Response to Reviewer 2

Abbreviations:

AR Author Response (Johannes Horak)

RC Reviewer Comment

> Note that in this response orange bold text**, such as this,** or a red equation, indicates a change in the manuscript text (either an addition or some rephrasing). Blue or crossed out text ~~(example)~~ indicates the removal of text. Additionally, each modification in the manuscript is complemented by the respective page number and line where the modification occurred (e.g. **P01L01**). In the case where a table was altered, the respective modification is highlighted by an orange box around the modification.

RC: 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:

> **AR:**
>
> We thank the reviewer for taking the time to read the manuscript and provide constructive criticism. Please find our detailed answers below!

RC: (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.

> **AR:** We have carefully revised Section 2.1 and Section 2.2.1 to better introduce and clarify the changes made to ICAR and which quantities (perturbed or background state) are employed for the calculations. Due to the larger revisions we did not copy all changes in our Author response, please refer to the revised Sections in the updated manuscript. However, please find below smaller changes that directly address specific criticisms.
>
> In particular Section 2.1 now shows the equations used to calculate the Brunt-Väisälä frequency and mentions early on that ICAR employs the quantities from the perturbed state of the domain (in contrast to the assumptions of linear mountain wave theory). This is now additionally indicated by the symbols used in the equations, e.g. the variables mentioned in context with the background state (given by the forcing data) are now distinct from those describing the perturbed state of the domain.

Similarly, Section 2.2.1 was revised to better indicate the difference between ICAR-N and ICAR-O and more clearly state which equations are employed and what state (base state or perturbed state) is the input for their evaluation.

Adjustment to the manuscript

P04L06: The forcing data set represents the background state of the atmosphere and must comprise the horizontal wind components **(U, V)** pressure **p**, potential temperature **Θ** and water vapor mixing ratio **q_{v0}**.

Adjustment to the manuscript

P04L26:

> Note that depending on whether a grid cell is saturated or not, either the moist, N_m **(Emanuel, 1994)**, or dry Brunt-Väisälä frequency N_d is ~~used~~ **employed in Eq. (4) and calculated as**
>
> $$N_d^2 = g\frac{d\ln\theta}{dz} \quad \textbf{and}$$
>
> $$N_m^2 = \frac{1}{1+q_w}\left\{\Gamma_m \frac{d}{dz}\left[(c_p + c_l q_w)\ln\theta_e\right] - \left[c_{\text{¿}} \ln T + g\right]\frac{dq_w}{dz}\right\} \quad ,$$
>
> **with the acceleration due to gravity g, the temperature T, the potential temperature θ, the equivalent potential temperature θ_e, the saturated adiabatic lapse rate Γ_m, the saturation mixing ratio q_s, the cloud water mixing ratio q_c and the total water content $q_w=q_s+q_c$, and the specific heats at constant pressure of dry air and liquid water c_p and c_l. Note that ICAR employs quantities from the perturbed state of the domain to calculate N even though in linear mountain wave theory N is a property of the background state (Durran, 2015). Statically unstable atmospheric conditions (i.e., $N^2 < 0$) in the forcing data are avoided by enforcing a minimum Brunt-Väisälä frequency of $N_{min} = 3.2\ 10^{-4}\ s^{-1}$ throughout the domain.**

Adjustment to the manuscript

P06-P07

> **2.2.1 Calculation of the Brunt-Väisala frequency**
>
> **From the initial state of θ and the microphysics species fields at t_0 (see Eq. 12), ICAR calculates the (moist or dry, Eq. 6 and 7) Brunt-Väisälä frequency N for all model times t_m smaller than the first forcing time t_{f1}. During each model time step, the θ and microphysics species fields in the ICAR domain are modified by advection and microphysical processes. Therefore, for model times $t_m > t_0$, θ and all the microphysics species q represent the perturbed state of the respective fields, denoted as**

$\theta = \Theta + \theta'$ and

$q = q_0 + q'$.

Note that in this notation, the perturbed water vapor field is denoted as q_v, the background state water vapor field as q_{v0} and the perturbation field as q_v'. Consequently, during all intervals $t_{fn} \leq t\_m < t_{fn+1}$, where t_{fi} are subsequent forcing time steps, N is based on the perturbed states of potential temperature and the microphysics species at t_{fn}. More specifically, all atmospheric variables ICAR uses for the calculation of N with Eqs. (6) and (7) are represented by the perturbed fields.

