Authors Response to Referees

Implementation and evaluation of diabatic advection in the Lagrangian transport model MPTRAC 2.6

Jan Clemens

March 2024

We thank Referee No. 1 and No. 2 for their reviews and for further guidance on how to revise our manuscript. We have revised the manuscript according to the reviewers' advice. Below are enlisted the detailed responses. Pages and lines are provided as they are in the revised manuscript.

1 Reply to review 1

Comment:
This article describes the adaptation of a Lagrangian transport model, MPTRAC, to use diabatic velocity coordinates, the use of which are beneficial to Lagrangian transport calculations. The main component of this is an interpolation scheme to compute these diabatic velocities at particle locations, which requires a vertical search.
The adapted model is benchmarked against another model, CLaMS, using a sensible set of case studies, where the errors are compared to those within MPTRAC upon variation of parameters. This is a good practical approach.
My main point about the paper is that it is motivated by performance on next generation HPC, but nothing is said about performance in this paper. I appreciate that performance optimisation is a second step after checking that the numerical approach is sufficient (that's the main content of this paper). However the general reader would appreciate a bit more discussion. So,
1. What is it about MPTRAC that makes it expected to perform better than CLaMS on GPUs etc? Are these properties preserved by the new adaptation? Can this be argued from a performance model?
2. Can you provide a performance profile of the new adaptation versus the standard MPTRAC to see where potential bottlenecks have arisen, and discuss (if necessary) how they would be addressed in performance optimisation?
I think that addressing these questions would provide a better documentation of this stage of the development of MPTRAC.

Answer:
One limitation of CLaMS is its lack of shared-memory multiprocessing capability with OpenMP, unlike MPTRAC. CLaMS can only use MPI parallelisation to distribute air parcels across multiple processes, which leads to a major problem as CLaMS reads the entire wind field for each individual MPI process. This results in an accumulation of high-resolution data (e.g. ERA5) on the compute node with every newly added process, strongly limiting the scaling of CLaMS on the compute nodes. However, the implementation of the diabatic scheme into MPTRAC has overcome this problem.

The new adaptation of MPTRAC preserves the data and loop structure of MPTRAC and is as compatible with the former OpenACC and OpenMP, i.e. shared memory multiprocessing and offloading, as before. The preservation of the structures is due to the fact that the components have mostly swapped roles. Pressure tendencies are replaced by zeta tendencies in the code, and air parcel positions are no longer expressed only in pressure but also in zeta. However, based on preliminary profiling, the advection kernel on the CPUs shows a performance loss of about 30%, due to a more complex interpolation required for the coupling between former modules and the zeta coordinate. As a way of speeding up the computations, we consider using a faster but less accurate interpolation method for intermediate interpolations within the advection steps. The precise interpolation could only be evaluated when entering and leaving the advection module. We would like to do this as a second step together with further performance optimisations.

We have discussed the performance aspects and aim of the paper now in the manuscript in more details as we agree that it needed clarification. We added: *"In this study, we evaluate the accuracy of the newly implemented scheme in MPTRAC through a detailed intercomparison with results from the CLaMS model, and by placing model differences in the context of the sources of uncertainty inherent in Lagrangian transport models."*

at page 2 line 54, as well as:

"The trajectory code of the CLaMS uses only the Message Passing Interface (MPI) to distribute air parcels across multiple processes."

at page 2 line 29, as well as:

"The new implementation in MPTRAC involves a role reversal between pressure and zeta coordinates, and pressure and zeta tendencies. This ensures that the data structures maintain the required structure for memory sharing, multiprocessing, and offloading to GPUs. In contrast, adapting the CLaMS code for parallelisation with OpenMP and OpenACC would require restructuring of loops and extensive rewriting of data structures to define proper data regions accessible for the shared memory or the GPUs.". at page 2 line 50. We close the paper with: *"Furthermore, now that the implementation has been validated, additional performance analyses and optimizations can be carried out."*

2 Reply to review 2

Comment: Page 2, line 41: What implies an improved "interoperability" between CLaMS and MPTRAC? Does this mean, that the advection of trajectories is calculated by MPTRAC and the mixing process by CLaMS? As far as I know, in CLaMS, the calculation of mixing is started by gathering the trajectories onto one CPU. If that is correct, doesn't this step slow down the performance of MPTRAC? What is the "coupled mode" in MPTRAC?

