www.defra.gov.uk

Evaluating the Performance of Air Quality Models

Issue 3/June 2010

Dick Derwent
Andrea Fraser
John Abbott
Mike Jenkin
Paul Willis
Tim Murrells

Llywodraeth Cynulliad Cymru
Welsh Assembly Government

DOE
Department of
the Environment
www.doeni.gov.uk

The Scottish
Government

defra
Department for Environment
Food and Rural Affairs

Executive Summary

This report provides basic advice to Defra's contractors concerning what should be considered as 'best practice' for air quality model evaluation. A model intercomparison protocol has been laid out which provides a framework for the Air Quality and Industrial Pollution Division of Defra to conduct a review of its current modelling activities and to ensure that the models used are fit-for-purpose and reflect current state-of-the-art.

The report provides some background to air quality models and discusses a number of issues that are central to air quality model evaluation. A critical step in the evaluation of model performance is the comparison of model results with observations. Annexes are provided giving detailed advice on the selection and availability of network data and the influence that the choice of chemical mechanism, dry deposition parameterisation, emissions and meteorological data can make on model comparisons with observations. However, models must be right for the right reasons and this implies further evaluation beyond a simple comparison with observations.

Protocols are provided for ground-level ozone, acidification and eutrophication and urban air quality modelling. Each evaluation protocol poses three general questions:

- Is the scientific formulation of the model broadly accepted and does it use state-of-the-art process descriptions? This is the scientific evaluation question;
- Does the model replicate observations? This is the operational evaluation question; and
- Is the model suitable for answering policy questions and fulfilling its designated tasks? This is the diagnostic evaluation question.

The protocols do not in themselves answer these questions but exist merely to elicit information from each air quality modelling team to allow Defra to form a view on whether a particular model is fit-for-purpose or not.

Table of contents

1 Introduction

1.1 Air Quality Modelling and Defra

1. The aim of this study is to provide an air quality model intercomparison protocol to assist Air Quality and Industrial Pollution AQIP Division of Defra in the evaluation of the air quality models employed in support of policy formulation and assessment. There is a continuing need for Defra to review current modelling activities to ensure that the models used are fit-for-purpose, reflecting the current state-of-the-science, to assess uncertainties and to ensure that the models are able to handle changes in the expected drivers of future policies, including climate change.

2. Defra relies on its air quality modelling contractors to provide evidence that each model is fit-for-purpose and to document its level of performance. The purpose of this air quality model intercomparison protocol is to provide basic advice to Defra’s contractors concerning what could be considered as ‘best practice’ for air quality model evaluation. This will assist Defra in its judgement of the performance of each model and whether it meets their requirements. The protocol is therefore seen as a tool to assist Defra and its contractors in building a long term programme of model development and improvement. By this means, Defra will continue to receive policy advice and support based upon up-to-date science and state-of-the-science modelling tools.

3. Defra and its contractors build air quality models to address a range of purposes and tasks and these must be taken into account when evaluating model performance and judging fitness-for-purpose. A number of general tasks have been categorised for environmental models by Beck (2002), including:

   - To encode succinctly an archive of knowledge;
   - Have an exploratory tool for exploring our ignorance,
   - To provide a means of communicating scientific notions to others; and
   - To provide a tool for prediction to support decision making or policy formulation.

4. Some or all of these tasks can be identified in the air quality modelling studies carried out by Defra and its contractors. The Master Chemical Mechanism, for example, provides an elegant and systematic means for compiling the many thousands of chemical reactions that drive the atmospheric chemistry of the oxidation of organic compounds and their role in ground-level ozone formation. The Photochemical Trajectory Model (PTM) model has been used to provide a first attempt at mass closure for $PM_{2.5}$ and to ascertain the extent of our current ignorance of the sources and distributions of $PM_{2.5}$ components. The same can be said of the studies to characterise the sources of emission of the organic compounds that contribute most to ground-level ozone formation. Dispersion models are often used in public and planning inquiries to visualise and map the likely pollution footprints from new and existing developments and processes. The use of models as tools for the prediction of future air quality, taking into account the impacts of current and future policies, has been the main reason for the development of air quality models such as FRAME and the Ozone Source Receptor Model (OSRM). Empirical models such as those used in the Pollution Climate Mapping (PCM) project are used for Defra compliance assessment reporting for EU Air Quality Directives. Defra also commissions daily air quality forecasts for dissemination to the general public as a contribution to public health protection. These different tasks will need careful assessment in the sections of the report below so that each model is evaluated on the correct basis.

5. Because each of these different tasks place different requirements on the modelling tools, air quality model evaluation cannot be carried out without reference to the purpose for which the model has been built and the requirements for accuracy and timeliness placed on the model results by Defra. In recognition of the different uses of air quality models and the different requirements of Defra, the model intercomparison protocol developed in this study will need to be first differentiated between the major air quality policy areas, such as: ground-level ozone, acid deposition and eutrophication and urban air quality. With each policy area model,

evaluations must take into account the context in which each model has been developed and the purposes for which the output is used by Defra.

6. It has not been possible in this first ‘strawman’ protocol to address all of Defra model applications covering all pollutants and all scales. We have not addressed the modelling of persistent organic pollutants (POPs), heavy metals and mercury, for example. At the global scale, we have not covered the modelling of the tropospheric ozone background and at the microscale, the dispersion of motor vehicle emissions. Equally well, no attention has been given to the modelling of pollutant effects on vegetation nor the modelling of human population exposure to air pollutants.

1.2 Layout of the Report

7. There are a number of issues that run as a common theme through air quality model evaluation that are independent of the major air quality policy areas listed in the paragraph above. These issues include the importance of atmospheric emissions and meteorological data and so these are addressed in Section 2 below. The important theme of the comparison of model results with observations which underlies most model intercomparison activities is discussed in Section 3. Section 4 deals with statistical and other metrics employed in the evaluation of model performance. Section 5 presents the model evaluation protocol which addresses the main air quality policy areas and provides detailed advice and guidance for air quality model evaluation.

8. Annexes provide more detailed discussions of the background issues relevant to air quality model evaluation. Annex A deals with the selection of monitoring network data, Annex B with the choice of chemical mechanisms, Annex C with deposition parameterisations and Annex D with emissions data.

9. Protocols are provided for ground-level ozone, acidification and eutrophication and urban air quality modelling and pose three general questions:

- Is the scientific formulation of the model broadly accepted and does it use state-of-the-art process descriptions? This is the scientific evaluation question;
- Does the model replicate observations? This is the operational evaluation question; and
- Is the model suitable for answering policy questions and fulfilling its designated tasks? This is the diagnostic evaluation question.

10. In the paragraphs below, the three steps of model evaluation, scientific, operational and diagnostic, have been addressed in some detail. There is a fourth step: probabilistic evaluation, which aims to capture the uncertainty or level of confidence in air quality model predictions (Dennis et al., 2009). Probabilistic evaluation is particularly relevant where air quality models are employed in support of air quality policy formulation and assessment. It requires the use of multiple model runs of the same model with different input parameter or process choices or multiple runs of different models to characterise the uncertain distribution in the model predictions, such as the impact of an emission policy scenario on air quality. Probabilistic evaluation is beyond the scope of this present study.

11. The protocol described in Section 5 does not itself answer the questions concerning whether the models used by Defra are state-of-the-science tools and are fit-for-purpose but exists merely to elicit information from the modelling teams to allow Defra to form a view on these issues.

