revised manuscript





3. The authors decided to take a simplified approach to the energy budget analysis, but I believe too many terms are left out to support the discussion. A more detailed or even full energy budget analysis seems warranted.

a. Line 539: “the extraction of kinetic energy by the wind turbines can be compensated for by two sources of energy: (1) Vertical turbulent flux ... and (2) Work done by the geostrophic pressure gradient.” This is not entirely correct, as the extraction of kinetic energy can also be compensated by a favourable perturbation pressure gradient induced by gravity waves, which acts on top of the fixed geostrophic pressure gradient.

This is true, we added it in the list:

The extraction of kinetic energy by the wind turbines can be compensated for by three sources of energy:

”

1. Vertical turbulent flux of kinetic energy at rotor top level,  $W_f$



2. Work done by the geostrophic pressure gradient on the flow below rotor top level,  
 $W_{\text{gpg,wt}}$

3. Work done by the perturbation pressure gradient on the flow below rotor top level,  
 $W_{\text{ppg,wt}}$

"

We also added a equation for  $W_{\text{ppg}}$  and a graph for the work done by the perturbation pressure gradient in Fig 10 / now Fig. 11 (see further below).

The power density of the energy input by the perturbation pressure ( $p^*$ ) gradient on the flow below rotor top level  $z_t = 270$  m is calculated as:

$$W_{\text{ppg,wt}}(x) = - \int_{z=0}^{z_t} \frac{\partial p^*(z)}{\partial x} \bar{u}(z) + \frac{\partial p^*(z)}{\partial y} \bar{v}(z) dz |_{y=120 \text{ km}} \quad (16)$$

b. Line 561: "The wind turbines extract approximately 70% of the total energy input  $W_{\text{kef}} + W_{\text{gpg,wt}}$ ." I have a problem with this statement as part of the energy might come from flow deceleration inside the wind farm (i.e., the advective term). By decelerating the flow releases kinetic energy that can be used by the wind turbines or that can be dissipated. The total energy input needs to take this advective source term into account. As shown by Allaerts and Meyers (2017, J. Fluid Mech.), the combination of advective sources and perturbation pressure can be up to 15% of the energy input.

Yes that is true. We added a note on this in the comparison with the infinite wind farms:

"The wind turbines extract approximately 70 % of the total energy input  $W_{\text{kef}} + W_{\text{gpg,wt}} + W_{\text{ppg,wt}}$ , which is a relatively large value. Johnstone and Coleman (2012) and Abkar and Porté-Agel (2014), who analyzed the energy budgets for an infinite wind farm in a NBL, reported that 35 % and 45 %, respectively, of the energy input by the geostrophic pressure gradient is extracted by the wind turbines. The difference could be explained by the fact that, for the finite wind farms in this study, additional energy is provided to the turbines by the divergence of horizontal advection of kinetic energy (flow deceleration). Additionally, the differences could be explained by the low Reynolds number of  $Re = 1000$  in the simulations of Johnstone and Coleman (2012) and the higher roughness length of  $z_0 = 0.1$  m in the simulations of Abkar and Porté-Agel (2014)."

Except for the case SBL-300-7D, the power input by the perturbation pressure averaged over the entire farm length is approximately zero. Thus the 70 % statement is still correct for the NBL case discussed here.

c. Line 601: “As this influx is proportional to the BL height, much more energy is available for the wind turbines in the case CBL-1400-7D than in the case CBL-700-7D. This results in a higher kinetic energy flux ...” I have difficulty with accepting that more energy available automatically leads to higher energy fluxes without actual data or equations to support this. It would for example be interesting to look at the kinetic energy content of the flow above the wind farm and see how that decreases.