However, in linear mountain wave theory N is a property of the unperturbed background state (e.g. Durran, 2015), an assumption that is not satisfied by the calculation method employed by the standard version of ICAR. This study therefore employs a modified version of ICAR that, in accordance with linear mountain wave theory, calculates N from the state of the atmosphere given by the forcing data set if the corresponding option is activated. In the following, the modification of ICAR basing the calculation of N on the background state is referred to as ICAR-N, while the unmodified version, that bases the calculation on the perturbed state of the atmosphere, is referred to as the original version (ICAR-O). If properties applying to both versions are discussed, the term ICAR is chosen.

RC: 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.

AR:

The authors believe that the previous version of the manuscript may have potentially caused a misunderstanding, in particular with regards to the governing equations underlying ICAR, the equation employed to calculate the $\Theta(z)$ profile of the background state employed for the idealized simulations, the symbol employed for the surface value of the potential temperature Θ_0 and the reference to Durran (2015). In the following paragraphs we attempt to clarify this and indicate separately which changes we made to the manuscript.

The reference to Durran 2015 was to underline that in linear mountain wave theory, N is a property of the background state and not the perturbed state of the atmosphere. However, Durran employs a definition of N yielding a linear $\Theta(z)$ profile, which is not employed by us. In the following we compare the equation for N used by Durran (2015) and that employed by us where we have explicitly stated when a variable depends on z:

Durran (2015) equation for N:

$$N^2 = \frac{g}{\Theta_0} \frac{d\bar{\Theta}(z)}{dz}$$

the equation for N employed by us:

$$N_d^2 = g \frac{d \ln \Theta(z)}{dz} = \frac{g}{\Theta(z)} \frac{d\Theta(z)}{dz}$$

In the denominator of the first fraction, Durran uses Θ_0 to denote a constant reference potential temperature, which does not occur in the equation employed by us, in our case it is $\Theta(z)$ – the vertical potential temperature profile of the background state which is not necessarily equivalent to Durrans $\overline{\Theta}(z)$.

We additionally want to note that our definition of the surface value of potential temperature (previously Section 3.2, P7L32: "...is characterized by a potential temperature at the surface, $\Theta_0 = 270K$, …") was denoted with the symbol Θ_0 which may have contributed to the misunderstanding. Note that Durrans Θ_0 is not the same as the potential temperature at the surface, which is what our Θ_0 denotes.

Furthermore in the previous version of the manuscript we did not state on which equation we based our calculation of $\Theta(z)$, which left more room for interpretation. These $\Theta(z)$ profiles are calculated by solving for θ in equation (6) (see the revised manuscript), resulting in $\Theta(z) = \Theta_0 \exp(N^2 z/g)$ where Θ_0 is the surface value of potential temperature. The manuscript now states the equation on which the calculation of our $\Theta(z)$ profiles is based. We additionally removed the symbols for the surface values of potential temperature and pressure since they do not occur anywhere else in the manuscript and may have contributed to the misunderstanding.

Adjustment to the manuscript

> We revised all of **Section 2.1** and **Section 2.2.1** to better clarify the underlying equations and which input quantities were used for which equation. Since these are larger changes please refer to the updated manuscript to view the revised Sections.

AR (continued): Section 2.1 now explicitly states the equation used by ICAR to calculate the dry Brunt-Väisälä frequency:

Adjustment to the manuscript

P04L29:

$$N_d^2 = g \frac{d \ln \theta}{dz} \quad (6)$$

AR (continued): Section 3.2 (Topographies and initial soundings) now refers to equation (6) as the basis for the calculation of the $\Theta(z)$ profiles. Note that we also removed the symbols for the surface values of potential temperature and pressure as to avoid misunderstandings resulting from these.

Adjustment to the manuscript

P10L01:

The vertical potential temperature profile of the base state **$\Theta(z)$** is characterized by a potential temperature at the surface~~, Θ_θ =~~ **of** 270 K, a constant Brunt-Väisälä frequency, $N = 0.01\ s^{-1}$ **and calculated by solving Eq. (6) for θ**. The horizontal wind components of the base state are chosen as $U = 20\ ms^{-1}$ and $V = 0\ ms^{-1}$, and the surface pressure as ~~p_θ =~~1013 hPa.

