Response to RC 1

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

1 Summary

Major Comments

SUMMARY
In this manuscript, Schwenk and Miltenberger present a North Atlantic warm conveyor belt (WCB) case study comparing moisture transport, hydrometeor properties, and radiative effects between convection-permitting and convection-parameterizing simulations. A key contribution is a novel algorithm that combines Lagrangian trajectory information with moisture fields to construct three-dimensional WCB masks, which are then used for an Eulerian analysis of WCB outflow regions.
The studies findings – higher moisture, fewer but larger hydrometeors, and reduced outgoing longwave radiation (OLR) in the WCB outflow of the convection-parameterizing simulation – are highly relevant for understanding WCB-related biases in upper-level dynamics and radiative budgets in climate models.
I particularly commend the authors on their innovative approach to create WCB masks and the generally clear and well structured presentation. The effort to diagnose and interpret the processes underlying the simulated differences is especially valuable, as it points toward potential pathways for improving convection parameterizations.
I look forward to seeing the study published after one major issue is addressed. This concerns the treatment of differing model resolutions. Using data at different resolutions both in the WCB detection algorithm and in the subsequent analysis raises the risk that some of the diagnosed differences reflect sampling effects rather than physical processes. As a result, it possible that parts of the results are influenced by sampling-related biases. Addressing this issue would substantially strengthen the line of argument in several parts of the results section.
I detail this resolution concern, along with additional general comments, below,

followed by specific comments on individual sections and figures.

Thank you very much for the thorough reading of our manuscript, the positive comments and the good suggestions for improving our work.

General Comments

As noted above, I have one major concern, which relates to the treatment of simulation output at differing horizontal resolutions. As I understand the setup, the authors analyse output at 3.5 km resolution for the nested and at 13 km resolution for the global. Consequently, one grid point in the global simulation corresponds to roughly nine grid points in the nested domain, with clear implications for the sampling and statistical representation of quantities that vary on small spatial scales.

The authors acknowledge this issue in the manuscript, but its implications are not addressed sufficiently in the analysis. In my view, this is problematic because resolution-related sampling differences may influence both the identification of WCB regions and the statistical properties diagnosed within them in an unintended manner. Below, I outline two specific ways in which this resolution mismatch could affect the results and interpretation, and which I believe should be addressed to strengthen the robustness of the study's conclusions.

1. Differences in WCB masks partly due to sampling?
   The WCB mask algorithm is elegant and provides a valuable way of evaluating a Lagrangian feature in an Eulerian framework. However, I am concerned that the use of original-grid data at different resolutions in the WCB detection algorithm may introduce sampling effects that influence the resulting WCB masks.
   In particular, I wonder why the authors do not use moisture fields (q_v) regridded to the resolution of the global simulation as a common basis for identifying WCB masks in both simulations, which would seem more consistent to me. The finer grid of the nested simulation inevitably exhibits greater small-scale variability in q_v. As a result, regions in which high- and low-q_v grid points coexist at high resolution may yield different mask extents than the same fields after regridding, even if the large-scale structures were identical between simulations. Under the current approach, this alone could lead to systematic differences in WCB masks. This concern is particularly relevant for high-level WCB masks, where a more spatially dispersed WCB with diffuse edges may lead to the inclusion of grid points with relatively low q_v in the high-resolution masks. While I am not certain to what extent this effect is quantitatively important, Figure 3 is at least consistent with this interpretation, as a small gap in the high-resolution WCB mask corresponds to a much larger non-WCB area in the global simulation. The systematic differences found in WCB-mask size make me particularly interested in this effect: Maybe, part of the difference in size

comes from the different grid spacing?

In addition, the differing resolutions necessitate the use of two slightly different versions of the mask-generation algorithm, with different numbers of iterations that are determined by trial and error for each resolution. Regridding the high-resolution data prior to mask generation would allow the same algorithmic settings to be applied to both simulations, thereby removing one subjective element from the procedure.

A final, more conceptual concern relates to the use of a threshold in a quantity that is itself analysed within the WCB regions. Since WCB impacts on humidity and hydrometeor properties are a central focus of the study, it is not obvious that defining WCB membership based on a fixed q_v threshold is entirely unproblematic.

In summary, I am concerned that the use of original-grid output for WCB mask generation may introduce resolution-dependent sampling effects. If the authors can convincingly argue that these differences are physically meaningful or even desirable, this concern may be alleviated. Otherwise, I suggest testing the sensitivity of the WCB masks to the use of regridded data and, if necessary, repeating parts of the analysis based on this more consistent masking approach.

