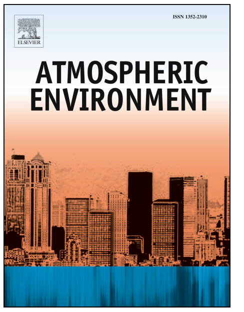

Analysis of UK and European NO_X and VOC emission scenarios in the Defra model intercomparison exercise

Richard Derwent , Sean Beevers , Charles Chemel , Sally Cooke , Xavier Francis , Andrea Fraser , Mathew R. Heal , Nutthida Kitwiroon , Justin Lingard , Alison Redington , Ranjeet Sokhi , Massimo Vieno

PII: S1352-2310(14)00383-5

DOI: 10.1016/j.atmosenv.2014.05.036

Reference: AEA 12980

To appear in: *Atmospheric Environment*

Received Date: 11 March 2014

Revised Date: 9 May 2014

Accepted Date: 12 May 2014

Please cite this article as: Derwent, R., Beevers, S., Chemel, C., Cooke, S., Francis, X., Fraser, A., Heal, M.R., Kitwiroon, N., Lingard, J., Redington, A., Sokhi, R., Vieno, M., Analysis of UK and European NO_X and VOC emission scenarios in the Defra model intercomparison exercise, *Atmospheric Environment* (2014), doi: 10.1016/j.atmosenv.2014.05.036.

ANALYSIS OF UK AND EUROPEAN NO_x AND VOC EMISSION SCENARIOS IN THE DEFRA MODEL INTERCOMPARISON EXERCISE

Richard Derwent[a,]* Sean Beevers[b], Charles Chemel[c], Sally Cooke[d], Xavier Francis[e], Andrea Fraser[d], Mathew R. Heal[f], Nutthida Kitwiroon[b], Justin Lingard[d], Alison Redington[g], Ranjeet Sokhi[e], Massimo Vieno[h]

[a]rdscientific, Newbury, Berkshire, United Kingdom

[b]Environmental Research Group, King's College, London, United Kingdom

[c]National Centre for Atmospheric Science (NCAS), Centre for Atmospheric and Instrumentation Research (CAIR), University of Hertfordshire, Hatfield, United Kingdom

[d]Ricardo-AEA, Gemini Building, Fermi Avenue, Harwell International Business Centre, Didcot, United Kingdom

[e]Centre for Atmospheric and Instrumentation Research (CAIR), University of Hertfordshire, Hatfield, United Kingdom

[f]School of Chemistry, University of Edinburgh, Edinburgh, United Kingdom

[g]Met Office, Exeter, Devon, United Kingdom

[h]Centre for Ecology and Hydrology, Bush Estate, Penicuik, Midlothian, United Kingdom

*Corresponding author: Tel: +441635569377

E-mail address: r.derwent@btopenworld.com (R.Derwent)

1. Introduction

Air quality models play an important role in air quality policy development by simulating and visualising the conversion of ozone precursor emissions into ground-level ozone levels. Policy makers formulate abatement strategies which aim to reduce ozone levels by reducing ozone precursor emissions. Strategies can be evaluated to determine whether any emission reductions have been stringent enough to achieve acceptable air quality in terms of internationally-accepted air quality standards, guidelines and targets. Strategies may not necessarily be judged as pass or fail but may be evaluated side-by-side with other strategies or against a do-nothing scenario. Increasingly policy makers are using cost-benefit analysis in which the costs of an abatement strategy may be set against the benefits of any environmental improvement as predicted by air quality models.

A huge range of air quality models address ground-level ozone and almost all of them have been used in Europe in a policy context (see Kukkonen et al., 2012). Here the ability of a number of the ground-level ozone models used by the Department for Environment, Food and Rural Affairs (Defra) for its policy support and development to respond to policy-relevant questions, is addressed. The model predictions for a given emission scenario differed widely and we try to explain why. For simplicity, we focus on an episode of peak ozone in southern England and two policy-relevant questions in the context of this one episode: is it better to reduce nitrogen oxide (NO_x) emissions or volatile organic compounds and is better for any reductions to be undertaken concertedly across Europe or unilaterally within the UK to reduce peak ozone levels? This study addresses the potential conflicts that may arise when several models are employed to provide support and advice to policy makers regarding emission control strategies to reduce episodic peak ozone in the UK. Potential conflicts are illustrated with reference to NO_x and VOC emission sensitivities and to controlling emissions from different geographical areas. This study does not try to formulate such policy advice and support but rather focusses on the difficulties inherent when conflicting results are available from eight air quality models.

2. Methodology

The models employed in this study have all been employed to describe photochemical ozone formation across north-west Europe and across the UK. Full details of the eight distinct models are given in the Supplementary Information. They include 3-dimensional Eulerian grid models, a Lagrangian atmospheric dispersion model and moving box trajectory-based models and employ a range of chemical mechanisms to describe photochemical ozone formation from VOC and NO_x emissions. A brief summary of the models is as follows:

- Community Multi-scale Air Quality (CMAQ) model (with 3 distinct implementations)
- Air Quality Unified Model (AQUM),
- European Monitoring and Evaluation Programme for the UK (EMEP4UK) model,
- Numerical Atmospheric dispersion Model Environment (NAME) model,
- Ozone Source Receptor Model (OSRM),
- Photochemical Trajectory Model (PTM).

To reduce the scope and complexity of the study to a level which was tractable, detailed attention was given to the behaviour of ground level ozone during July 2006 at the long-established EMEP rural air quality monitoring station at Harwell, Oxfordshire, UK. This station is located about 80 km due west of London and is surrounded by agricultural fields and a large campus of research establishments. The location of this site is considered typical of much of rural south-east England.

The weather across the UK generally during July 2006 was notable because of its high pressure and high frequency of southerly winds. It was very warm and increasingly humid during the first six days of July 2006, with temperatures of 30 – 32 ^{o}C recorded daily in southern England. From the 14th onwards, the weather was sunny and increasingly hot, with daily maximum temperatures above 32^{o}C from 16th – 27th (Eden, 2006). Ozone observations for Harwell were taken from the UK National Air Quality Archive (http://www.airquality.co.uk/archive/data_and_statistics.php) and converted from µg m^{-3} to ppb units using the factor 0.50. These data demonstrated the occurrence of photochemical ozone episodes producing hourly ozone levels in excess of 50 ppb on 1st – 4th, 6th, 15th – 22nd, 24th – 27th July. The peak hourly ozone level of 106 ppb was recorded on 18th July see Figure 1. Also shown on Figure 1 are the daily advection regimes (as compass bearings N through NW) based on Lamb Weather type (LWT) (Jenkinson and Collinson, 1977) where A refers to anticyclonic and C cyclonic, on NILU FLEXTRA trajectories (Stohl et al., 1995) for Harwell and on the NAME model (see Supplementary Information) air history maps (Manning et al., 2011) where EU refers to advection from a large area of north-west Europe.

Intentionally, no attempt was made to harmonise the input data to the models. Necessarily, the models have used comparable sources for the emission inventory data, for example, based on European Monitoring and Evaluation Programme (EMEP) emissions and the UK National Atmospheric Emission Inventory (NAEI) (for further details, see the Supplementary Information), with VOC speciation data from the NAEI. However, no attempt was made to harmonise the hourly, weekly and seasonal time profiles, gridding or speciation profiles assumed. The models have used different meteorological archives and descriptions of meteorological processes and meteorological models to drive the different parameterisations of boundary layer processes, deposition, atmospheric transport and dispersion. Again, no attempt was made to harmonise the chemical mechanisms employed despite the known sensitivity of ozone predictions for North America to chemical mechanism choice (Luecken et al., 2008) nor the biogenic VOC emission inventories and their speciation.

Each of the 8 models was set up with their respective base case conditions for July 2006 and run in their standard configurations as described in the Supplementary Information. The highest hourly ozone levels predicted each day by each model are plotted together with the corresponding observations in Figure 1. All of the models were able to account satisfactorily for the observed day-to-day variations in daily peak ozone levels in that they exhibited elevated levels during the periods 1st – 4th , 15th – 20th and 24th – 27th July with relatively lower, background levels between 7th – 14th and 28th – 31st July. Some of the observed episode days, however, were missed by some of the models. Individual normalised mean biases (NMBs) for daily ozone maxima for July 2006 spanned the range from -0.18 to -0.04. In the context of the simple evaluation criterion of NMB being in the range -0.2 < NMB <0.2, proposed by Derwent et al., (2010), model performance was considered entirely satisfactory for all eight models for July 2006 at Harwell.

NMBs were negative for all models for July daily maximum ozone levels at Harwell, largely because of poor model performance for July 18th and 19th, see Figure 1. Only one model simulated over 100 ppb for the daily maximum ozone level on these days and seven models gave less than 90 ppb. Model performance was therefore generally poor for these days with highest ozone levels. It is conceivable that the observations were strongly influenced by ozone precursor emissions associated with the 2006 heat-wave which are not adequately represented in the emission inventories employed in the standard model configurations. Air quality during much of the spring and summer of 2006 was influenced by wild-fires in the Russian Federation (Saarikoski et al., 2007; Witham and Manning, 2007; Anttila et al., 2008; Niemi et al., 2009) and it is possible that this influence specifically impacted upon the observed ozone levels at the Harwell station during July 18th and 19th.

Model performance against observations is the subject of further study (Carslaw, 2013) and is not considered further here. It is enough to note that the performance of all eight models during July 2006 as a whole was considered satisfactory and all of the models were able to account satisfactorily for the observed day-to-day variations in the daily peak ozone levels. Because the performance of each model was considered satisfactory, there was no reason to distinguish one set of model results from another and accordingly we have anonymised the models. Each set of model predictions was considered an equally plausible set of possible answers to the policy-relevant questions:

- Do the models agree on the sensitivities to peak O_3 levels to NO_x and VOC emissions?

