Geosci. Model Dev. Discuss., 5, C1395–C1407, 2013
www.geosci-model-dev-discuss.net/5/C1395/2013/

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.

N. H. Savage et al.

nicholas.savage@metoffice.gov.uk

Received and published: 30 January 2013

We would like to thank referee 2 for their comments which have helped improved the paper. We now reply to each of the comment in line.

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).

As we explain in the paper, this is an online model (unlike those in the MACC Regional

Full Screen / Esc

Printer-friendly Version

Interactive Discussion

Discussion Paper

Air Quality Ensemble). Since it is based on the same meteorological model as used for our NWP, we will be able to use this system to evaluate the potential benefits of including online composition for weather forecasting as well as the reverse. In addition, as we indicate in the introduction, this new model will enable us to address air quality policy-related questions, which is not a primary objective of MACC. For this purpose we need our own model which we can modify and configure as required.

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.

We do not agree with the reviewer on this point and feel that the use of the three different inventories is a strength, rather than a weakness of our model. The UK National Atmospheric Emission Inventory is clearly the most appropriate one to use over the UK as it is at high resolution (1km) and is updated on an annual basis. This means we have to use a different inventory elsewhere to cover the rest of our domain. As our main objective is air quality forecasting over the UK (and the data we use to validate the model are all from the UK, except Mace Head), we do not consider the lower resolution over the rest of the European domain to be a major issue. We also have improved our explanation of the temporal and vertical profiles we use. We have added text to Sect. 2.6 to make the benefits of our approach clear.

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

We have now added the results of some sensitivity studies on this issue and discuss the impact of the LBCs in a new Sect 4.3.

GMDD

5, C1395–C1407, 2013

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.

We wish to convey the fact that AQUM performs well in reproducing the variability encountered in ozone concentrations. We have modified the wording to help clarify this point and in the main body of the text we discuss model variability compared to observations in several places.

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.

We have revised the abstract to correct this and have extended our analysis of PM in the text to include PM2.5.

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.

We agree and have removed the duplication.

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).

We have modified the text to clarify that this is a climatology of biogenic SOA and hence there is no problem of double-counting. We have given more details on this climatology in Sect 2.3.

GMDD

5, C1395–C1407, 2013

Page 3136, l 20: " . . . it reacts initially with any sulphuric acid (H2SO4)." It is said before that SO_4^{2-} exists only as aerosol, so from where do you get sulphuric acid?

We have clarified the discussion of this to indicate that the formation of ammonium sulphate in the model depletes the ammonia and only if there is ammonia left can ammonium nitrate form.

Page 3137, l 1: "The mineral dust scheme has six size bins covering radii from 0.0316μm to 3.16μm." 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

We acknowledge the limitations of our CLASSIC aerosol scheme clearly in the text and have added a more detailed description in Sect. 5 of our plans to improve the model with the new UKCA-GLOMAP scheme.

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?

We have added more information about the CLASSIC aerosol scheme which now makes clear that the parameters of the log-normal modes are fixed, thus allowing us to calculate the scattering using 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 condi-

GMDD

5, C1395–C1407, 2013

tions 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

We have added a new section (Sect. 4.3) devoted to an analysis and discussion of model sensitivity to chemical LBCs.

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.

This was an omission. We have added a new section (Sect. 2.5) which gives details about the model initial conditions.

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.

We have clarified the text to explain that the emissions are taken 'as they are'. We have added more text later in that section to explain the benefits of our approach using different resolution datasets which have up-to-date data.

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).

We acknowledge this and in Sect. 5 we have added more information on our plans to develop a more sophisticated emissions scheme.

Page 3140, l 5: ". . . rather than being generated by fixed emission fields." I

GMDD

5, C1395–C1407, 2013

hope your emissions are not "fixed" but variable in time with seasonal, weekly and daily variations. Could you say something about this?

We have clarified in the text that our emissions account for seasonal, weekly and diurnal variability.

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

We have clarified this in the text.

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.

We acknowledge this and in Sect. 5 we have added more information on our plans to develop a more sophisticated emissions scheme.

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.

We acknowledge this and in Sect. 5 we have added more information on our plans to develop a more sophisticated emissions scheme.

Section 2.7: What about initial conditions for chemistry?

This was an omission. We have added a new section (Sect. 2.5) which gives details about the model initial conditions.

