General comments

The paper investigates energy production and wake losses of six hypothetical 4 GW offshore wind farm scenarios with a year-long large-eddy simulation. Due to the coupling with a large-scale weather model the dataset represents a large and realistic variety of wind speeds, directions and stabilities that many other LES-studies lack due to their idealized character. The results show in great detail how wake and blockage losses as well as the wind farm wake itself depend on wind speed, wind direction, stability and different wind farm layouts. A sensitivity study provides confidence in the robustness of the results by investigating the effect of e.g. different resolutions and domain sizes. The paper is well structured, the applied methods are comprehensively described and the results are analyzed and discussed thoroughly. In summary, I think that the presented work is very relevant and of great value for the wind energy community.

However, I believe that the LES study has one major shortcoming, that should be addressed: Some recent LES wind farm studies show that the flow through large wind farms is significantly affected by pressure gradients that are related to gravity waves in the stably-stratified free atmosphere which are induced by the wind farm itself (e.g. Allaerts (2017), doi:10.1017/jfm.2017.11, Lanzilao (2022), doi: 10.1088/1742-6596/2265/2/022043). The studies show that large domain heights (usually more than 10 km) and a well-tuned damping layer at the domain top is needed to capture these waves properly and to avoid reflection at the upper boundary. Your setup does not meet these requirements and that might distort the results (especially the first-row losses, but also the overall performance and the wind farm wake flow). I guess that a complete rerun of the simulations with a much larger domain height might not be feasible and that the overall effect will be only a few percent. Thus, I only ask you to comment on how this shortcoming might affect the results in the discussion section. A reference to Allaerts et al. (2018) doi:10.1088/1742-6596/1037/7/072006 (Annual impact of wind-farm gravity waves on the Belgian-Dutch offshore wind-farm cluster) might provide some plausibility/verification of your blockage / first-row loss results.

Additionally, I believe that some information needs to be added or corrected in the manuscript. I have commented on the respective parts below.

Specific comments

Line 42: Please highlight what the shortcomings of the cited studies are and why there is a need for performing a study like you did.

Line 46: I think that the term "aerodynamic losses" is not a well-known term in the wind energy community and that you have to define it before. Here, in the introduction, you might better explicitly name the two parts of it: wake and blockage losses, which are known terms.

Line 70: How is buoyancy calculated? Is a fixed reference temperature used or rather a vertical profile?

Line 72: "There is a wide variety of subgrid models, of which a significant part uses the concept of eddy viscosity:" Is there a wide variety of SGS models in GRASP or in all LES codes? Please only mention, which model is used in this study.

Line 76: K_m should also be introduced (as eddy viscosity).

Line 81: If the grid is anisotropic, how is the grid spacing defined then? Please add this information.

Line 88: If humidity is part of your prognostic equations, how is condensation and precipitation treated in the model? Most of other LES wind farm studies neglect humidity completely, so please add a few words on this topic.

Line 126: Is the nudging time scale at the boundaries also 6 h or less? I guess that 6 h results in a too weak nudging to damp out the wind farm wake.

Line 135: Which density is taken (anelastic approximation)? The one at hub height?

Line 136: Omit the "²" in M_D. Or is it a kind of geometric mean, i.e. the wind speeds in the disc are weighted by squaring?

Line 144: "In order to quantify wake losses, we compare the energy production of the wind turbines with the production of so-called thrustless turbines": You later call these losses "aerodynamic losses" and if I get it right these include wake AND blockage losses. Please use the term "aerodynamic losses" here.

Line 144-150: I wonder whether the thrustless turbines really output the correct power. If they do not exert drag on the flow then also the disc-averaged wind speed and thus power will be higher than for a real (thrustful) turbine (You do not use free-stream wind speed as input). How do you compensate for that?

Line 152: Why the year 2015? Is it a very representative year?

Line 156: "The horizontal domain size extends to 76800 m.": Please make a short reference to the sensitivity study section to show that this is sufficiently large.

Line 165 and Fig. 2: What is meant with the modeled wind speed at 92 m height? Is it a horizontal average over the precursor domain?

Figure 2: Because it is also relevant for your results, a histogram of the stability should be added here. Additionally, I suggest to include a histogram of the boundary layer height, because this is a relevant parameter for wake losses and blockage effect and it is often varied in other LES studies (see my comment on line 373).

Figure 3: Please insert the names of the scenarios as titles of the subplots.

Line 183: I think better than just mentioning the range of rated powers you should list all the used turbines here (name + power).

Line 199: "We designate the difference between the two as ‘aerodynamic losses’": Can these losses be further subdivided, e.g. in wake loss and blockage loss? Then please mention here.

Table 1: Please add the names of the scenarios that you use in the figures (e.g. Fig. 5) as a new column.

Line 206: "The aerodynamic losses vary between 12 % and 18 %": Here you provide relative values in % for the aerodynamic losses, but further above (line 199) you define it as a difference (i.e. a power). Please modify your definition accordingly and consider to add an equation for that.

