The manuscript “Under-resolved gradients: slow wake recovery and fast turbulence decay with mesoscale Wind Farm Parameterizations” by Radünz et al. compares LES of a regular-spaced idealised offshore wind farm in a conventially neutral boundary layer with the same wind farm under the same meteorological conditions in an idealised setup of the mesoscale model WRF.

In the mesoscale model different grid resolutions and parameters of the wind farm parametrization are studied. The authors compare these different setups to the LES results and conclude that the wake dissipates slower in the mesoscale model because of under-resolved vertical gradients of wind speed and a decay of TKE that is occurring in a smaller distance downstream of the wind farm compared to the LES.

The authors address a relevant topic as mesoscale modelling has become one of the cornerstones of studying the effect of expanding wind energy employment on power output and the state of the atmosphere. The manuscript is well written, and the different steps are easy to understand.

However, I do not think the authors provide enough evidence for their strong conclusion that the mesoscale model is under-representing the wake decay. The reason for my doubt lies in a fundamental flaw in the study setup.

Figure 4a clearly shows the difference in power production of the wind turbines in the LES and the mesoscale model. The difference in power production (in the BASE case it is about 20% more power) results in a significant difference in removed momentum between the models. I find it hard to be convinced that the large difference of wake deficit especially visible in Figures 6a and 8e is not mainly due to this difference in removed momentum. I would actually claim that Figure 6d reveals that the wake decay is quite comparable outside of one grid point just downstream of the wind farm.

Furthermore, the study is based on a very specific case of wind farm flow. The situation that all turbines are aligned in quite close distance to each other and experience full wake exposure of all upstream turbines is in reality very rare. Much more often the wind turbines are actually operating in partial wake or in the wake of a much more distant upstream turbine. In addition, wind farms with regular layouts like the studied one are rarely to be found. For this particular case the mesoscale models stands no chance to replicate the power production of the LES, as the momentum extracted of each turbine is equally distributed over the whole cross-section of the grid cell. Thus, I would not recommend to base general conclusions of the mesoscale model’s ability to model wind farm flow on this particular extreme flow case.

Further comments:

- Title: “Under-resolved gradients: slow wake recovery and fast turbulence decay with mesoscale Wind Farm Parametrizations” - I strongly oppose the idea that the wind farm parametrization should be responsible for the wake recovery and the calculation of turbulence downstream of the wind farm. This job should be solely on the PBL scheme. Thus, I find the title misleading.
- Description of the PBL scheme: A description of how the PBL scheme uses TKE for calculating mean momentum transport as well as a description how the PBL scheme itself is calculating TKE is missing. I think this is crucial to get an idea about what is happening in the mesoscale model. In fact, the used PBL-scheme is not even mentioned in the manuscript.
- L. 65 ff. – I don’t like this part of the introduction and Fig.1 – I think this should be entirely moved to the discussion and conclusion. It is making assumptions about the wake

recovery that are actually part of the study. As I mentioned earlier, I do not see enough evidence of a consistent underestimation of wake recovery of wind farms by the WRF-WFP. In fact, comparison with power data (Sanchez-Gomez et al. 2024) show a quite good performance.

- L.79: "overestimation of power losses often reported when using WFPs" – Montavon et al. 2024 relate the overestimation of power losses mainly to the error in calculating gross yield which was addressed with the correction in Vollmer et al., 2024. In fact, they find the patterns of external as well as internal wakes quite well replicated by the WRF WFP. In general I would recommend to treat any reported differences in energy yield before this correction carefully and not relate them to higher wake losses.
- L.231 ff: The reason for the different development of power production along the turbine rows needs to be explained more precisely. The consequences for drawing any conclusions about wake recovery need to be made aware of.
- Figure 5i – This figure clearly shows one of the weaknesses of the MYNN-PBL scheme at high grid resolution – there is no horizontal diffusion of momentum, so the individual turbine wakes persist throughout the whole domain. I think the manuscript should focus much more on this limitation. Also, I do not understand why the control volume has to be smaller for this particular case. I would argue it should be the same as in the other cases to allow for a fair comparison.
- Figures 6a & 7a – In the LES the wind recovers by the end of the domain, even reaching a higher wind speed than upstream of the farm. In the mesoscale model the wind never recovers to upstream conditions. Can any influences of the boundary conditions in the models be excluded as a reason for this discrepancy?
- L.301 ff: In the end, there is a pronounced difference of wake recovery at exactly one grid cell downwind of the farm (Fig. 6d). It would be interesting to understand what this "key mechanism" is.
- Figure 9: I think the comparison along one column of turbines is misleading. If a location exactly between the turbine rows would have been selected, the TKE in the mesoscale model would actually be higher than in the LES, as indicated by Fig. 8g. The meaningful comparison in my opinion would be the cross-stream averaged profiles.
- L. 510 ff. – I am not convinced that the difference occurring at exactly one grid cell is responsible for much of the observed differences in the wake deficit everywhere else.

The study has some interesting data to work with, that has the potential to identify key limitations in the PBL-scheme of the mesoscale model to model a wake-recovery comparable to higher fidelity models. I disagree with the conclusion of the manuscript, that a general slower wake recovery in the mesoscale model is and can be identified with these data. If the manuscript clearly addresses the key limitations of the comparability of the two data sets and focuses on the wake recovery processes related to the PBL scheme, it has the potential to add a valuable contribution to the discourse.