Answer:
The CLaMS mixing wasn't applied in the study. However, CLaMS modules make use of the trajectory calculations in zeta coordinates. Hence, the implementation of diabatic transport into MPTRAC is required to make use of this modules to the fullest extent. Hence, "interoperability" is improved, because MPTRAC now can fulfill more functionalities that the CLaMS trajectory module offered before.
Indeed, CLaMS gathered the trajectories onto one CPU. We agree that this might lead to a slow down of the calculations with MPTRAC. However, MPTRAC has its own mixing module, that is suitable for GPUs. The interoperability with legacy CLaMS code and concepts is relevant to evaluate newly implemented mixing schemes in comparison to the legacy code or if we decide to rewrite the CLaMS mixing for GPUs. With our approach we guarantee backward compatibility (i.e. also using inefficient but well validates modules) and the transition to GPU codes at the same time.
The coupled mode of MPTRAC is a calculation where the advection is performed in zeta coordinates, but the further parameterisations operate on pressure coordinates. For example, the module for turbulent diffusion adds random disturbances to the position of the air parcels given in pressure coordinates. To clarify what we meant, we rephrased to *" Thus, in MPTRAC, advection can be performed with the diabatic scheme in zeta coordinates as in CLaMS, while at the same time, modules based on pressure coordinates, such as the particle diffusion or convection module can be employed (This mode is also refereed to with "coupled mode" in this study) (See page 2, line 46)."*

Comment: Page 4, line 94: .. model differences are evaluated. Better say: ..model results are evaluated.

Answer: We agree and changed it.

Comment: Formulae 1 and the corresponding text do not match. The zeta coordinate at the sigma_r-level is the last value of a smooth transformation to potential temperature and not the beginning of the transformation. Please correct the sentence.

Answer: We have corrected the sentence to: *"Near the surface, the hybrid zeta coordinate follows the orography in the form of a sigma-like coordinate. With increasing altitude, the zeta coordinate is smoothly transformed into the potential temperature $\theta(p)$, which is reached at the reference level ($\sigma_r = 0.3$)."* (See page 137, line 5)

Comment: Page 5, Eq. 2: Is the last term of Eq. 2 correct? I mean that the inner derivation of the sinus function (of Eq. 1) should be: -sigma_dot(p)/(1-sigma_r)

Answer: We have corrected the equation.

Comment: Page 5, line 132: To my knowledge, diabatic heating rates are also calculated from convective and diffusion tendencies. Only in the stratosphere, radiative heating rates dominate.

Answer: We agree and have elaborated in more detail: *"While the diabatic rates are derived from the energy balance including among others radiation, latent heat and turbulent mixing, kinematic rates are calculated from the continuity equation."* (See page 5, line 44)

Comment: Page 5, line 138: You state that there is a need to smooth zeta profiles, that are not monotonic with height. My question is: Did you smooth the zeta profiles after you calculated them with Eq.1 ? If yes, how can you guarantee an exact transformation from zeta to pressure and vice versa? Don't you violate equation 1 then?

Answer: Yes, the smoothing is done after the calculation according to eq. 1. Since both the pressure and the smoothed zeta profiles are monotonic, an exact transformation is possible using interpolation, although Eq. 1 is no longer satisfied.

Comment: You mention that "many processes in the troposphere are not diabatic (e.g. convection)". How much is "many" and why is this statement necessary? To my knowledge, convection is a diabatic process in the troposphere.

Answer: We have removed this statement.

Comment: Chapter 2.1.3: The chapter on "interpolation" confuses me. The more I read the less I understand. Please, see my specific remarks below:

Answer: We understand that this chapter may have been difficult to follow. Therefore, we have made some changes to make it easier to understand. We have rephrased some sections and added a second example to help explain the general concepts of interpolations.

Comment: Page 6, line 166: "Interpolation with positions given only in eta coordinates therefore requires additional consideration". You should shortly describe, what the additional considerations are.

Answer: We have added *"Interpolation with positions given only in zeta coordinates therefore requires additional considerations, e.g. about how to find the vertical position of the box that includes an air parcel, when the data is not stored in the air parcels coordinate."*, to point out the major question to consider. (see page 7, line 180)

Comment: Page 6, line 168: You perform a time interpolation locally for each air parcel in MPTRAC, and in CLaMS you interpolate the wind field globally for the time step procedure. Maybe there is some information missing in this sentence? I do not understand what you want to say.