Indeed using two different resolutions for mask generation may lead to systematic differences between the masks. In general the edges of WCB outflow (and gradients in general) are sharper, i.e. representable gradients are larger (also well visible in Fig. 3), in high-resolution data (also well visible in Fig. 3). Therefore we would argue that the WCB outflow mask is better defined in the high-resolution data compared to the low-resolution data. Also the size of the masks in both data-sets is comparable with the deviation in area in the below-percent range. However, we have tested the impact of using the regridded q_v *data from the convection-permitting simulation with the masking algorithm from the global simulation. The differences to the original high-resolution mask (regridded to the grid spacing of the global run, see following comment) are small. As example we have included Fig. 3 from the original manuscript, but now also including rows for the regridded* q_v *and mask field (third row) as well as mask derived from regridded* q_v *field (fourth row) as Fig. 1 in this reply. The impact on the distribution of outflow properties is relatively small for all considered variables and likely arises from slight shifts of the mask edges.*

As special care as been taken with the mask construction in the high-resolution data including some hand-crafted masking, which we were not able to repeat for the re-application of the global mask, and because we think there is added physical value in the larger-gradient field from the high-resolution data, we have retained the original mask for the analysis (regridded to the global grid spacing, see next comment).

As to using q_v *for defining the WCB outflow, we have looked at other independent variables as e.g. temperature and potential vorticity, but did not find as suitable structures for the mask extraction.*

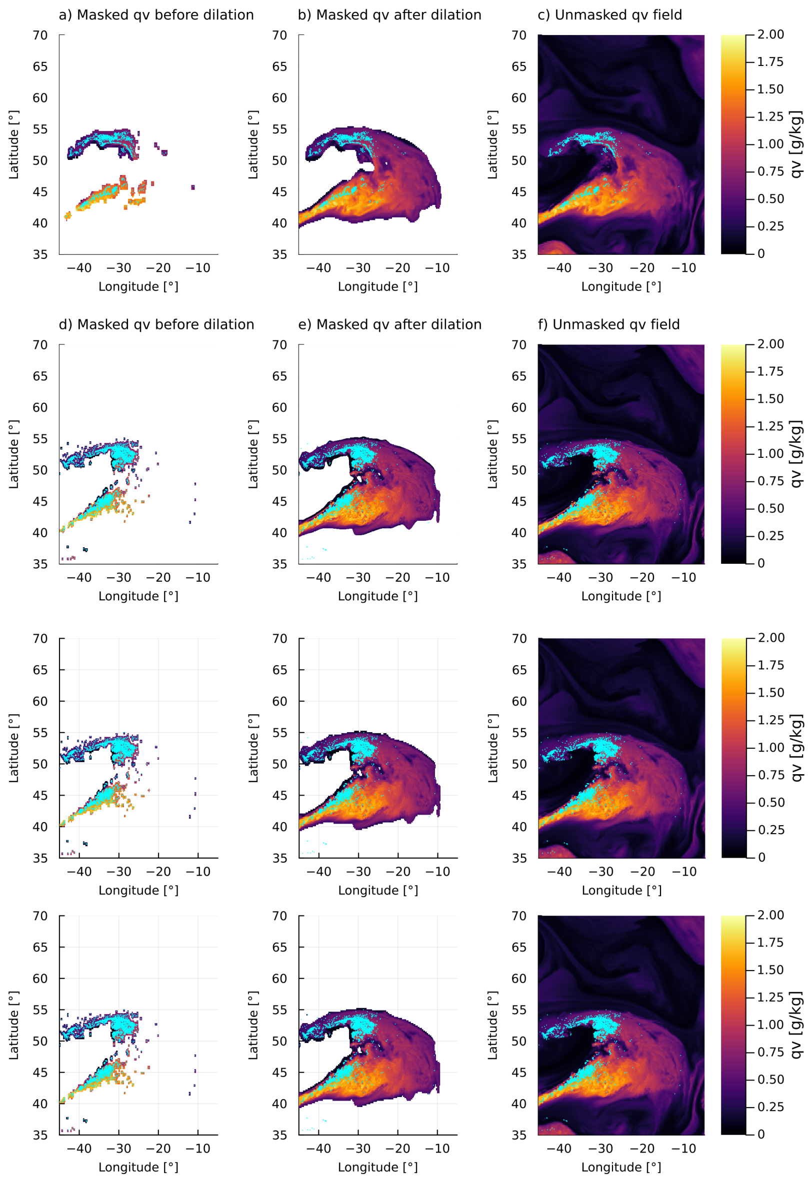

Figure 1: q_v at 400 hPa for the global simulation (top row) and nested simulation (bottom row) at 02:00 UTC on the 22. September, 2017. The area within the WCB mask is shown before (a, d) and after (b, e) the dilation. The the total q_v field in (c, f). Trajectory points within ± 25 hPa of this pressure level are plotted in cyan. The third and fourth row show the fields using (i) q_v and mask regridded to the global resolution and (ii) q_v regridded to the global resolution and re-derived mask (using the algorithm used for the global data).