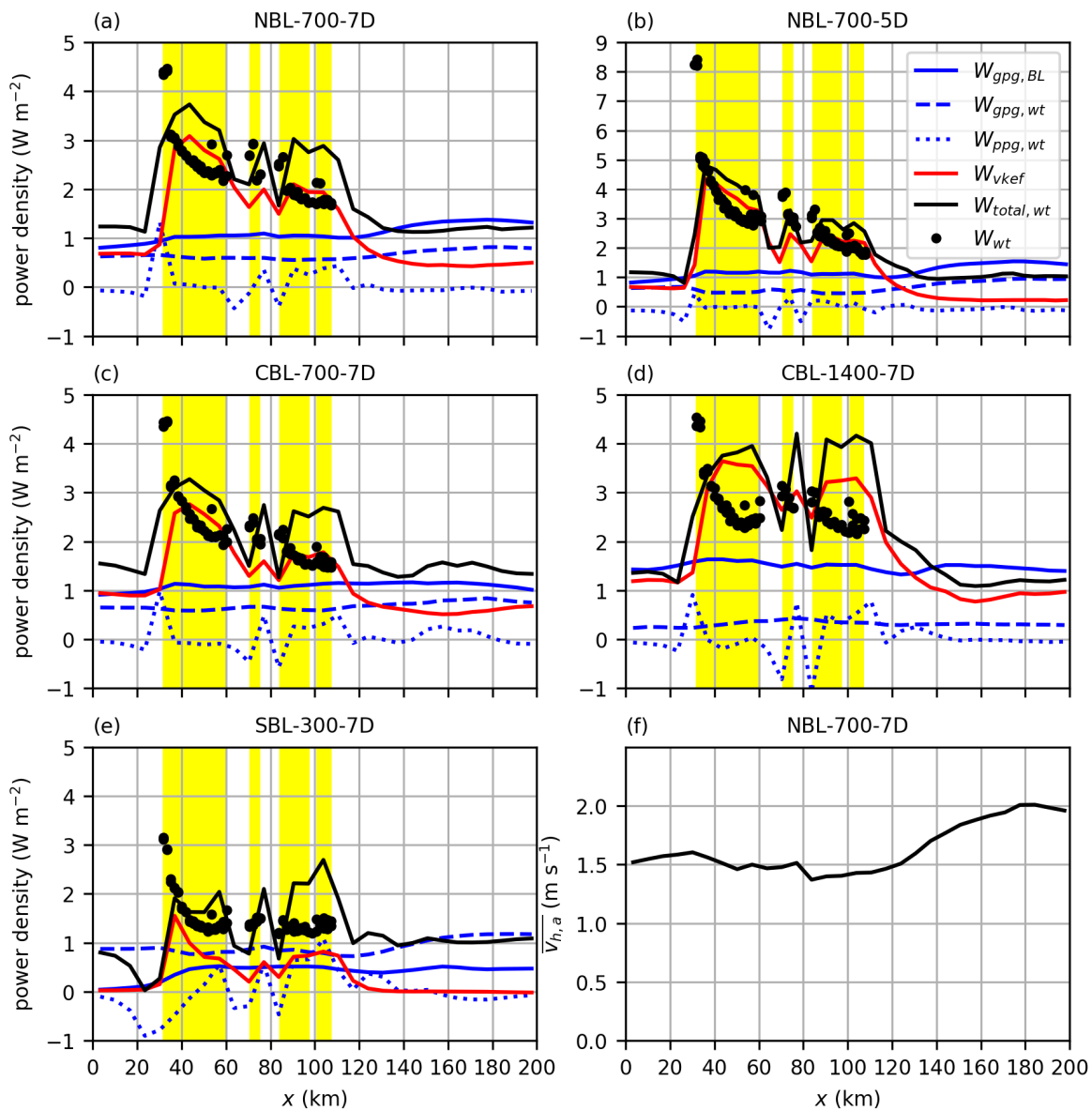
We deleted this reasoning and just state that a thicker BL results in a greater vertical kinetic energy flux and turbine power densities.

"However, for very large, but finite-size wind farms, as in this study, the power output depends significantly on the BL height, as it is shown in Fig. 10c and d. The vertical kinetic energy flux is greater and decays slower for the thicker BL (CBL-1400-7D), resulting in higher turbine power densities."

Line 612: “This redistribution is done by a favourable perturbation pressure gradient inside the wind farms and reaches power densities of approximately 1 W/m<sup>2</sup> (not shown in Fig. 10).” Energy redistribution due to pressure gradients in the case of wind farm blockage is a major topic nowadays, so showing the actual perturbation pressure gradient contribution could be really helpful in this ongoing debate. I see no reason to exclude it from the figure.

We have included the perturbation pressure contribution in Fig. 10 (now Fig. 11, see below).





