Final response to the Interactive comment by the anonymous referee #2 on “Quantifying the bias of radiative heating rates in NWP models for shallow cumulus clouds” by Nina Crnivec and Bernhard Mayer;

Referee’s comments in blue, our answers in black;

Overall assessment: Minor revision
Summary
This manuscript reports on radiative biases that can be expected to occur when NWP models neglect detailed solar and infrared 3D radiative transfer (i.e., all NWP models all of the time). Results for full 3D transfer are compared to the Independent Column Approximation (ICA), which an NWP model would employ if it resolved domains as well as the LES fields used here, and a highly parametrized radiation scheme which an NWP model would use if its resolution was substantially less than that of the LES fields. The extent of experiments and analyses is adequate for ACP. The authors are experts in this area of work, and there are no obvious signs of either technical or interpretative issues. The manuscript is, however, a bit long... but this might be of no concern for a largely electronic journal (though it might compromise the number of people who read the entire document?). I recommend this manuscript be published after minor revision. A few points follow.

We thank the anonymous referee #2 for a good assesment of our work and his/her many insightful comments. Taking into account many of these suggestions, further improved the quality of our manuscript. Below please find our reply to both “general” and “minor” comments.

Reply to “General comments”

1. Note that much of the issue addressed here is the ability of an NWP model, run at a resolution >> 0.025 km) to adequately provide the information needed to do the benchmark 3D radiative transfer computations. It was assumed that variables such as mean cloud water content and cloud fraction have been produced perfectly by the NWP model (i.e., in total agreement with the LES). This is unlikely to be the case, and those errors could easily rival, or exceed, errors due to use of either the ICA or the simple model. Moreover, my sense is that most NWP models with resolutions near 3 km would neither attempt to get sub-grid cloud fractions nor invoke the max-rand overlap (MRO) rule... I suspect they would assume grid-cells to be either full or free of cloud water? If this assumption is wrong, perhaps the authors could provide a reference to support their claim.

1. Yes, in general we agree that one of the major problems of NWP radiation schemes are the input parameters themselves, especially those related to clouds (cloud optical properties, cloud fraction). The uncertainty of these parameters, however, could be very model-dependent (depending on the complexity of microphysics and cloud schemes employed). An accurate quantification of these uncertainties and their effect on the uncertainty of a 1-D (NWP-type) experiment, is out of the scope of the present study (since the overall length of the manuscript is already long).

Regarding second part of your comment #1: Most current NWP models with resolution 2 – 5 km do not assume either fully covered or fully clear-sky boxes. Instead, they account for subgrid-scale cloudiness (assumed to be nonprecipitating, but still greatly impacting radiative transfer) with the aid of simple (often relative-humidity based) cloud schemes. An example pointed out in the manuscript is a former operational model of German Weather Service COSMO-DE (2.8 km resolution). In fact, the model has recently been replaced by a new version called COSMO-D2, which is running at 2.2 km resolution and still has a parameterization for subgrid-scale cloudiness. Also, The COSMO-D2 model, for example, employs a Delta-two-stream radiative solver of Ritter and Geleyn (1992), which accounts for partial cloudiness assuming max-rand overlap rule.

2. Following from point 1... It would have been interesting to see ICA results as one progressively backs-off full LES resolution (i.e., starting at 0.025 km and going up to ~2.5 km) without making assumptions about the nature of unresolved clouds; just cell-filling if some cloud exists. Obviously, cloud fractions will increase and mean cloud water contents will decrease, but is this what well-resolved NWP models do?

2. Following from point 1: The assumption of the present study is that an NWP model at several km grid spacing contains unresolved cloudiness (motivation for our 1-D experiment). We agree, however, that as the NWP resolution continues to improve, and in fact a few models have already started shutting down their cloud fraction parameterizations, the question how does the ICA bias scale as one progressively degrades grid spacing, is very relevant. In fact, we have already performed some initial experiments on that, which might be a topic of a separate study also. Otherwise, O'Hirok and Gautier (2005) already partially investigated this issue.

3. My expectation at first was that the authors were going to report on LES sensitivities when 3D radiative transfer models were used interactively. The analyses were, however, strictly diagnostic. Interactive studies were referenced at the top of pg. 3. Why weren't they performed and reported on here?