Answer: We further extended on this by *"A further difference is that time interpolation is performed locally for each air parcel in MPTRAC, e.g. wind data is collected around the position of the air parcel and subsequently interpolated in time. In contrast, CLaMS interpolates the wind field in time and globally in advance for the four time steps of the Runge-Kutta scheme, i.e. the entire wind data field is interpolated in time and subsequently used for all air parcels."* (see page 7, line 190)

Comment: Page 6: line 175: Please reformulate the sentence beginning with "At the beginning of ..." to for instance: "In CLaMS at first, the interpolation in time is performed".

Answer: We have changed this.

Comment: Page 7, Figure 1: A figure should make a description in the text more clearly, however, the representation of 6 in my view identical cubes do not help understand, how the interpolation is performed. Moreover, the description of the left upper cube is "zeta", with a head line "3-d data is in eta coordinates" (which seems to be the ECMWF data). This description confuses me. I would suggest to describe the procedure of interpolations in a bullet list and delete of the upper part of Fig. 1. Note: It would be easier to read figure 1 if it would be made larger. The same applies for Fig.2.

Answer: We have modified the figures to address the issues. The second part of the figures is now larger, and in the first part we changed the captions and focused on the most important information. Additionally, we added a 2-d example to introduce the basic interpolation concepts before explaining the detailed interpolation method.

Comment: Page 7, line 190: "If a different box is used for re-interpolation..". I would suggest to delete this part of the sentence, just say "Then, significant errors may..."

Answer: We agree, that this information can be removed in line 190.

Comment: Page 8, line 199: each of these columns....

Answer: Fixed.

Comment: Page 9, line 228: The statement of this sentence is not clear: You "compare" the time interpolation for the air parcel in MPTRAC with the meteorological field as in CLaMS? A similar sentence can be found on page 6. Please reformulate the sentence.

Answer: We removed the sentence here, because we now already explained it at page 6.

Comment: Chapter 2.3.1: I hope that I have not overlooked it in the text, but what I found missing here, is that you explicitly state and emphasize that you evaluate a MPTRAC simulation with a CLaMS simulation. I note this, because it is mentioned as the first result in the abstract! Further, a reference uncertainty source (I guess it is the "Transport" uncertainty source) should clearly be mentioned, otherwise the uncertainties have no (physical) reference value.

Answer: We now begin the chapter with *"The implementation of the diabatic transport scheme in MPTRAC, used with the ERA5 reanalysis, is evaluated by a detailed intercomparison with CLaMS trajectory calculations for a global ensemble of air parcels. To put the differences found in the trajectory calculations between CLaMS and MPTRAC in a broader context, the effects of, first, external sources (using different reanalyses, resolutions and vertical velocities) and, second, internal sources (e.g. interpolation and integration methods) were investigated."* to emphasize that we intercompare MPTRAC results with CLaMS results and want to compare the existing uncertainties with a bunch of reference uncertainty sources. We added the sentence *"Finally, we introduce the deviation between the initial position of the air parcels and their final position as a physical reference to compare with. This deviation is labelled "transport" in Table 2."* in addition at the end of the chapter to introduce this reference properly.

Comment: Page 15, figure 3c middle: Can you explain the pronounced daily cycle in the lower stratosphere only for CLaMS-default, MPTRAC-bestfit, MPTRAC-default?

Answer: The (apparent) daily cycle in the stratosphere could have many different causes, such as gravity waves or the diurnal tides.

Comment: Page 15, figure 3 caption, "theupper"

Answer: Fixed.

Comment: Page 16, figure 4: Colors are difficult to differentiate on paper!

Answer: We have modified the figures to be better readable.

Comment: Page 16, caption fig. 4: Please explain, why it is acceptable to exclude these outliers. I wonder, if the estimation of a maximum value is meaningful at all, if you ignore outliers? Moreover, could you please describe, why the selected value of 1.5 of the inter-quartiles is an appropriate value? Is it possible to show a figure, where the outliers are included?

Answer: We used the default values of pythons MATPLOTLIB.PYPLOT.BOXPLOT routine, which include 84% to 93% of all air parcels and hence the most part of the statistics. However, we have include the entire ensemble in the new version of the manuscript, as we agree that this gives the complete overview.