AR (continued)

Note that the linearized and Boussinesq approximated governing equations of ICAR (see Barstad, 2006 or equations 1-8 in the revised manuscript), do not explicitly depend on $\Theta(z)$. They only require N to be constant, which is satisfied by our $\Theta(z)$ profiles. Additionally, note that ICAR is able to handle $\Theta(z)$ profiles yielding variable values of N by splitting the domain into smaller subdomains (please refer to Gutmann (2016) for a description of how this is accomplished by ICAR).

In addition to Barstad (2006), Boussinesq approximated governing equations that do not require linear $\Theta(z)$ profiles have been documented in the literature. For instance, Markowski (2011, p. 166, equations 6.18-6.21) and Nappo (2013, p.31, equations 2.1 – 2.4). Markowski (2011) explicitly states $N^2 = g/\overline{\Theta}\ d\overline{\Theta}/dz$ (page 166, right column at the bottom) and in his notation $\overline{\Theta} = \overline{\Theta}(z)$ (page 166, left column first line).

Nappo (2013) uses $N^2=-g/\rho_0\ \partial\rho_0/\partial z$ and states explicitly, that ρ_0 is a function of z on page 32 in the first line below equation 2.17. Note that N^2 in Nappo (2013) and Markowski (2011) are essentially equivalent due to equation 1.65 in Nappo (2013): $g/\Theta_0\ \partial\Theta_0/\partial z = -g/\rho_0\ \partial\rho_0/\partial z$.

The authors expect that differences due to a version of ICAR based on a set of linear wave equations derived from the anelastic equations would manifest mainly above the troposphere (e.g. Doyle, 2021, particularly Fig. 3), thereby not affecting orographic precipitation in most cases. A modification to ICAR with larger effects would, in our opinion, be the inclusion of the amplification of the perturbations due to the decreasing density (see equation 21 on P11L18 in the updated manuscript and P11L19-21). Nonetheless, the authors agree that implementing and evaluating either of these modifications to ICAR are interesting and necessary avenues for future research. However, they consider it outside the scope of this manuscript. We added these aspects to the relevant paragraphs in the discussion.

Adjustment to the manuscript

P39L22:

Future work could implement and investigate whether the amplification of perturbations (see Eq. 21) due to the vertical density gradient yields ICAR-N results closer to that of WRF. Another conceivable avenue for future investigations in that regard could be the implementation and evaluation of a set of linear wave equations derived from the anelastic equations into ICAR-N.

Durran, D.: MOUNTAIN METEOROLOGY | Lee Waves and Mountain Waves, in: Encyclopedia of Atmospheric Sciences (Second Edition), edited by North, G. R., Pyle, J., and Zhang, F., pp. 95 – 102, Academic Press, Oxford, second edition edn., https://doi.org/10.1016/B978-0-12-382225-3.00202-4, 2015.

Barstad, I. and Grønås, S.: Dynamical structures for southwesterly airflow over southern Norway: the role of dissipation, Tellus A, 58, 2–18, https://doi.org/10.1111/j.1600-0870.2006.00152.x, 2006.

Gutmann, Ethan, et al. "The intermediate complexity atmospheric research model (ICAR)." *Journal of Hydrometeorology* 17.3 (2016): 957-973.

Doyle, J. D., Gaberšek, S., Jiang, Q., Bernardet, L., Brown, J. M., Dörnbrack, A., Filaus, E., Grubišić, V., Kirshbaum, D. J., Knoth, O., Koch, S., Schmidli, J., Stiperski, I., Vosper, S., & Zhong, S. (2011). An Intercomparison of T-REX Mountain-Wave Simulations and Implications for Mesoscale Predictability, *Monthly Weather Review*, *139*(9), 2811-2831. Retrieved Jan 23, 2021

Markowski, P., & Richardson, Y. (2011). *Mesoscale meteorology in midlatitudes* (Vol. 2). John Wiley & Sons.

