

Interactive comment on “Quasi-Newton Methods for Atmospheric Chemistry Simulations: Implementation in UKCA UM Vn10.8” by Emre Esenturk et al.

Anonymous Referee #1

Received and published: 10 March 2018

This is a well crafted paper describing the numerical methods and improvements thereon with the UK global chemistry model. I recommend its reading for anyone interested in just how such models work, or in the methods used to solve small-size, non-sparse, non-linear systems. The authors apply quasi-Newtonian methods to the UK global chemistry model and carefully diagnose the consequent improvement in computational costs. It is a valuable contribution to the research community and has some important ‘lessons learned’. As is the case with all such deep-model development that I have tried, the results are often not as optimistic as one hoped for when embarking, but still worth the journey. A speed-up of 13% in chemistry and 3% in total model run time is still valuable and, of course, if the other model components are worked on similarly,

C1

the overall speedup in the chemistry becomes more important. There is one mystery in the paper that I would like explained, and some minor suggestions below.

Overall, the manuscript is written vary carefully and is easy to read. I can find hardly any typos or awkward sentences. Please fix the Table 2 title, FN should be NR I believe?

What I cannot understand is how the quasi-Newtonian method (QN), which provides only an approximation of the true Jacobian use in the standard Newton-Raphson (NR) method, is able to find an answer within the radius of convergence and thus avoid the NR method ‘wandering around the wasteland of bad solutions’. I would expect this ability from Markov Chain Monte Carlo methods, but not QN. Do the authors have any explanation for this?

p.6/l.10 The deep-seated drawback of implicit schemes is that they are inherently only first-order in time. (At least all the ones I have played with). People are developing higher order tracer transport schemes (as in Lauritzen’s papers), but with first-order chemical solvers, we are left with first-order errors. In fact, the authors analysis of errors in QN vs. RN later in the paper might do better to compare these with halving the time step. I think the relative errors shown here are all trivial.

p.6/l.22 (This is an aside) I am very surprised that UKCA still uses the old tricks of putting emissions and deposition into the kinetics. This is highly unstable since it forces these terms to be absorbed into a single grid cell for the full chemical time step. It becomes totally unstable with increasing vertical (or horizontal) resolution. The boundary layer is mixed on time scales of 30-90 min, and high dep velocities will empty a 50 m layer in the chemical time step. It is good that UKCA works with this coupling, but it may fail depending on intensity of local emissions or deposition.

p.10. This is a very nice derivation. In our model, we stop re-evaluating and solving the full Jacobian when our relative errors drop below some threshold (0.1 or 0.03). Within the radius of convergence, just iterating on the right-hand side vector (your c-sup-k)

C2

converges almost as swiftly and much more cheaply. I had not seen the factor $(1 - a_{\text{sup-k}})$ before and will be interested to try it out. You discuss correctly about the lower operation count of QN vs NR, but it might be useful to note here the simple numerics that solving $J \cdot \Delta c = -f(c)$ requires $n^3/3$ operations (inverting J requires n^3) but that resolving this as in eqn 10 requires only n^2 ops.

p.12. Can you explain simply about the CPU time vs wall clock time. When we run, the CPU time is usually 6 times the wall clock time (effectively we are using 6 out of the 8 cpus on the board).

p.14/l.12. I would not have thought that 'stiffness' applied to any given iteration, but rather it was a property of the Jacobian (and the system as a whole).

p.16. Regarding the analysis of Fig 4, I would not exaggerate the 1st hour since the initial conditions (some mean values) are horribly off from the correct answer and this almost never happens in a continuously running model. The changes over twilight are a globally common problem and very important to compute efficiently.

p.18. Table 5 is a bit hard to figure out. I see that QN2-3 decreases chemistry, but its increase in convection is artificial since the total time decreased. Is there an easier way to do this? Also, it would be nice to see the timings for all the components.

p.19/l.1-14. I am surprised that you allow 40 iterations with NR. Since convergence is quadratic, relative errors for the last 4 iterations go as: e^{-1} , e^{-2} , e^{-4} , e^{-8} . Thus the problem is to get within the radius of convergence (say, e^{-1}). If our NR solver does not find a solution within 7 iterations, it is time to try a new starting guess or cut the time step as you note. It would be interesting to single out and diagnose what is going wrong when a few grid cells hold up all. We found that happened when our convection lofted very-high isoprene values to the upper trop. So the fundamental problem is finding a better starting guess.

p.21. Very nice.

C3

p.25 Table 7 is impressive, and I suspect dwarfed by the Δt errors.

Interactive comment on Geosci. Model Dev. Discuss., <https://doi.org/10.5194/gmd-2018-32>, 2018.

C4