2 Background to Air Quality Modelling

2.1 Purpose of Air Quality Modelling

12. In each of the policy areas considered here: ground-level ozone, acidification and eutrophication and urban air quality, models have the same underlying purpose. They exist to convert emission inventories into air concentrations and depositions. In the main, air quality policy-makers formulate strategies and policies to reduce emissions and models are used to predict their consequences in terms of the changes in air concentrations and depositions. Often, there is some form of air quality target, guideline or standard or some form of deposition target such as a critical deposition load that can be used to evaluate the strategy based on air quality model results. Strategies may not necessarily be judged as pass or fail, meeting critical loads and levels or not, for example, but may be evaluated side-by-side with other strategies or against a do-nothing scenario. Increasingly, policy-makers are using cost-benefit analyses in which the costs of the abatement strategy may be set against the benefits of the environmental improvement. In essence then, the air quality models are employed to convert atmospheric emissions into air concentrations or depositions.

2.2 Horses-for Courses

13. The complexity of the air quality models offered by Defra’s contractors varies considerably from one policy area to another. Some urban air quality models may treat air pollutants as inert species whereas ground-level ozone models necessarily have to treat quantitatively the atmospheric chemistry of a large number of organic compounds. Models for PM mass closure necessarily have to treat both primary and secondary PM components equally accurately and without bias. Ground-level ozone models have to treat summertime photochemical episodes whereas acid deposition models must be able to treat the wet scavenging of acidic pollutants during episodes of rain. Initially, this complexity has been addressed by Defra and its contractors on a single issue, ‘horses-for-courses’ basis, with models such as OSRM, for example, focussed entirely on ground-level ozone and FRAME focussed on acid deposition and eutrophication. More recently, the modelling community has give attention to the development of ‘one-atmosphere’ models, capable of addressing multiple issues, including many of the policy areas of concern to Defra. There are considerable long term advantages to Defra in switching from single issue to multiple issue modelling. Without some form of model evaluation protocol, it is difficult to elucidate and highlight these long-term advantages.

14. The degree of complexity will also depend on the number of simulations required for a particular application. A complex model will be required to assess the importance of a particular chemical reaction or meteorological process. However source-receptor studies used to correlate pollutant deposition to specific emissions sources can involve hundreds of model simulations whilst uncertainty studies typically require thousands of model simulations. For such applications, simpler modelling approaches with low simulation times are necessary. In addition, policy makers have to face the difficult challenge of managing finite financial resources and, inevitably, the development of complex models requires a higher level of funding.

2.3 Elements of Air Quality Models

15. The main elements of air quality models are designed to address:

- The emission of primary pollutants, including their spatial and temporal distributions;
- The dispersion of primary pollutants in plumes away from stacks, roadways, industrial premises and so on, with local influences from street canyons, buildings and topography;
- The advection and dispersion of primary pollutants into the downwind environment;
- The removal of primary pollutants by chemical reactions, deposition to underlying surfaces, scavenging by particles, cloud and rain droplets;
- The formation of secondary pollutants by chemical reactions;
- The removal of secondary pollutants by chemical reactions and their deposition to the underlying surface, scavenging by particles, cloud and rain droplets;
- The in-flow and out-flow of pollutants from and to local, regional, transboundary and intercontinental sources, and
- The exchange of pollutants with the free troposphere.

16. Each air quality model may contain some or all of the above elements and acts as a processing tool that converts an emission inventory into air quality concentrations and depositions. The policy strategy would generate a change in the emission inventory which could then be processed into a change in air concentrations and depositions.

2.4 Handling Complexity

17. Studies of air quality from the viewpoint of scientific understanding have demonstrated much more complexity than has been possible to represent in mathematical form in a model that could be used to make quantitative predictions. But a quantitative, mathematical description is usually held to be essential for making policy decisions based on the ‘best available’ science. In moving from the full complexity of real-world environmental systems to a tractable mathematical model requires the introduction of simplifying assumptions. Complexity is present in all the above elements of air quality models and simplifications are essential if these elements are to be represented in practical mathematical models. The complexities that are omitted from air quality models are usually forgotten in most model applications. This is fine if previous work has shown that they have a negligible effect on the system. More often, gross simplifications have been made because we lack the required understanding or adequate measurement techniques or because we are unable to identify the nature of the processes actually occurring in the real-world on the scales required for our models.

18. Models contain assumptions and simplifications that are known to contain problems. The assumptions and simplifications are nevertheless not arbitrary. Some may reflect physical intuition, they may be purpose-specific, they may have explanatory power or they may follow on from previous successful model approaches.

19. Models are always incomplete and efforts to make them more complete can cause problems (National Research Council, 2007). Increasing the complexity of models can introduce more parameters with uncertain values, decreasing transparency and increasing overall uncertainty. It is sometimes preferable to omit capabilities that do not improve model performance substantially. Some complex models are characterised by substantial uncertainties because they contain more parameters than can be reliably estimated with the available observations. Nevertheless, adding complexity can imply replacing arbitrary parameters with those that can be more closely tied to measurable processes.

2.5 Model Calibration and Scaling

20. Parameters are quantities that represent the intrinsic characteristics of the environment being modelled and are specified by the model user so that the main elements of the air quality model are set up accurately. Parameters generally belong to one of two types. There are those that are intended to reflect the specific characteristics of the processes and their dynamics. There are other parameters that reflect the characteristics of the locations where the models are applied. The distinction is sometimes blurred and some process-related parameters are used to adapt a model to a specific location. Some parameters have to be calibrated in some way for each application of the model.

21. Calibration and scaling is a real problem in air quality modelling. If air quality models could be defined in such a way that all model parameters were geophysical constants then there would be no need for model calibration. It is not possible to measure all the required parameters during a field campaign to minimise this problem because of the limitations of current measurement methods. Model parameters are usually calibrated on the basis of limited measurements, by extrapolation from applications at other sites or by inference from a comparison of model outputs at the location of interest. All these calibration methods have their problems.

22. Inference of parameter values by comparison of observed and predicted responses is generally carried out within an optimisation framework in which ‘best fit’ parameters are determined. The difficulty of finding optimum parameter values increases with the complexity of the model. As computer power has increased, it has been possible to add more complexity to models. More and more process understanding has been built into models with the aim of improving the science, at the cost of adding more and more parameters. The available observation base may not have improved at a commensurate rate, leaving some complex models with sufficient degrees of freedom to give a good fit to the observations. It does not necessarily follow that the parameters are robustly estimated or that the apparent optimised model is the only model that will give a good fit to the observations.

23. For most air quality models, the steps taken by model developers to calibrate and tune model parameters are lost in the mists of time. This is often the case where models have been brought in from abroad or where there is a substantial proprietary element or ownership or where the model is considered the intellectual property of a particular institute or institution. All air quality models contain important elements of calibration and tuning against observations. Without access to adequate records of tuning and calibration, models may be incorrectly applied or model developers may be unable to adapt the model for a new application. Understanding the impact of calibration and tuning is central to the process of evaluating model performance.

24. Defra’s contractors naturally want to improve their representations of real-world complexity but this can easily result in models with more parameters than can easily be specified for a given location. Parameters may represent real-world behaviour over some particular time period but can they be transferred from one model to another or from one time period to another without adjustment or from one location to another?

25. In some cases, particularly for local and urban scale modelling of long-term average concentrations, there are insufficient data to recalibrate model input parameters. Modellers then often scale the model outputs using lumped scaling parameters determined, for example, by regression analysis of the modelled and measured concentrations. The method is widely used for Local Authority Review and Assessment (Defra, 2009) and in the preparation of national maps of pollutant concentrations for reporting to the European Commission (Kent et al., 2007) and for auditing the contribution made by sources regulated by the Environment Agency. However, it should be recognised that the scaling factors are characteristic only of the data set used in their derivation and they may not be applicable to other locations or other time periods.