Nappo, Carmen J. *An introduction to atmospheric gravity waves*. Academic press, 2013.

RC: (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.

AR: The first paragraph in Section 3.4. now gives additional information and context for the tested boundary conditions. Additionally, we added the equations corresponding to the respective BCs to Table 3.

Adjustment to the manuscript

P11L24: To this end several alternative BCs to the existing zero gradient boundary condition are added to the ICAR code, their abbreviations**, mathematical formulation** and their numerical implementation are summarized in Table 2. **All BCs constitute Neumann BCs except for the zero value Dirichlet BC.** Per default ICAR imposes a ZG BC at the model top to all quantities**, corresponding to the assumption that, e.g. the mixing ratio of hydrometeors q_{hyd} above the domain is the same as in the topmost vertical level. A ZV BC imposed on, e.g., q_{hyd} avoids any advection from outside of the domain into it. The CG, CF and CFG BCs assume that a either the gradient, flux or flux gradient of ψ, respectively, remains constant at the model top, representing different physical situations. The respective discretizations of the equations given in Table 2 then determine the value of ψ_{Nz+1} .**

Please note that in order to show the adjustment right below, we inserted a screenshot of the updated table and highlighted the changes to the caption and table itself with orange rectangles.

Adjustment to the manuscript

P12:

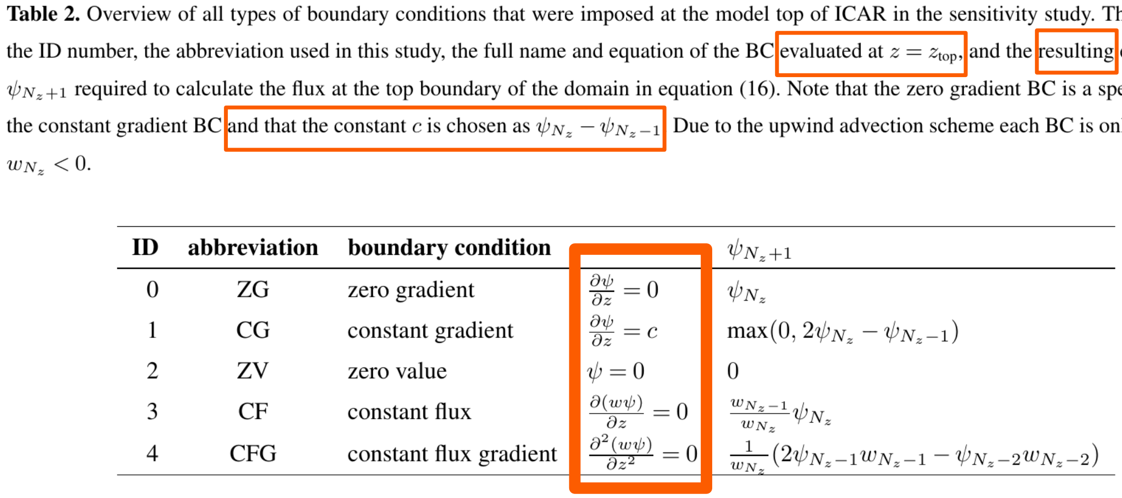

Table 2. Overview of all types of boundary conditions that were imposed at the model top of ICAR in the sensitivity study. The table lists the ID number, the abbreviation used in this study, the full name and equation of the BC evaluated at $z = z_{\text{top}}$, and the resulting equation for ψ_{N_z+1} required to calculate the flux at the top boundary of the domain in equation (16). Note that the zero gradient BC is a special case of the constant gradient BC and that the constant c is chosen as $\psi_{N_z} - \psi_{N_z-1}$. Due to the upwind advection scheme each BC is only applied if $w_{N_z} < 0$.

ID	**abbreviation**	**boundary condition**		ψ_{N_z+1}
0	ZG	zero gradient	$\frac{\partial\psi}{\partial z} = 0$	ψ_{N_z}
1	CG	constant gradient	$\frac{\partial\psi}{\partial z} = c$	$\max(0, 2\psi_{N_z} - \psi_{N_z-1})$
2	ZV	zero value	$\psi = 0$	0
3	CF	constant flux	$\frac{\partial(w\psi)}{\partial z} = 0$	$\frac{w_{N_z-1}}{w_{N_z}}\psi_{N_z}$
4	CFG	constant flux gradient	$\frac{\partial^2(w\psi)}{\partial z^2} = 0$	$\frac{1}{w_{N_z}}(2\psi_{N_z-1}w_{N_z-1} - \psi_{N_z-2}w_{N_z-2})$

RC: 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.

