General Comments

Description of paper
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 a 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.

Mayor 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.

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 signifigance 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.

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.

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

   I suggest to add 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?

More Comments

L 1	It's not true that only the first row of the plant is influenced
L 95	*with* a *smaller time step*
Table 1	should also have the height of the two domains
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?
Figure 5	Why are the lines not converging to zero at 20 D?
Figure 8 b	From Figure 8 a 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.
L 289	See comment for L1
Discussion & Conclusions	For me the chapter is too long and hard to read. I suggest to restructure the chapter. The part of the non-existent gravity waves for example could be written much shorter and more concise.