Response to RC 2

Effects of Model Grid Spacing for Warm Conveyor Belt (WCB) Moisture Transport into the Upper Troposphere and Lower Stratosphere (UTLS)—Part II: Eulerian Perspective

Cornelis Schwenk and Annette Miltenberger

August 13, 2026

General comments

This study uses ICON simulations to investigate the role of horizontal grid spacing in moisture transport by warm conveyor belts (WCBs) into the upper troposphere and lower stratosphere (UTLS). I find the Eulerian framework to define WCB ascent and outflow layers both interesting and novel. The results provide valuable insight into the radiative implications of WCBs when moving toward higher model resolution. Overall, the manuscript is well written, the figures are of high quality, and the topic is highly relevant for the weather and climate community and of clear interest to ACP readers. However, there are several points that, in my view, should be addressed before the manuscript can be considered for publication. These issues are detailed in the specific comments below. In particular, I am concerned about (i) the fact that this manuscript is presented as Part II of a two-part study, while Part I is still under review, and several key methodological aspects and interpretations are repeatedly referred to the companion paper, and (ii) the attribution of the results to the convection parameterization when the simulations differ simultaneously in grid spacing and in the treatment of convection. Given these points, I recommend major revision. *Thank you for the positive remarks and the careful suggestions for improvement.*

Major and specific comment

- Dependence on a companion paper under review: I have concerns about the fact that this manuscript is presented as Part II of a two-part study while Part I is still under review. The present manuscript relies on and repeatedly cites the companion paper for key methodological aspects and interpretation. This is unusual and, in my view, problematic for the evaluation of the present work as a stand-alone contribution. I would recommend either (i) making this manuscript substantially more self-contained,

(ii) postponing submission until Part I is accepted, or (iii) providing a clear justification and summarizing the essential elements from Part I that are required to assess the results here.
The paper in question has been accepted in the meantime and we refer to it accepted version throughout the revised manuscript.

- Description of the experiments: The description of the model experiments is rather limited. While I am aware that the experimental setup is described in a separate publication, I think the present manuscript should include more information to allow the reader to properly assess the results. In particular, I would expect at least a concise description of the vertical resolution and some indication of known biases or limitations of the simulations.
The experimental model setup is described in detail in the now published companion paper. We have added a comment about this to direct the interested reader to that publication. Additionally, we have added information on the vertical resolution and a statement to the limitation of the models.

- Convection versus grid spacing: The authors largely attribute the differences in WCB properties to the treatment of deep convection. However, the simulations differ not only in whether deep convection is parameterized or explicitly resolved, but also in horizontal grid spacing (13 km with parameterized convection versus 3.5 km with explicit convection). This makes it difficult to disentangle the respective roles of resolution and convection treatment. The current comparison therefore mixes two effects. Ideally, this would require experiments at the same horizontal resolution with different convection treatments, or at least a more careful and explicit discussion of this limitation. As it stands, it is not clear whether the reported differences should primarily be attributed to the convection scheme itself or to the change in resolution.
The reviewer is of course right that the use of the convection scheme is not the only difference in the simulation, although we have kept the other parameterisations identical some minor adjustments to settings are needed, that reflect explicit resolution of a different part of the motion spectrum. In addition, higher-resolution itself allows a more detailed resolution of underlying topography, steeper gradients and larger small-scale variability in the flow field all of which in turn may influence the simulated flow and cloud field. However, in an ideal model world one would expect convergence of the fields if the simulations are considered at the same temporal and spatial scales (with appropriate regridding applied). This is know to not be achievable with state-of-the-art models likely due to deficiencies in parametersations and their adaptability across the range of scales.
To ameliorate the concerns of the reviewer and RC1, we have regridded the high-resolution simulation data with a first order conservative method to the grid of the global simulation and repeated the analysis using the regridded data. Thereby we at least are able to remove the impacts of larger

variability in the high-resolution fields. Overall the main results are uneffected by this change. The main difference in the used model set-up (except for grid spacing itself) is the change in convection parametrization, we therefore think the largest part of the differences in the field considered on identical grids is due to the convection parameterisation. This is further supported by enhancement of differences in regions, where the convection parameterisation is active (as discussed in section 3.2 and shown in Figs. 10, A6 and A7 of the manuscript). While these are indications, a proof that the convection parameterisation drives the majority of the differences is not possible. We modified the discussion in the section 6 (conclusions) to clarify that the differences are not only due to the convection parameterisation, but that there are clear indications that the representation of water transport by the convection parameterisation plays an important role for the differences between the two simulations.

