Geosci. Model Dev. Discuss.,
doi:10.5194/gmd-2016-314-RC1, 2017

Geoscientific Model Development Discussions
Open Access
EGU

Interactive comment on "Regional and seasonal truncation errors of trajectory calculations using ECMWF high-resolution operational analyses and forecasts" *by* Thomas Rößler et al.

P. Seibert (Referee)

petra.seibert@univie.ac.at

Received and published: 20 March 2017

1 General

This paper investigates the performance of a number of numerical schemes to integrate the trajectory equation. This is done using the LPDM MPTRAC. As the code is prepared for parallel computing, the performance is investigated also as a function of the number of threads (for one of the schemes only). For the tests, 10-day simulations were carried out using ECMWF data with 16 km grid spacing, and results are presented for different regions of the globe, layers of the atmosphere, and seasons.

Printer-friendly version

Discussion paper

This study is a useful addition to the previous investigations, as it also tests higher-order methods rarely used in atmospheric transport modelling, and as parallel performance is included. I recommend publishing it after consideration of the following remarks. I think that the authors have some choices with respect to doing additional calculations and/or evaluations, and I hope that they would be able to consider my respective suggestions, as the value of this work could be siginificantly increased in that way.

2 Major remarks

1. In my opinion, there is one aspect in the setting of the numerical experiments which is not ideal. The regular LPDM code has been used, that is, including a stochastic wind component to represent turbulence. Existing similar studies have been carried out with simple trajectory models. It is not very clear what the consequence of adding stochastic wind components is for the deviations between the schemes tested. The authors propose turbulence as explanation for several of the observed variations in accuracy, but this remains hypothetic. I would strongly recommend to repeat at least a subset of the simulations with all kinds of stochastic influences (turbulence, mesoscale fluctuations, convection if it exists in the model) switched off, present and discuss these results as well.

2. Another open question is whether RK4 with 60 s time step is a suitable reference method. If one extrapolates the RK3 or RK4 curves in Fig. 8 (bottom), one would arrive at an AHTD value of about 100 km at 60 s (probably against a hypothetical perfect simulation). The time step has to be reduced until a further reduction does not reduce AHTD significantly in order to establish a reference simulation. (I see that Hoffmann et al. (2016) claim that convergence already was reached at 120 s, but this is in obvious contradiction with the results reported here.) This might change the apparent relative benefits of higher-order methods.

3. As the authors rightly point out, higher-order methods are unlikely to bring much gain if we use linear interpolation. This points to another option for a potentially optimal trajectory calculation, at least as a reference method: Linear interpolation should allow to solve the trajectory equation analytically within a grid cell and between two times of wind field availability (cf. Seibert, 1993). Admittedly, the need to bound each calculation step at grid-cell borders has a potential to make this method a bit cumbersome and computationally probably less efficient.

4. Another methodological issue is the questions on which transport times the final evaluation of schemes should be based. Even though not explicitly mentioned, Fig. 8 seems to be made with results after 10 days. I don't think this is the most appropriate choice. As discussed in Section 3.2, there is a strongly non-linear growth of the deviations with time. This growth has nothing to do with numerical errors, it is solely a function of atmospheric flow patterns (diffluent flows or bifurcations). Thus, a longer calculation mainly amplifies initial deviations which are due to the different truncation errors. The longer calculations only mean more calculational efforts, and the true truncation errors are obscured by the increasingly important atmospheric flow influences, probably exaggerating the difference between atmospheric regions or seasons (note also that for example polar-region trajectories mostly leave the polar domain within the 10 days). Please also look at results with much shorter transport times and consider replacing the 10-day results by them.

5. Finally, the results are certainly sensitive to the resolution of the wind field data. Results obtained for the specific case of 16 km / 3 h therefore cannot be generalised. Keeping in mind the conclusions of Stohl et al. (1995), Brioude et al. (2012), and Bowman et al. (2013), 3 h intervals for the wind fields are coarser than what would be desired at this horizontal resolution. As 1 h is provided by ECMWF, I am wondering why it was not used. This also diminishes the value of the results presented here, as most people would want to use the 1-h data if

they go to the highest horizontal resolution. There would be a number of ways to produce more general results, such as trying out different resolutions or to parameterise the recommended time step by flow field properties such as (local) spatial and/or temporal derivatives at different orders.