Comment: Page 19, line 362: Here you state for the first time in the description of Fig. 4-6 the "overall transport median". In the figures 4-6 it is shown as the transport AHTDs or AVTDs. I hope I have not overlooked it elsewhere, but what I miss in this chapter is that you clearly mention the reference value to which all of the calculated uncertainties are related to. As the "transport" is defined as the difference between start and end point of the trajectory, I would not see it as an uncertainty quantity, rather a (physical) difference reference value, where uncertainties should be compared with.

Answer: We mentioned it only in one of the Table, and agree that it should be mentioned more clearly. Therefore we added the sentence *"Finally, we introduce the deviation between the initial position of the air parcels and their final position as a physical reference to compare with. This deviation is labelled "transport" in Table 2."*, as mentioned before.

Comment: This statement is not represented in Fig.4-6. Please, include the respective values of uncertainty or a reference to another figure, where these results are presented. I think this is an important result, where you show, that your model development in MPTRAC is in accordance with CLaMS, which is a widely published and evaluated model!

Answer: We have references to the Figures 4-6 now where it is presented as "best-fit" (see page 22 line 415.)

Comment: Page 19, line 372: Is it really worth to mention differences in results due to compilation flags? The question, however, is: Are results still reliable or equal in a statistical sense, when model components or even a different computer architecture, compiler etc. is used?

Answer: The reference to the compilation flags has been removed. Although, in the study by Hoffmann (2022), minor differences were found between calculations performed with different compiler flags, these differences are likely negligible for our analysis. Additionally, we are currently conducting a larger benchmarking study with MPTRAC on multiple machines, which will provide more detailed answers to such questions.

Comment: The maximum of the normalized mean AVTD for "diffusion" can be found in the tropopause region. What can be concluded from the differences in vertical structure of normalized AVTD of "diffusion" in comparison with mean AVTD? Why are the profiles for diffusion of mean AVTD and normalized AVTD not equal in shape, if the normalization is the mean vertical path length of CLaMS, e.g. one value?

Answer: The normalization is done for each height level separately. Therefore the profiles for diffusion are not supposed to be equal in shape. The peak of normalized mean AVTD of diffusion over the tropopause is a result of a smaller vertical transport and normalization in this height range (15-25 km). The vertical transport over the troposphere region is therefore, in relative terms, strongly increased by the parameterized diffusion. We now have added the information that the normalization is done for every level and not for the entire profile at once. We also strongly reduced the smoothing as it deformed the normalized profiles inconsistently to the absolute profiles. Also note that we have corrected the effective height, which wasn't plotted correctly before.

Comment: Page 20, line 413: "largest source of uncertainty", isn't it the largest relative source?

Answer: Yes, we corrected the paragraph to reference to the correct uncertainties in the Figure (See p 24, line 473)

Comment: Page 20, line 413: Afterwards is not the correct adverb in this sentence.

Answer: We have removed this sentence.

Comment: Page 21 fig 7: Please add "average AVTDs for MPTRAC scenarios (see also Table 2)" in the figure caption.

Answer: We added a reference to the table.

Comment: Page 22, line 432: Say: vertical transport deviations remain smaller...

Answer: Fixed.

Comment: Page 25, lines 459: You write that the convergence of air parcels (Fig. 10) near the tropopause and near the surface is a consequence of up- and downdrafts in the troposphere including the tropopause and surface as a transport barrier. Though, I would suggest that the convergence of parcels near the surface/tropopause seems to be a result of missing mixing and convective transports, as the troposphere is a well-mixed region. Please comment.

Answer: You are right, considering the correct definition of up- and downdrafts, which are local an small scale per definition. We meant, however, the mean vertical upward and downward transport in the reanalysis data fields. The troposphere is a well-mixed region. The Lagrangian transport calculations without resolved mixing and convective updrafts, as we apply here, do not the follow the well-mixed criteria. Instead they show vertical transport according to the resolved mean flow. With fully resolved convection e.g. by parameterization, the agglomeration of the air parcels at the top and bottom of the troposphere presumably would be reversed. Therefore, we reformulated: *"Sub-grid scale process, such as convection, that would be required to reach a well-mixed troposphere are not parameterized in the calculations. Therefore, the accumulation of air parcels is a consequence of up- and downward transport limited to the resolved mean flow of the troposphere, combined with the tropopause as an upper transport barrier and the ground as the lower transport barrier. " (See page 28, line 524).*

Comment: Page 25, line 468: You describe the green lines in Fig. 11a as contour lines of air parcel frequencies. Are they absolute number of parcels of the CLaMS default simulation? If yes, please specify this also in the Fig 11 caption.