3. Perhaps this is a misunderstanding. The underlying aim of the present study was to quantify NWP radiative biases as **exactly** as possible, which implies the usage of an accurate 3-D Monte Carlo model as a benchmark (only possible offline). The systematic and accurate documentation of these NWP radiative biases (together with a discussion of the relative importance of their various causes: poor representation of cloud + neglected horizontal photon flow) could serve as a useful guideline for a novel radiation parameterization, which would be accurate (with respect to the 3-D Monte Carlo on resolved cloud), but simultaneously fast enough to be used interactively. At the moment an interactive study is not possible (within NWP framework).

On the other hand, interactive studies of the effects of ICA biases on shallow cumulus clouds within LES framework, have already been carried out by the second author of this manuscript (see the Introduction section and the references therein – Jakub and Mayer, 2017; Klinger and Mayer, 2017).

Reply to "Minor comments"

1. pg. 1; L24: I'm not sure that the work here "highlights the importance of an improved representation of clouds even at the resolution of today's regional (limited-area) numerical models". Perhaps the (conditional) radiative errors are minor in the overall energy budget and the NWP model won't even respond to them in a significant way?

1. In order to study how an NWP model would respond to the improvement of radiative biases quantified in this study, a proper parameterization of these biases should be developed and coupled to an NWP model. This is out of the scope of the present study.

2. It would be useful to mention the resolution of the LES in the abstract.

2. Done. "(25 m grid spacing)" added in the abstract.

3. pg. 2; L22: This should be "Baker", not "Barker" (reference list, too).

3. Done. "Barker" changed to "Baker" in the text and in the reference list.

4. Title of Section 2.2: The assumption is that the NWP model gets this correct. It probably won't, and that could represent the majority of error! Moreover, if it gets cloud fraction and mean water content correct, why not assume that it does better than MRO and use something better than that in the simply model (like decaying exponential overlap)?

4. Whereas some global models certainly use more sophisticated overlap schemes (like decaying exponential overlap), most regional NWP models still apply maximum-random overlap assumption for partial cloudiness. Since we want to quantify current radiative biases in regional NWP models, we apply radiative solver together with the maximum-random overlap.

5. pg. 6; L13: Given the length of the explanation for the delta-Eddington and the simple model, it seems that Bugliaro et al.'s equation could be stated explicitly.

5. Probably you are right, although we still prefer to omit the equation of Bugliaro et al. (2011) and the related explanation of the parameters, because this would easily add an extra paragraph (and the entire manuscript is already quite long).

6. Section 2.4: Note that the McICA method is equivalent to the ICA when the full LES field is available to be sampled. When it is not available, additional assumptions are needed for the stochastic cloud generator... it would be interesting to see how it performs using some standard settings.

6. This would indeed be interesting, but is out of the scope of the present study.

7. pg. 9; L10: Why were standard conditions used as opposed to the LES's temp, vapour, etc.?

7. Although the (evolving) LES conditions could be perceived as more realistic, the use of standard conditions throughout the entire set of ten cloud scenes helps to isolate the effects of clouds on radiative biases, the focus of this study.

8. pg. 9; L18: 65,536,000 photons seem like over-kill for domain-average fluxes! Perhaps you could mention typical Monte Carlo uncertainties for HRs in cloudy layers?

8. We agree. The reason for the large number of photons traced is that we simultaneously aimed to investigate the differences between 3-D and ICA calculations locally in individual LES grid boxes (not for this study, but in particular for the above-mentioned study of the ICA biases as a function of model grid spacing, see General Comment #2). The highly-accurate results of these experiments, however, were used also in the present analysis. As a consequence, the Monte Carlo noise of the domain-averaged quantities discussed in the present study, is irrelevant for the overall ICA / 1-D radiative biases (relative to 3-D Monte Carlo).

9. pg. 9; L23: I presume that "cloud cover of 52.3% placed over land" was a typical condition given that it was highlighted often?

