The paper deals with numerical modifications to the ICAR model made and tested by the authors. In general, it is a very extensive work with lots of statistical tests and well-selected metrics. As such, the paper certainly fits to the scope of GMD. It has, however, weaknesses that are shortly mentioned before more specific remarks are added. The general remarks are listed according to the three main modifications discussed in the paper:

(a) The usage of an undisturbed potential temperature to calculate the background thermal stability N is a very useful new option to ICAR. The respective sub-section 2.2.1 should be written in a clearer way. I would recommend to use equations specifying the calculation of N, e.g. N^2=g/Theta_0 dTheta/dz with Theta = Theta_b + Theta'; here, Theta(x,y,z,t) is the full field consisting of a background Theta_b(z) and wave perturbation Theta' and Theta_0 is the constant surface value defined on page 7 line 33. By specifying if only Theta_b or the full Theta is used to calculate the background N, the reader can easily comprehend what is done.

Writing this equation, another thought came to my mind. It seems to be that the numerical simulations were conducted with another equation for calculating Theta_b from a given N-value, namely, N^2=g d(ln(Theta_b))/dz which would explain the values of the isentropes in Figure 1. I hope, my guess is right as this is not specified in the paper. Here is my general concern: as the linear wave equations are derived under the Boussinesq approximation (assuming linear Theta_b profiles with constant N) how the exponential increase of Theta_b fits to the given assumptions?

In other words, does the Boussinesq approximation per se limits the vertical extent and depth of the numerical model simulations? Linear numerical ICAR simulations covering the whole troposphere and lower stratosphere should be made with another set of linear wave equations that is derived from the anelastic equations. So, parts of the observed deviations might be related to the violation of the applicability of the Boussinesq approximation. A discussion of this aspect is highly appreciated in the parts relating to the tests of the model top.

(b) There is an extensive testing of different boundary conditions applied at the models top. Generally, there are three types of boundary conditions for finite difference schemes: Dirichlet boundary conditions (prescribes the value), Neumann boundary conditions (describes the derivative), and mixed boundary conditions. It would be beneficial for the reader to obtain a structured and - this is the main point - physically-motivated description of the various boundary condition settings as listed in Tables 2 and 3 based on the knowledge of finite-difference schemes.

Furthermore, each boundary condition leads to a different numerical problem to be solved and differences in the presented solutions are not surprising and obvious. However, they were never discussed in the frame of physical arguments. Only, an "optimization" based on the extensive tests was presented that might be feasible but I learnt not much and I have doubts if it can lead to general conclusions.

The other main issue with the presentation of the boundary condition is the missing information about the boundary conditions for the velocity components. Obviously (again I have to guess as this information is not provided in the paper), the perturbations of the velocity components are calculated everywhere in the model domain and at the boundaries. This would explain, why finite values of w' appear at the uppermost model level. Here, I would recommend a very simple, well-approved solution that avoids all the extensive testing: set w'=0 at the model top and relax all PERTURBATION variables u', v', w', Theta', q’, … to zero in a shallow sponge layer beneath the model top. In this way, you avoid any flux of material in and out of the model top and the wind, Theta, and moisture fields only consist of the background values.

(c) The following comments relate to the general aspects in the paper. I had a hard time reading the text. The text is full of many details that are hard to track, sometimes repetitions hinder the reading, and there is information missing that initiates thoughts if I as a reader have missed something in the previous text or if this information simply not given in the text. I really would appreciate a careful editorial revision of the whole text as the whole writing is in contrast to the high-quality figures prepared for this submission.

I just go through some sections to provide examples (only selected examples - I didn't scan the whole text):

page 1, line 2: "As a consequence, ..." Of what?? The sentence further says that a model may yield correct results for the wrong reasons. So, is this a consequence of the content written in the sentence before?

page 1, line 8: Is evaluation the same as the above-mentioned verification?

page 2: Parts of the Introduction are rather general that (in my view) only have a slight or limited relation to the content of the paper, especially, the parts belonging to the thoughts about verification. Either they should be deepened or omitted.

page 2, 1st para: It is also the imperfection of observing systems (especially, when you consider remote-sensing systems and combinations of them) that lead to a fragmentary and often unsatisfying verification.

page 2, line 10: A good reference is Stensrud, D. (2007). Parameterization Schemes: Keys to Understanding Numerical Weather Prediction Models. Cambridge: Cambridge University Press. doi:10.1017/CBO9780511812590. Furthermore, it is often the lack of knowledge about essential processes that limit predictability (e.g. gravity wave parametrizations).

page 2, line 12: I have problems with the saying "right, but for the wrong reason". Most often, only one selected diagnostics is picked. The Zhang paper makes it clear: if the authors would have looked at vertical winds instead they had realized that there is an essential mechanism not represented in the model, namely the convection. So, the story with the causal chain (see also page 1 line 3) can be misleading. It is too much linear thinking in it - at least for my taste.

page 2, line 28/29: I could imagine that a more educational verification would be the comparison with a linearized version of WRF or another NWP models. This would provide a real one-to-one comparison. I wonder why this option is not considered.

Section 2

I recommend to rewrite the whole section (especially, Section 2.1) totally and add all parts that appear later in the text regarding the model set up (essentially, 2nd and 3rd para from page 7, Section 3.3, and maybe more). Section 2 should provide the reader with all information to understand the numerical integration of ICAR. This is probably best done by presenting the applied linear wave solutions, the advection equation (for which variables?? It was not clear to me when I read the paper first), and by specifying the initial and boundary conditions for all quantities by means of equations. The authors might argue this is done in the Gutmann paper but the above example of the calculation of the Brunt-Väisälä frequency shows that clarification is necessary as much as is possible.

Regarding advection: Do you advect full Theta or Theta'? Do you advect specific variables? For example, is Psi in Equation (1) rho times, say Theta?

There are probably more issues but I stop here. Altogether, I'm not convinced by the stated advantage of using ICAR (less computer time – this was not documented) when one has to spend massive resources (time and man power) to optimize a model for applications (microphysical and moist processes) that are rarely linear. Also, the back-link to the verification theme in the Introduction could be added!

Andreas Dörnbrack