

Dear authors,

I think the paper now reads very well and I would like to thank you for considering many of the comments I proposed. I only have two minor comments that I need you to consider because I think it is only fair to do so:

Dear reviewer,

Thank you for your continued help on this manuscript. We have integrated your feedback into the manuscript and summarize our changes below.

1 Line 79: you are being politely by saying "the results from these studies should be interpreted with caution". I am sorry but in this case one cannot be polite. The bug in the code has enormous implications and surely all results from previous studies are wrong. So it is not about interpretation; it is about providing results and conclusions that are wrong.

Thank you for underscoring the severity of this bug. As far as we are aware, the effects of this bug have only been studied in one steady-state, idealized neutral simulation for one background wind speed. Without additional analysis, we are hesitant to make claims about impacts in idealized simulations in unstable or stable simulations, as well as real simulations. However, we agree that the magnitude of the impact of this bug may be large, and we have updated the manuscript to reflect this sentiment. The text now reads "the results from these studies should be interpreted with caution, as it is possible that this bug may have significant impacts." (L80).

2. Line 213: "TKE are weaker in the unstable MYNN..." They are not weaker; they are more than halved! They are also half the values of the stable ones. Here the reader would like to know if this is what one might expect or if this is an issue of the simulation. You do not really try to explain it and I am kind of ok with that but then I would like you to state something more... could it be the way you run the simulations (domain-wise)? could it be a bug in MYNN? could this be true and then 3DPBL be wrong? I think this point is quite important because Fitch explicitly introduces TKE at the turbine grid cell and so I would think that the inflow TKE is quite important for the understanding of the abilities of the scheme. You are obviously not able to match TKEs between the schemes (which is ok as they are different) but the issue here is that you have two completely different trends of TKE with stability with major differences in the values.

While we also find the behavior of MYNN's TKE odd, we are confident that we do not have a bug in our simulations, and we provided the namelist so that others may also investigate this behavior. Our current explanation for the weak TKE is that the MYNN constants were calibrated against the land-based Wangara experiment, which has substantially stronger heat flux values (as high as 0.21 K m/s, Nakanishi and Niino 2009) than the offshore-based heat flux that we use (20 W/m² ≈ 0.015 K m/s). We conclude the paragraph "Contrary to what might be expected, we note that hub-height values of TKE are weaker in the unstable MYNN simulations than in the

neutral MYNN simulations. We hypothesize that these low TKE values occur due to the weak heat fluxes.” (L216)