**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 #2

Received and published: 19 December 2012

Review of the paper

"Air quality modelling using the Met Office Unified Model: model description and initial evaluation"

by

N. H. Savage, P. Agnew, L. S. Davis, C. Ordonez, R. Thorpe, C. E. Johnson, F. M. O'Connor, and M. Dalvi.

The paper describes the development of an online air quality modelling system, the AQUM, which is based on the Met Office Unified Model combined with the UKCA

chemistry and aerosols model. The model performance is tested as well in forecast mode as in hindcasts of specific air pollution episodes. The model performance is satisfying demonstrating that the model can be used for the specified purpose of air quality forecasts. This is a solid paper, well structured and with results that illustrate the model performance quite well. No specific scientific question is treated, however this might not be particularly necessary for a paper to be published in Geoscientific Model Development.

Nevertheless there are some important points that need to be considered by the authors in a revised version. I list the most important ones here, they will be mentioned again in the additional specific comments below.

1.It can be questioned if another air quality model is needed because there are already numerous AQ models operated in Europe. The authors should point out more clearly why they developed this model and what the central application will be. They need to make clear why it is not suitable for their application to use one of the existing model systems (e.g. from the MACC community). 2.Another weak point is the representation of emissions that stem from a number of different inventories with different resolution and vertical distribution. Not much is said about the temporal variability. The authors clearly need to improve this part of the model system. 3.The model domain is not very big and AQUM is not applied on a larger domain to provide boundary conditions for the central domain. Therefore the boundary conditions will play some role for the results, in particular for aerosols and for ozone. This influence needs to be investigated (e.g. by sensitivity runs) and discussed in the paper

The paper can be published if the authors can make strong arguments for their model development and when the specific comments below are considered.

Additional specific comments:

Page 3132, l 12: "The results demonstrate that AQUM has a large dynamic range of modelled ozone levels and has a good level of responsiveness to elevated ozone

episode conditions." It is not clear what this exactly means, maybe it can be expressed a bit more precisely. The abstract mentions only ozone but I think today's air quality models should also have PM in focus. Nothing is mentioned about that in the abstract.

Page 3134, l 6: "AQUM has a horizontal resolution of around 12 km and 38 vertical levels up to a model top height of 39 km." This is mentioned again later in section 2.5. I think it is not necessary to double it.

Page 3136, l 7: "In addition there is a diagnostic aerosol scheme for sea salt and a fixed climatology of secondary organic aerosols (SOA)." Could you explain this in a bit more detail? What is the temporal and spatial resolution? What is included in this climatology? I see a problem with double counting (because FFOC and biomass burning might include secondary aerosols) and with inconsistencies (because SOA from the climatology might appear at different times and places as e.g. FFOC primary emissions).

Page 3136, l 20: " . . . it reacts initially with any sulphuric acid (H2SO4)." It is said before that SO42- exists only as aerosol, so from where do you get sulphuric acid?

Page 3137, l 1: "The mineral dust scheme has six size bins covering radii from 0.0316 μm to 3.16 μm." The entire aerosol scheme is mixed out of different approaches. Mineral dust is treated in different size bins, ammonium nitrate is in "accumulation mode" and some other species are given in climatologies. All this looks very inconsistent and difficult to treat from an emissions' perspective. Additionally you use an aerosol scheme that was built for a climate model for air quality purposes. Anthropogenic and biogenic SOA might not be well represented in your climatology. Please comment on this

Page 3137, l 11: "The direct radiative effects of all aerosols are included in the model by use of wavelength dependent scattering and absorption coefficients calculated off-line according to Mie theory." How do you know number and size of your particles which is necessary to apply the Mie theory?

Page 3137, l 28: "For pragmatic reasons the reanalyses and forecasts produced by the GEMS and MACC projects have been used to provide boundary conditions for the composition fields." How much do your lateral boundary conditions influence your results? Your domain is not very big and the boundary conditions play a major role for some of the species, e.g. ozone. From this point of view it is not surprising that your model results agree well with the GEMS/MACC results.

Page 3137, l8: This paragraph should say something about the initial conditions for the chemistry fields, but I cannot find this information. For a two days forecast it is quite important from what you start.

Page 3138, l 24: "Outside of the UK, emissions are taken from the European Monitoring and Evaluation Programme (EMEP) emissions datasets, which cover Europe at 50 km resolution." Are these emission fields further downscaled with surrogates like population density or taken as they are? It would be important to discuss the impact of a very coarse resolution of the emissions in one part of the domain and the very fine resolution in another part.

Page 3130, l 22-28: In fact this vertical distribution is over-simplified. There are other simple but better ways available that you could have used without much effort (e.g. Bieser et al, Environmental Pollution 159, 2935).

Page 3140, l 5: " . . . rather than being generated by fixed emission fields." I hope your emissions are not "fixed" but variable in time with seasonal, weekly and daily variations. Could you say something about this?

Page 3140, l 10: " . . . emissions of SOx required by CLASSIC are split into high- and low-level components . . . ." In which altitudes are they put?

Page 3140, l 26: "The vertical distribution of aerosol sources are split into high (320 m) and surface sources . . . ." This looks unrealistic to me. Why is there nothing in between? There should be a number of chimneys with heights that are in between

these values.

Section 2.6: The emissions part is a weak point in your model system. It looks very much inconsistent and with rather crude assumptions. You should say more about the future developments than just stating that an improved representation is necessary. This is an obvious shortcoming of your study and you need to present a concept how to solve that.

Section 2.7: What about initial conditions for chemistry?