2. Sampling effects in the evaluation of WCB-area properties
   Beyond the identification of WCB regions, similar sampling issues may also affect the evaluation and interpretation of WCB-area properties. For much of the analysis, the authors rely on output at the original model resolutions, even when discussing not only mean values but also spatial structure and full distributions (see, for example, the discussion around line 265 ff.). For quantities with pronounced small-scale variability, differences between the simulations may therefore partly reflect differences in sampling density rather than genuine differences in physical behaviour.
   In particular, differences in spatial structure or in the width and tails of distributions may arise simply because the higher-resolution simulation resolves more small-scale variability, rather than because the underlying processes are fundamentally different. This is relevant not only for the interpretation of distribution tails, but also for dynamically important variables for which scale interactions play a role. One example is vertical motion, which is discussed in connection with Figure 4, where finer resolution alone could lead to apparently stronger variability or more extreme values. A possible way to address this issue would be to complement the existing analysis with diagnostics based on regridded data, in addition to the original-resolution fields for those cases where sampling effects are likely to influence the conclusions, as already acknowledged by the authors around line 265 ff. This would allow the authors to assess how the higher-resolution simulation behaves at spatial scales that are also represented in the coarser simulation, and to better separate resolution-related sampling effects from physically meaningful differences.
   Such an approach could also strengthen the interpretation of results for hydrometeor number concentrations, Since the central argument concerns aggregated effects of convection-permitting versus convection-parameterizing behaviour – rather than purely stochastic or noisy variability associated with explicit convection.
   Thank you for rising the important and indeed under-discussed aspect of different sampling statistics and represented variability at different grid spacings. We have now regridded the high-resolution simulation to the grid of the global simulation using a first order conservative method and use the regridded fields for the majority of the analysis. While some changes in the tails and width of the distribution emerge (the mean / median is as expected largely uneffected), the key features of the WCB outflow property distribution remain intact. We use the analysis based on the regridded fields for the majority of the paper and retain some distributions based on the high-resolution fields in the Appendix, where we think additional insight from the resolved variability from the high-resolution simulation aids / adds to the physical insight.

Some more general, but less critical points:

1. This is more of a framing-issue: You consistently take the convection-permitting simulation as a kind of ground truth. While this is under-

standable, I would consider wording the comparison a bit more neutral due to the lack of an actual ground truth to which to compare the simulations.
We adjusted the wording in the abstract and conclusions accordingly.

2. Relevance of the findings for NWP: I find the results on UTLS humidity very interesting from an atmospheric dynamics point of view. Maintaining tropopause sharpness hinges on producing accurate humidity gradients and representation of radiation at the tropopause, such that the presented results are relevant not only for climate- but also weather modelling. I encourage the authors to mention this in their abstract as well as discussion to point colleagues in the dynamics community toward the relevance of this work.
Thank you for this excellent suggestion. Indeed this is an important aspect, which we currently pursue in follow-up studies. We have added some words / sentences to the abstract / conclusions, respectively.

3. My last general question concerns the choice of analysed domain: The authors choose only nested-domain trajectories, specifically for TF3. I don't quite understand why the resolution is so critical here – if the aim is to study the impact of the convection parameterization on WCB-outflow. If the air parcels are mostly just travelling at high levels in this time frame, it doesn't seem too important that the grid size is the same at this stage? Shouldn't the effects of explicit convection (even if in a mixed manner) survive the air parcels travelling outside the high-res domain?
The reviewer is right to point out that the impact of convection should persist outside of the innermost nest due to the two-way nesting strategy. This was indeed originally the motivation to choose a two-way nested configuration. However, an in-depth analysis of the physical and nudging (towards the higher-resolutions simulation) tendencies shows some compensating effects in the outer domains (not shown), which may arise due to inconsistencies between e.g. parameterised convective transport and resolved convective transport and the attempt of nudging to imprint the signatures from the higher-resolution nest onto the coarser resolution simulation. Therefore we the analysis of the global and intermediate nest becomes very involved, partly subject to specifics of the ICON nesting strategy and less focussed on physical processes. Note that for this reason the global simulation used for the analysis here has been conducted without any high-resolution nests included. Therefore we think the paper benefits from restricting the analysis on the innermost model domain.