- Do the models agree on the relative importance of UK precursor emissions to those in the rest of Europe?
- Do the levels of agreement improve if those models and days that had poorer matches between models and observations were excluded?

3. NO_x- versus VOC-sensitivity

An important issue in developing strategies for amelioration of ground-level O_3 is whether to reduce NO_x emissions or VOC emissions or both. To address this issue, attention has been focussed in the modelling on the impact of four simple NO_x and VOC emission scenarios, keeping all other emissions constant:

- S1: 30% reductions in man-made NO_x emissions across Europe,
- S2: 30% reductions in man-made VOC emissions across Europe,
- S3: 30% reductions in man-made NO_x and VOC emissions across Europe,
- S4: 30% reductions in man-made NO_x and VOC emissions across the UK.

The choice of 30% is arbitrary. It is nevertheless comparable to the scale of emission reductions that policy-makers commonly consider. It has been chosen because it is neither too small nor too large and to be consistent with a large literature on photochemical ozone model sensitivity to VOC and NO_x emissions, see for example, Sillman (1999) and Sillman and He (2002). To assess the impact of 30% across-the-board reductions in man-made NO_x and VOC emissions relative to the 2006 base case, each model ran the S1 and S2 emission scenario cases. The maximum hourly ozone levels simulated for the base case and the two scenario cases for each day of July 2006 were determined for each model.

<u>Overall impacts on July-mean O_3 levels</u>

The impact of the 30% reductions in NO_x emissions carried out across the UK and the Rest of Europe (RoE) (Scenario S1) on the July mean daily maximum ozone levels varied considerably between the eight models. O_3 responses (base case minus scenario case) covered the range from -2.0 ppb to +2.0 ppb, with three models producing an increase (-ve response) and five models producing a decrease (+ve response). Figure 2 presents a 'box and whisker' plot of the eight model responses. The interquartile range, shown as a shaded box, confirms that the median model response of +0.4 ppb was not statistically different from zero.

In contrast, Figure 2 shows that the impact on the July mean daily O_3 maximum of 30% reductions in man-made VOC emissions (Scenario S2) was a decrease (+ve response) for all eight models, with responses spanning the range from +0.4 to +3.2 ppb. The median response of +1.2 ppb was statistically significantly different from zero. These model simulations showed that VOC reductions always produced an improvement is air quality, in contrast to the mixed results for NO_x reductions using the July mean daily maximum O_3 as an index.

Daily assignments of NO_x- versus VOC sensitivity

The responses to the 30% NO_x emission reduction and the 30% VOC emission reduction carried out across the UK and the RoE were analysed by considering the model responses on individual days rather than for the month as a whole. If the O_3 response to a 30% NO_x reduction was greater than that to a 30% VOC reductions, then that day was assigned as NO_x-sensitive. Conversely, if the O_3 response to a 30% VOC reduction was greater than that to a 30% NO_x reduction, then that day was assigned as VOC-sensitive. Table 1 shows the VOC- versus NO_x-sensitive assignments for each day of July for each of the eight models. There was complete agreement on the assignments on only six days, with differing levels of disagreement on the remaining 25 days. However, all models showed how the NO_x- versus VOC-sensitivity could switch on a daily basis from NO_x-sensitive to VOC-sensitive and back again during the month. The question is which model is giving the correct assignment when there are apparent contradictions.

Figure 3 presents a scatter plot of the O_3 responses to 30% NO_x reductions against the O_3 responses to 30% VOC reductions for the eight models and for the 15 50-ppb episode days. Also shown is the 1:1 correspondence line representing the locus of equal responses. Points above the line have responses to 30% VOC reductions that are greater than to 30% NO_x reduction and so have been assigned as VOC-sensitive. Points below the line have been assigned as NO_x-sensitive. The vast majority of points are located above the x-axis showing that almost all of the points show positive responses to 30% VOC reductions and hence that air quality improves. In contrast, there are a small but significant number of points to the left of the y-axis, showing that some models show negative responses to 30% NO_x reductions, implying that air quality deteriorates.

The majority of the points in Figure 3 form a 'wedge-shaped' pattern. The apex of the wedge is at the right-hand side of the plot, at the high NOx-response – low-VOC response and widens towards the left-hand side of the plot. There is a tendency for VOC-responses to be smallest when NOx-responses are greatest and VOC-responses to be greatest when NOx-responses are negative. This characteristic tendency has its origins in the theory underpinning NO_x- and VOC-sensitivity as demonstrated by Sillman (1999) and Sillman and He (2002). Superimposed on this characteristic tendency is the impact of model uncertainty which is manifest in terms of the relative scatter between the sets of model points. The axis of the wedge-shaped pattern is almost perpendicular to the 1:1 correspondence line. As a consequence, the characteristic tendency and the model uncertainty strongly impact on the location of the points relative to the 1:1 correspondence line and hence on the NO_x- versus VOC-sensitivity assignments. There are 62 points out of the 120 that are

VOC-sensitive and 58 points that are NO_x-sensitive, indicating a slight preponderance in favour of VOC-sensitivity for the episode days.

The above analysis has shown that there can be a considerable level of disagreement between model assignments of policy-relevant characteristics for O_3 during July 2006. Policy-makers expect that all models used in their support are able to reproduce real-world behaviour. So now we check to see if, by setting a benchmark for such comparisons, we are able to disregard some model results and to focus only on those that deliver good model performance against observations (for this particular test case). Accordingly we set a benchmark of ± 0.1 for the NMB for each day and disregard model results outside this range. This benchmark is set at an arbitrary level and has been tightened to ± 0.05 specifically for the PTM model because some information about observed O_3 levels has been used in the selection of the results from multiple replicates using different back-track trajectories, (see the Supplementary Information for further details). The setting of the benchmark level is a compromise: set too low and all model results would be filtered out and set too high and the situation would not substantially change from that in Table 1.

Table 2 presents the NO_x- versus VOC-sensitivity assignments for only those models that achieved the benchmark NMB of ± 0.1 (± 0.05 for the PTM) on a given day during July 2006. Comparing Tables 1 and 2 shows how setting a benchmark for model performance on each day could drastically reduce the number of table entries. However, there was also a marked reduction in the number of contradictory assignments. Those models that performed better against observations on particular days appeared to give more robust assignments in terms of VOC- versus NO_x-sensitivity. The refinement process in moving from Table 1 to Table 2 has led to a decrease in the proportion of assigned days from 25 out of 31 to 7 out of 21, thereby increasing the level of consensus between the models.

Nevertheless, Table 2 shows that selecting for better model performance did not remove all conflicts. Of the 31 days in July 2006, no conflicts were recorded for 20 days, conflicts were recorded on 7 days and no assignments were possible for 4 days. Of the days with conflicts, 4 days were non-episode days with observed maximum hourly O_3 levels below 50 ppb, leaving only 3 days where the conflict in assignment may have some policy significance. Of the days when a clear-cut assignment could be made, twice as many days were assigned to the NO_x-sensitive category than to the VOC-sensitive category. Generally speaking then, the 'best' models indicated that actions to control NO_x emissions rather than VOC emissions would be the more effective approach to reducing episodic peak O_3 levels at Harwell during July 2006.

All of the four days at Harwell during July 2006 when no assignments were made were episode days, including 18th July on which O_3 levels exceeded 100 ppb. All models had difficulty in simulating O_3 mixing ratios approaching these levels. It is possible that the observed O_3 levels on this and on the other three days were strongly influenced by O_3 precursor sources that were omitted from or were inadequately included in the emission inventories. Possible candidate sources include agricultural burning and forest fires as explained in Section 2. Equally well, there may be difficulties in describing meteorological conditions during these episode days. In any case, filtering by model performance removed the NO_x- versus VOC-sensitivity assignments that may have been based on possibly inadequate evidence.

The conclusion from Table 2 is that there are fewer contradictory NO_x- versus VOC-sensitivity assignments when model performance is used to select the 'better' or 'best' models on each day. The 'best' model changed from day to day and no single model was 'best' on all days. The choice of benchmark based on a daily NMB in the range ± 0.1 (and ± 0.05 for the PTM) was arbitrary and the selection of a different benchmark would change the character of Table 2. However, two conclusions would still stand, namely: selecting 'best' models reduces apparently contradictory assignments and no one model would always be the 'best' model on all days.

4. UK- versus Rest of Europe dominance

A further important issue for UK policymakers has been whether the balance of effort in terms of O_3 precursor emission reductions should be focussed on UK emissions or on emissions from the Rest of Europe (RoE). To assess this issue, attention has been directed to the simple emission scenarios S3 and S4, which focus on the influence of O_3 precursor sources in the UK versus those across Europe as a whole. Figure 2 presents a 'box and whisker' plot of the eight model responses to precursor emissions reductions carried out across Europe as a whole (S3) and across the UK (S4).

Since the UK emissions were included in the European emissions, an estimate of the impact of the RoE emissions could be obtained by subtraction of the UK impacts from the European (UK+RoE) impacts. Therefore if, for a given day and given model, the O_3 response to the 30% reduction in UK-only VOC and NO_x emissions was greater than the difference in response between European emissions reductions and UK emission reductions, then that day was assigned as UK-dominant. Conversely, if the response to the reductions in UK emissions was less than the difference in responses between the European and UK emissions reductions, then that day was assigned as RoE-dominant. This subtraction assumes that O_3 responses are linear and additive, a reasonable working assumption for these relatively small percentage reduction in precursor emissions.