Page 3142, l 17: I see the advantages of the error metrics you introduce but I see the necessity to compare results from different model evaluations against each other. Therefore I ask you to think carefully if you really need another measure to describe your model results. Why is the normalized mean bias not enough?

Page 3143, l 12: same comment as above: Why don't you use the MAE which is much more common that the FGE?

Page 3143, l 23: Why did you choose the mentioned bias/error combinations? Is there something special with these values?

Page 3145, l 20: One problem with web pages is that they might be disappeared or renamed in a few years. If you cite web pages you should at least say when you accessed them. However I ask you if there is other literature where this information can be found and to use these sources.

Page 3146. l 9: same comment as above.

Page 3147, l 24: "This pattern of over-estimation of NOx at rural sites and under-estimation at urban ones is consistent with the model resolution being too coarse to properly resolve sources of NOx." I agree that NOx will be underestimated by the model in source regions but why should it be overestimated in remote regions. Here, your model resolution should be appropriate to capture the measured concentrations.

C1089

Page 3148, l 19: ". . . the available evidence suggests that this component of PM has a smaller negative bias in AQUM." Unfortunately this evidence is not available for the reader, so I cannot judge whether you are right or wrong. I find it therefore difficult to follow your conclusions. You might get additional information on your hypotheses by looking at the available speciated PM (nitrate, sulphate, ammonium) at some EMEP stations.

Page 3149, l4: "Whilst most models can be tuned to give reasonable monthly or annual averages, a more discriminating test is whether models can respond in episode conditions and demonstrate a wide dynamic range, ..." You can easily test this by comparing the variances of observations and model results. However you do not give numbers for this statistical measure.

Page 3149, l 25: "The model generally agrees well with the observations at this site throughout the month, ..." To me this looks like an underestimation. You should give some statistical data for this station and period in the figure in a readable way.

Page 3150, l 20: see my previous comments about referencing web sites.

Page 3151, l 16: "This trend to over-forecast ozone levels is continued . . .": Can you speculate about the reason for this? I wonder to what extent your domain is influenced by advection from the west outside the model domain. In fact, the question is to what extent you test your model and to what extent the boundary conditions are tested.

Page 3152, l4: see my previous comments about referencing web sites.

Page 3512, l 12: "This is likely to be because the ozone concentrations were close to the threshold value for much of April, so that small errors in the model forecast concentration values could often result in incorrect classification as a hit or false alarm." Yes this is true. But it is also true for your (good) hit rates in July 2006 when you don't have many values around 100 micrograms/m**3. You don't mention this sometimes problematic property of the hit rate at that place.

Page 3152, l 18: "AQUM predicts the overall evolution of the episode well, ..." Why is this the case? You mention before that you typically underestimate PM10 by 50% (as many models do) and you give reasons for this. Now you meet the measurements rather well, so why is the general picture not true for this case?

Page 3152, l 21: "Provisional speciated PM observations . . .": Are these measurements reliable or not? If not, you should not use them. If yes, they are not provisional. Aren't there other speciated PM observations available for the UK?

Page 3152, l 26: "The modelled speciated PM2.5 concentrations are shown in Fig. 14 for comparison with the observations . . .": It would be much easier to compare model and observations if they would be in one plot or in plots that are separated by species and not by model/observation.

Page 3153, l 1: ". . .and significantly over-predicts values on 23 April and other days." Can I draw the conclusion that inorganic secondary aerosols are typically overestimated? This could explain why you sometimes have good agreement with PM10 observations although you seem to miss large parts of other aerosol components mass (e.g. wind blown dust, resuspended material, organics?). Could you comment on this?

Page 3153, l 15: "There were several periods of elevated ozone during June–October: in both early and late July as well as at the end of September and early October.": It is quite hard to see this in Fig. 15.

Page 3153, l25: "However the difference between the two model systems is not large.". This looks like not a big surprise provided that AQUM uses MACC boundary conditions. Throughout the entire section 4.3 the question that comes to the reader's mind is why AQUM is needed if MACC (and several models in MACC) is already there. You need to answer this more clearly.

Page 3155, l14: "This is likely to be the result of emissions which are not represented in the annual average inventories, such as re-suspension of deposited coarse PM, sea

C1091

salt or wind blown dust." I could also imagine other sources of error, e.g. organic aerosols which are partly only given in climatologies of unknown quality.

Page 3155, l19: " an improved representation of emissions, allowing different vertical and temporal profiles to be applied according to sector and interactive emissions of biogenic VOCs": How exactly do you want to solve this problem?

Table 2: Please add mean measured values and the number of stations considered.

Table 3: same as table 2.

Table 4: same as table 2.

Table 5: same as table 2.

Fig. 1; What means "oper"? Refer to urban and rural in the caption

Fig. 2: same as fig. 1.

Fig.3: same as fig. 1.

Fig. 5: What is "shvnd"? What is written in yellow in the lower right?

Fig 6: What is "shvnd"?

Fig. 7: What is "oper"? What is written in yellow in the lower right?

Fig. 8: What about model output on 2 and 5 July?

Fig. 9: It is very difficult to see the colors of the different stations and compare them to the model results. It looks like you have clear boundary effects in the North. Can you say something about it?

Fig. 10: What is "oper"? What is written in yellow in the lower right?

Fig. 11: same as fig. 1

Fig. 12: What is "oper"? What is written in yellow in the lower right?

C1092

Fig 13/14: combine observations and model in one figure or distinguish between species if one figure is too busy.

Fig. 15: It is difficult to see the episodes you mention in the text. What is written in yellow and green in the lower right?