- Reference dataset / evaluation against reanalysis: Related to point (b), while the focus is on comparing two ICON configurations, it would strengthen the paper to include a reference against a reanalysis such as ERA5. This would help to put the results into context and to assess the magnitude and sign of possible biases in the different simulations, rather than only discussing differences between model configurations.
  This suggestion is beyond the scope of this paper, since it would include an analysis of an entirely different dataset. Firstly, in order to reproduce the WCB masks we would have to run trajectory simulations based on ERA5 data, which need to be calculated offline, resulting in the first departure from the methods in this paper (online trajectory calculation). The resulting analysis would further be muddied by the fact that the ERA5 dataset is different to our ICON data in a variety of ways. ERA5 data is produced by simulations that utilize an approximate resolution of 25 km, versus our resolutions of 13 km, 6 km and 3 km, and the underling model framework is inherently different to ICON (physics parametrization schemes, advection calculations, etc.). Finally, the reanalysis ingests observational data over 12 h intervals, which results in partly unphysical tendencies but ensured the data is closer to the observed state of the atmosphere than our free-running simulation. The difference between ERA5 meteorology and our ICON simulations is larger than the difference between our two different free-running ICON simulations. Hence, Mixing these datasets and comparing the results would be a large amount of work if done properly, and would introduce many major caveats that would shift the focus of this paper. We therefore refrain from including a comparison to ERA5 in the present paper. However, we mention in the conclusions, that this would be an interesting line of further inquiry (along with the comparison to observations themselves).

Minor and technical corrections

- Line 5-6: It may be worth mentioning explicitly that a 6.5 km simulation is also produced.
  In order to make clear that intermediated nests are used, but still keep the sentence brief for the abstract, we now state: "a convection-parameterizing run at ∼13 km and a nested simulation with a convection-permitting innermost domain at ∼3 km km grid spacing." This should provide enough context to make clear that multiple resolutions are used, but that our focus lies on the highest resolution domain.

- Line 17-18: In "the most important positive climate feedbacks...", consider giving one or two concrete examples.
  "[...] one of the most important positive climate feedbacks, alongside cloud and surface albedo feedbacks ([...], Forster et al., 2021)

- Line 25: There is an extra pair of parentheses in "((Guo and Miltenberger, 2025))".
  corrected

- Line 59: Part of the question "question: Can low-resolution, convection-parameterizing simulations (which are typically used for global forecasting and climate predictions) adequately represent WCBs,..." has already been addressed in previous studies (e.g., Dolores-Tesillos et al., 2025; Joos, H., et al., 2023). This should be acknowledged and briefly discussed.
  This is a good point, our open ended question indicates more novelty than necessary. This question has been addressed and also partly answered by the studies mentioned by the reviewer (some aspects are lacking).
  We now state: "While previous work has demonstrated that convection-parameterizing weather and climate models are generally able to reproduce many large-scale aspects of WCBs, they still exhibit systematic biases in their representation (Joos et al. 2023, Dolores-Tesillos et al. 2025). Together with the insights from recent high-resolution, convection-resolving studies, this raises the question of whether convection-parameterizing simulations can accurately represent the microphysical processes governing WCB moisture transport into the UTLS".

- Line 71: It is uncommon to cite a paper that is still under review; I would not recommend this. See major comment a).
  The paper in question has been published in the meantime.

- Line 75: What is the vertical resolution of the simulations? This likely matters for airparcel ascent and should be discussed.
  Information about the vertical grid used and the vertical grid spacing has been added to the description of the ICON set-up (section 2.1)

- Section 2.1: The description of the simulations lacks detail. Since the study focuses on vertical transport, I would expect a clearer discussion of

the vertical resolution and how it changes (if at all) with horizontal resolution. Is the vertical resolution scaled with the horizontal grid spacing? If not, how is this choice justified?
The information and a justification has been added to the ICON setup description.

- Line 116: This is unclear to me. Do you first run a global simulation and then perform additional simulations with two nested domains? Please clarify the experimental design.
To clarify, we now state: "Two separate simulations of this WCB case with identical setups but different different grid spacings have been conducted."

- Line 131: The "online-trajectory module" is mentioned, how does this differ from LAGRANTO (Sprenger and Wernli, 2015)? A brief comparison would be helpful.
We add: " Unlike offline trajectory tools such as LAGRANTO (Sprenger and Wernli, 2015), which calculate trajectories from archived model output, the online module computes trajectories during the simulation using the instantaneous model winds at every model time step, thereby avoiding temporal interpolation errors."

- Line 138: "within a visually defined region at a specific time..." is unclear. What exactly is meant by "visually defined"? Please clarify or rephrase.
We now state: "and be within a region defined by visual inspection of the cloud and moisture fields at a specific time."

