Interactive comment on “GEM-AQ, an on-line global multiscale chemical weather system: model description and evaluation of gas phase chemistry processes” by J. W. Kaminski et al.

Anonymous Referee #3

Received and published: 22 November 2007

This is a well-written description of the development and evaluation of a new model integrating chemical and dynamical processes over a range of scales. The paper describes the various compartments of the model and evaluates its performance against atmospheric observational data from a range of different sources. Although the paper is essentially a technical report rather than a presentation of new scientific results, it serves an important purpose in documenting and characterizing the model and is therefore worthy of publication.

Although the authors have done an impressive job of developing a model with unique features and great potential, I do not find the evaluation they present here convincing.
For selected chemical species they present examples of comparisons against observations, but they do not attempt to justify their choice of location or month, to explain the associated limitations, or to put the comparisons in a global context to provide a more thorough assessment of model performance. Specific discrepancies between the model and observations are highlighted and are shown to be consistent across the different comparisons, but the origin of these is only speculated at rather than properly explored. In general, the comparisons are not drawn together in such a way as to convince the reader that despite a number of imperfections (which all models have) the overall performance of this model is good. The evaluation aspect of the paper is therefore much weaker than it should be.

I believe that the paper is appropriate for publication in ACP, but that significant revisions are required in the evaluation section before it is ready for publication. I would recommend that the authors reassess their approach to this section in light of the suggestions for improvements below.

Recommendations:

(1) While examples of comparisons at individual locations and seasons are welcome, they need to be drawn together with some assessment of performance over global scales. The simplest way to do this is to use a statistical approach over a larger number of measurement locations to show that mean or median mixing ratios and variability are in accord with observations. This has been done in part for SCIAMACHY NO2, but could usefully be done with surface, sonde and aircraft measurements too. An advantage of this approach is that it is necessarily climatological in nature, and thus biases due to geographical and meteorological sampling are minimized. In general, the paper would clearly benefit from a more numerical assessment of model performance. How much do the mean O3, CO or NO2 columns or burdens differ based on Figs 3, 5-7? The pattern-matching exercise presented here is useful but not sufficient.

(2) Model performance can partly be assessed with non-observables, for example
ozone budgets, trace gas burdens and lifetimes, etc. At present, one or two of these global diagnostics are covered very briefly at the end of the conclusions. I believe these diagnostics deserve their own section in the main body of the paper, and this would also provide an opportunity to compare results with previous studies, as summarized in, e.g., IPCC 2001, or Stevenson et al. [JGR, 2006].

(3) Clear justification is required for the choice of observations used in the comparisons. This will make the comparisons more meaningful and will avoid any accusation of "cherry picking".

(4) Clear justification is required for the choice of emissions used. In particular, the scaling of lightning emissions and omission of aircraft emissions appear somewhat arbitrary, and the methods used for distributing lightning and biomass burning emissions need a clearer introduction. While it is appropriate to use 1990 emissions for comparison with ozonesonde climatologies, they are not ideal for comparison against recent satellite measurements of short-lived tracers. If the focus is 2001-2005, why not use 2000 emissions?

Specific Comments

The abstract is insufficient: it outlines the approach taken, but does not report any results. This should be addressed by adding a couple of sentences that summarize model performance, ideally in a quantitative manner.

p.14898 l.18: It would be helpful to enumerate meso-gamma scale for readers outside the meteorological community by adding, e.g., "(2-20 km)".

p.14902 l.14: The origin of the biomass burning data needs to be described here - move the explanation from the conclusions (p.14911 l.1-5) to this introductory section.

p.14902 l.6: the phrase "archived in 2000" is irrelevant and potentially misleading, and should be removed.

p.14902 l.20: Are monthly-mean lightning emissions applied uniformly throughout the
month (based on the mean convective cloud distribution), or are they only applied when convective clouds are present (i.e., they are event-specific)? This is an important distinction, and further clarification is required here. In addition, it would be interesting to know why emissions were scaled down to 2 Tg/yr, and why aircraft emissions were omitted.

p.14903 l.2: The sentence starting "To better account for stratosphere/ troposphere exchange in polar regions..." is puzzling and should be rephrased clearly. Does it mean "To reduce excess stratospheric influx..."?

p.14903 l.22: How often are these climatological variables updated? Surface roughness, in particular, would benefit from frequent update (monthly at the very least). Are inter-annual variations considered? What is the source of this data?

p.14904 l.25: The over-prediction of tropical upper troposphere ozone may also be influenced by the stratospheric boundary condition imposed above 100 hPa. Has the sensitivity of the results to the location or magnitude of this upper boundary been assessed?

p.14906 l.15: CO is a good tracer of transport, but note that primary emissions contribute less than half the total source [see, e.g., Shindell et al., JGR, 2006], so chemical processes also contribute to the comparisons here.

p.14908 l.21: Is GEM-AQ sampled at the same time of day as the SCIAMACHY measurement?

p.14909 l.18: "may be too low": underestimation of emissions over China is to be expected if the EDGAR emissions for 1990 are used. One of the coauthors has previously pointed this out clearly [Richter et al., 2005], so it should come as no surprise here. Is the underestimate consistent with the trends estimated in these earlier studies?

p.14909 l.20: The significance of section 3.4 is greatly weakened by the sentence starting on line 28, and by the poorly-justified choice of August 2001 results to compare
with October 1992 data. It is not clear that we can learn anything from this. It would be better to choose a more recent measurement campaign, or one less heavily influenced by biomass burning (or anthropogenic) sources.

p.14910 l.16: HNO3 is high and remarkably uniform in the upper troposphere here. Might this be responsible for elevated NOx and hence overestimated ozone production in the upper tropical troposphere rather than lightning, which at only 2 Tg/yr is probably underestimated?

p.14911 l.8: The meteorological biases need to be assessed and at least partly understood before the chemical biases can be explained. Has any previous study examined these?

Table A2: The product yields denoted by betas aren’t explained.

The vertical coordinate in Fig 1 is pressure, but in Figs 2 and 10 is altitude. For consistency it would be helpful to use altitude in Fig 1.

Fig 10: Identify model and observations in caption, or add a legend.

Typos, etc.

p.14897 l.23: rephrase, or replace "for" with "of the importance of an"

p.14898 l.7: replace "will allow for introducing" with "allows"

p.14906 l.12: incidents -> episodes

p.14906 l.17: insert "is" between "but" and "also impacted"

p.14912 l.17: Zang -> Zhang