2.6 Best or Optimal Models

26. The idea of searching for a single, optimal set of parameters to drive an air quality model is central to the concept of research leading to a realistic description of real-world behaviour (Beven and Freer, 2001). It is rarely considered that such an optimal description may not be possible or uniquely identifiable. To reject the concept of an optimal model or model parameter set in favour of multiple possible or plausible parameter sets is called equifinality (Beven, 1993). One implication of rejecting the concept of an optimal parameter set and accepting the concept of equifinality, is that the uncertainty associated with the use of models as predictive tools might be wider than hitherto thought. If there are many different acceptable parameter sets, all of which are consistent with the calibration data, the range of model results is likely to be greater than might be suggested by the optimal model. Model tuning and calibration may serve to hide uncertainty and lull the policy-maker into a false sense of security about the adequacy and robustness of model predictions.

27. Intuitively, it is recognised that the description of some of the above major elements of air quality models may be far superior in some models and parameterisations compared with others. The logical consequence of this would be the concept of a ‘super-model’ where all the superior elements were linked together in one model application. Setting aside the issues of commercial propriety and computer capacity, there are some real-world limitations to the concept of such a ‘super-model’ in the above context.

2.7 Real-World Limitations

28. These real-world limitations arise because each parameterisation, whether elegantly superior or manifestly simple, has to work in concert with all the model elements to generate an air quality prediction. As an example, when it comes to the comparison of atmospheric dispersion model and observed 15-minute SO_2 concentrations from the Defra AURN network, model performance may sometimes appear to be poor if 15-minute SO_2 emissions data are not routinely available for all important sources of SO_2. In reality, rarely are such high time- and spatially-resolved emission data available and the apparent performance of an atmospheric dispersion model may often seem less than optimal.

29. There are many circumstances when the performance of the inherently superior elements of air quality models are compromised. The PTM model predictions of ground-level ozone concentrations along 5-day trajectories are remarkably similar whether calculated using the Master Chemical Mechanism with all its thousands of chemical reactions and the Carbon Bond Mechanism with its inherent simplicity. Simple chemical mechanisms can be highly tuned and calibrated against laboratory data so that they deliver good model performance for ozone. All other model species may differ widely between the two chemical mechanisms. Indeed, even the model ozone responses to NO_x and VOC precursor emission changes also may differ widely. In this case, tuning and calibration has served to hide uncertainty and to compromise the adequacy and reliability of model predictions of ozone responses to NO_x and VOC controls. These are precisely the model results that are important to policy-makers and they need to be as robust as possible. Comparison of model results against ozone observations would not necessarily demonstrate a lack of model robustness and would not allow an optimal choice between the two chemical mechanisms.

30. Of all the major elements that are assembled into an air quality model, emissions modules and inventories have been the source of most debate and controversy. Rarely has assessment of model performance focussed attention on issues other than problems and inadequacies with emissions data. There is a tendency to see the problems with emissions data as outside issues and not part of the model framework and formulation. This tendency has been reinforced as the task of the assessment of emissions has been taken away from modelling groups and given to independent agencies or institutions. With the policy focus now on emission inventories, inventory development is seen as a task in itself. In an attempt to avoid duplication of effort,

modelling groups are encouraged to see emission inventory data as given, just as if they were geophysical quantities. They are far from this ideal, in reality.

31. If there were systematic biases in emissions inventory data, leading to the over- or under-estimation of particular pollution sources, then tuning and calibration could lead to distortion and the favouring of parameter sets that in some way compensate and reduce the bias in the model outputs that would otherwise have been expected. It is straightforward to see how this may lead to the situation where the model is correct for the wrong reason due to the compensation of errors. Underestimation of the emissions from one source may have been compensated by adjustments to the deposition and loss processes for that pollutant, for example. Also, underestimation of the emissions of highly reactive ozone precursor species could be compensated by a small increase in ozone productivity across a large number of less-reactive ozone precursors. Sometimes, there is a specific policy requirement from Defra to use a given emission inventory despite any reservations that there may be concerning its adequacy and fitness-for-purpose.

32. Presently, there is little opportunity to check the reliability of emissions data without resorting to air quality modelling. This implies that our view of air quality model performance is distorted by our view of the adequacy of current emissions data. We could take concerns about emissions out of the equation by performing a model evaluation in which identical emissions data are supplied to a number of models and then comparing the predicted values from the range of models. If there were enough models, then the mean of the model results would imply some form of ‘best’ result and the distance of a given model from the mean could be taken as a measure of model performance. But how far is the ‘best’ model from reality? Without a comparison of model results with observations, this important question is difficult to answer. Model comparisons using harmonised emissions data have the tendency, over the long term, to reduce the spread in model results through the harmonisation of parameterisations and the judicious selection of common input data. Such model intercomparisons may hide uncertainty and lull policy-makers into a false sense of security.

33. So we must take emissions data for what they are, uncertain model input parameters that are full of short-comings, inadequacies and pitfalls, just as all the other model input parameters. Much attention has been given by policy-makers to the continual improvement of the reliability of national emission inventory data and the modelling community has benefitted from this through the years. Rarely, if ever, do the modelling studies carried out by Defra contractors use these national emissions annual totals. Pollution episodes may involve particular emission sources and for short periods of time, far removed from the concept of a national source strength over the annual time-frame. In general terms, air quality models require emissions from individual, specific sources such as road links, power stations and large industrial complexes and with high time resolution, say at least 15-minutes or so, in some situations. This same level of information is required for the entire UK and for all important upwind sources. Emission inventory data at these levels of spatial and temporal resolution are not always available to Defra and its contractors and will not be available in the foreseeable future, although it has to be said that there are many elements in the mapped National Atmospheric Emissions Inventory (NAEI) used in models that do include local source information, for example power station reported emissions and observed traffic counts which can be used to estimate traffic emissions on specific major roads. But there are, for example, natural limits to what can be known about traffic emissions because of the influence of random events and fluctuations such as crashes, traffic light failures, queues, road repairs and congestion and local deviations from other nationally averaged parameters used in the calculation of traffic emissions. Model requirements go well beyond these limits. Defra and its contractors therefore have to replace highly spatially and temporally disaggregated emission inventory data with assumptions and simplifications concerning temporal and seasonal profiles and spatial distributions that ultimately can be traced back to the national emission inventory totals.

34. Emission inventory parameters are not available with the spatial and temporal resolution required for the air quality models in all situations. Model performance is always compromised by the adequacy of emissions data but the level of compromise is model dependent. It may be a major issue for some models but only a minor issue for others. Some of the issues associated with handling emission inventory data in air quality models are addressed in some detail in Annex D below.

35. It goes without saying that an inadequate treatment of any of the major elements of an air quality model can exert a large impact on air quality model predictions. Errors in meteorology are a major source of error in air quality models. If the model wind does not advect air pollutants correctly from pollution sources to air quality monitoring network stations, then the model will not predict the observed pollution episodes, irrespective of how well specified are other features of the air quality modelling system. Deposition episodes during rain events are even more difficult to model accurately because not only must wind speeds and directions be accurately predicted but also so must the occurrence and magnitude of precipitation.

36. The main meteorological parameters whose errors affect the accuracy of air quality model predictions following Seaman (2007) and WMO (2008) are as follows:

- Boundary layer depth;
- Surface and boundary layer wind speed and direction;
- Surface and boundary layer temperatures;
- Turbulence;
- Surface parameters;
- Cloud cover and solar radiation;
- Cloud microphysics, and
- Precipitation.

The winds, boundary layer depth and turbulence are critical because they control the atmospheric dispersion and transport of air pollutants. Precipitation, in addition, is critical for those air quality model applications addressing acidification and eutrophication.

37. Large amounts of meteorological data are required for air quality model applications and so air quality model development has gone hand-in-hand with the development of meteorological models. Historically, air quality models used climatological data but over the years, following the steadily increasing model complexity, stronger links have been forged between air quality and numerical weather prediction models. Much current attention is now being given to building coupled air quality and numerical weather prediction models, providing a seamless interface between the two sets of coding, obviating the need to transfer large amounts of data between modelling systems.

