Interactive comment on “Global model simulations of air pollution during the 2003 European heat wave” by C. Ordóñez et al.

C. Ordóñez et al.
carlos.ordonez@metoffice.gov.uk

Received and published: 2 January 2010

RC: Referee comments – AR: Author replies

RC1: This article presents a diagnostic study of the skill of global chemistry-transport models in the simulation of pollutant concentrations during the Summer of 2003 and specifically the heat wave period. The article also presents the results of sensitivity studies concerning emissions, deposition, resolution, coupling. Results presented are among the outcome of the GEMS FP7 project. The work represents a big effort to understanding the specific difficulties of global models to simulate extreme regional episodes. It pinpoints the weaknesses and points to improve. What we learn from the article is mostly the deficiencies of global models, although it is difficult from this study to have general conclusions because the three models used here have different
problems. It is not clear that we learn much more, especially about the processes, but the paper is still a valuable contribution. Thus I recommend the study to be published but I have a number of points below that must be addressed beforehand.

AR1: We appreciate the positive comments of the reviewer. We have tried to address most of the points indicated by him/her. However the model runs presented here were set up a long time ago and because of that it is not possible to do a new run for every single suggestion.

Specific comments:


AR2: Reference added.

RC3: P 16860 L22: It is not clear how CTMs are forced from wind fields. Are wind fields linearly interpolated from 6-hour analysis? How does this rough interpolation influence mass conservation between interpolation times? How are vertical wind fields processed (diagnosed or forced)? What are the differences between the “mass flux method” for TM5 and the methods for other models? References to previous work should be added, at least, but it would be better to have details here.

AR3: The main difference between the mass-flux method used in TM5 and the approaches used in MOCAGE and MOZART is that TM5 calculates horizontal wind speeds from 3-h ECMWF analysis fields while the other two models linearly interpolate horizontal wind speeds from 6-h ECMWF analyses. Vertical winds are calculated in all models by imposing mass conservation. More details are provided below and
summarised in section 2.1 of the revised version of the manuscript:

- Hourly horizontal winds on MOCAGE and MOZART grids are obtained by interpolation in space and time from ECMWF 6-hourly analyses. Vertical wind in MOCAGE is calculated from the horizontal components by imposing mass conservation for each atmospheric column with a zero vertical velocity boundary condition at the top of the atmosphere. Vertical velocities in MOZART are derived from the divergence of the horizontal velocity fields, using a flux form semi-Lagrangian transport scheme based on Lin and Rood (1996) which allows for tracer mass conservation.

- As mentioned in Section 2.1, “In the case of TM5, wind fields are derived from mass fluxes through the grid cell boundaries”. A reference to “Krol et al. (2005) and references therein” has been added to that sentence. Some explanations on the TM5 mass-flux method, copied from that paper, are included here but not in the main text to avoid extending it unnecessarily:

"With TM5 being an Eulerian grid box model, the input required for advection should consist of mass fluxes through the boundaries of each grid box cell. The procedure is described in detail in Segers et al. (2002); here a brief outline is given. The produced mass fluxes are valid for time intervals of 6 h. The vertical mass distributions (kg air) at the begin and end of an interval are computed from the surface pressures and the hybrid coefficients of the vertical layer structure. By this, the mass change per model grid cell (kg/s) during the time interval is defined. The mass fluxes (kg/s) through the boundaries should describe how air mass is flowing from grid box to grid box, explaining the mass changes that are dictated by the changes in surface pressure. First, the vertical fluxes through the bottom of the grid boxes are computed by integration of horizontal divergence, and taking into account the horizontal pressure gradients on the hybrid grid (Segers et al., 2002). Second, a first guess of the horizontal fluxes is computed from horizontal winds, which in turn are computed from the ECMWF divergence and vorticity. Finally, these first guess values are slightly modified in a way that the modified horizontal fluxes and the vertical fluxes together explain the observed mass
Several tests have revealed that this new pre-processing algorithm significantly improves the vertical transport in the tropopause region (Bregman et al., 2001, 2003)." For the TM5 runs used in this paper the mass fluxes are valid for time intervals of 3 h, unlike in Krol’s paper (6 h).

