Geosci. Model Dev. Discuss., 5, C926–C931, 2012
www.geosci-model-dev-discuss.net/5/C926/2012/

Geoscientific
Model Development
Discussions

Interactive comment on "Air quality modelling using the Met Office Unified Model: model description and initial evaluation" *by* N. H. Savage et al.

Anonymous Referee #1

Received and published: 21 November 2012

This paper describes the air quality modeling system AQUM as operated routinely by the MetOffice, and evaluates it against a set of ground-based air quality monitoring stations in the UK, mainly assessing surface ozone, but also PM and NO2. To evaluate the model, a set of metrics are used that are not (yet) very common in the scientific literature, and therefore they are introduced and explained in considerable detail. 'Climatological' performance, as well as performance during events of high pollution are evaluated. The model performance is furthermore compared against the MACC ensemble and 2 day forecasts are compared to 1 day forecasts. This paper is well written, and valuable to serve as a description and reference of the modeling system, and therefore well suited for publication in GMD. Nevertheless, I have a couple of comments

which should be addressed by the authors.

General comments:

Considering the chemistry scheme the authors refer to O'Connor et al., 2012, which is yet unpublished. Therefore it is to some respect difficult to assess the chemistry mechanism. The authors include a list of reactions in the supplementary information. Specifically it would be good to include rate constants and references to this list. It is unclear whether dry deposition is calculated online using actual meteorology, or whether a climatology is used. The model seems to extend into the stratosphere (39 km altitude), which is rather high for an air quality model and possibly needs some explanation. I wonder what is the performance of stratospheric ozone, and how this may affect the surface concentrations.

Additional information of the general model performance (and possible biases) of the chemistry scheme as well as the aerosol would help to assess reasons of specific performance of the model for RAQ purposes, and to assess which discrepancies can be related to issues regarding emissions and which to (specific) model assumptions. It would put the current model performance into perspective of the global chemistry modelling activities that have been performed with the same chemistry / aerosol scheme.

With respect to the evaluation, I find the assessment of the model performance rather rudimentary. Partially this is a consequence of the nature of this paper. Nevertheless, I think the paper would become more valuable if some more specific information, on, e.g., seasonal cycle, and/or more spatially discriminated performance would be presented. This can relatively simply be achieved. Also an attempt to explain specific model performance in Sect. 4 and all subsections would be helpful. The evaluation would further benefit from additional evaluation of other quantities and/or trace gases (e.g. AOD, ozone sondes, CO, aircraft observations, NO2 columns, . . .).

Finally, to my opinion the authors should be more critical in selecting appropriate figures to support their messages. I miss spatial maps of, e.g., O3, while there are a couple

GMDD

5, C926–C931, 2012

of soccer plots that do not yield helpful information. Time series are very helpful and clarifying, but not if one cannot see the data (as in Fig. 15).

Specific comments:

P3134, l 8: "(NAE)": Please note the boundaries of the domain (E/W/N/S), and refer to fig. 9 for graphical interpretation P3135 l 4: "40 transported species": Please provide a table with all species as used in the model, and additional tables which species are subject to wet/dry deposition respectively. P3137 l 16:"Indirect effects are as described by Jones et al. (2001)." Do I understand correctly that aerosol composition is influencing meteorology? Is this assessed anywhere, and how does this relate to possible biases in the aerosol module? P. 3138, l 5: GEMS/MACC data are used as LBC in AQUM. Biases in the LBC have been reported elsewhere (Schere et al., 2012), and could potentially influence the system. What is the impact of the use of these (aerosol/gas-phase) boundary conditions for the AQUM model performance?

Section 2.6: Emissions. It would be good to state explicitly that no diurnal or seasonally varying emissions are applied as yet, if I understand this correctly. Would this possibly contribute to differences in ozone biases for urban and rural stations?

P3141, L4: The use of GFED aerosol emissions for the year 2000 seems rather arbitrary: It might be better to refrain from using such emissions, or otherwise use actual emission estimates. Also CO and NOx fire emissions are missing.