Table 3 shows the UK- versus RoE-dominance assignments for each model and for each day in July 2006. There was complete agreement on UK- and RoE-dominance on only five days and some disagreement on the remaining 26 days. Again, it was apparent that assignment of the major source regions, whether UK or RoE, varied from day to day and so again the question is which of the model assignments is correct for each day.

A detailed analysis of UK- versus RoE-dominance is hampered by a lack of simple rules such as those that exist for NO_x- versus VOC-sensitivity. However, a simple scatter plot provides a suitable introduction to the UK- versus RoE-dominant assignments. Accordingly, Figure 4 presents a scatter plot of the O_3 responses to 30% reductions in both NO_x and VOC emissions carried out across the UK and RoE versus the responses to 30% reductions carried out across the UK only, for all models and all 15 episode days. Also shown is the 1:1 correspondence line which represents the locus of points where the responses across the UK and the RoE are equal to those across the UK only. Figure 4 shows that a small fraction of points lie above the line and that the vast majority of points lie below the line. That is to say, most models indicate that the O_3 levels on most episode days at this location are dominated by ozone precursor sources in the RoE and that the levels on only a few days are dominated by precursor sources in the UK. Subtracting the O_3 responses to the emission reductions in the UK only from the responses to the reductions carried out in the UK + RoE, yields an estimate for the response to the emission reductions carried out in the RoE only. The greater the response to emission reductions carried out across the RoE, the further the points move below the 1:1 line in Figure 4. Responses to RoE-only emission reductions are thus seen to be relatively large compared with responses to UK-only emission reductions on all episode days and with all models. Nevertheless, the considerable amount of scatter in this figure mean that it is not straightforward to draw robust conclusions about UK- versus RoE-dominance on specific days using specific models.

Over all the episode days and all the models, the average O_3 response to 30% emission reductions in both NO_x and VOC in the UK was 0.0 ± 1.5 ppb. Whereas, that to reductions carried out across the RoE was considerably greater at 2.7 ± 0.7 ppb. Episode days were highly likely to be RoE-dominant and this conclusion was robust to choice of model. It was associated with the preponderance of transport from north-west Europe during July 2006 as noted in Figure 1.

To reduce the conflicts between UK- versus RoE-dominance assignments, filtering by model performance against observations was undertaken as shown in Table 4 using the same benchmarks as for Table 2. Again, the number of contradictory assignments has been drastically reduced. Of the 31 days in July 2006, cross-model agreement as to UK- versus RoE-dominance has been reached on 18 days, contradictory assignments on 9 days and no assignments on four days. The possible reasons for the lack of assignments on the four episode days have been highlighted above.

Contradictory assignments were found on nine days compared with seven days for NO_x- versus VOC-sensitivity. This suggests that UK- versus RoE-dominance is somewhat less robust compared with NO_x- versus VOC-sensitivity. Nevertheless, on the basis of Table 4, it is concluded that the 'best' models gave less contradictory assignments, that the 'best' model changed from day to day and that no model was designated as 'best' model on all days. Generally speaking, the 'best' models indicated that daily maximum O_3 levels at Harwell during July 2006 were impacted more by precursor emission sources in the RoE than by sources within the UK.

5. Implementing an ENSEMBLE approach

In the field of atmospheric dispersion modelling, conflicting realisations of air quality forecasts are increasingly being resolved through the use of ensembles (Potempski and Galmarini, 2009). In the field of air quality modelling, Van Loon et al. (2007) and Vautard et al. (2009) employed ensembles extensively in their study of O_3 and nitrogen dioxide (NO_2) levels across Europe using seven regional air quality models. Following their lead, the arithmetic mean of all eight sets of model results and their sensitivity cases were calculated to develop a synthetic set of model results, ENSEMBLE, which were processed in an analogous manner as the set of eight model results. The benchmark of NMB of ± 0.1 for each day was applied and the results for the ENSEMBLE were added to Tables 2 and 4.

Looking at the ENSEMBLE results in Table 2 for NO_x- versus VOC-sensitivity, there appeared to be no clear advantage from the ENSEMBLE results over the individual models A – H in terms of the number of days with NMBs passing the benchmark. The models A – H showed between 4 and 14 entries, whereas the ENSEMBLE showed 12 entries. The ENSEMBLE confirmed the assignment to NO_x-sensitive on five days and added to the conflicting assignments on the remaining seven days. On this basis, it was concluded that the ENSEMBLE approach did not add significantly to the assignment of NO_x- versus VOC-sensitivity for episodic peak O_3 at Harwell, Oxfordshire during July 2006.

The ENSEMBLE results for UK- versus RoE-dominance following the implementation of the NMB benchmark, confirmed the assignments based on the individual models A – H on five days and added to the conflicting assignments on the remaining seven days. The ENSEMBLE approach did not add significantly to the assignment of UK- versus RoE-dominance.

6. Discussion and conclusions

One of the main purposes of air quality modelling in Europe is to assist and support policymakers in the formulation of robust and cost-effective strategies for the control of the transboundary formation and transport of O_3. Because a number of O_3 precursor emission sources have already been effectively controlled, the remaining policy options tend to be expensive or complex. Options for the further control of VOC emissions involve tackling solvent emissions, industrial emissions or evaporation from the gasoline distribution chain. Those for NO_x emissions involve tackling diesel vehicle exhausts and large and small combustion sources. Policymakers in the UK can reasonably ask the modelling community whether the balance of future effort should be focussed on VOC or NO_x emissions, or both, and, in view of the evidence for transboundary O_3 formation and transport, whether future efforts should be focussed on domestic precursor sources or on foreign sources. These considerations have driven the formulation of this present study and its focus on the categoric assignments as to whether the episodic peak O_3 levels in south east England in July 2006 are NO_x- or VOC-sensitive and whether they are dominated by precursor sources within the UK or in the RoE.

In this study, attention has been focussed on the EMEP monitoring station located at Harwell, Oxfordshire in the rural south east UK. This location was chosen because of its relative remoteness from large population centres. Other candidate stations were ruled out because of their coastal locations (Lullington Heath, Rochester, St Osyth and Sibton) which would have unduly biased the results in favour of transboundary sources rather than local formation and transport. Some stations are too close to London (Teddington, Hillingdon and London Eltham) and roadside stations were ruled out because they would be subject to the influence of local O_3 destruction rather than local formation.

However, the focus on a specific station for the analysis may not necessarily play to the strengths of the Eulerian models compared with the Lagrangian or moving parcel models. A strength of grid-based models is that they can yield maps showing how O_3 levels and O_3 responses vary spatially over entire regions, such as the south east UK. However, because of potential uncertainties in defining horizontal transport within a spatial resolution of a few km, spatial mismatch may occur between gridded model output and the actual grid square containing an individual monitoring station, i.e. the model may fail to reproduce high O_3 at one particular site on a given day (the criterion used in this study to define good and poor model performance) for a reason unrelated to its skill in general at capturing VOC-NOx-ozone photochemistry over larger spatial and temporal domains. This will potentially be an issue where a large spatial gradient in O_3 occurs in the vicinity of the monitoring site chosen for observation-model comparison. Figure 5 shows the simulated daily maximum hourly O_3 level for the 6th July across the whole of southern England from one of the grid models in this intercomparison which illustrates the strong spatial gradient in maximum ozone across the location of the Harwell monitoring station (marked by the black circle in the figure). We therefore note that our approach of utilising data from a single monitoring station to evaluate model performance may somewhat have favoured Lagrangian over Eulerian model approaches if our sole aim had been to evaluate model performance. However, the aim of this study has been to illustrate the issues involved in using models in support of air quality policy formulation rather than the selection of the 'best' model.

By setting a benchmark in terms of model performance against observations, we have been able to filter the policy-relevant assignments made with eight air quality models of NO_x- versus VOC-sensitivity and UK- versus RoE-dominance to obtain a more robust understanding of the origins of the O_3 episodes observed in the south east of England during July 2006. There were fewer contradictory assignments when model performance against observations was used to select the 'best' model out of the eight models on each day. The 'best' model changed from day to day and no one model was always designated the 'best' model on all days. The choice of benchmark for the daily NMB was arbitrary and selection of a different benchmark could change the character of the analysis.

In this study, the use of an ensemble approach has been assessed following the suggestions of Van Loon et al. (2007) and Vautard et al. (2009). Both studies reported advantages of using ensembles for the assessment of long-term O_3 levels using annual mean and SUMO35 metrics. There appeared to be little advantage in using ensembles for the assessment of NO_2 levels because the ensemble failed to represent the highest peak values. Our conclusion is that the ensemble approach did not add significantly to the analysis of emission sensitivities at Harwell during July 2006. Our focus was on episodic peak O_3, a metric that is generally underpredicted in models. This may go a long way towards explaining why the ensemble approach offered little advantage in this study.

These conclusions will need to be extended by further work in the future to cover different regions of north-west Europe and to different months and years with their different advection regimes and hence source-receptor relationships. We urgently need to understand the differences in model formulation that have led to the observed conflicts in model responses, whether these lie in meteorological datasets, biogenic VOC emissions or different temporal profiles in emissions. This work shows that we currently do not have access to a single air quality model that is guaranteed to deliver the most likely outcomes to policy makers in terms of emission sensitivities on each day. It is important to maintain a diversity in model approaches to further the development of our understanding of O_3 transboundary formation and transport in north west Europe. We need a wide diversity of models, not because it would guarantee a more accurate ensemble, but because it would give more chances for model results to be acceptable and robust for policy purposes.