AR: The authors agree that differences in the solutions were to be expected. We think that we did not state the underlying motivation for our tests clear enough. As pointed out by the reviewer and our manuscript, the problem of optimizing boundary conditions could potentially be circumvented mostly by e.g. employing a relaxation layer or just setting the model top high enough. Both methods have the “disadvantage” of requiring high model top elevations. However, particularly for ICAR a low model top elevation would be desirable due to the associated lower computational cost. At the beginning of our investigation it seemed feasible that an optimal boundary condition could potentially be found that would, overall, lead to satisfactory results for low model top elevations. The tested BCs are numerically easy to implement and correspond to comparatively simple physical situations, such as keeping the flux or the flux gradient at the model top constant, all of which could assumed to be physically reasonable assumptions at the model top under certain (but not all) circumstances. While we were able to reduce errors (depending on the boundary condition combination), for low model top elevation simulations these reductions weren’t enough to more closely reproduce the fields in the respective reference simulations. Therefore we consider this less of an optimization: There simply is no optimal combination of the tested BCs that achieves the goal of lowering the minimum possible model top elevation Z_{min} (apart from the discussed exception occurring for θ). Therefore, from our work it follows that future research should either explore more sophisticated options for BCs or entirely different strategies, such as the relaxation layer suggested by the reviewer.

We updated Section 2.2.2 and the discussion accordingly to better clarify our motivation and consequences for future studies.

Adjustment to the manuscript

> **P08L08: A solution to address both issues would potentially be to include a relaxation layer directly beneath the model top (see, e.g. Skamarock et al., 2019). Within this relaxation layer vertical wind speeds would tend towards zero with decreasing distance to the model top and perturbed quantities would be relaxed towards their value in the background state. Another potential solution is employed by** full physics models such as the Integrated Forecasting System (IFS) of the European Center for Medium-Range Weather Forecasts (ECMWF, 2018), the COSMO model (Doms and Baldauf, 2018) or the Weather Research and Forecasting (WRF) model (Skamarock et al., 2019). **These models** place the location of the upper boundary at elevations high enough where moisture fluxes across the boundary are negligible. While applying **either** treatment to ICAR is, in general, an option, it is undesirable since **both necessarily result in higher model tops and therefore** would severely increase the computational cost of ICAR simulations. **Hence**, this study investigates whether the application of **computationally cheap** alternative boundary conditions is able to reduce errors caused by, e.g., the unphysical mass influx **and loss** described above.

Adjustment to the manuscript

> **P39L31: Another possible venue for future research that aims to mitigate the influence of the upper boundary could be the implementation of a relaxation layer directly underneath the model top. In this layer perturbed quantities could, as they approach the model top, gradually be relaxed towards their background state values, while w is relaxed towards zero.**

RC: 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.

AR: We have emphasized the information that w' is calculated at half levels (e.g. also at the top of the topmost vertical level) at the relevant locations in the manuscript.

The authors agree that setting w' = 0 at the model top and adding a shallow relaxation sponge layer beneath the model top is potentially viable solution that warrants further exploration. We added the sponge layer as a potential solution to a relevant paragraph in Section 2.2.2 and added it to the relevant part in the discussion **(see our previous AR and the associated adjustments to the manuscript directly above)**.

However, considering the state of the art with regards to ICAR before our investigation, boundary conditions seemed the more likely approach to allow ICAR simulations to run without compromising physical fidelity AND retain at least some computational advantage of the low model tops. Our results now show that BCs, while correcting some errors in the potential temperature and microphysics fields, mostly play a negligible role in the required model top elevation. In light of these results a sponge layer could – depending on its depth – potentially allow the choice of lower model tops than obtainable with the BC approach. Nonetheless it would still require substantial testing and development efforts to include and evaluate the functionality in ICAR. We considers this a very interesting and promising approach but outside of the scope of the presented manuscript.