References:


Krol et al. (2005): Already cited in the ACPD version of the manuscript.


RC4: About time interpolation: The time interpolation may be critical in particular because it may introduce noise in mass conservation with effects hard to predict. Time interpolation in morning hours between 6UTC and 12UTC may also introduce very large errors in downmixing of residual layer. If mixing is too strong too early this inhibits precursor concentration build-up with fresh morning emissions to favour chemical reactions. 6-hour interpolation can be very bad for other meteo variables like for instance radiation (if used). It is probably too late to do that but it would have been interesting to
consider 6-hour analyses followed by 3-hour short-term forecasts, and interpolations in between. I strongly encourage the authors to provide an analysis of the effect of interpolation by comparing coupled and uncoupled model results if possible (that can be a case study).

AR4: 6-h meteorological feedback has been used for the MOZART and MOCAGE standalone simulations while 3-h meteorological fields have been provided to TM5. To test the influence of the time resolution of the meteorological feedback used in the CTMs we have compared output fields from two MOZART simulations with similar configuration: the MOZART standalone run (6-h meteorological feedback) and an additional MOZART forecast run coupled to IFS with 1-h meteorological feedback (originally called “MOZART eywm” but referred to as “MOZART 1h-met” in the manuscript). Unlike in the coupled MOZART run used throughout the manuscript (COUPL), the chemistry tracers are fully transported in MOZART, with no feedback from IFS, in the MOZART 1h-met simulation. This makes it much more appropriate than the COUPL run for testing the effect of the time interpolation of IFS meteorological fields.

Three meteorological fields (pressure, specific humidity and wind speed) as well as two chemical fields (CO and ozone) from these simulations have been interpolated to the location of the three European MOZAIC airports (Frankfurt, Paris and Vienna) for different pressure levels in the proximity of the PBL (950 hPa, 900 hPa and 850 hPa). Very similar results have been found for all airports and pressure levels. As an example, 3-hour output fields from both model simulations at 850 hPa above Frankfurt are compared with each other and to MOZAIC measurements in a new figure. Correlations (R), average differences (d) and ratios of standard deviations (s) for 3-h output fields from both models are shown in the plots. Reasonably high correlations and small differences as well as ratios s close to 1 are found in the case of the meteorological parameters not only for 850 hPa above Frankfurt but also for the other airports and pressure levels analysed. The differences between both simulations are larger in the case of CO and O3. The figure also shows mean biases (b) and correlations (R) be-
between fields from both runs and measurements. For this latest calculation, model output has been interpolated to those times with MOZAIC observations, and both daytime and nighttime data have been considered. Although the results from this comparison of model output on a coarse grid with in-situ measurements at variable temporal resolution should be interpreted with caution, no noticeable improvement in the biases of the meteorological fields can be observed with an increase in the frequency of the meteorological feedback. However similar reductions in the biases as those in the figure have been consistently found for CO and O3 in the case of all airports and pressure levels, with the only exception of CO at 950 hPa above Vienna. The comparison of average diurnal cycles of CO and O3 from both simulations during August confirms the somewhat higher CO and O3 mixing ratios in the run with 1-h meteorological feedback (not shown in the paper). The correlations with measurements at the three pressure levels are higher for the 1-h meteo run with the exception of temperature (no noticeable difference between both runs) and specific humidity above Vienna.

Overall, these results suggest that the 6-h meteorological feedback and subsequent hourly interpolations used in two of the CTMs do not have a very large impact on the simulation of O3 and CO. However hourly meteorological input generally improves the correlation of modelled meteorological parameters and chemical tracers with measurements and also reduces biases for O3 and CO. We understand the concerns of the reviewer on the time interpolation, but the different parameterisations of transport – vertical turbulent tracer flux in the PBL (diffusion), convection, and advection – and chemistry used are expected to contribute more than the time interpolations to the differences found among the CTMs.

A short summary of the above paragraphs and a new figure have been included in the new Section 5.5 of the revised manuscript.

RC5: P 16861 L4-5: Injecting fire emissions at ground level is a very bad choice (see e.g. Hodzic et al 2006 cited). It could induce wrong PM concentrations in situations like the heat wave. Again it is probably too late to change that but the choice of, at least,
spreading fire emissions throughout the PBL would be better.