Other scientific questions:

- Abstract and result sections: When talking about power density for very large wind farms, the work by Enrico Antonini comes to mind (see, e.g., Antonini and Caldeira, 2021, papers in Proc. Nat. Sci. Acad. and Applied Energy). How does your work compare to their estimate of the power density limits for large wind farms?

We thank the reviewer for mentioning this article, which we haven't been aware of. Antonini's results show a power density of 1-1.5 W/m<sup>2</sup> for G=12 m/s and lat = 46.1°, which is consistent with our findings. We referred to this article in line 586ff in the results section:

"The energy input by the pressure gradient ( $W_{pg,wt} + W_{pg,BL}$ ) achieves power densities of only 1 – 2 W m<sup>-2</sup>, which is consistent with the geophysical limits to power densities of large wind farms found by Antonini and Caldeira (2021), who reported approximately 1.5 W m<sup>-2</sup> for a latitude of 46° and a geostrophic wind speed of 12 ms<sup>-1</sup>. This power density is much smaller than the power density achieved by the first-row wind turbines. As the case NBL-700-5D shows (Fig. 11b), a reduction of the turbine spacing from  $s = 7 D$  to  $s = 5 D$  approximately results in a doubling of the power density of the first-row wind turbines (from 4.5 W m<sup>-2</sup> to 8.5 W m<sup>-2</sup>), but the power density of the last row wind turbines is as low as for  $s = 7 D$ ."

Abstract and result sections: the clockwise flow deflection above the boundary layer is stated as one of the major findings occurring above very large wind farms and not for small wind farms. However, as mentioned on line 464, the effect might be overestimated due to conservation of mass flux in the FA. Moreover, the sensitivity to the Rayleigh damping layer is said to be out-of-scope, so there is no way of knowing how important this effect is, if it occurs at all. This seems quite a controversial result to me, so I wouldn't highlight it without additional sensitivity analysis or explicitly mentioning the uncertainty about this effect.

In line 477 we state that this effect (as well as the speed up above the BL) might be overestimated. Our current investigations show that this effect also occurs if the vertical distance to the Rayleigh-damping layer is much greater. So we are confident, that this effect is real.

Line 196: "While steady-state turbulence is reached after only a few hours, achieving a steady-state mean flow can take several days." How did you assess stationarity of turbulence and mean-flow quantities? Which quantities did you track and when did you consider it to be steady state (e.g., time rate of change less

than 1% per hour)?

We added the following lines to clarify how we define steady-state:

“While steady-state turbulence is reached after only a few hours, achieving a steady-state mean flow can take several days. Here, we declare the mean flow as steady if the oscillation amplitude of the hub height mean wind speed is less than 0.5 % and declare the turbulence as steady if the change in friction velocity is less than 2 % in 4 h.”

Figure 7 and 8: It is interesting to see that the IL returns to its original height for CBL cases but not for the NBL cases. Is this due to faster wind farm wake recovery in the CBL cases, allowing horizontal convergence?

With this setup and the available data we can only guess why this happens so that we decided to not further comment in this effect. It is difficult to separate different mechanisms, that could be responsible for this effect. For example the heterogeneity in y-direction makes an analysis difficult. In the follow-up study we use a more idealized setup that will allow us to investigate this effect much better.

Figure 8:  $x = 120\text{km}$  is identified as the near wake. Could you say how far downstream that is from the wind farm trailing edge, i.e., what is the x-position of the last turbine at this y position?

We added this information in the figure caption:

"The wind farm trailing edge is located at  $x=108\sim\text{unit}\{\text{km}\}$ "

Figure 8: The wind speed excess in the FA for NBL-700-5D is greatest at 120km and smaller (but not zero) at 180km. I'm surprised to see that the wind direction at 180km is more ageostrophic than at 120km. This means that there should be some other force at play other than the force balance between pressure gradient and Coriolis force. Or is the flow not in a geostrophic balance due to some kind of inertial oscillation in space (almost like a Rossby wave I suppose, a spatial wave supported by the Coriolis effect?) Can you elaborate? Same goes for NBL-700-7D and SBL case.

Yes exactly, it is some kind of inertial oscillation in space. We added a sentence on this in line 475:

"Note that the highest wind speed excess occurs at  $x=120\sim\text{unit}\{\text{km}\}$  but the highest deflection angle occurs at  $x=180\sim\text{unit}\{\text{km}\}$ . This effect can be interpreted as an inertia oscillation in space (along  $x$ ), with the deflection angle being  $90^\circ$  phase shifted relative to the wind speed excess."