Adjustment to the manuscript

> **P04L10: In particular mass based quantities such as water vapor are stored at the grid cell center while the horizontal wind components u and v are stored at the centers of the west/east or south/north faces of the grid cells and the vertical wind component w at the center of the top/bottom faces of each grid cell.**

Adjustment to the manuscript

> **P07L16: Note that the vertical wind components are calculated at half levels with Eq. (8) and that, in particular, no boundary condition is required to determine w at the model top.**

RC: (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.

> **AR: We have carefully revised the manuscript, in particular Section 2** with a view on the reviewer's criticism. Particularly the introduction of all relevant equations and a better differentiation between background state variables and perturbed state variables is aimed to address the issues raised.

RC: 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?

> **AR: P01L01-02:** We rephrased to "The **evaluation** of models in general is a non-trivial task and can, due to epistemological and practical reasons, never be considered as complete. **Due to this incompleteness,** a model may yield correct results for the wrong reasons, i.e. by a different chain of processes than found in observations."

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

> **AR:** Thank you for pointing this out, indeed it is not and we have made modifications in the manuscript to clarify the difference. The sentence now reads:

Adjustment to the manuscript

> **P01L01:** The **evaluation** of models in general is a non-trivial task and can, due to epistemological and practical reasons, never be considered as complete.

Adjustment to the manuscript

> **P02L22:** In acknowledgment of the fundamental limitation of verification, **models are evaluated rather than verified, and** ...

RC: 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.

AR: We deepened and rephrased the first two paragraphs of the introduction to better establish the connection of our specific research question to the general topic of model evaluation. The authors believe that the wider view (including the epistemological and practical problems of model verification that translate to model evaluation) is particularly useful to emphasize that the comparison of model output to isolated measurements alone is not enough to sufficiently evaluate the reliability of a model.

Adjustment to the manuscript

> **P02L03:** All numerical models of natural systems are approximations to reality. They generate predictions that may further the understanding of natural processes and allow the model to be tested against measurements. However, the complete verification **or demonstration of the truth of such** a model is impossible for epistemological **and practical** reasons (Popper, 1935; Oreskes et al., 1994). **While the correct prediction of an observation increases trust in a model it does not verify the model, e.g. correct predictions for one situation do not imply that the model works in other situations or even that the model arrived at the prediction through what would be considered the correct chain of events according to scientific consensus. In contrast, a model prediction that disagrees with a measurement falsifies the model, thereby indicating, for instance, issues with the underlying assumptions. From a practical point of view, the incompleteness and scarcity of data, as well as the imperfections of observing systems place further limits on the verifiability of models. The same limitations apply to model evaluation as well, however, evaluation focuses on establishing the reliability of a model rather than its truth.**

RC: 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.

AR: The authors agree that this should be addressed explicitly as well. When extending the paragraph we included the reviewers suggestion.

Adjustment to the manuscript

> **P02L10: From a practical point of view, the incompleteness and scarcity of data, as well as the imperfections of observing systems place further limits on the verifiability of models.**

RC: 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).

AR: We agree and included this as an additional reference.

Adjustment to the manuscript

> **P02L15:** **Those** models approximate and simplify the world and processes in it by discretizing the governing equations in time and space and by modeling subgrid-scale processes with adequate parametrizations **(e.g. Stensrud, 2009)**.

RC: 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.

AR: The authors agree that this is a rather obvious example, however, it serves to illustrate the problem. Nonetheless, human error, such as picking the wrong diagnostic, is arguably also a limitation to model evaluation and the fundamental source of error in epistemology. It may be reduced by measures introduced to, among other things, deal with this source of error such as peer review, careful study design or guidelines for model setups (e.g. see reference Warner, 2011 in the manuscript), but never entirely mitigated, and result in the model yielding predictions matching observations despite arriving at them through a possibly overlooked different process. Of course once the mistakes are identified it is usually easier to see how it could have been avoided in the first place.

However, it is also possible, for instance, that observations revealing that some current model arrives at a certain result through a different casual chain than what actually takes place, have not been performed or cannot be performed with high enough accuracy. It is also possible that the required observational techniques simply have not been developed or even envisaged yet.