Acknowledgements

This study and the modelling work carried out with the OSRM and PTM model was supported by the United Kingdom Department for Environment, Food and Rural Affairs (Defra), Atmosphere and Local Environment under contract AQ0704 and the NAME modelling at the Met Office under contract AQ0902. The NERC Centre for Ecology & Hydrology also supported the development of the EMEP4UK model. The WRF/CMAQ modelling approach used by the University of Hertfordshire was developed as part of the TRANSPHORM (FP7 funded collaborative project), CREMO project (Environment Agency) and the AQMEII Initiative. The help from David Carslaw, King's College London with the collection of the model data is gratefully acknowledged. Discussions with Paul Agnew, Met Office are also much appreciated. NILU is acknowledged for providing FLEXTRA trajectories (www.nilu.no/trajectories) used in this study.

References

Anttila, P., Makkonen, U., Hellen, H., Kyllonen, K., Leppanen, S., Saari, H., Hakola, H., 2008. Impact of the open biomass fires in spring and summer of 2006 on the chemical composition of background air in south-eastern Finland. Atmospheric Environment 42, 642-6486.

Appel, K. W., Chemel, C., Roselle, S. J., Francis, X. V., Hu, R.-M., Sokhi, R. S., Rao, S. T., and Galmarini, S., 2012. Examination of the Community Multiscale Air Quality (CMAQ) model performance over the North American and European domains. Atmospheric Environment, 53, 142-155.

Derwent, R.G., Witham, C.S., Redington, A., Jenkin, M., Stedman, J., Yardley, R., Hayman, G., 2009. Particulate matter at a rural location in southern England during 2006: model sensitivities to precursor emissions. Atmospheric Environment 43, 689-696.

Eden, P., 2006. Weather log July 2006. Weather 61, doi:10.1256/wea.150.06.

Guenther, A., Wiedinmyer, C., 2007. User's Guide to the Model of Emissions of Gases and Aerosols from Nature (MEGAN), Version 2.01.

Jenkinson, A.F., Collinson, F.P., 1977. An initial climatology of gales over the North Sea. Synoptic Climatology Branch Memorandum No. 62, Meteorological Office, Bracknell.

Kukkonen, J., Olsson, T., Schultz, D.M., Baklanov, Klein, A.T., Miranda, A.I., Monteiro, A., Hirtl, M., Tarvainen, V., Boy, M., Peuch, V.-H., Poupkou, A., Kioutsioukis, Finardi, I.S., Sofiev, M., Sokhi, R., Lehtinen, K.E.J., Karatzas, K., San Jos´e, Astitha, R.M., Kallos, G., Schaap, M., Reimer, E., Jakobs, H., Eben, K., 2012. A review of operational, regional-scale, chemical weather forecasting models in Europe. Atmospheric Chemistry and Physics 12, 1–87

Luecken, D.J., Phillips, S., Sarwar, G., Jang, C., 2008. Effects of using the CB05 vs. SAPRC99 vs. CB4 chemical mechanism on model predictions: Ozone and gas-phase photochemical precursor concentrations. Atmospheric Environment 42, 5805-5820.

Manning, A.J., O'Doherty, S., Jones, A.R., Simmonds, P.G., Derwent, R.G., 2011. Estimating UK methane and nitrous oxide emissions from 1990 to 2007 using an inversion modelling approach. Journal of Geophysical Research 116, D02305, doi:10.1029/2010JD014763.

Niemi, J.V., Saarikoski, S., Aurela, M., Tervahattu, H., Hillamo, R., Westphal, D.L., Aarnio, P., Koskentalo, T., Makkonen, U., Vehkamaki, H., Kulmana, M., 2009. Atmospheric Environment 43, 1255-1264.

Pouliot, G., Pierce, T., van der Gon, H. D., Schapp, M., Moran, M., Nopmongcol, U. Comparing emission inventories and model-ready emission datasets between Europe and North America for the AQMEII project. Atmos. Environ., 53, 4-14.

Potempski, S., Galmarini, S., 2009. Est modus in rebus: analystical properties of multi-model ensembles. Atmospheric Chemistry and Physics 9, 9471-9489.

Redington, A.L., Derwent, R.G., Witham, C.S., Manning, A.J., 2009. Sensitivity of modelled Sulphate and nitrate aerosol to cloud, PH and ammonia emissions. Atmospheric Environment 43, 3227-3234.

Ryall, D.B., Maryon, R.H., 1998. Validation of the UK Met Office's NAME model against the ETEX dataset. Atmospheric Environment 32, 4265-4276.

Saarikoski, S., Sillanpaa, M., Sofiev, M., Timonen, H., Saarnio, K., Teinila, K., Karppinen, A., Kukkonen, J., Hillamo, R., 2007. Chemical composition of aerosols during a major biomass burning episode over Northern Europe in spring 2006: Experimental and modelling assessments. Atmospheric Environment 41, 3577-3589.

Sakulyanontvittaya, T., Duhl, T., Wiedinmyer, C., Helmig, D., Matsunaga, S., Potosnak, M., Milford, J., Guenther, A., 2008. Monoterpene an sesquiterpene emission estimates for the United States. Environmental Science and Technology 42, 1623-1629.

Schere, K., Flemming, J., Vautard, R., Chemel, C., Colette, A., Hogrefe, C., Bessagnet, B., Meleux, F., Mathur, R., Roselle, S., Hu, R.-M., Sokhi, R. S., Rao, S. T., Galmarini, S., 2012. Trace gas/aerosol boundary concentrations and their impacts on continental-scale AQMEII modeling domains. Atmospheric Environment, 53, 38-50.

Sillman, S., 1999. The relation between ozone, NO_x and hydrocarbons in urban and polluted rural environments. Atmospheric Environment 33, 1821-1845.

Sillman, S., He, D., 2002. Some theoretical results concerning O_3-NO_x-VOC chemisty and NO_x-VOC indicators. Journal of Geophysical Research 107, D22, 4659, doi10.1029/2001JD001123.

Simpson, D., Benedictow, A., Berge, H., Emberson, L.D., Fagerli, H., Flechard, C.R., Hayman, G.D., Gaus, M., Jonson, J.E., Jenkin, M.E., Nyiri, A., Richter, C., Semeenai, V.S., Tsyro, S., Tuovinen, J.-P. Valdebenito, A., Wind, P. , 2012. The EMEP MSC-W chemical transport model - technical description. Atmospheric Chemistry and Physics 12, 7825-7865.

Sofiev, M., Vankevich, R., Lotjonen, M., Prank, M., Petukhov, V., Ermakova, T., Koskinen, J., Kokkonen, J., 2009. An operational system for the assimilation of satellite information on wild-land fires for the needs of air quality modeling and forecasting. Atmospheric Chemistry and Physics 9, 6833-6847.

Stohl, A., Wotawa, G., Seibert, P., Kromp-Kolb, H., 1995. Interpolation errors in wind fields as a function of spatial and temporal resolution and their impact on different types of kinematic trajectories. Journal of Applied Meteorology 34, 2149-2165.

Van Loon, M., Vautard, R., Schaap, M., Bergstrom, R., Bessagnet, B., Brandt, J., Builtjes, P.J.H., Christensen, J.H., Cuvelier, C., Graff, A., Jonson, J.E., Krol, M., Langner, J., Roberts, P., Rouil, L., Stern, R, Tarrason, L., Thunis, P., Vignati, E., White, L., Wind, P., 2007. Evaluation of long-term ozone simulations from seven regional air quality models and their ensemble. Atmospheric Environment 41, 2083-2097.

Vautard, R., Schaap, M., Bergstrom, R., Bessagnet, B., Brandt, J., Builtjes, P.J.H., Christensen, J.H., Cuvelier, C., Voltescu, V., Graff, A., Kerschbaumer, A., Krol, M., Roberts, P., Rouil, L., Stern R., Tarrason, L., Thunis, P, Vignati, E., Wind. P., 2009. Skill and uncertainty of a regional air quality model ensemble. Atmospheric Environment 43, 4822-4832.

Vautard, R., Moran, M. D., Solazzo, E., Gilliam, R. C., Matthias, V., Bianconi, R., Chemel, C., Ferreira, J., Geyer, B., Hansen, A.B., Jericevic, A., Prank, M., Segers, A., Silver, J.D., Werhahn, J., Wolke, R., Rao, S.T., Galmarini, S., 2012. Evaluation of the meteorological forcing used for the Air Quality Model Evaluation International Initiative (AQMEII) air quality simulations. Atmospheric Environment 53, 15-37.

Vieno, M., Dore, A. J., Stevenson, D. S., Doherty, R., Heal, M. R., Reis, S., Hallsworth, S., Tarrason, L., Wind, P., Fowler, D., Simpson, D., Sutton, M. A., 2010. Modelling surface ozone during the 2003 heat-wave in the UK. Atmospheric Chemistry and Physics 10, 7963-7978.

Witham, C., Manning, A., 2007. Impacts of Russian biomass burning on UK air quality. Atmospheric Environment 41, 8075-8090.

Table 1. Assignments of NO_x- or VOC-sensitivity for each model A-H for each day of July 2006.