Answer: No, we refer to the air parcels per cell and specify this now in the caption.

Comment: Page 25, line 473: This sentence is unclear. What is meant by "patterns", or "smoothed peaks"? green contours shrink?

Answer: We have removed the term "patterns" with *"air parcel distribution as found without diffusion"*. We specified that the circumference of the contours shrink. We keep 'smoothed peaks', but specified that it means that we find the air parcels to be scattered more strongly around the maxima of the air parcel distribution. (See page 29, line 541)

Comment: Page 26, lines 490: What did you want to say with "climatological findings"? That more air parcels are found at higher altitude around latitude of 45° with ERA-interim? I would suggest to replace "climatological findings" with "the results by Ploeger et al." Please comment.

Answer: We agree that this statement needs clarification. Therefore we wrote: *"These results are in agreement with findings of Ploeger et al. [e.g. 2021] who studied the climatological zonal structure of diabatic velocities in ERA5 and ERA-Interim."* (See page 29, line 569)

Comment: Page 28, line 521: Say:.. differences between scenarios are small...

Answer: Fixed.

Comment: Page 29, line 525: "to assess the implementation of the zeta coordinate...", shouldn't it be "to evaluate the implemented zeta coordinate in MPTRAC, we conducted simulations using..."?

Answer: Yes, we reformulated the sentence.

Comment: Chapter 4: The conclusion section is rather long and repeats the description of results in detail. I would suggest to shorten the conclusion section (more a summary and conclusion section) and focus on the main results of the evaluation of the new model and the main results of the analysis (scenarios, uncertainties, effect on parcel distribution ..)

Answer: We shortened the conclusion section by removing some details and focusing on the main results.

3 Further remarks

We have revised many parts of the paper to address the remarks and questions. In addition to the reviewers' comments, we revised the manuscript as listed in the following. Typos and smaller revisions can be found in the file highlighting track changes.

In Table 1, the integration methods were mixed up. MPTRAC always uses the mid-point scheme with the default set-up. This is corrected in the new manuscript.

We added the great-circle distance as a measure of the AHTD, because the Euclidean distances looses to much of accuracy for the global runs after 90 days, see Section 2.3.2 (Eq. 8).

In Figure 8, we mixed up the Southern and Northern Hemisphere in the description and text, what is corrected now in the revised manuscript.

We corrected the effective height everywhere, as it was plotted incorrectly in all plots.

Fig. 13 was slightly adapted, because the last data point wasn't plotted correctly.

The appendix B was extended to support the statement that the difference between kinematic and diabatic transport calculations is lower in ERA5 than in ERA-Interim. Furthermore, Fig. B2 was plotted again with the correct diagnostic that also was used by Sparling et al. [1997], and Ploeger et al. [2011].

The discussion of the numerical stability was removed entirely as it is not required for our analysis.

References

Felix Ploeger, Mohamadou Diallo, Edward Charlesworth, Paul Konopka, Bernard Legras, Johannes C. Laube, Jens-Uwe Grooß, Gebhard Günther, Andreas Engel, and Martin Riese. The stratospheric Brewer–Dobson circulation inferred from age of air in the ERA5 reanalysis. *Atmos. Chem. Phys.*, 21 (11):8393–8412, June 2021. ISSN 1680-7316. doi: 10.5194/acp-21-8393-2021. URL https://acp.copernicus.org/articles/21/8393/2021/.

L. C. Sparling, J. A. Kettleborough, P. H. Haynes, M. E. McIntyre, J. E. Rosenfield, M. R. Schoeberl, and P. A. Newman. Diabatic cross-isentropic dispersion in the lower stratosphere. *J. Geophys. Res.*, 102(D22): 25817–25829, 1997. ISSN 2156-2202. doi: 10.1029/97JD01968. URL https://onlinelibrary.wiley.com/doi/abs/10.1029/97JD01968.

F. Ploeger, S. Fueglistaler, J.-U. Grooß, G. Günther, P. Konopka, Y. S. Liu, R. Müller, F. Ravegnani, C. Schiller, A. Ulanovski, and M. Riese. Insight from ozone and water vapour on transport in the tropical tropopause layer (ttl). *Atmos. Chem. Phys.*, 11(1):407–419, 2011. doi: 10.5194/acp-11-407-2011. URL https://acp.copernicus.org/articles/11/407/2011/.