Overall we do acknowledge the criticism and modified the manuscript in cases where it contributed to better clarity.

Adjustment to the manuscript

> **P03L21:**This study aims to improve the understanding of the ICAR model and develop recommendations that maximize the probability that the results of ICAR simulations, such as the **spatial** distribution of precipitation, are correct **and caused by the physical processes modelled by ICAR and not by numerical artifacts or any influence of the model top (correct** for the right reasons**)**.

Adjustment to the manuscript

> **P03L25:** For a given initial state, a correct representation of the fields of wind, temperature and moisture as well as of the microphysical processes are a necessity to obtain the correct distribution of precipitation ~~for the right reasons~~.

Adjustment to the manuscript

> **P15L18:** Section 4.6 additionally investigates whether this seemingly optimal result, as suggested by the lowest mean squared errors, was achieved ~~for the wrong reasons.~~ **due to the low model top potentially influencing the microphysical processes within the domain and the calculation of N being based on the perturbed fields.**

Adjustment to the manuscript

> **P36L27:** Hence, it seems that the underestimation in precipitation near the crest and to its lee of an ICAR simulation with reasonably high model top compared to WRF (Fig. 9) is partly compensated in an ICAR simulation with a too low model top (ICAR-$O_{4\,\mathrm{km}}$ in Fig. 14) by spurious effects introduced by the upper boundary conditions. ~~It follows that the seeming imrpovement in the latter case is right but for the wrong reasons.~~ **Note that this seeming improvement is not due to a more realistic representation of cloud formation processes.**

RC: 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.

> **AR:** We agree that this approach would be of educational value, however, our aim was not to provide a one-to-one comparison. The intent behind using the standard version of WRF was to also infer differences due to non-linearities, wave amplification and the density decrease with height. This is stated in the introduction on **P03L31** and the methods Section on **P08L25-L26**. Note that the capability of ICAR to approximate the exact linear solution is inferred from comparing it to the analytical solution.

RC: 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.

> **AR:** The authors agree that as much clarification as possible is needed, even if some equations from the cited references are repeated. Therefore, we followed most of the recommendations by the reviewer and revised Section 2, with a particular focus on a rewrite of Section 2.1. We included all the relevant equations from Gutmann 2016 to provide better context for the presented modifications, specified equations for the initial and boundary conditions and indicated which quantities are advected with the advection equation. **Please refer to the revised Section 2 and partial revisions in Section 3.4 in the new manuscript.**

However, we retained the basic structure since it follows, in our opinion, a logical and intuitive pattern: Providing a description of ICAR as it is in Sect. 2.1, describing the motivation and the modifications to ICAR in Sect. 2.2. and then detailing the specific model setups in the method subsection 3.1. This additionally allows readers to specifically jump to a given section to reread details of, for instance, the setup instead of having to search through one large section for these details. While we agree that other ways to structure the manuscript are possible, for the presented study this approach appears suitable to the authors.

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

AR: ICAR advects Theta' (or θ in the notation employed in our manuscript). In the former equation (1) (now equation 9), $\psi = \theta$ or $\psi = q_v$, and so on – density is not included. We clarified this in the rewrite of Section 2.1.

Adjustment to the manuscript

P05L21:

The microphysics species, n_i, n_r and θ are advected with the calculated wind field according to the advection equation (Gutmann et al., 2016):

$$\frac{\partial \psi}{\partial t} = -\left(\frac{\partial (u\psi)}{\partial x} + \frac{\partial (v\psi)}{\partial y} + \frac{\partial (w\psi)}{\partial z}\right),$$

where ψ denotes any of the advected quantities.

RC: 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!

AR: The authors once again thank the reviewer for his constructive criticism and for sharing his opinion. Please note that the computational advantages of ICAR are documented in Gutmann 2016.

Regarding the back-link to the verification theme, please note the final paragraph in the conclusions:

P41L17: *This study highlights the importance of a process-based in-depth evaluation not only with respect to ICAR but for models in general. Particularly for regional climate models (RCMs) and numerical weather prediction (NWP) models, the results of the case study demonstrate a potential pitfall when model parameters are inferred solely from comparisons to measurements, potentially leading to situations for which model results are more prone to be right but for the wrong reasons.*