38. Meteorological institutes and national weather services have always focussed their greatest attention on the accuracy of their predicted wind, temperature, pressure, humidity and precipitation data in their routine numerical weather prediction models (Seaman, 2007). These variables are readily measured by the meteorological observation networks and there has been a steady improvement in forecast accuracy over the years. However, in comparison, the accuracy of numerical weather prediction estimates of boundary layer depth and cloud cover have made little improvement. The lack of accurate observations of boundary layer depth and cloud properties has hampered the improvement and enhancement of meteorological model parameterisations for these critical input parameters to air quality models. There are also significant problems associated with the generation of high spatial- and temporal-resolution rainfall fields that are required for air quality model applications addressing acidification and eutrophication.

39. In addition to the use of meteorological data to drive meteorology and physics parameterisations, a considerable amount of such data are required for emission and deposition algorithms. For example, emissions of isoprene from natural biogenic sources require a great deal of meteorological data concerning plant canopy temperatures and radiation, see Annex D. Furthermore, deposition algorithms (Annex C) require information on turbulence and atmospheric structure close to the ground.

2.8 Air Quality Forecasting Models

40. An important application of air quality models is the forecasting of future air quality and the preparation of air quality bulletins on behalf of Defra. Air quality bulletins and warnings are issued when air quality is forecasted to deteriorate either under wintertime or summertime episode conditions. The former episodes are usually driven by NO_2 or PM and the latter by O_3. The model intercomparison protocol for air quality forecasting models therefore follows exactly the protocols developed for the O_3, NO_2 and PM models developed in Section 5 below. However, because the general public may modify their behaviour in response to the air quality bulletins and warnings, stricter critieria should be set in terms of air quality forecasting model performance compared with the above models. Acceptable air quality forecast model performance requires that the prediction of false episodes and the underprediction of real episodes are both minimal. Further discussion of evaluation procedures and metrics for air quality forecast models are given in Agnew et al. (2007).

3 Comparisons of Models Against Observations

41. A policy-maker will expect that any model used in support of policy formulation is able to reproduce real-world behaviour. This might imply reproducing the observed diurnal or seasonal cycles in pollutant concentrations or spatial contrasts implied by roadside or urban concentration increments or contrasts driven by meteorological conditions such the concentration differences between episode and non-episode days or spatial gradients between the south and east versus the north and west of the British Isles. Whilst this is a reasonable expectation, there are a number of outstanding issues and questions. How is this comparison to be carried out? How is good agreement in one part of the model to be set against poor agreement elsewhere? How accurate are the observations and what volume of the environment do monitoring site observations represent? In view of the importance given to the comparison of models with observations in most model intercomparison exercises, these are important questions in the present context. A survey of the available UK monitoring network data is given in Annex A below for use in these model intercomparison activities. This Annex also includes a discussion of the data quality that can be expected from each monitoring network.

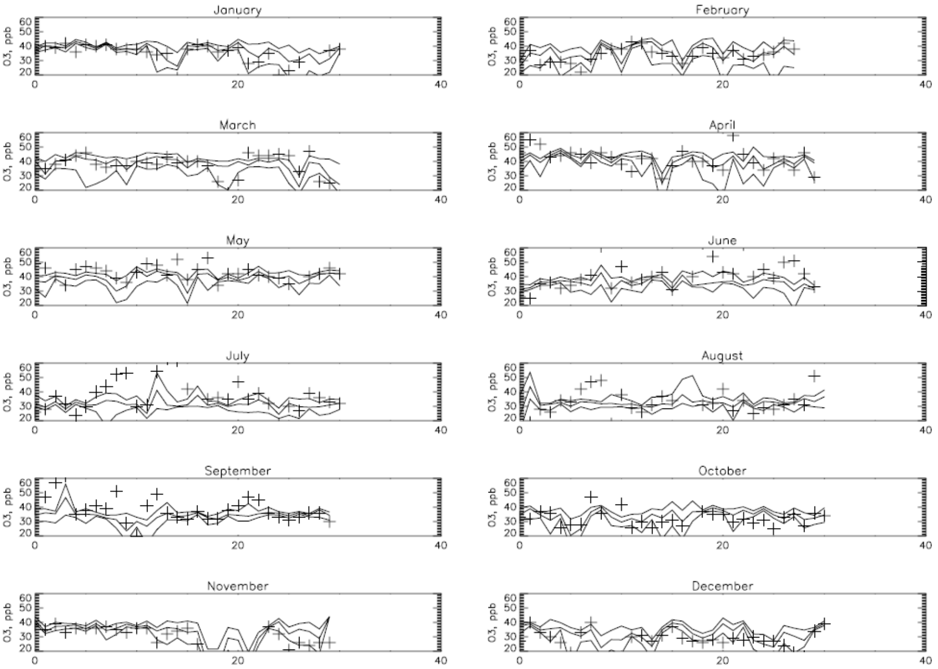
3.1 What is Involved in the Comparison of Models with Observations?

42. Before these questions can be answered, consideration needs to be given to the process of comparison of model results with observations. In the air quality model applications carried out by Defra and its contractors, these comparisons are usually made against monitoring network data and field campaign data obtained at sometime in the past. For this reason, comparisons with observations are sometimes called ‘history matching’. Model evaluation is then a description of how well the ‘lines’ of the model pass through the ‘crosses of the historical observations, see Figure 3.1 for example.

43. History matching is an inherently difficult process (Oreskes et al., 1994; Beck, 2002). Rarely will all the relevant information be available with which to set up a model optimally so that its intrinsic level of performance can be assessed. Inevitably, gaining access to sufficiently accurate historic emissions and meteorological data presents real problems. Ideally, a first set of past observations is used to calibrate and tune the model. Then the calibrated model, with no further adjustment, is tested against the second, independent set of past observations which it should match acceptably. However, often these calibration and tuning steps were carried out elsewhere or by other institutes and any records may have been lost over time. Most model intercomparison exercises therefore omit the first stage of tuning and calibration and perform the second stage only, for example, see Van Loon et al., (2007). An agreement will be made between the modelling groups concerning which pollution episodes, campaigns or time periods from a month to a year are to be studied and efforts made to compile and harmonise historic emissions data and to collect observations from network databases. If some model results appear manifestly wrong, then often the opportunity will be given for models to be rerun, leading ultimately to an improvement in the overall comparison with observations.

3.2 Model Verification and Validation

44. In describing the process of history matching and what has been learnt from it in the above paragraphs, the term evaluation has been used in the context of judging model performance. This term evaluation and the phrase evaluating model performance have been used as alternative to the widely used terms of verification and validation. The differences between evaluation and verification or validation are important (Oreskes et al., 1994), particularly in the context of air quality modelling. To say that a model is verified is to say that its truth has been demonstrated which implies its reliability as a basis for policy and decision making. The term validation is commonly used interchangeably with verification, indicating that model predictions are consistent with observations. Validation is also used to suggest that a model is an accurate representation of physical reality. Policy-makers will look to see if models have been verified or validated.



January
February
March
April
May
June
July
August
September
October
November
December

Figure 3.1. Evaluation of the PTM model for ozone at Harwell, Oxfordshire for 2005 showing the 'crosses' of the AURN observations and the 'lines' of the model. The observations are the daily maximum 1-hour concentrations. The lines show the estimated model 50-%ile (middle line), 95-%ile (upper line) and 5-%ile (lower line) concentrations.

45. The term validation is widely used to describe evaluation of model performance against measured data. Oreskes et al. (1994) argue that validation and verification of models of natural environmental systems is not possible. This is because there are always input parameters that are poorly known, fine scale details of crucial importance which are inadequately understood, and assumptions and inferences which may not be valid under all circumstances. Such views apply to the complete validation of models, applicable to all possible circumstances and scenarios. In practice, the term validation is frequently used in a more restricted sense to describe the evaluation of model predictions against measured data, which might more appropriately be described as partial, or conditional validation, as it applies only to a limited range of input parameters and, hence, outputs. The distinction between complete and partial model validation is an important one, as a model giving excellent predictions from within a limited range of input parameters may perform poorly when used beyond that range of inputs.