- Line 158: The term "patchy" is vague here; please consider using a more precise term.
"patchy" is a term that many people will be familiar with, but it is not very precise. We still feel that it explains the resulting first-order mask field well, but now write: "The resulting mask is sparse ("patchy") [...]" instead.

- Line 166: I am not familiar with the "dilated" procedure. Please explain what is meant.
Dilation refers to including all neighboring points of all points in the mask. Multiple dilations allow the field to grow into the surrounding space, like a balloon being blown up. We now state: "The mask is repeatedly dilated (inclusion of neighboring points for all points in the mask) ca. 60 times in the global simulation and 300 times in the nested simulation (optimal number of iterations determined through trial and error)."

- Line 169: The word "crawl" is unclear in this context; please rephrase or explain more precisely.
The sentence is precise without "crawl" so we have removed this part: "This allows the mask to slowly ~~'crawl' forward and~~ fill in the WCB area without reaching into regions outside of the WCB [...]"

- line 175: TF2 and TF3 are not defined when first introduced.
  Good catch. The sentence works well without referring to TF2 and TF3: "The masks that we produced for both simulations ~~during TF2 and TF3~~ are visually very similar in terms of area"

- Line 184: How exactly are these fields "collapsed"? Please clarify the procedure.
  A 3D mask ist collapsed by taking the max value along (in this case) the z-dimension, thus giving the 2D "shadow" of the mask on the ground. This is now made clear in the text.

- Line 193: The word "muddy" is unclear; please replace it with a more precise term.
  We now write: "The qv field is therefore difficult to attribute to the WCB, which in turn makes it difficult to prevent the masking procedure from dilating too far."

- Line 207: I suggest changing "WCB detection" to "Eulerian WCB detection" for clarity.
  Good suggestion. Implemented.

- Line 209:
  - "rely on pre-training a machine learning algorithm on…" the wording suggests the model needs to be trained repeatedly; please clarify or rephrase.
  - "variables, but instead only requires qv." However, trajectories are still required. This sentence should be clarified.

  We now write: "It also does not rely on training a machine learning algorithm on a specific combination of variables and large amounts of similar data, but instead only requires qv and the trajectory data from one experiment."

- Line 16-17: It might be useful to compare the 3 km and 6 km simulations directly to better isolate the role of resolution.
  In this paper we discuss the role of scale in so far as it allows for the representation of convection, instead of having to be parameterized. The 6 km simulation uses a convection parameterisation. In addition, the 6 km simulation is affected by the two-way nesting adopted to ensure that the WCB outflow can propagate beyond the confines of the innermost nest and avoid issues at the lateral boundaries of the innermost nest. The 6 km simulation is hence simulataneously impacted by the convection-permitting cloud structure in the innermost nest and the coarse global simulation. In particular in regions of convective clouds this can lead to compensating tendencies from explicit and parameterised vertical motion (not shown). Therefore, we think the data is not a good comparison, but a re-run of the

model with only the 6 km nest would be required. This is beyond the scope of the revision. Note that for exactly the two-way nesting issue we also repeat the global simulation without and with the higher-resolution nest. Nonetheless include a larger spectrum of grid-spacings and the impact on the representation of convection would make an interesting follow-up study. We therefore add this to the discussion: "Additionally, it would be interesting to conduct a similar study, but with different horizontal grid-spacings that each allow for convection (i.e., 2km vs 4km resolution), to determine if similar differences in the vapor and hydrometeors fields can be found as well."

- Line 228: I suggest replacing "greater" with "larger".
  Done.

- Line 232: The parameters Ni, qi, and ri are not defined here.
  Have added descriptions

- Line 293: "convective precipitation over the preceding 15 min exceeds 0.01 kgm-2." The choice of this threshold needs either a reference or a brief justification.
  Have added: "[...] a threshold selected based on sensitivity tests and visual inspection to provide a robust separation between convective and non-convective regions."

- Line 394: "10-7 kgkg-1", please clarify where this threshold comes from.
  We add: "This threshold was chosen based on visual inspection of the hydrometeor distributions and is mostly consistent with values adopted in previous studies (e.g. Luettmer et al. 2025)."

- Section 6: I found this section somewhat confusing, as it reads more like a conclusion. I suggest splitting it into two sections: one for discussion (comparison with previous studies, implications, limitations) and one for conclusions.
  We have rewritten the conclusion to improve the clarity of the structure and to reflect the comments made by other reviewers.

- Line 459: Could your approach be used to validate ELIAS 2.0, or conversely, could ELIAS 2.0 be used to validate your method? This might be worth discussing.
  We write: "[...] and could serve as an alternative to or validation of the [...]"

- Line 469: A closing parenthesis is missing at the end of the sentence.
  Added.