3 Specific and Minor Remarks

1. The title could be rephrased for example as “Truncation errors of trajectory calculations using ECMWF high-resolution data diagnosed with the MPTRAC Lagrangian particle dispersion model”

2. Page 1, line 1: *Abstract*. The abstract could be shortened by removing non-essential background and more concise wording.

3. Page 1, line 4: *kinematic equation of motion* (comes also in other places). I don’t feel comfortable with this wording. “Equations of motion” for me would refer to the Euler or Navier-Stokes equations. Why not call this the *trajectoy equation*?

4. Page 2, line 6: *Lagrangian particle dispersion models have proven*. Under this chapeau, next to real LPDMs, LAGRANTO is listed which is a simple trajectory model and not an LPDM. I think it does no harm to enumerate it here, but not under a category that doesn’t fit (and there is no reason to focus specifically on LPDMs here, as the truncation error problem occurs in the same way in trajectory models).

5. Page 3, line 3: *The T1279L137 ECMWF operational analysis data used here have 16 km effective horizontal resolution, about 180 – 750m vertical resolution at 2 – 32 km altitude, and are provided at 3 h synoptic time intervals*. “Provided at 3 h . . . ” is not entirely correct – it is your choice. Analyses are available every 6 h

and forecasts at steps of 1 h. It would be useful if you indicate what composite of AN and FC fields you were using here and not on the next page.

6. Page 3, line 7: *LPDM studies using this new data set*. It is not clear what you mean by "this new data set". Obiously, ECMWF rules will not allow to make the ECMWF data set that you have used here available for general use.

7. Page 3, line 26: *meteorological wind fields*. Just 'wind fields' should good enough. If the model uses other fields as well (e.g, thermodynamic or surface fields), please explain in more detail. I am also wondering wether the model considers convection – it is invoked as a possible explanation later, but in Hoffmann et al. (2016) I did not find a reference to convection being a simulated process (if it isn't, it should also not be invoked). Maybe you want in general to provide a little bit more information about the model, especially considering that the only paper published so far is not open-access.

8. Page 3, line 31: *While atmospheric reanalyses ... typically have a horizontal resolution of ∼ 100 km or less, the resolution of operational forecast products has been continuously improving during the last decades*. Reanaysis products' resolution has improved as well! And better write "≈ 100 km" (`\approx`) or "ca. 100 km" to not confuse with symbol for proportionality (symbols appear also on p. 7 and p. 10).

9. Page 4, line 4: *For usage with MPTRAC, the wind fields have been interpolated horizontally to a longitude-latitude grid*. Have they really been interpolated (from another, e.g. reduced Gaussian, grid), or were they just extracted at the given grid through MARS (by evaluation of the spectral data)?

10. Math vector notation: You are using upright bold letters for vectors. Standard notation would italic bold, accessible (with the amsmath package) for example through `\boldsymbol{text}`.

11. Page 6, line 4 ff.: $\boldsymbol{k}_1 = \ldots$.It seems that you define certain velocities as $\boldsymbol{k}$. It is very unusual to denote a velocity by $\boldsymbol{k}$ and not with a letter such as $\boldsymbol{u}$ or $\boldsymbol{V}$, upper- or lowercase, and even more difficult as you don't give an explanation in words of these variables.

12. Page 8, line 5, 8: *5 latitude bands, 3 altitude layers*. According to standard typesetting rules, numbers less or equal to twelve in running text should, in general, be written out (same for "2nd/3rd-order" elsewhere).