Line 210: "Both the higher production and the higher aerodynamic losses for the IEA10 scenario can be related to behavior of the respective power curves (see Figure 4)." I think the primary reason for the larger power of the IEA 10 turbine is the larger rotor diameter (which results in a higher power levels in region 2 (partial load) of the power curve). Please point out this relation here.

Is it true that the aerodynamic loss is the opposite of what is often called "wind farm efficiency" (loss = 1 - efficiency)? If yes, you should mention it here and also use this term in the abstract and as a keyword because I think that "aerodynamic loss" is not a well-known term in the wind energy community.

Figure 6:

6 a + d: Are these curves averaged over the entire year?

6 b + e: This is a density function, so the unit should be GWh / (m/s) or you should mention the bin width (in this case a bar plot would be more appropriate)

6 c + f: Please mention that the values are normalized (although it is obvious). Another option: Do not normalize, so that the plot shows the yearly production/loss of each scenario.

Please add this information in the figure caption.

Line 249: "This can be understood by interpreting the total energy production as a function of wind speed as the convolution of the wind speed probability density (Fig. 2b) and the power curves." It is rather a multiplication of the two functions than a convolution, because they are not shifted relative to each other (i.e. the power curve value at a certain wind speed is multiplied with the probability at that same wind speed).

Equation (17): Is the temperature theta_l an absolute temperature in Kelvin or a relative temperature in °Celsius? Based on equation (8) I guess that it is in °C (because 273.15° is subtracted), but it should be an absolute temperature that can

not become negative (refer e.g. to the textbook of Stull "Boundary layer meteorology", chapter and equation 5.6.3, in which the virtual potential temperature is used).

Line 260: See comment on Figure 2 (add histogram of R_b).

Figure 7: - Please add explanation of dashed lines in the caption or in the legend.
How is fractional aerodynamic loss defined and calculated? How is relative loss defined? Is there a difference between relative and fractional loss? If not, please use one term consistently.

Line 267: "Here, for the most stably-stratified conditions, relative losses are roughly 10% larger than for convective conditions." I guess you mean the step from 0.3 to 0.4 between unstable and stable cases. But this is a difference of + 33 % OR 10 percentage **points**. See also line 451 (Conclusions).

Line 297: "Interestingly, the reduction of the first-row wind speed *deficit* (?) continues for much higher wind speeds."

Figure 9: a) please use a nice vertical axis label. b) Please define "first-row loss", here it seems to be normalized (relative loss ?)

Line 344: "The impact of the wind farm on the year-round, omni-directional wind field is in the order of 20 km, after which a velocity deficit of less than 1 % is observed.": Since the domain boundary has only a distance of about 25 km to the wind farm edges, I wonder whether the boundaries affect this result. Or have you also looked at the wakes in the sensitivity study? If not, then please add a critical note here on that.

Line 373: Wind speed, direction and stability are investigated but not boundary layer height. Many other LES wind farm studies vary the boundary layer height and show that it can have a significant effect on the wake and the wind farm performance (e.g. Maas and Raasch, 2022). Thus, I highly recommend to add a subsection and a figure on that topic, if your dataset provides the required information.

Line 409:"In our main simulation, the resolved fraction of the momentum flux is larger than 80 % for 70 % of the time.": At which height have you evaluated the momentum flux?

Technical corrections

line 33: "example" instead of "exampe".
line 39: "cluster of offshore wind farms"
line 60: GPUs instead of GPU's
Line 88: Use consistently subgrid or sub-grid (or maybe even better subrid-scale).
Line 110: "enough" instead of "enought"
Line 120: "where alpha is ..." (wrong order)
Line 124: I am not a native speaker, but I would write "setup" instead of "set-up". Please also search for other instances.
Line 125: "Stevens et al." in brackets (use citep instead of citet)
Line 139: "manufacturer" instead of "manufactorer"
Line: 147: thrustless, exert
Line 176 and others: "5 MW.km^{-2}" There should be a space between units: " MW km^{-2}"
Line 183: "Power and thrust curves"
Line 193: "10 MW/km2" should be "10 MW km^{-2}". Please check all other units in the manuscript (line 238, 241,276...)
Line 246: "the impact **of** this effect"
Line 248: "the total energy production peaks at a wind speed of approximately 12 m s^{-1}"
Line 257: "over the rotor ~~blade~~ of the IEA15 turbine, i.e. between heights of 270 and 30 m"
Line 267: Decide to use "unstable" **OR** "convective" consistently in the entire manuscript (in the figures and the text).
Line 269: "the impact of stab**il**ity"
Line 296: "e.g."
Line 413: "production"
Line 429: "Overall, we argue that the sensitivity experiments presented here, ~~do no in~~validate the reasoning and conclusions discussed in the Results Section."
Line 473: "The authors would like to acknowledge funding from the Topconsortia for Knowledge and Innovation (TKI) *[missing word here]* funded Winds of the North Sea in 2050 (WINS50) project."