

Geosci. Model Dev. Discuss., referee comment RC2
<https://doi.org/10.5194/gmd-2021-382-RC2>, 2022
© Author(s) 2022. This work is distributed under
the Creative Commons Attribution 4.0 License.

Comment on gmd-2021-382

Anonymous Referee #2

Referee comment on "Massive-Parallel Trajectory Calculations version 2.2 (MPTRAC-2.2): Lagrangian transport simulations on graphics processing units (GPUs)" by Lars Hoffmann et al., Geosci. Model Dev. Discuss., <https://doi.org/10.5194/gmd-2021-382-RC2>, 2022

The manuscript by Hoffmann et al. presents an impressive piece of work. I can only congratulate the authors on the development of MPTRAC and its parallelization on GPUs, which is the main topic of the study. The manuscript is well written and structured and the methods and results sections are easy to follow.

I thus only have a few minor comments, suggestions, and corrections that the authors should consider before publication:

The introduction is quite MPTRAC-centric. Since the focus is on code parallelization, it would be good to include references on parallelization approaches in other Lagrangian dispersion models, e.g. Brioude et al. (2013) for FLEXPART-WRF, Jones et al. (2007) for NAME, Pisso et al. (2019) for other versions of FLEXPART (with MPI or OpenMP parallelization and asynchronous I/O in case of FLEXPART-COSMO). There is actually also a GPU version of (parts of) FLEXPART developed many years ago (<https://db.cger.nies.go.jp/metex/flexcpp.html>), but unfortunately it was never published in peer-reviewed literature to my knowledge.

The introduction should also explain more clearly, what the main areas of application of MPTRAC are. It seems to be designed primarily to study large scale atmospheric transport in the free troposphere and stratosphere but not for transport and mixing in the atmospheric boundary layer (ABL). This is important to mention, because Lagrangian models are increasingly being used for inverse emission estimation, for which e.g. a proper representation of turbulent mixing in the ABL is critical.

The manuscript convinced me that MPTRAC is a technically carefully designed, flexible and computationally efficient model. However, I was less convinced that it is also doing a good job in terms of accurately representing atmospheric transport. A key criterion for Lagrangian particle dispersion models, for example, is the well-mixed condition of

Thomson (1987): A tracer well-mixed in the atmosphere should not un-mix due to the simulated transport. This is challenging to achieve but is critical for simulating mixing in the ABL or inversely estimating emissions, for example. Simple mixing schemes (e.g. without density correction term) as implemented in the model lead to un-mixing. It would be good to know the magnitude of un-mixing generated by the model in long simulations (un-mixing likely saturates at some point). This could be studied in a simulation similar to those presented in Section 3.3, but where particles with uniform mass are initialized proportional to air density. Particle densities should ideally remain proportional to air density throughout the simulation.

The synthetic tracer simulations presented in Section 3.3. are suitable to study differences between the CPU and GPU versions, but they are not sufficient to demonstrate that transport is generally well represented in the model. A much more challenging diagnostic for stratospheric transport, for example, would be age of air, which is known to be underestimated by many transport models.

I thus strongly encourage the authors to focus on such critical aspects in future studies to provide a thorough scientific benchmark for future applications of the model. This is more a comment than a suggestion for modifying the current publication.

Small points:

Page 7, line 184: What exactly do you mean by "pushed back"? The standard approach in Lagrangian models is that particles are reflected. "Pushing back" likely leads to accumulation of air parcels at the surface or upper boundary of the model.

Page 9, line 250: Shouldn't it be $|\phi| > \phi_{\max}$? Same issue on the next line on page 10.

Page 15: Convection is parameterized in an overly simplified way, since e.g. deep convection does not at all lead to uniform vertical mixing. It would be good to mention (and to consider) more advanced approaches such as Forster et al. (2007, <https://doi.org/10.1175/JAM2470.1>).

Page 18: Also dry deposition is described in a highly simplified way. Dry deposition does not only depend on particle or gas properties but also on the state of the atmosphere (in addition to surface properties). Also here it should be mentioned that more advanced approaches for Lagrangian models exist, e.g. Webster and Thomson (2012, <https://doi.org/10.1504/IJEP.2011.047322>).

Page 30: Which number of compute cores of the GPU is the most relevant number for MPTRAC? Is it the number of FP32 or FP64 cores? Later it becomes clear that it is the

latter. Is double precision really needed? Did you test MPTRAC with single precision?

Figure 7: The differences between GPU and CPU simulations presented in panels b), d) and f) are likely due to statistical noise. This could be shown by performing multiple CPU simulations with different random seeds and evaluate the differences in the same way as the differences between CPU and GPU.

Section 3.7: I didn't quite understand this scaling test. Why does the runtime shown in Fig. 11 not decrease with the number of MPI tasks? What is the difference between a weak and a strong scaling test?

Small corrections and typos:

Page 11, Line 272: Change to "The following choices are made .."

Page 23, Line 500: shouldn't it be "interpreting" rather than "interpolating"?

Page 30, line 678: "MPTRAC was build" -> "MPTRAC was built"

Page 40, line 830: "33% if the overall runtime" -> "33% of the overall runtime"

Page 40, line 857: It should be Figs. 10a and b rather than 9a and b.