Figure 8: For some cases, the vertical axis includes part of the Rayleigh damping layer. This is particularly visible for the neutral cases. The vertical profiles show non-physical behaviour in the Rayleigh damping layer. I think it would be useful to remind the reader of that by adding a line or a gray zone to indicate the location of the damping layer, or by only showing results below the damping layer.

We now set the upper limit of the z axis to below the damping layer.

Line 505: Why not obtain the reference power for a single turbine using the same inflow profiles of the corresponding case, rather than using the neutral case as a reference for all? Now the reference power does not account for differences in wind shear or wind veer under non-neutral stability conditions, so you are attributing effects due to stability at individual turbine level to wind farm efficiency.

We now also ran simulations for CBL 700 CBL, 1400 and SBL 300 cases to obtain reference powers for these cases. Only the SBL case shows a significantly different reference power and thus higher wind farm efficiency (see table below). We added the numbers in Table 2 and in the text.

Reference power (corrected for power overestimation):

NBL: 12.56 MW

CBL 700: 12.51 MW

CBL 1400: 12.53 MW

SBL: 11.45 MW

--> Reference power is smaller for SBL case. This results in higher wind farm efficiency:

case	Wind farm efficiency (old/new)	
	N-1	Zone 3
NBL-700-7D	0.87/0.87	0.58/0.58
NBL-700-5D	0.77/0.77	0.41/0.41
CBL-700-7D	0.86/0.86	0.54/0.54
CBL-1400-7D	0.86/0.88	0.63/0.64
SBL-300-7D	0.61/0.66	0.42/0.46

We also corrected the values in the text. The overall message and conclusion stays the same.

Line 591: “Figures 10c and d show that a doubling of the BL height has approximately no effect on the energy input by the pressure gradient on the undisturbed inflow.” I’d say there is a clear difference:  $W_{gpg,BL}$  is 50% higher in CBL-1400-7D than in CBL-700-7D. Are you maybe referring to the sum of  $W_{gpg,BL}$  and  $W_{gpg,wt}$ ? Please clarify.

Yes, we meant the sum.

"Figure 10c and d show that a doubling of the BL height has approximately no effect on the energy input by the geostrophic pressure gradient ( $W_{gpg,wt} + W_{gpg,BL}$ ) on the undisturbed inflow."

Line 629: “Some tuning of the domain height and the boundary conditions was necessary to capture this phenomenon correctly.” How did you assess whether you capture gravity waves correctly?

L441 - 449

We rephrased that: "Some tuning of the domain height and the boundary conditions was necessary to obtain stable simulation results." We did not investigate the interaction of waves with the top and inflow boundary in detail but in the following sentence we highlight that it is an important task to find best practice rules for simulation setups that capture this phenomenon as realistically as possible.

Line 640: “..., the turbine spacing of very large wind farms should be at least 7 rotor diameters to achieve an acceptable wind farm efficiency.” What value of wind farm efficiency do you consider acceptable? Isn’t this up to the wind farm developer? Similar statements were made in the abstract, so please adjust accordingly.

That is true. We rephrase this to a more general statement:

“Because this power density is only as small as  $2 \text{ W m}^{-2}$ , high wind farm efficiencies can only be achieved by large turbine spacings.”

Minor/technical•••••comments:

Line 44: For your information, the Lillgrund wind farm has also been simulated with LES by several authors. Consider adding references to, e.g., Churchfield et al. (2012,

AIAA) and Nilsson et al. (2015, Wind Energy) to make this list more complete.

We added the mentioned articles and thank the reviewer for this hint.

Line 60: I'm curious as to how much computational resources (in terms of core hours, e.g.) were needed for simulations of this scale. Could you comment on this in the paper?

We added information about number of cores and computing time, also due to a comment of the other reviewer.

"The simulations were carried out on 5120 cores on one of the supercomputers of the North-German Supercomputing Alliance (HLRN). A simulation required a Wallclock time of 25 to 50 h."

Equation 2: The buoyancy term makes use of angular brackets, but the use of these is not defined in the text. Do they indicate horizontal averaging? How do you do that in the main domain which is horizontally not homogeneous?