Specific Comments

Abstract

- Line 6: Here, we present ...
  Corrected

- Line 17 ff: "variations in water vapor concentration drive one of the most important positive climate feedbacks (Li et al., 2024; Held and Soden, 2000; Dessler et al., 2013, e.g.,)," Why not briefly mention which feedback mechanism this is?
  Done

- Line 35 ff: There is a mixing of singular and plural formulations regarding WCBs here. Consider rewriting this for clarity?
  Done

- Line 90: The sentence starting with "Therefore" has some grammar problems and is rather lengthy. Maybe split into two?
  Done

Methods

- Line 134: 'The starting regions were selected ..' or both singular
  Done

- Line 139: 'visually defined region' does not really help me understand your criterion here. It would help if that would be a bit more specific (which I think would be good, even though you describe the method in a detailed manner in Part I.)
  Done

- Line 184: There is a 'that' too much.
  done

- Line 214: this should be 'for both simulations'.
  Done

Results

Here, I thought whether it might be useful to generally clearly delineate between things that are in line/ in (apparent) contradiction with findings from the Lagrangian study, as well as completely new aspects that could only be found in the Eulerian analysis. You do this in parts of this section, but might be a bit more consistent in doing so.
We reworded parts of the text to make this more clear. Also this is aspect is emphasized more in the conclusion section now.

- Line 228: With Results being consistent across TF2 you mean spatially / temporally/ both?
  Both: The shifts in the vertical velocity distribution are similar irrespective of the pressure level considered (spatial consistency) and the time instance within TF2 (temporal consistency). The sentence has been rephrased to make this clearer.

- Line 234: You talk about Figure 8 here, before Figure 5-7. Consider moving Figure 8 up, so the reader doesn't need to scroll that much?
  Figures have been reordered.

- Line 232: How about briefly defining N_i, r_i, q_i?
  Done

- Line 260: 'while they are almost identical in both simulations ...' ?
  Corrected

- Line 321: 'Which are much more abundant' ?
  Corrected

- Line 410: 'The large differences found for OLR'
  Corrected

- Line 415: I think there is a minus sign missing in front of the temperature.
  Corrected

- Line 434 f.: 'These values are chosen'
  Corrected

- line 424: 'That the effect described here is stronger in the global than in the nested'
  Corrected

Conclusions and Discussion

- Line 449: Remove 'the' ?
  Corrected

- Line 468: Paranthesis missing after (Fig. 12)
  Corrected

- Line 497 ff: I would appreciate if this comparison to the first part was more concrete. How about briefly stating the difference again, so people don't need to jump between papers to understand (especially the) discussion? Additionally I find the argument you make in line 500ff somewhat hard to follow. Maybe you try to rephrase it such that the connection between the first and second sentence of the argument becomes more clear?
  We restructured the part of the conclusion that summaries the results and puts them into context of the part I of the present study to enhance clarity and to reflect some of the comments from the other reviewers.

- Line 505 ff: I think you somewhat understate the relevance of your results for climate modelling. The models are tuned to be balanced for present-day climate, but as you mention later, if the overall role of convection changes in a warmer atmosphere, the mis-representation of radiative effects of convection might lead to biases in the energy budget of the planet. Maybe you could consider merging your last paragraph, where you discuss this, to this paragraph.
  We enhanced our statements with respect to climate modelling.

- Lastly, I think that your findings are relevant not only for climate modelling, but also for NWP, where we know that the humidity-gradient at the tropopause is relevant for accurately forecasting jet intensity and RWB evolution (e.g. Gray et al., 2014; Saffin et al., 2017) and hence forecast fidelity.
  We added a few sentences with respect to the role of our findings for synoptic-scale weather evolution.

Figures

- Figure 1: Caption misses description of colored contours, likely SLP. The dashed black lines are kind of hard to see, especially given the also black coastlines. Maybe a different color (a light grey?) could distinguish it more clearly from other features?
  The caption has been extended accordingly. The figure has been added to improve visibility of the line indicating the trajectory starting region.

- Figure 3: Sentence 3 has a 'the' too much in the beginning. You could also consider noting 'global' and 'nested' simulation beside the respective rows, to make the figure even more self-explanatory.
  Caption corrected. The figure has been edited as suggested.