46. One of the main objectives of the COST Action 732 has been the determination and improvement of model quality for the application of micro-scale meteorological models to the prediction of flow and dispersion processes in urban or industrial environments (Britter and Schatzmann, 2007). With their focus on short-range dispersion modelling, they lay out an evaluation protocol that has several distinct elements: a scientific evaluation process, a verification process applied to the model code, the provision of field and physical modelling data that allow model validation and an operational evaluation process that reflects the needs and responsibilities of the model users. Their usages of the terms verification and validation are consistent with the processes that they describe, particularly those involved with model validation. In terms of the distinctions drawn in the paragraph above between complete and partial model validation, the COST 732 evaluation process specifically is aimed at complete validation. Whilst the language of the COST 732 evaluation protocol is not appropriate in the present study because the detailed field and physical modelling data are not available, the processes envisaged are definitely relevant and have strongly influenced the approach adopted here in the paragraphs below.

47. Local Authorities are required to review and assess air quality in their areas from time to time and to develop plans for local air quality management in order to achieve air quality objectives in their areas. Technical guidance is provided to help and support Local Authorities whilst conducting their reviews and assessments, particularly on the application and implementation of urban air quality models (Defra, 2009). This Technical Guidance defines model validation as the process by which the original model developers compare model results with observations. Because these validation studies are unlikely to have been undertaken in the Local Authority area, further comparisons are required of the model results versus observations at relevant locations. Model verification is seen as the process by which any differences between model and observation are investigated and where possible minimised. In some cases, the Technical Guidance requires that appropriate adjustment factors are applied to ensure that model predictions are representative of monitoring information from the Local Authority area. Linear regresssion, for example, may be used to generate a functional relationship of the form: $y = mx + c$, which allows the tuning of the model predictions, y, to fit the observations, x. The intercept, c, might, for example, adjust for regional and urban background contributions that may be inadequately represented in the model or unavailable from observations. The slope parameter, m, may represent the influence of meteorology or local building influences on atmospheric dispersion or local site emissions which differ from the national picture. Once the urban air quality model is tuned in this way, it is acceptable for use in the Local Authority review and assessment process.

48. There are therefore clear differences in the language and modelling procedures between those adopted in the Local Authority Technical Guidance documentation (Defra, 2009) and those developed in this report. There is no straightforward way to reconcile these differences. The evaluation protocol described below is therefore not appropriate for the evaluation of the urban air quality modelling carried out by Local Authorities. These activities are considered to be outside of the scope of this report as laid out in the terms of reference for this study by Defra and descibed in the Introduction above. Generally speaking, urban air quality models within the Local Authority review and assessment regime are used by individual Local Authorities to assess air quality at a few specific locations. They are thus not supported by enough observations to allow both the required local tuning and the more general evaluation of model performance. This is not to say that these urban air quality models have not been evaluated, far

from it, merely that they have not been evaluated for the individual application within the Local Authority review and assessment regime. By way of example, Righi et al., (2009) provide a statistical and diagnostic evaluation of the ADMS-Urban model compared with carbon monoxide observations from an urban air quality monitoring network in Ravenna, Italy. In completing their evaluation of the performance of ADMS-Urban, Righi et al. (2009) use the same statistical metrics and evaluation procedures as those adopted below and do not follow the procedures and language of the Local Authority Technical Guidance (Defra, 2009).

49. In reviewing the regional-scale photochemical air quality modelling systems used in North America by the United States Environmental Protection Agency, Dennis et al., (2009) begin by agreeing with Oreskes et al., (1994) that these models cannot be validated in the sense of being proved true, since truth is in principle inaccessible to us. They assert that these models have both predictive and diagnostic value and that this value must be demonstrated through model evaluation exercises. Russell and Dennis (2000) define model evaluation as the assessment of the adequacy and correctness of the science represented in the model through comparison against empirical data, laboratory and in situ tests and the analysis of natural analogs. They see model evaluation as a process of model confirmation relative to current understanding. Multiple confirmatory evaluations can never demonstrate the veracity of a large photochemical modelling system as confirmation is a matter of degree. However, evaluation can raise doubts about the science in a model. These views have strongly influenced the approach to model evaluation adopted in this report.

3.3 Model Evaluation

50. Assuming that the model code faithfully represents the model specification there are three general questions to be answered when evaluating environmental or mathematical models, which can broadly be expressed as follows:

- is the scientific formulation of the model broadly accepted and does it use state-of-the-art process descriptions? This is the scientific evaluation step.
- does the model replicate observations adequately? This is the operational evaluation step.
- does the model reflect the needs and responsibilities of the model user and is it suitable for answering policy questions and fulfilling its designated tasks? This is the diagnostic evaluation step.

These questions lie at the heart of the protocols proposed here in Section 5 for the evaluation of the performance of the air quality models used by Defra and its contractors for policy formulation and assessment.

4 Metrics for Evaluating Model Performance

51. Comparisons against observations give an invaluable guide to overall model performance. Understandably, policy-makers expect modellers to establish the trustworthiness of their models. For air quality models, this almost always involves some form of comparison of model predictions against measured concentrations. However, the ability of an air quality model to reproduce measured concentrations from the past does not guarantee its adequacy for the future or for predicting the response to pollution control strategies. Agreement with observations is inherently partial. Models agree with some observations but not all. A model can certainly perform well against historic observations and the precision and accuracy of the fit can be quantified. The performance of models can be evaluated relative to past observations, relative to other models or against our own theoretical expectations, but the performance of a model, especially for future projections of concentrations, cannot be ascertained precisely. Nevertheless, the comparison of model predictions against past observations is a good first step in the evaluation of model performance.

52. A review of the air quality modelling literature yields a plethora of metrics that have been used and recommended to quantify the differences between model predictions and observations and so to support model evaluation (Chang and Hanna, 2004; Yu et al., 2006; US EPA, 2007). Irwin et al., (2008) discuss these metrics and propose other new unbiassed metrics as replacements for the traditional evaluation metrics. Metrics appropriate to the evaluation of regional air quality models have been jointly discussed by the COST Action 728 and the WMO-GURME programmes (WMO, 2008). Further metrics have been proposed for the evaluation of air quality forecast models by Agnew et al., (2007). Operational evaluations of different air quality models in the past have yielded an array of statistical metrics that are so diverse and numerous that it is difficult to judge the overall performance of the models.

53. Here, we recommend some form of simplification and rationalisation by adopting a limited number of statistical metrics so that there is a large degree of harmonisation between the evaluation approaches adopted by the modelling teams. Each of these metrics assumes the existence of a number, N, of pairs of model, M_i, and observed, O_i, concentrations. The index, i, might run over a time series at a given location encompassing a pollution episode, field campaign or a convenient time period from a week, say, to a year. Equally well, the index, i, might run over the same time period for all locations within a monitoring network. The pairing of the model results and observations implies that each pair covers the same time period and the same averaging time. The choice of pollutants, locations and time periods is left entirely to the modelling teams.

54. It is convenient to start a model evaluation with a comparison of the model predictions and observations as a simple time series for a single location. This shows straightforwardly whether the major observed pollution events have been predicted or not and shows up any false episodes that were predicted but not observed. The bias is calculated as the model result with the observation subtracted and the error as the absolute value of the model result with the observation subtracted. Figure 4.1 shows an example of bias and error for each day of 2005 for Harwell, Oxfordshire from the PTM model. Such time series plots show whether poor model performance is concentrated on particular days and seasons or whether it is distributed throughout the year. In the particular case of Figure 4.1, poor PTM performance for ozone is noticeably concentrated during the summer months from June to September where the lines diverge markedly indicating model underestimation of the episodic ozone peaks.