13. Page 11, line 30: *land surface ratio*. I guess that "land-surface fraction" is meant.

14. Page 11, line 30: *The tropospheric mid-latitudes were expected to cause the largest errors, because the most complex wind systems occur in this region due to a larger land surface ratio and more complex orography*. The distribution of continents and orography is relevant for the difference between the midlatitudes of the two hemispheres, but not for differences between midlatitudes and elsewhere – this latter effect is due to the structure of the global circulation which in the end is caused by the poleward increase of the Coriolis parameter, allowing for Rossby waves and baroclinic instability to occur there.

15. Page 12, line 1: *The south pole has the smallest errors*. Probably you want to say that the smallest errors were found over Antarctica / the southern polar region.

16. Page 12, line 5: *The relative high errors in the tropics are probably caused by a stronger turbulence in that region. The lower bound of the stratospheric region of our test cases is 16 km, since the tropopause reaches an average altitude of 16 km near the ITCZ, turbulent movements due to deep convection can occur more frequently in the lower stratosphere above the tropics*. Is turbulence due to convection resolved in MPTRAC? If not, it can't be invoked as an explanation here.

17. Page 12, line 9: *During northern hemisphere wintertime land-sea temperature differences as well as the temperature gradient between the North pole and the equator are largest, which allows for more intense and complex dynamic patterns to occur than in summer*. I would not refer to the meridional temperature gradient as the pole-equator temperature gradient – the pole is a single point and neither the pole nor the equator typically represent the locations of the extreme temperatures. Furthermore, the baroclinicity in mid-latitudes rather depends on the subtropical region temperatures than on equatorial ones.

18. Page 12, line 12: *We need to stress that each simulation lasts only 10 days, which is a relatively short time interval to analyze seasonal effects. Fast temporal variations and changes in medium-range weather patterns can blur out the impact of seasons that is observed here*. To better resolve the seasons you don't need longer trajectories, but more frequent starts or more years. I any case, I don't think that the seasonal effects are so interesting, you could discuss this just briefly. It is obvious that stronger variations in the wind fields will lead to larger truncation errors, and the dependence of the variability of wind fields on the seasons is well known.

19. Page 12, line 27: *Vertical transport deviations are about 800 – 1000 times smaller than the horizontal transport deviations*. As the atmosphere in general is anisotropic ($L \approx$ 10,000 km, $H \approx$ 10 km), this is trivial and not worth mentioning.

20. Page 12, line 35: *The median error gets somewhat larger in the troposphere, where particle paths are more likely being affected by atmospheric turbulence*. Hoffmann et al. (2016) says that MPTRAC uses the same diffusivity throughout troposphere and stratosphere. How is this compatible?

21. Page 13, line 15: *As an example, Fig. 7 shows results of scaling tests using the midpoint scheme with a time step of 120 s for different numbers of particles and*

OpenMP threads. It would be useful to explain why you are only testing OpenMP and a single node if MPTRAC is capable to work on distributed-memory systems as well.

22. Page 13, line 18: *the computing time is limited by an offset of . . . s, which is due to the overhead of the OpenMP parallelization*. Language-wise, I would prefer to speak about showing a plateau rather than "being limited by an offeset". Do these times refer only to the time spent in the trajectory calculation, or to the model as a whole? In the latter case, there is not only overhead from parallelisation but also from other parts of the model (the – minor – plateau even with a single a single thread seems to indicate some contribution.) One is also wondering here about your parallelisation strategy – is there a barrier after each time step? Is that needed?

23. Page 13, line 23: *It is also found that the code provides additional speedup if the simultaneous multithreading capabilities of the compute nodes are used, in particular for very large numbers of particles (on the order of* 10^6 *to* 10^7*). For smaller number of particles (*10^4 *or less) the speedup is limited due to the overhead of the OpenMP parallelization and by the limited work load of the problem itself*. This is an interesting part of your results, but I don't agree completely with your description and interpretation. There is always a drop at first when the number of threads exceeds the number of 24 cores, which is quite typical (see also the indications given in your footnote source). The interesting feature is that for a large enough number of particles, it then rises again. Maybe your computing specialists have more detailed insights for this behaviour? Also, I wondering why for the largest number of particles the first maximum is reached with 20 threads. Is this a plotting error, or is this related to memory access? We should also note some irregular behaviour for moderate numbers of particles toward the maximum number of threads.