We added the information that angular brackets indicate horizontal averaging. This rule is also applied for the main domain. The (in our opinion worse) alternative would be to use a fixed (constant in time) reference temperature profile.

Equation 3: The transport equation for internal energy is stated in terms of the potential temperature, while the buoyancy term in equation 2 uses virtual potential temperature. How did you relate these two quantities? Note that neither of these quantities/symbols are defined in the text. Please define all symbols in the text.

Since humidity is not considered in the simulations,  $\theta = \theta_v$ , so we changed  $\theta_v$  to  $\theta$  in equation 2. We also added a definition in the text.

Line 78: You mention that overbars are used to indicate filtered quantities, but the LES equations 1-3 contain no overbars for the velocity components nor for the potential temperature. The notation is inconsistent. Either add overbars consistently in the equation, or mention explicitly that overbars are not shown for these parameters (so only shown for the subgrid-scale stress and heat flux) and that  $u_i$  etc. correspond to the filtered quantities.

We now mentioned that these are filtered quantities and did not add overbars.

Line 142: Please describe what you mean with radiation boundary conditions. Does this simply mean zero velocity gradients and a fixed pressure value, or do you impose more intricate boundary conditions?

We added a reference:

"Details about the radiation BC can be found in Miller and Thorpe (1981) and Orlanski (1976)."  
The flow field is advected past the boundary with the maximum allowed speed according to the CFL criterium.

Section 2.4: What boundary condition did you set for the potential temperature at the top of the domain? Please specify in the text.

We added this info together with the pressure BC in line 145:

"At the domain top a Neumann boundary condition is set for the perturbation pressure and the vertical potential temperature gradient is kept constant."

Line 225: Basically, subsidence velocity is chosen such that  $w_{\text{sub}} * \Gamma = d\theta_0/dt$ . Maybe add the equation to clarify how  $w_{\text{sub}}$  is set.

We added this equation.

Line 241: "..., as this is the correct surface forcing method for SBLs." The word correct sounds quite strong and implies that applying a surface heat flux is incorrect. Rather, Basu et al. show that applying a surface heat flux can only represent weakly stable conditions and the surface forcing method is preferred. I wouldn't go as far as saying that there is a correct method for applying boundary conditions for SBLs.

We rephrased that to:

"A Dirichlet-condition is applied for the surface temperature, because prescribing a surface heat flux can lead to unphysical results (Basu et al., 2008)"

Section 2.5: This section doesn't add much, so I would try to integrate it elsewhere. Particularly the discussion on how to calculate resolved turbulent fluxes is, as far as I know, standard practice in LES, so I don't think it needs to be mentioned here. I'm not familiar with the term temporal eddy-correlation method, is that something you



introduced yourself? If not, add references to where this term is used as well.

We think that it is a good idea to clearly show how the fluxes and  $v_h$  are calculated and would like to keep this small section. We skipped the term "temporal", because "eddy correlation method" usually means correlation of timeseries and not correlation of 2d or 3d spatial fields.

Line 335: "Because the wind farms in this study also have a finite size also ..." The word "also" appears twice in this sentence, I think one of them can be removed.

Yes, is corrected.

Caption of figure 5: "The forces are normalized ... and are horizontally averaged over one turbine spacing along x and y." The main text speaks about horizontal averaging over 4 spacing in x and 2 in y. Please clarify.

One spacing is correct. We changed it in the main text.

Line 361: Could you state by how much percent the streamwise geostrophic pressure gradient force has increased? It is hard to see in the figure and a quantitative value would be instructive.

It's 80%, We added it.

Line 406: "The convective velocity scale is greater in the case CBL-1400-7D than in case CBL-700-7D ..." It would be useful to mention the definition of the convective velocity scale here (or add the actual value to table 1) to show why it differs between these two cases.

We added the definition in the text.

Equation 13 and fig 10: Use consistent naming:  $W_f$  or  $W_{v_{kef}}$ . There is also a typo on line 561  $W_{kef}$ .

We changed all terms to  $W_{\{v_{kef}\}}$

Line 637: "The achieved power density of turbines in the upstream part if the wind farms ..." This should be "... of the wind farms ..."

We corrected that.

We first like to thank the reviewer for the comprehensive, fair and constructive comments. Please find our answers to the comments below. We marked the reviewers comments blue and our answers black.