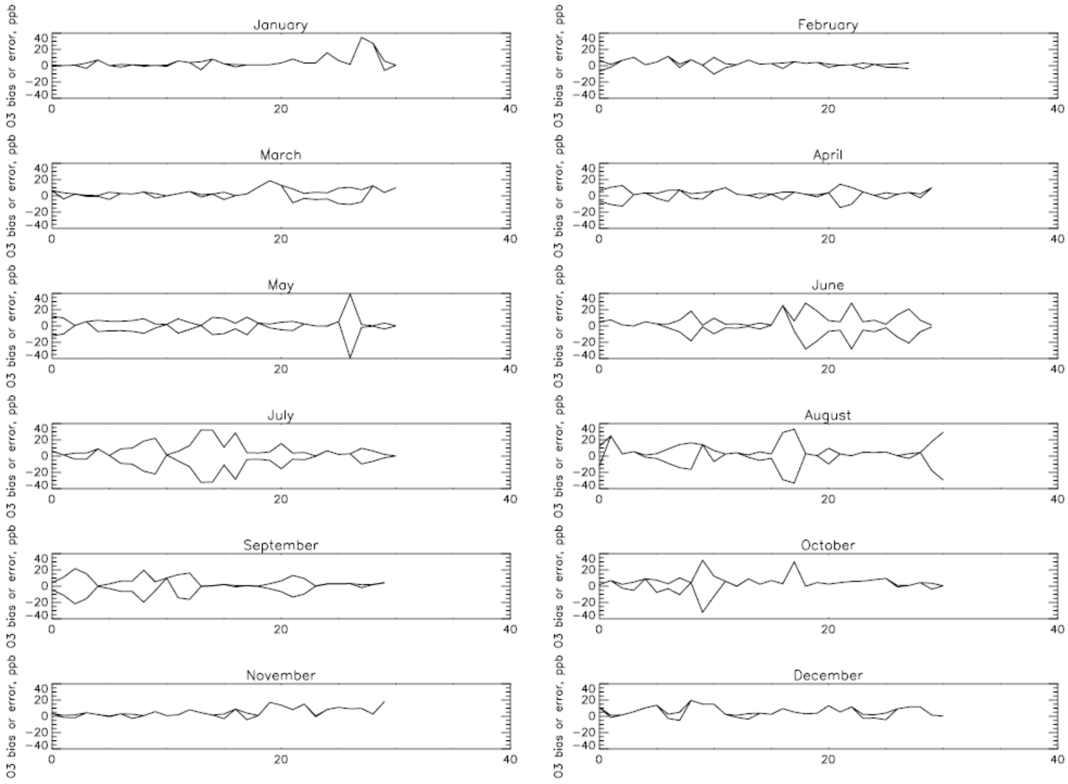


Figure 4.1. Bias and error for the PTM model for ozone for each day of each month of 2005 for Harwell, Oxfordshire. The upper line shows the bias and the lower line the error in the model predictions.

55. A scatter plot loses the time dimension but allows a quantitative comparison between model predictions and observations. It is helpful to overplot the scatter plot with the 1:1 correspondence line, together with the 1:2 and 1:½ lines. Such an example plot is shown in Figure 4.2 for the mid-afternoon ozone levels at Harwell, Oxfordshire for 2005 observed and predicted by the PTM model. A count of the fraction of points within ½ and 2 times the observations, FAC2, is a useful evaluation metric.

FAC2 is defined as the fraction of model predictions that satisfy:

$0.5 \leq M_i/O_i \leq 2.0$

It is recommended that an air quality model is considered acceptable if more than half of the model predictions lie within a factor of 2 of the observations and faulty if not. There is no justification available in the literature to underpin such a recommendation. It is, however, based on many years modelling experience over a wide range of pollutants of policy relevance. As formulated, it is independent of pollutant and averaging period. It represents the minimum acceptable level of model performance to be achieved for policy application. Further model development may improve performance relative to this metric.

56. Because correlation coefficients, R^2, of the scatter plots vary significantly with the quality of the model predictions, the presence of outliers and the number of paired points in the comparison, they are not recommended as evaluation metrics. Taylor (2001) proposed a graphical method based on correlation coefficients as a convenient means for summarising multiple aspects of model performance. However, they have not seen widespread adoption because of the influence of outliers when there is a small number of paired points.

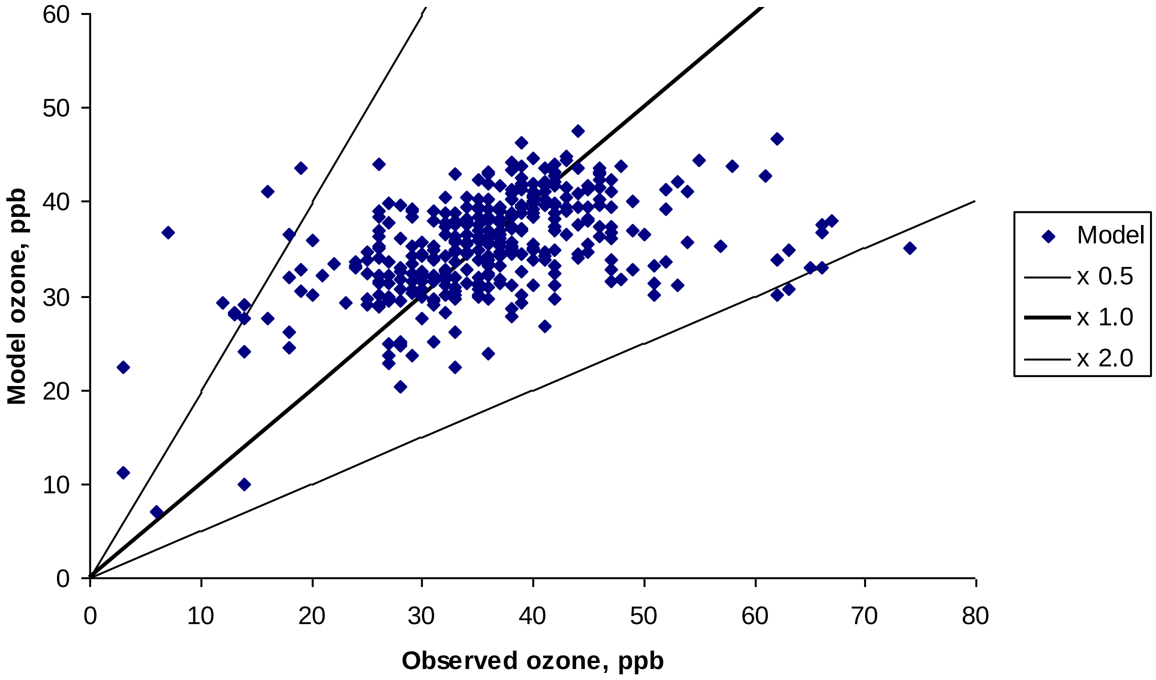


Figure 4.2 Scatter plot of PTM model vs observed mid-afternoon ozone levels for Harwell, Oxfordshire for 2005, showing the 1.0 to 0.5, 1.0 to 1 and 1 to 2.0 lines.

57. The two most commonly used metrics to quantify the departure between modelled and observed concentrations are the mean bias, MB, and the mean gross error, MGE (US EPA, 2007).They differ only in whether the sign of the difference is taken into account, allowing under-estimations and over-estimations to cancel or not. Mean bias and mean gross error are defined as follows:

$$MB = 1/N \sum_{i=1}^{N} M_i - O_i$$

$$MGE = 1/N \sum_{i=1}^{N} | M_i - O_i |$$

Mean bias and mean gross errors are useful measures of the over- and under-estimation by the model. The units of mean bias and mean gross errors are the same as the observations and so are readily understood.