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

GMDD

5, C1395–C1407, 2013

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?

As we explain in the text, the MNMB and FGE metrics we have used have the benefits of performing symmetrically with respect to under and over prediction and are bounded. When comparing different species which may be under or over- predicted, these properties are useful. The normalised mean bias does not have these properties.

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

As above

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

We have clarified these choices in the text.

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.

Where possible we have replaced web pages with other sources of information or provided more detail of the reference. In addition, all web pages now include the date they were last accessed.

Page 3146. l 9: same comment as above. Page 3150, l 20: see my previous comments about referencing web sites. Page 3152, l4: see my previous comments about referencing web sites.

These three questions have been replied to 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

GMDD

5, C1395–C1407, 2013

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.

We have pointed out in the text that this follows from the overall conservation of emitted mass. Note that in the UK there are few remote locations which are sufficiently far removed from sources for this effect not to be a factor.

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.

We have added an analysis of PM2.5 model performance to better illustrate this in the manuscript (see Sect. 4.1.4 and revised Table 2).

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.

Thank you for this valuable suggestion. We have added this metric (actually the standard deviation of model and observations) to Tables 2 to 7 and have used this in our discussions in the text.

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

GMDD

5, C1395–C1407, 2013

readable way.

We have amended the text to better describe the fit of model to observations.

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.

We now discuss this issue in a new section (Sect. 4.3) dedicated to the 'Sensitivity to chemical LBCs'. In addition, we acknowledge in the revised version of the text that the positive bias of AQUM for ozone derives partly from the positive bias found in MACC but is also partly due to other processes within the AQUM domain (see e.g. comments in Sect. 4.1.2).

Page 3152, 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.

We mention now at the end of Sect. 4.2.1 (July 2006) that the hit rate is sensitive to the threshold value chosen and to the overall pollution levels during a given period, and we refer there to the results found by this study for April 2011.

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?

We have added a discussion of PM2.5 to the paper and have made it clear that the

GMDD

5, C1395–C1407, 2013

reason AQUM models this episode well is because it is dominated by PM2.5

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?

Yes, these measurements are reliable and we have removed the qualifier 'provisional'. Note that it is rather difficult to obtain speciated PM measurements in the UK and we have no access to any other similar datasets at the moment.

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.

A merged figure would have too many lines to read easily and a different plot for each species would give too many plots. We feel this is best left as two separate figures.

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?

We have added an analysis of PM2.5 to the paper and this helps to make clear that the coarse component of PM is under-represented in the model.

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.

We have modified this figure (now renamed to Fig. 17) so that we now plot the daily ozone maxima, making the plot easier to read. The episodes in earl July, late July and

GMDD

5, C1395–C1407, 2013

late September / early October are easy to identify now.

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.

As discussed now somewhere else in the revised version of the manuscript, the MACC boundary conditions have an impact on AQUM. However, we have extended this specific part of the text to highlight some of the differences between both systems. While the MACC ensemble performs better than AQUM for mean value metrics such as bias and RMSE, the opposite conclusion can be drawn for other statistics such as the hit rate or the overall variability (given by the standard deviation). The reasons for these differences are discussed in detail in Sect. 4.4.

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.

Please see our first response above (item 1).

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 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.

The analysis of PM2.5 (added in a revised version) makes it clear that the greatest errors in PM10 are in the coarse component PM2.5-10 due to the sources mentioned.

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?

We have given more detail in the revised manuscript (Sect. 5) about our plans for

GMDD

5, C1395–C1407, 2013

improving the representation of emissions.

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.

We have done this for all relevant tables

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. 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?

We have improved all figures and addressed all the points above.

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

The model output is missing on those days due to some failures in the early stages of our operational forecast suite.

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?

We have now modified the colour and width of the boxes to improve this figure. Regarding the question about boundary effects, as a limited area model AQUM will obviously be subject to boundary effects; these are very clearly seen in this particular case because polluted air masses are being advected to the North.

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

We have decided to keep these figures as they were (please see reply above).

GMDD

5, C1395–C1407, 2013

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?

The previous Fig. 15 (hourly ozone) has been replaced by the new Fig. 17 (daily ozone maxima), where those issues have been addressed. It should be clear now that the lower right text corresponds to some of the error metrics described in Section 3.1.

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

GMDD

5, C1395–C1407, 2013