24. Page 13, line 33: *Among the 2nd-order methods the Petterssen scheme has the lowest computational efficiency, which is due to the fact that we tuned the convergence criteria for this method for accuracy rather than speed*. So, it is not "the Petterssen scheme" but your implementation of this scheme for which the statement holds! That is a bit of a pity, so we don't know how the Petterssen scheme would do with a more reasonable cut-off of the iterations. As this is quite a relevant issue, and some people might only look at the figure without reading the full thext, I suggest to mention that also in the figure caption (or better do some more realistic tests for a revised version).

25. Page 14, line 1: *The best efficiency, i. e., the best accuracy at the lowest computational costs, is mostly obtained with the midpoint and RK3 methods*. This wording is not providing an operational definition of "best efficiency", as best accuracy and lowest computations cost are mutually exclusive and you are not defining how exactly you want to measure the efficiency. A suitable measure would be the computation time to achieve a given AHTD. Do this for a value that is reasonable and then quantify the computation times, as just reading them out from a log-log diagramme is not so easy (note also the unexplained minor tick intervals better – use a full set of them). Thus, you may want to combine this paragraph with the following one. For the rating of Petterson (vs. midpoint), see above. Another question which needs to be answered is with how many threads this result was obtained, and whether there is any difference between schemes with respect to speed-up.

26. Page 14, line 17: *with an effective horizontal resolution of about 16 km*. Mention also the 3 h here!

27. Page 14, line 18: *The truncation errors of the schemes were found to cluster into three groups that are related to the order of the method*. Add "for a given time step".

28. Page 14, line 25: *We attribute this to larger small-scale variations caused by atmospheric turbulence and mixing in the troposphere*. The first part of the explanation is correct, but the second part not. These variations are not caused by turbulence (16 km is not turbulence scale !!) and certainly not by mixing (this would reduce and not amplify variability!).

29. Page 14/15, line 33–6: *[whole para]*. I suggest to rephrase this paragraph in line with the remarks made above for Section 3.4, making sure it clearly conveys the relevant facts and definitions.

30. Page 15, line 7–9: *The study of Seibert (1993) . . . . To achieve truncation errors that are smaller than overall trajectory uncertainty, they found that the time step should fulfill the CFL criterion as a necessary condition for convergence*. The recommendation there for a sufficiently small truncation error was 15% of the time step needed for convergence of the Petterssen scheme. If we assume that the reference accuracy has also improved in the meantime, an even smaller value would result. The CFL criterion is recommended to make sure that no small-scale features are skipped, not for convergence of the iterations in the Petterssen scheme.

31. Page 15, line 19: *However, the large variability of regional and seasonal truncation errors found here suggests that applications may benefit from more advanced numerical techniques. Adaptive quadrature could be an interesting topic for future research*. Note that adaptive time steps have been recommended by Seibert (1993) and were used already in the 1980ies for atmospheric trajectories by Maryon and Heasman (1988) and Walmsley and Mailhot (1983).

32. References: For Hoppe et al. (2014), quote the final paper and not the discussion version.

33. Figure 1: I would suggest to use the same scale for all pressure levels. I am

wondering why odd pressure levels are used (32.6, 180, 488 hPa) instead of standard levels. And I would suggest to reverse the colour coding for vertical velocity – meteorologists would find it more natural letting blue denote subsidence and red upward motion.

34. Figure 2: I don't deem this figure necessary. If you want to keep it, use an appropriate viewing position in the projection for the Northern hemisphere, presently we are looking from a point located somewhere above the South pole, like peeking through the ground, not down from space! Also, use hollow symbols of different shapes so that we can easily recognise coincident positions as such.