It should be noted the very large computational demand to achieve these results: each simulation comprises 2088 wind turbines with mesh sizes ranging from  $\sim 7.4 \times 10^9$  to  $\sim 1 \times 10^{10}$  cells, which is indeed remarkable (so it will be very informative to reveal the computational expense).

In line 134 we added a sentence on that:

"The simulations were carried out on 5120 cores on one of the supercomputers of the North-German Supercomputing Alliance (HLRN). A simulation required a Wallclock time of 25 to 50 h."

I recommend this work for publication but please provide an answer to the following questions and comments:

L89 "it is possible to prescribe a surface heating . . ." is this just a choice or a necessity, in view of the issues with the direct prescription of the surface flux as described in Basu et al. (2008)?

The importance of prescribing a surface heating rate instead of a heat flux for SBLs (and CBLs) is highlighted later in sections 2.4.2 and 2.4.3.

In section 2.4.3 (SBL) we rephrased line 241 to:

"A Dirichlet-condition is applied for the surface temperature, because prescribing a surface heat flux can lead to unphysical results (Basu et al., 2008)"

L94, Could you please briefly comment why the overestimation and how this is fixed? Also, would the use of a different distribution length along the smearing direction help to remedy this (a different value of the standard deviation than  $1\Delta x$  in the Gaussian distribution)? I think it is useful to provide these details here (even crucial in the case of the overestimation) as the reference you provide is not in English. Also note that in the archive of Meteorologische Fortbildung at the following link, the issue of the reference you cite appears as from 2015 instead of 2016: [https://www.dwd.de/EN/ourservices/pbfb\\_verlag\\_promet/archiv/archiv\\_promet.html](https://www.dwd.de/EN/ourservices/pbfb_verlag_promet/archiv/archiv_promet.html)

Yes, we kept this part short. We have now added some information and changed the year to 2015:

"The ADM is described in detail by Steinfeld et al.(2015) and Wu and Porté-Agel (2011). To avoid numerical instabilities, the disc element forces are distributed to the neighboring grid points by a three-dimensional Gaussian smearing kernel, which is approximated by computationally less expensive 4th-order polynomial. The smearing kernel has a default radius of  $2\Delta x$ , reaching approximately 78 grid points. The otherwise two-dimensional actuator disc is enlarged in the axial and radial direction by the smearing, resulting in a power overestimation of 26.8 %. The power overestimation can be reduced to 12.5 % by setting the kernel radius to  $1\Delta x$ , reaching approximately 10 grid points, without any numerical instabilities. The thrust coefficient is overestimated by 2% for  $2\Delta x$  and underestimated by 4% for  $1\Delta x$ . As a compromise, the smearing kernel radius is set to  $1\Delta x$  for this study. The wind turbine power output is corrected for the power overestimation by a factor of 1/1.125 before entering the wind farm power output analysis."

L127 and Table 1: should one assume that 225 deg is the wind direction at hub height obtained with a  $u_g$  at direction  $\alpha$  as shown in the table?

Yes, it is hub height wind direction. We added that:

"The wind direction at hub height is set to 225° by tuning the geostrophic wind direction  $\alpha$  appropriately (cf. Table 1). Southwest-wind is one of the most common wind directions in the German Bight."

L133, why have you chosen to use a uniform cell distribution in the horizontal plane? This question arises from the fact that if cell stretching was used away from the turbines, the savings could have been used to achieve a higher refinement either around the rotor and wakes or for the near ground zone, where the mesh requirements for a well-resolved LES are higher.

PALM allows only for a vertical stretching. In the horizontal directions the cells have to be uniform. The nesting feature of PALM could have been used for sparing resources, but we decided against this option in order to avoid any numerical effects on our results that might be introduced by the nesting. Thus we decided to use this uniform grid with 20 m grid spacing.

L134, why is this resolution enough? In the absence of a numerical or experimental evaluation, please provide an argument. This element is crucial as the recovery of the individual wakes depends on the amount of TKE resolved within and around.

There is a numerical evaluation in Steinfeld (2015) that shows that 5D behind the rotor (which is the smallest turbine spacing used in our study and due to the staggered configuration the first wake-turbine interaction happens at 10 D) the wind speed profiles for a grid spacing of below 10 grid points per rotor diameter are the same as for finer grid spacings. They used the 5MW NREL turbine with  $D=126\text{m}$  and results with 16 m grid spacing were still good at 5D, Fig. 4-7.