P3141 L5: Biogenic isoprene emissions: What about soil NOx emissions? Furthermore, would the use of climatological emissions be a reason for model biases during specific events.

Section 3. I miss a discussion of the temporal and/or spatial aggregation of model data. Why not include an assessment of the seasonal cycle? For ozone it could additionally be interesting to assess the diurnal cycle, and analyse daytime and nighttime biases separately.

GMDD

5, C926–C931, 2012

Section 4.1.2: Ozone: it is good to see the general performance statistics as presented in Table 2 and the soccer plot, which are essential for operational air quality model purposes. Nevertheless, I miss a seasonal evaluation, a spatial map of the model performance and/or an assessment of daytime and nighttime O3 biases separately. Such additional information help to appreciate better why the model is performing as it is. Table 2 shows a relatively low hit rate for ozone on annual basis (0.57), while the model is biased high by about 10%. Contrary to the conclusions of the authors, this suggests that the dynamical range in model is not very large. In other words, the model ozone variability is relatively flat. Could you comment on this?

P. 3147, l17. "molybdenum converters may overestimate the true concentration". Would it be possible to estimate a correction factor for this, based on (modelled) NOy concentrations, see, e.g., Lamsal et al., (2008)? Additionally it would be interesting to evaluate the model against satellite NO2 observations.

Section 4.1.4: The authors blame the biases in the model basically to uncertainties in the emissions. What are the reasons for this? There will be biases in the model too. Can this be quantified to some respect? How did the aerosol model perform in previous studies? Additionally it would be interesting to look at aerosol optical depth.

P. 3150, l. 4: "Figure 6 summarizes the good model performance. . ." I believe for this assessment it is more critical whether the model has desired hit and false alarm rates, and captures the event to a decent degree, as presented in Fig. 5. Statistics in Figure 6 is dominated by daily "background" performance rather than the event. Therefore I don't think this figure adds necessary information to this situation, and could be removed.

P.3150, l. 6: "bias is particularly low": but the rms remains rather high, so this result is only low by change.

P3150, l 26: "The model captures this first episode well": Is this not just by chance? In Fig. 7, other (smaller) events are not well captured, except for 26-27 July. What are

GMDD

5, C926–C931, 2012

differences?

P 3153, l. 1: “over-predicts values”: Are there any reasons for this? How to align this with the results that the model in general is biased low?

P 3154, l. 2: Interesting results that the hit-rate in AQUM is significantly better than MACC, while the false alarm rate is not significantly worse. Is it possible to find a reason for this? Does this for instance suggest that the ensemble approach is mostly suitable for background conditions?

P. 3155, l. 14: “Likely due to emissions”: I’m not convinced here. Are you sure that there are no model biases? In the outlook you mention the implementation of a new aerosol model. What do you expect from this upgrade?

Figure 11 may be removed, as this does not yield any relevant information that helps the reader.

Figure 13/14: Very interesting to see this type of evaluation. Why not merge the two figures into one, so that an assessment of the model performance is more easy?

Figure 15: Please modify this figure, to improve readability. E.g. show daily max. O3 only.

References:

K. Schere, J. Flemming, R. Vautard, C. Chemel, A. Colette, C. Hogrefe, B. Bessagnet, F. Meleux, R. Mathur, S. Roselle, R.-M. Hu, R.S. Sokhi, S.T. Rao, and S. Galmarini, Trace gas/aerosol boundary concentrations and their impacts on continental-scale AQMEII modeling domains, Atmos. Env., 53, 38-50, 10.1016/j.atmosenv.2011.09.043.

Lamsal, L., Martin, R. V., van Donkelaar, A., Celarier, E. A., Bucsela, E. J., Boersma, K. F., Dirksen, R., Luo, C., and Wang, Y.: Indirect validation of tropospheric nitrogen dioxide retrieved from the OMI satellite instrument: insight into the seasonal variation of nitrogen oxides at northern midlatitudes, J. Geophys. Res., 115, D05302,

GMDD

5, C926–C931, 2012

doi:org/10.1029/2009JD013351, 2010.

Interactive comment on Geosci. Model Dev. Discuss., 5, 3131, 2012.