58. It is also useful to provide a measure of the relative or fractional difference between the model predictions and the observations and this is generally achieved through normalisation. Relative measures are particularly useful when model performance for one pollutant is being compared with that for another pollutant for which concentrations are generally quite different. Two commonly used measures of relative difference are the normalised mean bias, NMB, and the normalised mean gross error, NMGE (US EPA, 2007). These are defined as follows overleaf:

$$\mathrm{NMB} = \frac{\sum_{i=1}^{N} M_i - O_i}{\sum_{i=1}^{N} O_i}$$

$$\mathrm{NMGE} = \frac{\sum_{i=1}^{N} \left| M_i - O_i \right|}{\sum_{i=1}^{N} O_i}$$

It is recommended that air quality models are considered acceptable if NMB values lie within the range between -0.2 and +0.2 and faulty if not. Again, there is no justification available in the literature to underpin such a recommendation. It is, however, based on many years modelling experience over a wide range of pollutants of policy relevance. It represents the minimum acceptable level of model performance to be achieved for policy application. Further model development may improve performance relative to this metric.

59. Furthermore, we recommend that air quality modellers generate a set of evaluation metrics along the lines described above for a range of air concentrations and depositions for a standard time period, episode or campaign to benchmark model performance. Should any significant improvement or embellishment be made to the modelling system, such as model maintenance, upgrade or change to coding procedures, then the benchmarks should be rerun for the standard time period so that any improvement in model performance can be registered. Where modelling teams make changes to hard-wired or default parameters or change parameterisations, new sets of evaluation metrics should be generated and placed alongside those generated with the initial parameters and parameterisations. Annex F presents a standard model evaluation analysis template to benchmark model performance.

60. The evaluation metrics detailed above, based on the FAC2 and NMB statistics, allow the comparison of model predictions with observations and yield valuable information and commentary on model performance. If a model fails these tests then we know that the model is faulty in some way. However, the reverse is never the case. If a model passes these tests, then we cannot say that the model is verified or valid. The model may be right for the wrong reasons. There may be cancellation of errors. All we can say is that the model is merely confirmed by the observations. Matching history cannot give us the verification and validation that modellers and policy-makers may desire**.**

61. Short-range dispersion models, operating over scales from a few metres to a few tens of kilometres, have in general been evaluated by comparison with measurements made near individual point emission sources. Small errors in the assumed wind direction and atmospheric conditions can lead to large errors in the modelled concentrations at individual monitoring locations for specific time periods. There is a risk that comparison of paired measured and modelled concentrations for specific locations and time periods would lead to the rejection of otherwise satisfactory models. For example, for regulatory purposes, it is often sufficient that

the model predicts correctly the maximum ground level concentration at any place and at any time or the maximum long-term average concentration at any place. Model evaluation studies for short-range dispersion models usually condition the data in various ways. Techniques adopted include the comparison of arc-maximum concentrations and quantile-quantile comparisons (e.g. US EPA, 2003). For the arc- maximum comparison, the maximum measured concentration along an arc of monitoring stations at approximately the same distance from the source is paired with the maximum modelled concentration along the arc. For the quantile-quantile comparison, the modelled and measured concentrations are listed separately in order from largest to smallest: the largest measured and modelled concentrations are then paired, followed by the second largest, etc. Modellers should provide details of any data conditioning procedures adopted prior to the calculation of the FAC2 or NMB statistics. For further information concerning the performance evaluation of short-range atmospheric dispersion models see the COST Action 732 programme (Britter and Schatzmann, 2007).

62. A model can certainly perform well against observations and it may well give satisfactory results when evaluated using the FAC2 and NMB evaluation criteria. Judgmental terms such as excellent, good, fair and poor are useful descriptions of model performance. We can talk about the model performance for mid-afternoons compared with night-times or summer *vs* winter or for rural *vs* urban background sites, for example. We can also talk about the performance of one model against another using these criteria or whether changing one of the major elements or parameterisations of a model improves model performance. The issue is that two or more errors in model input may cancel each other out. There is no way of knowing whether this cancellation has happened or not. A faulty model may pass the history matching tests and satisfy the FAC2 and NMB criteria and thus may appear correct. However, if a model fails to meet the FAC2 and NMB criteria then we can say that its performance is in need of improvement and that it is in some way faulty.

63. These difficulties with verification and validation arise because there may well be many different combinations of model parameters that allow the model to fit the observations and to match history. Air quality models are inherently non-unique. If different model constructions fit the observations then there is no way to chose between them other than to invoke other considerations. These other considerations could include simplicity, elegance, computer run time and convenience, for example. They could include the scientific quality of the parameterisations employed as well as whether the models are fit-for-purpose.

64. Comparison with observations and history matching is an example of an operational evaluation. Operational evaluations are seen as the best place to start model evaluations but are not seen as an end in themselves. Since a major application of air quality models is the investigation of the impact of future emission scenarios, it is of interest to evaluate the ability of models to simulate the effect of emission changes on air pollutant concentrations and depositions. These evaluations are termed diagnostic evaluations (US EPA, 2007). They provide a useful means of choosing between different model constructions that meet the evaluation criteria set for operational model evaluations. Historically, much effort has been devoted to operational evaluations but the requirement to provide advice and support to policy-makers on model performance raises the importance of diagnostic evaluations relative to operational evaluations.

65. In Section 5 below, detailed discussions are provided concerning operational and diagnostic evaluations for ground-level ozone, acidification and eutrophication and urban air quality policy models, respectively.

5 Evaluation Protocol for Air Quality Models

66. There are three general questions that underpin any evaluation of air quality models and they can be broadly expressed as follows:

- Is the scientific formulation of the model broadly accepted and does it use state-of-the-art process descriptions? This is the scientific evaluation step;
- Does the model replicate observations? This is the operational evaluation step; and
- Is the model suitable for answering policy questions and fulfilling its designated tasks? This is the diagnostic evaluation step.

The protocol described below does not in itself answer these questions but exists merely to elicit information from each air quality modelling team to allow Defra to form a view on whether a particular model is fit-for-purpose or not.

5.1 Basic Information Questionnaire

67. The basic information required from the modelling teams should include:

- Names and contact information for the modelling team;
- Name, version number and release date of model, if appropriate;
- Version history and provenance;
- Output variables predicted;
- Model type: eg. Eulerian grid, Lagrangian trajectory, Gaussian plume, box model, statistical;
- Chemical mechanism: Carbon Bond-04, -05, SAPRC-99, -07, MCM, (see Annex B);
- Dry and wet deposition scheme: (see Annex C);
- Origin of emissions data: eg. EMEP, NAEI, and version of inventories used (see Annex D):
- Meteorological data: eg. NCEP, MM5, WRF, Met Office, climatologies
- Quality assurance and standards for coding;
- Computer hardware and software requirements;
- Computer run times: eg. is the model suited to the analysis of multiple policy scenarios?
- Web-site for model code.

This information need not be evaluated and is required solely for the purposes of documentation.

5.2 Scientific Evaluation

68. The scientific evaluation needs to elicit enough information so that independent reviewers can reach a consensus about the appropriateness of the content of a model and its fitness for the intended purpose.

69. The information required here for ozone models includes:

- what is the purpose of the model? For example, the estimation of hourly ozone concentrations across the British Isles and their response to NO_x and VOC precursor emission controls?
- what level of scientific treatment is given to the major elements of the model?

How are national emission totals broken down to give hourly time resolved emissions at 1 km spatial scale, for example, or on whatever are the appropriate model time and spatial scales?

How are total VOC emissions split up into individual species?

How are natural biogenic emissions handled and what is the source of the meteorological and vegetation data employed?

What is the chemical mechanism employed and what sensitivity does the model show to this choice?

How are time variations in photolysis rates taken into account?