We added this information:

"This grid spacing yields a density of 12 grid points per rotor diameter, which is enough to resolve the most relevant eddies inside the wind turbine wakes. As Steinfeld (2015) showed, even 8 grid points per rotor diameter are sufficient to obtain a converged result for the mean wind speed profiles 5D behind the turbine."

L134 (and L204), also with regard to the cell resolution, please comment on its the adequacy in terms of the amount of subgrid TKE with respect to the total. In other words, how "well resolved" is the LES, especially for the wake region? L204, please describe how the subgrid convergence study demonstrated that 20 m was the adequate cell resolution in the horizontal directions.

The ratio of SGS-TKE to total TKE and SGS-momentum flux to total momentum flux is smaller than 10% for the SBL precursor run.

We added this information:

"Test simulations with 10 m and 20 m grid spacing showed that a grid spacing of 20 m is sufficient, if this SGS-model is used (less than 1 % difference in wind speed maximum and less than 5 % difference in BL height), whereas the results are more grid spacing sensitive (2 % difference in wind speed maximum and 20 % difference in BL height) if the standard-SGS model of PALM is used. For the SBL precursor run the ratio of SGS-TKE to total TKE and SGS-momentum flux to total momentum flux is smaller than 10%, except for the lowest grid point."

For SGS-TKE in the wake see the comment on line 386.

L165 and results section: it is mentioned that blockage from the parks is observed to reach the recycling plane, did the authors observe any dependence of the blockage with respect to the size of the domain (regardless of the fact that the boundary conditions at the sides are periodic)?

We have not tested this for this setup. But we have carried out tests for different domain length ( $L_x$ ) and width ( $L_y$ , periodic) for smaller wind farms and domains. The effect on the wake flow and the flow blockage was negligible.

L168, it's not totally clear how the height limitation set to the recycling of turbulence helps to avoid the increase in the BL height, is the subsidence mentioned in L90 not sufficient? However, it is clear that this region represents only a small portion of the total domain length, albeit still large (10 km).

We think that this sentence is clear. If no turbulence is recycled to the inflow above a certain height, then the flow is laminar above that height and thus the (turbulent) BL is limited to that height. In the CBL- case the BL growth is compensated by the subsidence. However, the BL could potentially slightly grow before the recycling plane due to the blockage effect of the wind farms. This growth would be mapped to the inflow and reinforce itself without the above mentioned recycling limit.

L194, please clarify the usage of the concept "steady state". It seems that in some instances the authors use these words when referring to a flow that has achieved statistical convergence such as in a developed BL which could be confusing with a flow that is indeed steady-state and as such, lacks transient features such as those reproduced by LES.

Also due to a comment of the other reviewer we rephrased and added information:

"While steady-state turbulence is reached after only a few hours, achieving a steady-state mean flow can take several days. Here, we declare the mean flow as steady if the oscillation amplitude of the hub height mean wind speed is less than 0.5 % and declare the turbulence as steady if the change in friction velocity is less than 2 % in 4 h."

L230,235, could the issues pointed at by Basu et al. (2008), such as those related to the estimation of the friction velocity, be added to these arguments?

If we understood it correctly, then the issues pointed out by Basu et al. (2008) are only valid for SBLs. This section (L230 ff) deals with a CBL.

L378, is it possible to provide an example (or a threshold) of a park size where, following these arguments, the turning would occur in the clockwise direction?

We think that deriving such a rule is not possible without performing additional simulations. It might also depend on stability and BL height, because the clockwise turning is caused by wind veer. Respective studies will be addressed in a follow up study.

