Review of 'Model development in practice: A comprehensive update to the boundary layer schemes in HARMONIE-AROME cycle 40' by de Rooy et al.

gmd-2021-184

This paper presents a detailed description of a package of changes to the representation of physical processes in the HARMONIE-AROME NWP model. The authors do, as they suggest, provide an "honest" description of their development process that is also informative and likely to be of widespread interest (beyond users of that particular model). Extensive analysis of the impact of the changes is given from idealised single-column model simulations to objective verification in NWP trials.

Overall the paper is well written and strikes a good balance in terms of the level of detail, given the breadth of schemes being altered in the package, and does a good job of explaining the motivation for the changes, be they from theorectical considerations, detailed analysis of LES or pragmatic changes to improve performance. I have got quite a few detailed questions and comments below, the most significant of which concerns a lack of clarity in the logic underlying the various updraughts used - I recommend this requires a careful review of the text and perhaps the addition of a flow diagram. A more general comment concerns the lack of attention given to momentum mixing, despite concern over wind speed forecasts prior to cycle 40. In particular, no mention is made of whether there is any momentum mixing by the massflux schemes. At least some comment on this aspect is required.

Further details and more minor comments are given below.

1. line 43, "cloud, turbulence and cloud scheme": two cloud schemes, one of which should be convection!
2. lines 138 and 202: good that you ackowlegde this inherent dependency on the height of the lowest model level in the entrainment rate but the dependence of the initial parcel properties feels like a much stronger one that you don't seem to worry about. Building in an unnecessary resolution dependence always seems like a bad idea. A physical height, such as the top of the surface layer would seem much better, given you are adding perturbations scaled by the surface fluxes which will quantify the near-surface gradients well enough
3. line 141: are the updraught area fractions really constant with height or does Table 1 show the initial fractions? Later on you are explicit that this is the case for the dry updraught (line 236) but for the moist updraught, given you have separate w and M profiles at least in the cloud layer, then that would imply you have an updraught area that varies? I can't see any use made of it, though, such as in the cloud scheme (see additional comment below)?
4. line 143: I don't understand the need for an a priori diagnosis of regime. Why not initialise both dry and moist plumes at the surface, calculate their evolution with height and from that diagnose if clouds are possible (based on the moist parcel reaching its LCL)? Is it just to save cpu time?

5. section 2.1.1: I'm confused by this description of the different parcels. Perhaps a logic flow diagram would help but several questions arise. (i) you say the test parcel is used to determine an estimate of the inversion height, so how can the moist updraught LCL come into this (line 179)? (ii) in line 190 you say "this iteration process converges very rapidly" but don't say anything about what steps are iterated and how the iteration is monitored or convergence measured. (iii) can you really be confident that if the test updraught doesn't reach its LCL then the moist parcel will not reach its own? Given the different formulations this doesn't seem certain. (iv) around line 200 you suggest the dry parcel cannot reach $z_{i,dry}+a_2$ but isn't $z_{i,dry}$ just an estimate from the rather different test parcel (as you can't know the inversion height for the dry parcel without knowing ϵ_{dry}), so what stops the dry parcel below $z_{i,dry}+a_2$ in practise? (v) in (10) is z_{lcl} the LCL of the test parcel as described in line 183? If so it would be good to make that definition of z_{lcl} explicit. But I'm concerned that you specify a change in the moist updraught's entrainment rate at z_{lcl} even though its LCL is likely to be different, not least because of the different fractional areas and sub-cloud entrainment rates. Does this not matter that the moist parcel's LCL may not conincide with where you change its entrainment rate?

6. line 216:, "deeper boundary layers will contain larger updrafts with relatively small entrainment values": the references supplied don't actually show that this dependence is wrong - perhaps the specific formulation in the REF scheme is not good for the ARM case, or do you have other concerns?

7. line 217, "Eq. (10) shows an inverse correlation between updraft vertical velocity and entrainment magnitude": it doesn't actually, but I think you are motivating its shape from expected vertical velocities?

8. line 269, "from there mass flux decreases linearly to 0 at cloud layer top": this sounds like a rather crude assumption and no massflux profiles are shown to illustrate its success or otherwise. For example, I might expect the massflux decrease to be constrained much closer to a sharper inversion, eg one maintained by radiative cooling of a stratiform layer at the cumulus top (as in the transition SCM cases shown in 3.3) than in the ARM case where shallow cumulus detrainment into a stable stratification is probably what determines the inversion thickness

9. line 357, "presume a Gaussian PDF": this seems like a big assumption, especially for cumulus clouds where the pdf is quite likely to be skewed?

10. Fig 4: I think you could usefully and safely (ie without cluttering the plot) add the dry and moist massflux scheme components (eg as dotted and dashed lines) for the cy40 simulations, so we can see the relative contributions. I think that would give valuable insight into the workings of the overall parametrization

11. line 460: you give the horizontal resolution but not vertical. Including the height of the bottom grid-level is also clearly important for the parcel initial properties.

12. line 494, "inclusion of the energy cascade": this does clearly improve the ultimate fluxes but it would be good to know it also improved the TKE profiles, which are not currently shown, and so improving performance for the right reason. For example, you also speculate (line 517) that "a plausible explanation for the presence of diffusive transport...are (dry) updrafts terminating around the inversion height" but

could you not be underestimating the transport by those dry updraughts themselves (again, showing the break down in Fig 4 would be instructive here)?

13. line 550, "underestimation of low values of cloud fraction in the upper part of the cloud layer": it is hard to work out the contour interval from the colourbar but the LES looks only to have a cloud fraction of a few percent, which could be similar to your moist updraught area. Are you not missing this (highly skewed, see above) contribution to cloudiness?

14. line 579: please could you give more detail of how the SCM is forced from RACMO (horizontal advection, surface fluxes or interactive land surface?), when are the forecasts initialised and how long are they for?

15. line 601: as noted in Beare and MacVean for GABLS, many NWP centres find they have to bias their turbulence scheme away from LES in order to improve objective verification of the forecasts. Is this not the case for you and, if not, do you have any insights as to why?!

16. line 623: I don't see why you would invoke a different set-up in this case only, especially for the REF scheme? Isn't the purpose to illustrate operational performance in simple test cases? It is then not clear, when you say (line 628) "Based on the considerations above, the stratocumulus regime with only a wet updraft is removed in cy40NEW", whether that is just in the SCM simulation or is this the motivation for this change in all tests?

17. line 636, "Key aspect of the large improvement with cy40NEW is again the better preservation of inversion strengths": this statement would be much stronger if backed up by a sample profile or cross section showing this sharper inversion.

18. line 638, "removal of the HARATU updates": these were reported (line 278) as being needed to alleviate problems with wind speeds so how do those look in cy40NEW? You do (finally, line 698) say the performance is maintained but in what metric (diurnal cycle of wind speed bias would be a good one to show)?

19. Table D1, Shallow convection scheme, cy40REF: I suspect the formula for ϵ_{lcl} has too many layers of subscript in $z_{z_{lcl}}$?