9. This is perhaps a misunderstanding. We started examining the dependence on SZA on a single cloud scene (section 3.1) and then subsequently incorporated the examination of one additional input parameter: albedo in section 3.2 (where the dependence on SZA is examined as well), cloud cover in section 3.3, by examining the entire set of ten cloud scenes (where the dependence on SZA and the dependence on albedo are examined as well) etc.
The cloud scene with cloud cover of 52.3 % does not necessarily reflect a "typical case", since in general there is a strong sensitivity of radiative biases to cloud cover. How do the radiative biases

scale with cloud cover on a full range of cloud cover between 0 and 100 % is described in section 3.3 (and section 3.4).

10. Fig. 4: Would this be a reasonable thing for an NWP model to do for SW radiation: Use the ICA (with a stochastic cloud generator) for SZA < 45 deg. and switch to asimpler (faster) 1D model for SZA > 45 deg.? Judging from this figure that might yield possibly tolerable errors???

10. This could be a solution for the cloud scene presented in this section (intermediate cloud cover of 52.3 %), but we don't want to make it into a general conclusion (the relative performance of ICA and 1-D depends on cloud cover etc.).

11. Section 3.3: Since 1D and ICA errors depend weakly on srf alb, perhaps this section could be reduced much in length (especially if you're looking to reduce overall length)?

11. We find the finding that the 1-D and ICA biases exhibit only a weak dependence on albedo (which could therefore be eliminated from the possible future radiation parameterization at a zero order) equally important as other findings regarding stronger dependencies of radiative biases on other parameters. Therefore we want to keep the discussion about the albedo effects in section 3.3 and elsewhere throughout the manuscript.

12. Fig. 7: Are these RMSE values for every grid-cell in the "cloud layer"?

12. Yes, this is the RMSE between the two averaged profiles (1-D, 3-D) within the cloud layer (excluding clear-sky regions above and below). I see that this was not explicitly mentioned in the text before (although it was mentioned that we examine only heating rates within the cloud layer, "in order to highlight the effects of clouds on radiative biases..."). Nevertheless, we have added this information in section 3.3: "The RMSE between the pair (1-D, 3-D) of heating rate profiles, ..." is extended to "The RMSE between the pair (1-D, 3-D) of heating rate profiles in the cloud layer, ...".

13. Fig. 8: For better comparison, the y-axes on the two rightmost columns should match:
-70:30 for the centre column and -20:50 for the right column.

13. Done.

14. Fig. 9: While the RMSE values look disconcerting, at times, it might not be an issue for the dynamics (cf. the abilities of GCMs to consume substantial radiative "noise" produced by the McICA method).

14. We agree. Nevertheless, in order to properly assess the impact of these biases on the dynamics, an appropriate interactive study should be carried out. We decided to keep the present discussion strictly diagnostic.

15. pg. 19; L8: "...demonstrates the general need for statistics." seems like odd terminology?

15. It might sound odd, but we think it is clear from the overall context what we meant.

16. Discussion near pg. 21; L15: Could there also be an underlying effect involving LWC variability as a function of mean LWC? What if the 1D model made some reasonable assumption about variance of LWC across layers (in addition to assuming that the NWP predicted both cloud fraction and mean LWC perfectly with the MRO)?

16. We are not sure if we understood the question correctly. The variance of the LWC certainly increases with increasing LWC, and this could partially explain the increase in the RMSE although we don't see a direct link between both. If the 1D model made some reasonable assumption about the variance of the LWC, this information could be used, however, to reduce the bias between 1-D and 3-D, see e.g. Shonk and Hogan (2008).

17. pg. 23; L13: I think that "Kablick III et al. (2011)" can be simply "Kablick et al. (2011)"... the III need be used only in the full reference.

17. Done. Kablick III et al. (2011) changed to Kablick et al. (2011) everywhere throughout the text.

18. pg. 23; L15: McICA is equivalent to the ICA when the subgrid-scale generator is perfect. It never is, but increasing levels of complexity can be easily added with confidence (e.g., accounting for correlations in sfc alb and LWP or Re and LWC)... this cannot be said for Tripleclouds.

18. Correct. The Tripleclouds radiative solver, on the other hand, has other advantages compared to the McICA: Tripleclouds does not produce any random noise at all. It is therefore justified to leave both in the outlook.