Model	A	B	C	D	E	F	G	H
1st	VOC	VOC	VOC	VOC	VOC	VOC	VOC	VOC
2nd	NO_x	NO_x	NO_x	NO_x	NO_x	NO_x	NO_x	NO_x
3rd	NO_x	VOC	VOC	NO_x	VOC	VOC	NO_x	NO_x
4th	NO_x	VOC	VOC	VOC	VOC	NO_x	NO_x	VOC
5th	NO_x	VOC	VOC	VOC	NO_x	NO_x	NO_x	VOC
6th	VOC	VOC	NO_x	VOC	NO_x	NO_x	NO_x	NO_x
7th	NO_x	VOC	VOC	VOC	VOC	NO_x	NO_x	VOC
8th	NO_x	VOC	NO_x	NO_x	NO_x	NO_x	NO_x	NO_x
9th	NO_x	NO_x	NO_x	NO_x	NO_x	NO_x	NO_x	VOC
10th	NO_x	VOC	VOC	NOx	NOx	VOC	VOC	VOC
11th	NO_x	NO_x	NO_x	VOC	NO_x	NO_x	NO_x	NO_x
12th	NO_x	VOC	VOC	VOC	VOC	VOC	VOC	VOC
13th	VOC	VOC	VOC	VOC	VOC	VOC	VOC	VOC
14th	NO_x	VOC	VOC	NO_x	NO_x	VOC	NO_x	NO_x
15th	NO_x	VOC	VOC	VOC	NO_x	VOC	VOC	NO_x
16th	VOC	VOC	VOC	VOC	VOC	VOC	VOC	NO_x
17th	VOC	VOC	VOC	VOC	VOC	VOC	NO_x	VOC
18th	VOC	VOC	VOC	VOC	VOC	NO_x	NO_x	VOC
19th	NO_x	VOC	VOC	NO_x	VOC	NO_x	VOC	NO_x
20th	NO_x	VOC	VOC	NO_x	NO_x	NO_x	NO_x	NO_x
21st	NO_x	NO_x	NO_x	NO_x	NO_x	NO_x	NO_x	NO_x
22nd	VOC	VOC	VOC	NO_x	NO_x	VOC	NO_x	NO_x
23rd	VOC	VOC	VOC	NO_x	VOC	NO_x	NO_x	VOC
24th	VOC	VOC	VOC	VOC	VOC	VOC	NO_x	VOC
25th	NO_x	VOC	VOC	VOC	NO_x	NO_x	NO_x	NO_x
26th	VOC	VOC	VOC	VOC	NO_x	NO_x	NO_x	NO_x
27th	VOC	VOC	VOC	NO_x	NO_x	VOC	NO_x	NO_x
28th	NO_x	VOC	VOC	NO_x	VOC	VOC	NO_x	NO_x
29th	NO_x	NO_x	NO_x	NO_x	NO_x	NO_x	NO_x	NO_x
30th	NO_x	NO_x	NO_x	NO_x	NO_x	NO_x	NO_x	NO_x
31st	NO_x	VOC	VOC	NO_x	NO_x	NO_x	NO_x	NO_x

Notes: highlighting denotes days when all assignments agree.

Table 2. Assignments of NO_x- or VOC-sensitivity for each model A-H for each day of July 2006 having filtered the results on the basis of model performance for each day using a NMB threshold of ± 0.1 (± 0.05 for the PTM model), together with the observed maximum hourly mean ozone concentration.

Model	Obs, ppb	A	B	C	D	E	F	G	H	ENS [a]
1st	82							VOC		
2nd	80			NO_x						
3rd	81							NO_x		
4th	79									
5th	38	NO_x								
6th	60	VOC		NO_x		NO_x	NO_x		NO_x	VOC
7th	29			VOC						
8th	34	NO_x				NO_x				NO_x
9th	32	NO_x		NO_x		NO_x				
10th	21			VOC						
11th	39							NO_x		NO_x
12th	35	NO_x		VOC				VOC		VOC
13th	33	VOC			VOC	VOC		VOC		
14th	42							NO_x		
15th	51						VOC	VOC		
16th	75							VOC		
17th	76									
18th	106									
19th	103	NO_x								
20th	58	NO_x						NO_x		NO_x
21st	61				NO_x			NO_x		
22nd	56				NO_x			NO_x	NO_x	VOC
23rd	43	VOC					NO_x	NO_x		VOC
24th	72									
25th	69		VOC			NO_x				NO_x
26th	65			VOC	VOC	NO_x				VOC
27th	63		VOC							
28th	43		VOC		NO_x				NO_x	VOC
29th	36					NO_x	NO_x	NO_x	NO_x	NO_x
30th	36	NO_x	NO_x		NO_x	NO_x				NO_x
31st	43		VOC		NO_x		NO_x	NO_x		

[a] ENS refers to the ENSEMBLE

Table 3. Assignments of UK- or Rest of Europe-dominance for each model A-H for each day of July 2006.

Model	A	B	C	D	E	F	G	H
1st	RoE	UK	RoE	UK	RoE	RoE	RoE	RoE
2nd	RoE	RoE	RoE	RoE	RoE	RoE	RoE	RoE
3rd	RoE	RoE	RoE	RoE	RoE	RoE	RoE	RoE
4th	RoE	UK	RoE	RoE	RoE	RoE	RoE	RoE
5th	RoE	RoE	RoE	RoE	UK	RoE	RoE	RoE
6th	UK	RoE	RoE	RoE	UK	RoE	RoE	RoE
7th	UK	RoE	RoE	RoE	RoE	UK	UK	RoE
8th	RoE	RoE	RoE	RoE	RoE	RoE	RoE	UK
9th	UK	RoE	RoE	RoE	RoE	UK	RoE	RoE
10th	UK	UK	RoE	RoE	RoE	UK	RoE	RoE
11th	UK	UK	UK	RoE	UK	UK	UK	UK
12th	UK	RoE	RoE	RoE	RoE	RoE	RoE	RoE
13th	RoE	RoE	RoE	RoE	RoE	RoE	RoE	RoE
14th	UK	RoE	RoE	RoE	UK	RoE	UK	UK
15th	UK	UK	RoE	RoE	UK	RoE	RoE	UK
16th	RoE	RoE	RoE	RoE	RoE	RoE	RoE	RoE
17th	RoE	RoE	RoE	RoE	UK	RoE	RoE	RoE
18th	RoE	RoE	RoE	RoE	RoE	RoE	RoE	RoE
19th	RoE	RoE	RoE	RoE	RoE	RoE	RoE	UK
20th	RoE	RoE	UK	RoE	RoE	RoE	RoE	RoE
21st	RoE	RoE	UK	RoE	RoE	RoE	UK	UK
22nd	UK	UK	RoE	RoE	RoE	RoE	RoE	RoE
23rd	UK	RoE	RoE	UK	RoE	UK	UK	RoE
24th	UK	UK	RoE	RoE	RoE	UK	UK	RoE
25th	RoE	UK	RoE	RoE	RoE	RoE	RoE	RoE
26th	RoE	RoE	RoE	UK	RoE	RoE	UK	RoE
27th	UK	RoE	RoE	RoE	RoE	RoE	UK	UK
28th	UK	RoE	RoE	RoE	RoE	RoE	UK	UK
29th	RoE	UK	RoE	RoE	RoE	RoE	RoE	RoE
30th	UK	RoE	UK	UK	RoE	UK	UK	UK
31st	RoE	RoE	RoE	RoE	RoE	RoE	UK	RoE

Notes: highlighting denotes days when all assignments agree.

Table 4. Assignments of UK- or RoE-dominance for each model A-H for each day of July 2006 having filtered the results on the basis of model performance for each day using a NMB threshold of ± 0.1 (± 0.05 for the PTM model).

Model	A	B	C	D	E	F	G	H	ENS [a]
1st							RoE		
2nd			RoE						
3rd							RoE		
4th									
5th	RoE								
6th	UK		RoE		UK	RoE		RoE	RoE
7th			RoE						
8th	RoE				RoE				RoE
9th	UK		RoE		RoE				
10th			RoE						
11th							UK		UK
12th	UK		RoE				RoE		RoE
13th	RoE			RoE	RoE		RoE		
14th							UK		
15th						RoE	RoE		
16th							RoE		
17th									
18th									
19th	RoE								
20th	RoE						RoE		RoE
21st				RoE			UK		
22nd				RoE			RoE	RoE	RoE
23rd	UK					UK	UK		RoE
24th									
25th		UK			RoE				RoE
26th			RoE	UK	RoE				RoE
27th		RoE							
28th		RoE		RoE				UK	RoE
29th					RoE	RoE	RoE	RoE	RoE
30th	UK	RoE		UK	RoE				RoE
31st		RoE		RoE		RoE	UK		