AR5: We understand the reviewer’s concerns on this. We recognise the importance of the injection of fire emissions for the simulation of CO during this episode and because of that we have commented this in the manuscript. The same as other emissions, fire emissions are injected in the lowest model level for MOZART, in the two lowest levels for TM5, and in the eight lowest levels (layer of 600 m) in the case of MOCAGE. Work to spread fire emissions in more model levels is being done in the follow-on project of GEMS (MACC). This should improve the simulation of PM, which is not the subject of this paper, and CO during biomass burning episodes. Unfortunately, at this stage it is too late to perform that work with similar model configurations as the ones analysed here.

RC6: P 16861 – Bottom: It is not clear what the run “with assimilation” will help to understand, in addition to the IFS-coupled run. What do we expect to deduce? This run should not tell much about the model deficiencies which is the strongest focus of the paper. Why not simply omitting this simulation in the rest of the article?

AR6: In this manuscript we have evaluated standalone simulations of three CTMs, some process-oriented sensitivity runs to understand some of the deficiencies in these models as well as a coupled run with improved meteorology. Unlike the sensitivity simulations, the coupled run with assimilation might not help understand processes. However in this context it is useful to evaluate whether some of those deficiencies (e.g. problems in initial or top boundary conditions, emissions, chemistry and transport schemes) can also be overcome by data assimilation procedures. We have found that the assimilation does not help improve the modelling of O3, for which the main deficiencies seem to be mainly due to the chemistry and dry deposition parameterisations but also to other processes. In the case of the simulation of CO, the deviations with respect to the measurements seem to be due, to a large extent, to the known deficiencies in the temporal resolution and injection of biomass emissions, but also to the parameterisations of chemistry and other physical processes. We have shown that CO
biases above MOZAIC airports can be partly corrected thanks to the assimilation. We cannot draw clear conclusions on which unresolved processes can be better overcome by data assimilation procedures, mainly because of the difficulties in assimilating tropospheric O3 with UV measurements. We think that these are important points that are worthwhile to mention and therefore have kept the coupled run with assimilation in the revised version of the manuscript.

RC7: P 16865, on observations used to evaluate models: Only the EMEP sites are used. However these are quite sparse over large parts of France where the ozone concentration was largest. It would have been nice to include more surface observations, in particular from French monitoring networks in rural areas.

AR7: As already explained in the reply to the first referee’s questions, the focus of this paper is not on the modelling of surface ozone but on the evaluation of tropospheric profiles of O3 and CO by global models. We consider that near-surface ozone measurements from the two data sets (EMEP and GAW) collected in the GRG subproject of GEMS are enough for this kind of analysis. We admit that these data are sparse over France, but they present reasonably good coverage e.g. over the South of Germany and surrounding countries where ozone levels were as high as in France during August 2003. The high ozone levels over those areas can be clearly seen in the EIONET observations presented in Figure 2 of Guerova and Jones (2007) and to less extent in some figures of Vautard et al. (2005). Both papers have already been cited in this manuscript.

RC8: P 16869 L5-6: the responsibility of higher reactivity is a statement (it is detailed later also), but only qualitative arguments are given. Please add a reference of a specific study which demonstrates this point. It is not clear why higher reactivity would lead to large ozone concentrations all over the Mediterranean Sea.

AR8: We agree that so far we have provided qualitative arguments and have not proved that higher reactivity in MOCAGE compared to other models should lead to large ozone
concentrations over sea. We have not provided a new reference but have done further analyses.