35. Figures 3 ff.: It would help the reader if you annotate subfigures or at least columns of subfigures.

36. Figures 5 and 6: This figure should be simplified. You don't need to show the two years separately, and I think you also don't need to show seasons separately. Thus you could have just three subfigures (three levels) and the five regions inside of each one. Then use a log scale for the AHTD, and symbols instead of bars (which will bring out the median also more clearly).

37. Figures 2, 3, 4, 7, 8: Please make sure that line width, colour intensity and marker size are sufficient to read all the content easily.

38. Using an enlarged printout of the lower part of Fig. 7, I tried to figure out the number of cores which works fastest as a function of the number of particles. I arrived at something like this:

#particles	#threads	remark
$<$50	1	
50 – 200	4	
200 – 300	8	very small interval!
300 – 1000	16	
1,000 – 50,000	24	= #cores!
$>$ 50,000	48	= max. #threads

I think that this evaluation would be useful for users. What is really striking is the fact that only (integer) powers of two show up as recommendable number of threads until 16. Then we can add 8 to arrive at the maximum number of cores (the question is open whether on a 32-core machine, 24 would show up or not), and then we can double once with hyperthreading. This is really a lesson for users, and if you have IT colleagues who are able to relate this behaviour to the hardware layout of your nodes, it would be even more useful.

39. Page 15, line 20: *Code and data availability*.

    - ECMWF data (of the kind used here) are not simply “distributed” by the centre. In general they would be available only for member-state NMS (or institutions authorised by them) and special-project holders. I suggest that the limited availability of these data is indicated. (I also thought that data provision could be mentioned in the acknowledgements.)
    - It would be useful to indicate the availability of the preprocessor which transforms ECMWF data to MPTRAC input data.
    - Does the version of the MPTRAC code available on github include the variety of integration schemes used here? If not, please make a statement about their availability.
    - It would be useful to provide the starting points of the trajectories as supplementary material so that the calculations become more reproducible.

References

Bowman, K. P., J. C. Lin, A. Stohl, R. Draxler, P. Konopka, A. Andrews, and D. Brunner (2013), Input data requirements for Lagrangian trajectory models. *Bull. Amer. Meteor. Soc.* **94**, 1051–1058, URL: http://doi.org/10.1175/BAMS-D-12-00076.1.

Brioude, J., W. M. Angevine, S. A. McKeen, and E.-Y. Hsie (2012), Numerical uncertainty at mesoscale in a Lagrangian model in complex terrain. *Geosci. Model Dev.* **5** (5), 1127–1136, URL: http://www.geosci-model-dev.net/5/1127/2012/, doi:10.5194/gmd-5-1127-2012.

Hoffmann, L., T. Rößler, S. Griessbach, Y. Heng, and O. Stein (2016), Lagrangian transport simulations of volcanic sulfur dioxide emissions: Impact of meteorological data products. *J. Geophys. Res.* **121** (9), 4651–4673, URL: http://dx.doi.org/10.1002/2015JD023749.

Maryon, R. and C. Heasman (1988), The accuracy of plume trajectories forecast using the U.K. Meteorological Office operational forecasting models and their sensitivity to calculation schemes. *Atmos. Environ.* **22** (2), 259—-272, URL: http://dx.doi.org/10.1016/0004-6981(88)90032-7.

Seibert, P. (1993), Convergence and accuracy of numerical schemes for trajectory calculation. *J. Appl. Meteorol.* **32**, 558–566.

Stohl, A., G. Wotawa, P. Seibert, and H. Kromp-Kolb (1995), Interpolation errors in wind fields as a function of spatial and temporal resolution and their impact on different types of kinematic trajectories. *J. Appl. Meteorol.* **34**, 2149–2165.

Walmsley, J. L. and J. Mailhot (1983), On the numerical accuracy of trajectory models for long-range transport of atmospheric pollutants. *Atmos.-Ocean* **21** (1), 14—-39, URL: http://dx.doi.org/10.1080/07055900.1983.9649153.

Interactive comment on Geosci. Model Dev. Discuss., doi:10.5194/gmd-2016-314, 2017.