[a] ENS refers to the ENSEMBLE

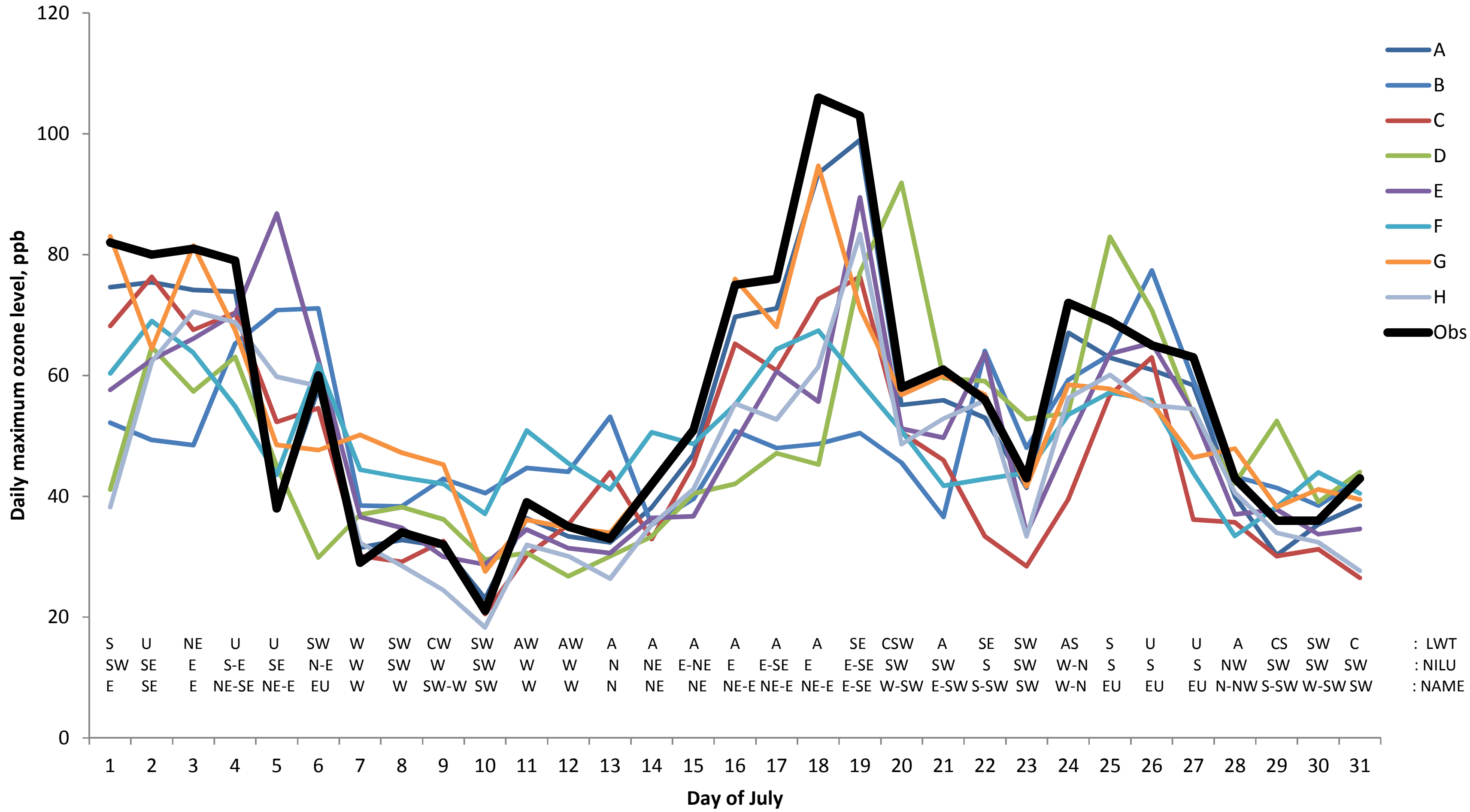

Figure 1. Daily maximum hourly ozone concentrations for all eight models A-H and observations for July 2006 at Harwell, Oxfordshire. Also shown are the daily advection regimes as Lamb Weather types (LWT), NILU FLEXTRA trajectories (NILU) and NAME air history maps (NAME), see text.

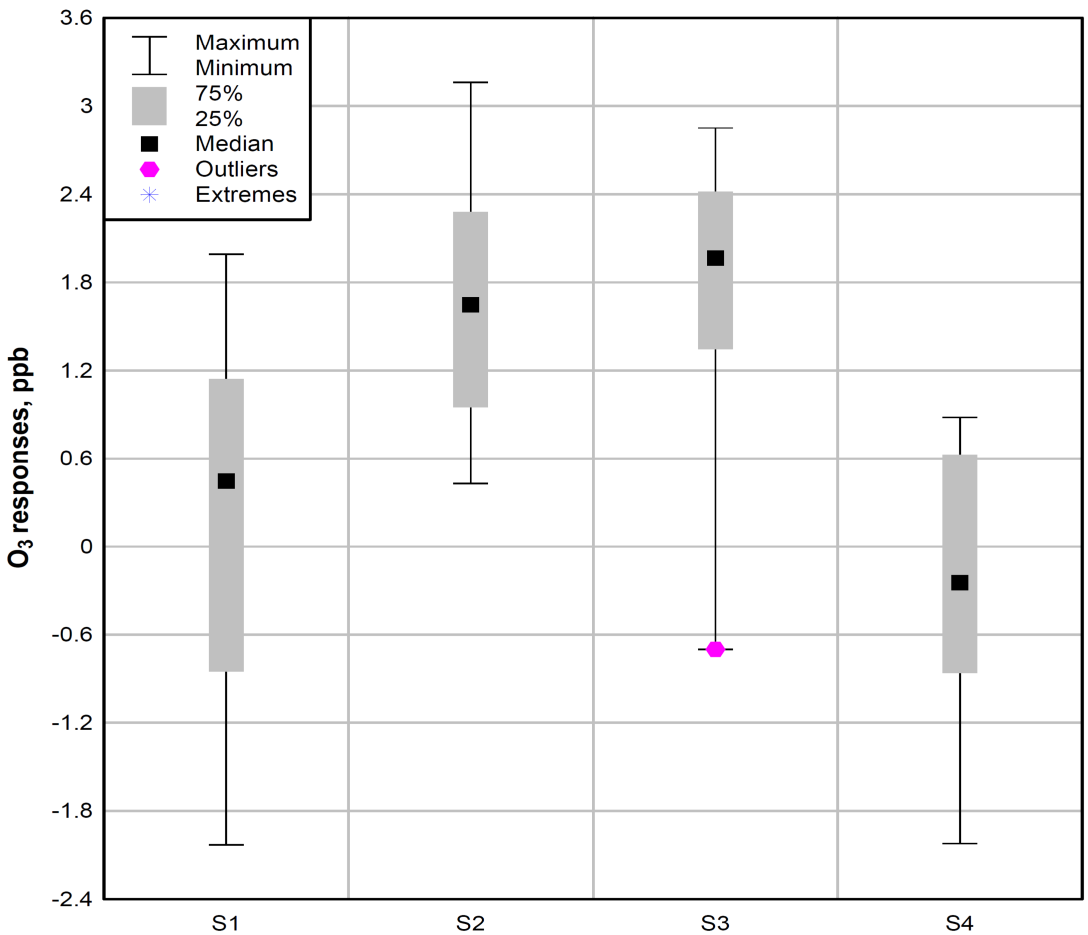

Figure 2. Box-whisker plots of the changes in July mean daily maximum ozone concentration across the eight models, for the S1 – S4 scenarios. Shaded box: interquartile range, black square: median.

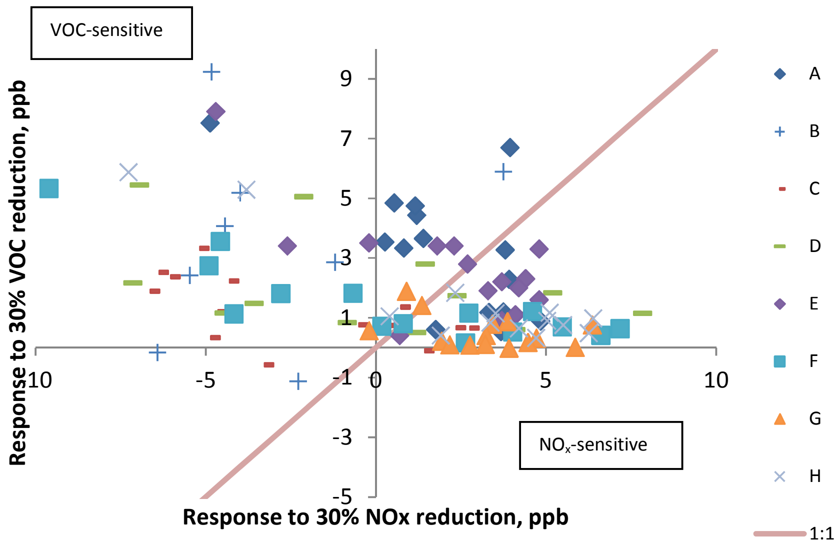

Figure 3. Scatter plot of the eight model O_3 responses to 30% NO_x reductions versus 30% VOC reductions for the episode days of July 2006. Also shown is the 1:1 correspondence line above which points indicate VOC-sensitive model simulations and below which they indicate NO_x-sensitive simulations.

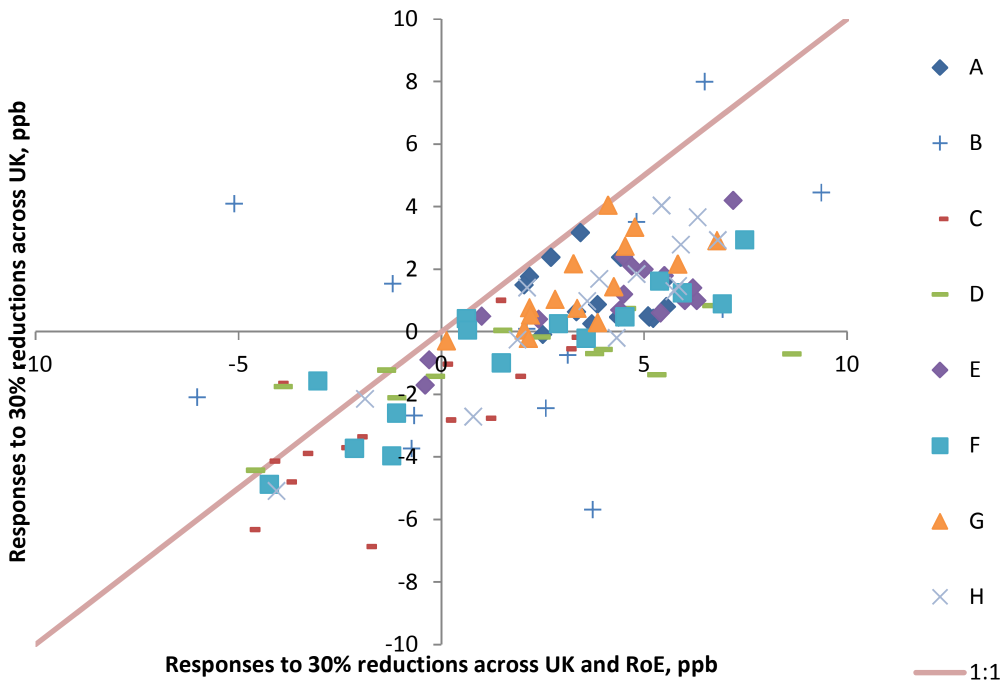

Figure 4. Scatter plot of the O_3 responses on episode days for the eight models to 30% reductions in NO_x and VOC emissions carried out across the UK and the RoE versus the O_3 responses to 30% reductions in NO_x and VOC emissions carried out across the UK only.