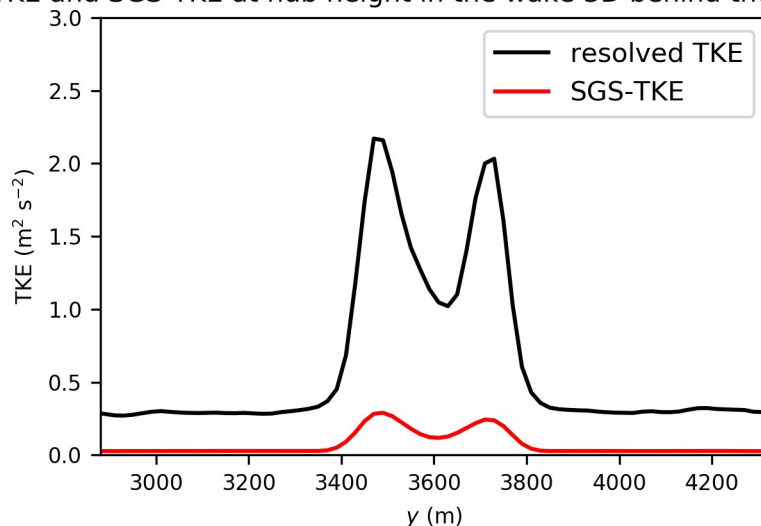
In Figs 5 and 6, scale is inverse with respect to the other (red to green and green to red), please maintain consistency to denote low to large values in scaling. Furthermore, it is rather common to see blue to red to depict small to large values (e.g. subfigure (f) in figures 4, 6 and 7), so to use blue in the middle can be somewhat disconcerting. Is this perhaps a choice to highlight the wakes behind the clusters?

This choice is intentional. Blue and especially red should indicate the wake, which is smaller wind speed but higher TI. We put blue in the middle to avoid a green-red color transition which might be difficult to see for red-green colorblind people.

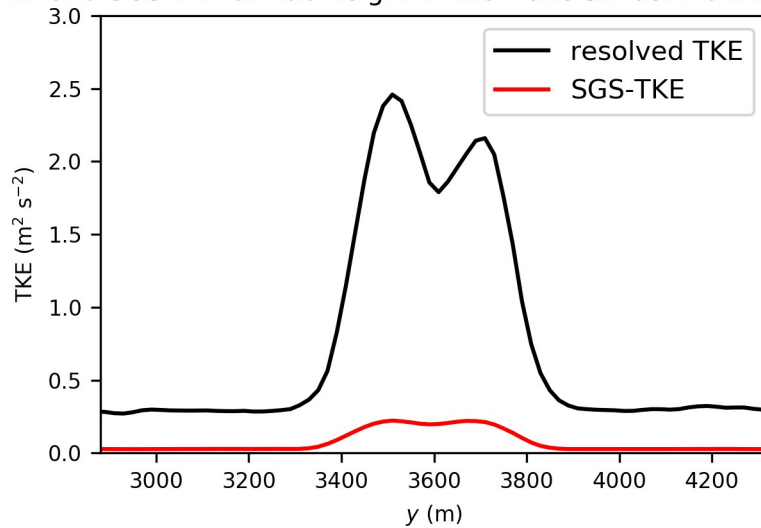
L386, the TKE does consider the subgrid scales. Since the resolution is relatively coarse to resolve the flow around the rotors, a sizable part of the TKE is potentially found within the subgrid scales. Please provide an argument regarding why only resolved scales should be considered in this analysis.

Resolved and SGS-TKE is displayed in the following figures and justifies that the SGS-TKE has no significant contribution here.

TKE and SGS-TKE at hub height in the wake 3D behind the turbine



TKE and SGS-TKE at hub height in the wake 5D behind the turbine



In line 387 (equation for TI and TKE) we added:

“The SGS-TKE is neglected, because it is smaller than 10 % of the resolved TKE, inside the wind turbine wakes at a distance of 3D or more.”

Figure 7, label in vertical scale should be z.

We corrected that.

L505, how is the power computed for each turbine?

The power is part of the output quantities of the advanced actuator disc model. We added a reference to an article of Wu and Porté-Agel (2010), which describes details of this model (in English language) in line 94 so that the interested reader can refer to this article.

DOI: 10.1007/s10546-010-9569-x

L608, it is remarked that for the SBL-300-7D case, the  $W_{gpg,wt}$  component is dominant but in the Fig. 10(e) it seems that  $W_{vke f}$  is about the same value throughout except only within the two gaps between farms.

Yes that is true. We rephrased that to:

“For the first 10 km of the wind farms the vertical kinetic energy flux dominates but further downstream, the energy input by the pressure gradient below rotor top level is greater or equal to the vertical kinetic energy flux”

Technical corrections

- L121, typo, replace “can not” for “cannot”

- L162: I suggest to remove first comma
- L308, typo, it should be “effect” (it says affect)
- L637, typo, “of” instead of “if”

We thank for these hints, and have considered them.