We have compared net chemistry tendencies and dry deposition velocities of ozone in MOCAGE and the TM5 run with zoom to $1^\circ \times 1^\circ$ (TM5-HWHR) for a number of grid cells over the Mediterranean during a period around the first fortnight of August. For those locations, net chemistry tendencies of O3 at daytime in the PBL are only 0–0.4 ppb/h higher in MOCAGE than in TM5-HWHR, while dry deposition velocities are very small for both models (lower than 0.05 cm/s) and around 10% higher in TM5. The analysis of three new sensitivity simulations (MOCAGE-STD, MOCAGE-VOCUT and MOCAGE-CUTALL, described in detail in AR10) for the period 16 July - 15 Aug 2003 indicates that the introduction of different initial conditions with lower tropospheric background O3 leads to a more or less sustained reduction in O3 somewhat lower than 10 ppb for the whole period over a location in the Mediterranean, which was not found for locations over land (see AR10). This suggests that there might be a stronger accumulation of O3 over sea than over land because of the weak sinks (e.g. dry deposition) over sea. In addition, the extreme reduction in emissions in the MOCAGE-CUTALL simulation introduces an additional 10 ppb decrease in O3 for the same location. The interpretation of these results is particularly difficult because of the smaller ozone changes due to different sources (e.g. generally chemistry at daytime) and sinks (e.g. dry deposition) over sea than over land. However it looks like a combination of factors such as somewhat higher reactivity and lower dry deposition velocities as well as the treatment of emissions and initial conditions might be responsible for the high O3 modelled by MOCAGE above sea.

These comments have been included in the new Section 5.4 of the revised manuscript.

RC9: P 16872 L6-10: There is already a discussion later about resolution, so this point is discussed twice. Probably the discussion should be removed at this place.

AR9: We agree with the reviewer. We have kept the text showing the comparison of the
high resolution MOZART run to MOZAIC profile measurements, but we have removed
the short discussion. The sentence “This is most probably due to a better simulation of
the accumulation and recirculation of pollution within the lowest levels and subsequent
photochemical production with the improved horizontal resolution” has therefore been
removed from this section. Discussions on this issue are now included only in Section
5.1 (Sensitivity to horizontal resolution).

RC10: P 16873. Nonlinearities and lack of resolution are invoked for explaining the
overestimations of MOCAGE. The argument reads like “because the RACM mecha-
nism has more detailed NMVOCs it has overestimations at coarse resolution”. This
is hard to believe, since several models (not in this study) use mechanisms similar to
RACM and do not show overestimations. At least to confirm such a behaviour an ex-
periment should be done by cancelling emissions in reactive VOCs and replacing them
with emissions in less reactive ones, in order to mimic the chemistry setting of the other
models and compare. This can be done for a case study and does not necessitate a
long-term run.

AR10: We have found that the main differences between MOCAGE and other mod-
els in terms of ozone are related to higher reactivity (resulting in higher net chemical
production of O3) and to lower deposition velocities in MOCAGE at daytime. Net chem-
istry tendencies and dry deposition velocities of O3 were compared between MOCAGE
and the TM5 run with 1°×1° zoom over the European domain (TM5-HWHR). These
two parameters have been evaluated during the periods for which they were available:
chemistry tendencies in the proximity of the PBL (950 – 800 hPa) for 1–12 Aug 2003
and deposition velocities for 1–20 August 2003. At nighttime, chemistry tendencies at
850 and 950 hPa above the MOZAIC airports are close to 0 in TM5-HWHR and slightly
negative in MOCAGE (on average of smaller magnitude than -1 ppb/h). At daytime,
they are clearly more positive in MOCAGE, particularly at 950 hPa (0.7 ppb/h higher
for Paris, 1.7 ppb/h higher for Vienna and 2 ppb/h higher for Frankfurt in MOCAGE than
in TM5 from 9 UTC to 18 UTC). The amplitude of the diurnal and day-to-day variabili
of the dry deposition velocity of O3 is larger and most probably more realistic in TM5-HWHR than in MOCAGE, with 3-h dry deposition velocities of 0.1–1.1 cm/s for TM5 and 0.2–0.6 cm/s for MOCAGE most of the time for the location of the MOZAIC airports, although in Section 5.3 we also indicate that this mechanism of ozone loss might be overestimated in TM5 during the heat wave. The average values of the dry deposition velocities differ by less than 0.1 m/s between both models, but they are clearly higher in TM5 than in MOCAGE at daytime, with the exception of some days during the heat wave. These analyses indicate that the higher net chemical production of ozone and smaller removal by dry deposition at daytime in MOCAGE are at least partly responsible for the higher O3 levels in MOCAGE compared to other models. Other mechanism not sufficiently well represented in MOCAGE might also contribute.