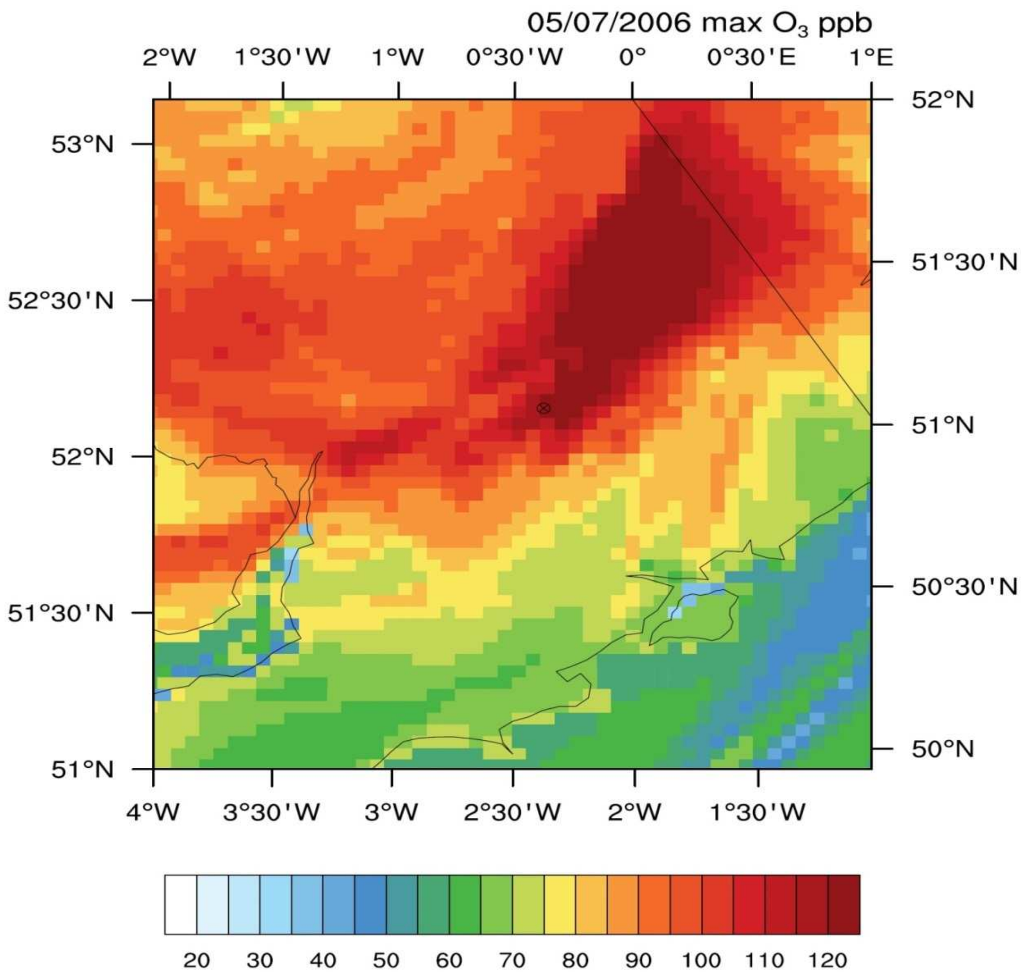

Figure 5. Simulated maximum hourly ozone across southern England on the 6th July 2006 from one of the Eulerian grid models in the model intercomparison. The black circled cross symbol marks the location of the Harwell monitoring site.

Highlights

- Emission scenarios were implemented in eight ozone air quality models
- NOx- and VOC sensitivities for peak ozone levels were highly variable between days
- Filtering by model performance minimised apparent conflicts between models

SUPPLEMENTARY INFORMATION

ANALYSIS OF UK AND EUROPEAN NO_x AND VOC EMISSION SCENARIOS IN THE DEFRA MODEL INTERCOMPARISON EXERCISE

Richard Derwent[a,]* Sean Beevers[b], Charles Chemel[c], Sally Cooke[d], Xavier Francis[e], Mathew R. Heal[f], Nutthida Kitwiroon[b], Justin Lingard[d], Alison Redington[g], Ranjeet Sokhi[e], Massimo Vieno[h]

[a]rdscientific, Newbury, Berkshire, United Kingdom

[b]Environmental Research Group, King's College, London, United Kingdom

[c]National Centre for Atmospheric Science (NCAS), Centre for Atmospheric and Instrumentation Research (CAIR), University of Hertfordshire, Hatfield, United Kingdom

[d]Ricardo-AEA, Gemini Building, Fermi Avenue, Harwell International Business Centre, Didcot, United Kingdom

[e]Centre for Atmospheric and Instrumentation Research (CAIR), University of Hertfordshire, Hatfield, United Kingdom

[f]School of Chemistry, University of Edinburgh, Edinburgh, United Kingdom

[g]Met Office, Exeter, Devon, United Kingdom

[h]Centre for Ecology and Hydrology, Bush Estate, Penicuik, Midlothian, United Kingdom

1. Details of the Models

1.1 CMAQ – AEA

The CMAQ – AEA model is an application of the United States Environmental Protection Agency (EPA) Community Multiscale Air Quality (CMAQ) modelling system which is a third-generation air quality model available online at www.cmaq-model.org . CMAQ is designed for applications ranging from regulatory and policy analysis to understanding the complex interactions of atmospheric chemistry and physics. It is a three-dimensional Eulerian (i.e., gridded) atmospheric chemistry and transport modelling system that simulates ozone, particulate matter (PM), toxic airborne pollutants, visibility, and acidic and nutrient pollutant species throughout the troposphere. Designed as a "one-atmosphere" model, CMAQ can address the complex couplings among several air quality issues simultaneously across spatial scales ranging from local to hemispheric. The CMAQ source code is highly transparent and modular to facilitate the model's extensibility through community development by members of the air quality modelling community. CMAQ was first developed in the late 1990's, the latest version 4.7.1 released in 2010.

In the CMAQ – AEA implementation, the model has been run at horizontal resolutions of 48km (Europe) and 12km (UK) for this study. A new version at 50km and 10km is currently used for the forecast. The 48+12km simulation uses a 26 layer vertical structure with 12 layers below 800m and a lowest layer of 9 m. The 50+10km forecast uses 19 layers, the lowest at 18m this increases the stability of the weather forecast. For limited studies the resolution was reduced to 4km.

European emissions are based on the 2006 EMEP emissions. UK emissions are based on the 2006 NAEI. Temporal profiles were used for the main emission SNAP sectors. Natural emissions are based on the Biogenic Potential Inventory. Numerical weather data are produced using WRFv3 on the same scale as CMAQ. Boundary and forcing conditions are provided by ECMWF for 2006 and GFS forecast is used for the daily AQ forecast. The chemical mechanism used for the AQ forecasting is Carbon Bond 05 with extensions for Cl, aqueous and aerosol chemistry. The alternative chemical mechanisms available in CMAQ v4.7 is SAPRC-99. CB-IV and RADM2 are available in earlier versions. Dry deposition currently runs within the MCIP (Meteorology Chemistry Interface Processor) and uses a surface exchange aerodynamic method using surface resistance, canopy resistance, and stomatal resistance to compute dry deposition velocities.

1.2 CMAQ – King's College London

The CMAQ – King's College London is an application of a 3-D Eulerian grid air quality model. CMAQ was released to the public in June 1998 by the United States EPA. The primary goals of the model are to improve 1) the environmental management community's ability to evaluate the impact of air quality management practices for multiple pollutants at multiple scales and 2) the scientist's ability to better probe, understand, and simulate chemical and physical interactions in the atmosphere. The CMAQ modelling system is set up at the ERG for both current and future policy assessment. Currently, the model is used as part of health impact assessment research at the ERG (MRC centre).

Domain setting: Domains with 4 nested level (23 vertical levels)

Dom1: 81km grid spacing, 47 x 44 cells

Dom2: 27km grid spacing, 39x39 cells

Dom3: 9km grid spacing, 66x108 cells

Dom4: 3km grid spacing, 72x72 cells

Dom5: 1km grid spacing, 62x51 cells

In the present study, European emissions were based on EMEP and UK NAEI emissions. Meteorological data were based on WRF3.1. The chemical mechanism used was Carbon Bond-05 with aerosol and aqueous chemistry. The dry deposition scheme was based on a surface exchange aerodynamic method which uses surface resistance, canopy resistance and stomatal resistance to compute dry deposition velocities.

1.3 CMAQ – University of Hertfordshire

The CMAQ modelling system configuration is as used by Appel et al. (2012) for AQMEII for the European domain using a horizontal grid spacing of 18 km. A detailed description of the anthropogenic emissions used is available in Pouliot et al. (2012). Biogenic emissions of isoprene and terpene, calculated using the Model of Emissions of Gases and Aerosols from Nature (MEGAN; Guenther and Wiedinmyer, 2007; Sakulyanontvittaya et al., 2008), are included on the same resolution as the anthropogenic emissions. The fire emissions were bases on 2006 daily fire estimates from the Moderate Resolution Imaging Spectroradiometer (MODIS) fire radiative power product (Sofiev et al., 2009).The calculations used 34 vertical layers. Model options employed include the CB05 chemical mechanism with chlorine chemistry extensions, the AERO5 aerosol module, the Asymmetric Cloud Model 2 (ACM2) PBL scheme. The simulations utilised boundary concentrations from the GEOS-Chem global model (see Schere et al., 2012). The meteorological fields were obtained from the Weather Research and Forecasting (WRF) model (see Vautard et al., 2012). For the WRF model run, the initial conditions and lateral boundary conditions were derived from the European Centre for Medium-range Weather Forecasts (ECMWF) gridded analyses.

1.4 EMEP4UK – Centre for Ecology and Hydrology

The EMEP4UK model (Vieno et al., 2010) is a Eulerian grid model based on the EMEP Unified model (Simpson et al., 2012). The development of the EMEP4UK model first started in 2006 by Massimo Vieno (University of Edinburgh, CEH Edinburgh), and Peter Wind and David Simpson (Norwegian Meteorological institute).

EMEP4UK is a nested model run at a spatial resolution of 50 km x 50 km (170 x 133 grid) over the full EMEP extended European regional domain and at a finer resolution of 5 km x 5 km (222 x 260 grid) over a British Isles domain for the main model results.

NAEI emissions data have been used for the UK and EMEP emissions data have been used everywhere else. Meteorological data have been obtained from the WRF model versions 2.2, 3.1.1, and 3.2. The EMEP Unified model chemistry scheme has been used although more chemical schemes are going to be available with the new version of the EMEP Unified model. The EMEP Unified model deposition scheme has been used to treat dry deposition.

1.5 Ozone Source Receptor Model (OSRM)

The OSRM is a Lagrangian trajectory model whose development has been led by AEA working through an enduring consortium of leading UK experts under contract to Defra (and previous Departments) since 1999. Following the initial design of the model in a research and development stage, various features of the model were enhanced to improve model performance, to take account of further developments in the underlying science and to make the model more suitable for direct application to Defra air quality policy. Since around 2005, the emphasis has shifted from development to maintenance and application of the model as a policy tool for examining the response of the UK ground-level ozone climate to changes in precursor emissions in the UK and Europe.

OSRM uses NAEI 1x1km emissions data for NO_x, VOCs, CO and SO_2 grouped into 8 source sectors for the UK. Over the Rest of Europe in the EMEP domain: EMEP 50x50km emissions data are used in combination with country totals for scaling to years up to 2020. Temporal profiles for man-made emission sources are employed for the different sectors. The NAEI VOC speciation profile is used and the assignment of the 664 individual VOCs in the NAEI speciated inventory to the 13 VOCs in the OSRM is based on reactivity and structural considerations. Gridded emissions for shipping are based on the Entec studies. An emission term is added to the emission rate of isoprene to represent the natural biogenic emissions from European forests and agricultural crops. The emission estimates can either be the same as those used in the UK PTM or from the biogenic inventory produced using the PELCOM land cover dataset and the TNO tree species inventory.

The UK Met Office provides meteorological datasets derived from the NAME model. 30 boundary layer meteorological parameters are provided at 6-hourly resolution over a year, covering a domain from 30°W to 40°E and 20° to 80°N at 1° spatial resolution. These data are used to derive 96-hour back trajectories to specified receptors. The OSRM now has meteorological data in this form for each calendar year from 1999 to 2009.

The current version of the OSRM uses an updated version of the mechanism in STOCHEM: 70 chemical species involved in 195 thermal and photochemical reactions. An experimental version of the OSRM (Version 25) has been using the most reduced form of the latest CRIv2 (CRIv2-R5) mechanism (196 chemical species, 555 reactions) which is linked to the Master Chemical Mechanism.

Dry deposition processes are represented using a conventional resistance approach, in which the rate of dry deposition is characterised by a deposition velocity. Different deposition velocities are used over land and sea. The ozone deposition velocity over land has an imposed diurnal and seasonal cycle. The OSRM works in conjunction with a surface ozone flux model which has been updated recently with the latest parameterisations from the SEI DO_3SE model and treatment of dry deposition is currently being modified to give reduced deposition during dry periods.

1.6 NAME – Met Office

The NAME model was originally developed by the Met Office's Atmospheric Dispersion Group following the Chernobyl incident to simulate medium and long range transport and wet and dry deposition of radionuclides. NAME is three-dimensional a Lagrangian dispersion model that simulates the dispersion, chemistry and deposition processes occurring in the atmosphere. The model runs employs three-dimensional meteorological fields from the Met Office Unified Model. The model is well documented and has numerous applications, for example modelling volcanic eruptions, accidental releases of radionuclides, the spread of foot and mouth disease and air quality. A detailed description of the NAME model physics can be found in Ryall and Maryon (1998) and a description of the atmospheric chemistry model applications can be found in Redington et al., (2009).

Pollutant emissions are represented by releasing millions of air parcels, each able to represent the released mass of many different species. The air parcels are carried by the three-dimensional wind field obtained from the Unified Model (UM). Local turbulent motion is simulated using a random walk technique which requires a diffusion coefficient calculated from the local turbulent velocity variance and the local turbulent timescale. Above the boundary layer these two quantities are fixed, but within the boundary layer they are defined in terms of the local atmospheric stability and local surface quantities. The UM provides direct output of boundary layer height for use in NAME.

NAME's chemistry scheme is based on that of the Met Office's global STOCHEM model. NAME's dry deposition scheme is based on the concept of a deposition velocity and has various degrees of sophistication. In its simplest form, a fixed deposition velocity for a given species is specified. More generally, a resistance analogy is used to calculate a species dependent deposition velocity. The surface resistance term, denoting the resistance to capture by the surface itself, for a given species can either be a simple fixed value or a more explicit parameterisation dependent on land surface properties. The laminar sub-layer resistance term, representing the resistance to transport through the thin quasi-laminar layer adjacent to the surface, is parameterised according to gaseous or aerosol species, and for aerosol species is dependent on the particle size. The deposition scheme is applied to all air parcels within the boundary layer.

The model domain was 14^{o}W-19.9^{o}E, 42^{o}N-62^{o}N with chemistry and output grid set to ~10km x 10km (0.15^{o} longitude, 0.09^{o} latitude). The model was run using emissions data for 2006 from the National Atmospheric Emissions Inventory (NAEI) over the UK (http://www.naei.org.uk) and from the European Monitoring and Evaluation Programme (EMEP) over the rest of Europe (http://www.emep.int). All emissions were assumed to be constant throughout the year at the annual rate. A daily cycle, varying according to the day of the week, was applied to pollutants emitted by road traffic. Over the UK the NAEI emissions were split into large point sources (containing specific release height information) and small area sources (4 km x 4 km) and large area sources (20 km x 20 km) with release heights of 0–20 m for traffic sources and 0–50 m for other sources. The EMEP emissions data was released from 0-100m.

NAME was run using meteorological data provided by the Met Office Unified Model in the form of three dimensional three hourly met fields, with a horizontal resolution of 0.375° latitude by 0.5625° longitude (~40 x 40 km over the UK), and thirty three vertical levels.

1.7 Air Quality in the Unified Model AQUM – Met Office

AQUM is a limited area configuration of the Met Office Unified Model (MetUM) which uses the UKCA chemistry scheme. The MetUM is a sophisticated system capable of modelling regions from limited areas to globally and with timescales from less than hourly to climate scales. UKCA development first began in 2003 as part of a joint project initially comprising the Met Office and the universities of Cambridge and Leeds, with the aim of building a chemistry and aerosols sub-model within the Met Office's Unified Model for use in climate modelling. Since 2005, AQUM (Air Quality in the Unified Model) has been developed by the Met Office as a configuration of UKCA for modelling regional air quality. AQUM is run online, as part of the Met Office Unified Model, which is an Eulerian meteorological model.

For modelling air quality in the United Kingdom, the following emissions data sets are typically used: NAEI emissions at 1km x 1 km resolution over the UK, ENTEC - 5km x 5km emissions (2007) for shipping surrounding the UK and EMEP emissions at 0.5^{o} x 0.5^{o} over the remainder of Europe. AQUM uses the RAQ (Regional Air Quality) scheme, which is an updated version of the STOCHEM chemical mechanism. Dry deposition is based on a Wesely scheme.

1.8 Photochemical Trajectory Model (PTM) – rdscientific

The PTM model is a moving air parcel trajectory model that is used to describe photochemical ozone and fine particle formation in north west Europe. The PTM model is used to quantify the contribution made by each VOC species and each VOC source category to the long-range transboundary formation and transport of ozone and PM across North West Europe. These contributions are defined in terms of Photochemical Ozone Creation Potentials POCPs and SOAPs. This is the only European model able to evaluate the role of a wide range of VOCs and their sources in ozone policy formation. POCPs are widely used in a wide range of policy analyses and in life cycle analyses.

The PTM uses SO_2, NO_x, NH_3, VOCs, CO and CH_4 emissions taken from 2010 version of NAEI for the UK and SO_2, NO_x, NH_3, VOCs, CO and CH_4 emissions for the rest of Europe were taken from the EMEP webdab (2010). Isoprene emissions were taken from EMEP. Terpene emissions were taken from Hope Stewart and Nick Hewitt for UK and GEIA for Europe.

4-day 3-D back-track trajectories from Met Office Unified model providing latitude, longitude, altitude, boundary layer depth, temperature were used to describe the meteorological processes. Between 30 and 1,000 equal probability trajectories arriving at each arrival point between 15:00 and 15:15 z each day from Met Office NAME model were used in the present study. A Wesely dry deposition velocity scheme was used but no treatment was given for wet deposition. All model results were obtained using the CRIv2 chemical mechanism. Details of the model description are given in Derwent et al., (2009).

The PTM was run with each of the 30 equal probability trajectories for each day. The trajectory that gave the closest results to the observations for ozone for each day was selected and these results were used in the Defra model intercomparison.