In addition, following the reviewer’s suggestion, we have performed some sensitivity runs with MOCAGE that we have compared with the MOCAGE base run for the period 16 July - 15 Aug 2003:

- STD: Continuation of the MOCAGE base run on 10 July 2003 but with initial conditions characterised by lower background O3 in the troposphere.

- VOCUT: As STD but with emissions and initial conditions of highly reactive VOCs (HC8, HC5, HCHO, OLI, OLT, TOL, XYL, CSL) set to 0 (see Stockwell et al., 1997, for details on these VOCs).

- CUTALL: Like VOCUT but also dividing the emissions of all the other VOCs by a factor of 2.

Results from MOCAGE-STD indicate that on the very first days of the simulations the lower tropospheric background O3 (also imposed on VOCUT and CUTALL) reduces O3 by up to 20 ppb in the lower troposphere for the location of the MOZAIC airports, but the differences become much less noticeable with the time and are nearly zero after around one month. At 850 hPa above Frankfurt, the run VOCUT reduces the fractional gross error in O3 from 27.6% to 17.6% and the modified normalised mean bias
from 26.2% to 9.3% during the period 16 July - 15 Aug 2003. The CUTALL simulation reduces the fractional gross error and modified normalised mean bias to 16.1% and -0.3%, respectively. Very similar results are found for Paris and Vienna, although biases for Paris are around 4% more positive than for Frankfurt, and both error statistics are generally of smaller magnitude for Vienna. With these tests we cannot mimic the behaviour of other chemical models that group the VOCs differently and that do not include some highly reactive VOCs, but we have shown that large part of the differences in the simulation of O3 by different models can be attributed to the way VOCs are treated and to some extent to differences in emission datasets. Both sensitivity runs, in particular the extreme CUTALL simulation, reduce the biases much more than the fractional gross errors. This simply indicates a compensation of overestimation and underestimation of O3 during different periods in these runs, and confirms that there are other mechanisms than treatment of VOC chemistry and emissions (e.g. removal of ozone by dry deposition, as shown above) that induce further uncertainty in the modelling of near-surface ozone.

We have included the above comments in the new Section 5.4 of the revised manuscript. We still attribute the higher O3 in MOCAGE to the more explicit chemistry scheme (combined with a coarse horizontal resolution) in that model, but also indicate that other important mechanisms such as dry deposition play an important role. We avoid any comparison with other regional models that use a chemistry scheme similar to RACM because they have higher horizontal resolution than the MOCAGE simulations used here.

RC11: P 16873 last line: “biogenic CO emissions” should be “fire CO emissions”

AR11: We will keep the text as it is. There are indeed missing biogenic emissions in the TM5 run used in this paper and that is partly responsible for the underestimation of CO by this model.

RC12: P 16874 L18: the fact that resolution increase does reduce the negative ozone
bias is somewhat contradictory with the argument for the overestimation for MOCAGE (coarse resolution induce positive bias).

AR12: The fact that a resolution increase reduces the negative ozone bias for MOZART is not necessarily contradictory with the argument for the ozone overestimation in MOCAGE (coarse resolution inducing positive bias). Depending on where both models tend to be (for a specific location and period) on the isopleth diagram for ozone production, the dilution of precursors might affect ozone production in different directions. The way horizontal resolution changes ozone production will depend not only on the chemistry scheme but also on other parameterisations of the models (e.g. boundary layer mixing, convection, effects of aerosols on photolysis rates) and on the chemistry time step. As an example, a zero-dimensional photochemical model study by Liang and Jacobson (2000) showed that ozone production efficiency of a box with air masses of different origin may be either underestimated or overestimated by a model that assumes uniform mixing within the box. They found that, under certain conditions, integrated ozone production may be over-predicted by as much as 60% in a 3-D model at a variety of scales. Under other conditions, such as in finely resolved urban airshed models, ozone production can be under-predicted by 20% at mid-latitudes during summer. Their results suggest that large-scale global models may have difficulty in correctly predicting ozone concentrations near urban/free tropospheric boundaries, which is consistent with results from other studies (e.g. Wild and Prather, 2006, and references therein).

Having said this, in the revised version of the manuscript we attribute the high O3 in MOCAGE not only to the high reactivity in that model, which is related to its chemistry scheme and horizontal resolution, but also to low dry deposition velocities at daytime in that model (see Section 5.4).

References (already present in the ACPD manuscript):

Liang, J., and Jacobson, M. Z.: Effects of subgrid segregation on ozone production


RC13: P 16878 L5-10. It is not clear what is exactly done with the switch off of off-European emissions. When are emissions switched off exactly? The results should be dependent of the switch-off time.

AR13: As already mentioned in the replies to the first referee, the TM5 sensitivity run without European emissions (TM5-HWEN) covers the 3-month period from 1 June to 31 August 2003. European anthropogenic emissions are switched off in this run from the very first day. Since we do not show any results until 16 July then the model was spun-up for a considerable number of days (45). This run is therefore equivalent to the TM5 run with zoom over Europe (TM5-HWHR) but without European anthropogenic emissions. The length of the TM5 runs is indicated now in the revised version of the manuscript (see Section 2.1.1) to make everything clearer.

RC14: L 17, same page. Can we really say that “influence of non-European emissions is small”, as background concentrations are important and mainly due to these emissions?

AR14: The influence of non-European emissions can be very important in situations when the relative contribution of the background concentrations to the total CO and O3 is large. However non-European emissions had a relatively small impact on European pollution within the boundary layer during the period of analysis, particularly during the heat wave period. The influence of those emissions on O3 and CO above Europe is larger for the upper tropospheric levels, as indicated in Figure 10 of the ACPD version of the manuscript. This is due to the stagnation of air masses in the mid and lower tropospheric levels over the European continent during that period. Our results are consistent with those from Lagrangian simulations in backward mode performed by
Tressol et al. (2008). Such simulations indicate suppressed long-range transport in the mid- to lower troposphere over Europe (Figure 4 of that paper) and small contribution of anthropogenic CO of North American origin to CO levels over Frankfurt (Figure 10 of that paper) during the heat wave.

RC15: P 16879. The resistance formula (4) is wrong, it should be $V_d = 1/(R_a + R_b + R_c)$. This could influence the discussion in the following page.

AR15: We thank the reviewer for finding this error. We have corrected equation (4). We have also modified the discussion in the following page accordingly, but the main conclusions do not change.

RC16: P 16882 L20: The statements concerning MOCAGE should be toned down, given the above remarks.

AR16: Following our replies to the above remarks concerning MOCAGE, we have toned down our statement on the comprehensive chemistry scheme and coarse horizontal resolution as being responsible for the overestimation of ozone by this model. Now we also mention the importance of dry deposition.

RC17: P 16885 L16-17: Whether “the development of global CTMs and the coupled system is not aimed at achieving such good performances” depends on the ambition we have for those systems. To improve emission diurnal variations does not make the computation times longer and is not difficult. Improving the fire emissions and injection heights should improve the performance with no additional computational cost. These are only examples. I seriously disagree with the statement, and believe such models should, in a few years time, give simulations at least as accurate as present days’ regional models, while regional models will describe refined meso-scale effects.

AR17: We agree that improving diurnal and weekly variations of emissions as well as implementing injection heights should improve the performance of the global models used here. We also agree that it is possible to use these models to model pollution at a
regional scale and that in a few years their simulations should be more accurate. However the large grid cells of global models usually contain air masses of very different origin which are treated as if they were uniformly mixed within the box. This, together with the difficulties in representing accurately some processes such as boundary layer mixing and convection at a coarse resolution, will certainly affect the mixing of primary pollutants and subsequent net ozone production. In addition, the large grid cells contain a mixture of land use types, which will affect the parameterisation of one of the most important ozone sinks: dry deposition. Because of the above mentioned difficulties, we still consider it appropriate to say that AT PRESENT TIME such models are not aimed at achieving such good performances that enable them to be compared on an absolute basis with surface measurements during episodes.

Interactive comment on Atmos. Chem. Phys. Discuss., 9, 16853, 2009.