What description and treatment is given to atmospheric dispersion, for example, is a full treatment of turbulence and diffusion given, or is it based on an eddy diffusion type parameterisation or is it neglected altogether through the use of a boundary layer averaged approach?

What is the source of the meteorological data used to provide surface and boundary layer winds, temperatures and humidities, boundary layer depths, cloud cover and radiation? What evaluation has been conducted on these input data? What is the evidence that any climatological data are appropriate for the present purposes?

What is the formulation of the ozone deposition scheme adopted and what is the source of the meteorological, vegetation and soils data employed?

What assumptions are made concerning the initial and boundary conditions for ozone and other pollutants?

70. The information required here for models addressing the deposition of acidic and eutrophying agents includes:

- what is the purpose of the model? For example, the estimation of annual mean concentrations and depositions of acidic and eutrophying species across the British Isles and their response to SO_2, NO_x and NH_3 emission controls?
- what level of scientific treatment is given to the major elements of the model?

How are national emission totals broken down to give time resolved emissions at 1 km scale, for example, or on whatever the appropriate model time and spatial scales? How are the annual emissions of ammonia split into hourly intervals, taking into account agricultural practices and meteorological data? What is the evidence for any assumptions concerning the speciation of sulphur emissions into SO_2 and H_2SO_4?

What is the choice of chemical mechanism employed and if a simple transformation coefficient approach has been adopted, how are the choices justified for the different coefficients and what sensitivity does the model show to these choices? How are background and initial concentrations set for the acidic and eutrophying species? What data are used to fix the concentrations of oxidants such as ozone, OH and H_2O_2, if they are not calculated within the chemical mechanism itself from emissions and background concentrations?

What description and treatment is given to atmospheric dispersion, for example, is a full treatment of turbulence and diffusion given, or is it based on an eddy diffusion type parameterisation or is it neglected altogether through the use of a boundary layer averaged approach?

What is the source of the meteorological data used to provide surface and boundary layer winds, temperatures and humidities, boundary layer depths, cloud cover and radiation and precipitation? What evaluation has been conducted on this input data? What is the evidence that any climatological data are appropriate for the present purposes?

What is the formulation of the dry deposition scheme and what is the source of the meteorological and surface data employed? Does the deposition of ammonia take into account the bidirectional nature of the surface exchange processes?

How are the scavenging processes described for soluble trace gases and aerosol particles? How are the rates of these processes linked to the meteorological data? What is the evidence that any climatological data such as ‘scavenging ratios’ or ‘constant drizzle scavenging rate coefficients’ are appropriate?

71. The information required here for urban NO_2 air quality models includes:

- What is the purpose of the model? For example, the estimation of hourly, daily and annual NO_2 concentrations across the UK and their response to emission controls; the estimation of annual average NO_2 concentrations within specific urban areas and their response to local emission controls; and
- What level of scientific treatment is given to the major elements of the model?

How are national emission totals broken down to give emissions at the appropriate spatial and temporal resolution?

Which emission sources are modelled explicitly and which are treated as background sources?

How are primary NO_x emissions split up into individual species: NO, NO_2 and HONO?

What is the chemical mechanism employed and what sensitivity does the model show to this choice? If oxidant partitioning schemes are used, how have they been derived, tuned and calibrated? What has been assumed about background ozone and NO_x concentrations and which sites and networks have been selected to set these values?

What description and treatment are given to atmospheric dispersion, for example, is a full treatment of turbulence and diffusion given, or is it based on an eddy diffusion type parameterisation or Gaussian dispersion coefficients, or is it neglected altogether through the use of a boundary layer averaged approach?

What is the source of the meteorological data used to provide surface and boundary layer winds, temperatures and humidities, boundary layer depths, cloud cover and radiation? What evaluation has been conducted on this input data? What is the evidence that any climatological data are appropriate for the present purposes? How are spatial variations in these parameters handled across large urban areas?

Has any allowance been made for the dry deposition of NO_2?

72. The information required here for urban PM air quality models includes:

- what is the purpose of the model? For example, the estimation of daily and annual $PM_{2.5}$ and PM_{10} concentrations across the UK and their response to emission controls; the estimation of annual average PM_{10} concentrations within specific urban areas and their response to local emission controls.
- what level of scientific treatment is given to the major elements of the model?

How are national PM emission totals broken down to give emissions at the appropriate spatial and temporal resolution?

Which components of PM concentrations are modelled and which components are treated as background? Primary emissions; secondary formation of sulphates and nitrates; organic aerosols; natural emissions; re-suspended emissions?

How are primary PM emissions split up into individual species such as elemental and organic carbon?

How are natural PM emissions handled, including forest fires, biomass burning, wind-blown dusts and soils, Saharan dust and sea spray?

What are the chemical mechanisms employed to model the formation of secondary inorganic and organic aerosols and what sensitivity does the model show to these choices? What is assumed about the direct emission of PM sulphate?

What description and treatment is given to atmospheric dispersion, for example, is a full treatment of turbulence and diffusion given, or is it based on an eddy diffusion type parameterisation or Gaussian dispersion coefficients-or is it neglected altogether through the use of a boundary layer averaged approach?

What is the source of the meteorological data used to provide surface and boundary layer winds, temperatures and humidities, boundary layer depths, cloud cover and radiation? What evaluation has been conducted on this input data? What is the evidence that any climatological data are appropriate for the present purposes? How are spatial variations in these parameters handled across large urban areas?

What is the formulation of the wet and dry deposition schemes adopted, how are variations with particle size taken into account and what is the source of the meteorological, vegetation and soils data employed?

What has been assumed about background PM levels?

5.3 Operational Evaluation

73. Operational evaluation is based on the FAC2 and NMB metrics. It needs elicit enough information so that independent reviewers can reach a consensus about how well the model reproduces real-world behaviour.

74. The information required here for ozone models includes:

What are the FAC2 and NMB metrics for the comparison of model ground-level ozone with observations for the chosen time periods? These may cover a field campaign or at least cover a pollution episode. They need to cover at least a week in duration and should extend to a pollution year, if possible. Explanations need to be given concerning the choice of the time period, the choice of averaging time, the selection of sites and monitoring networks, see Annex A. Comparisons with observations should be provided for as many sites as is practical and for as many time periods as is feasible. Comparisons can be performed on a pooled site basis (rural, suburban or urban background) or with a single site basis. The modelling teams should explain how good agreement in one part of the model or region of the model domain is to be set against poor agreement elsewhere. They should explain what significance they give to the situations where the FAC2 and NMB criteria are not met. Comments should be made about the model performance for mid-afternoons compared with night-times or summer vs winter or for rural vs urban background sites, for example. They should give a statement of what they view as the overall level of model performance and should comment on whether this is an improvement, or not, on previous model versions and other evaluations of their model in the published literature. Use of graphical methods is recommended, for example, using bivariate polar plots to show up areas of poor model performance (Carslaw et al., 2006). Operational evaluation of ozone modelling should be extended beyond ozone if at all possible. There may be the opportunity to perform comparison between model results and observations using VOC and NO_x observations for specific sites. During specific field campaigns, measurements have been made of a wide range of trace gases and free radical species, providing opportunities for a more robust operational evaluation than can be achieved with ozone alone.

75. The information required for models addressing the deposition of acidic and eutrophying agents includes:

What are the FAC2 and NMB metrics for the comparison of the model annual mean concentrations of SO_2, NO_x, NH_3, HNO_3, particulate sulphate, particulate nitrate, particulate ammonium and for the annual wet deposition of sulphate, nitrate and ammonium with observations at acid deposition monitoring site locations (see Annex A) for chosen time periods? The chosen time periods should cover at least a month, should extend to a year and, if possible, to several different years. Explanations need to be given concerning the choice of the time period and the selection of sites and monitoring networks. Comparisons with observations should be provided for as many sites as is practical and for as many years as is feasible. Because of the high modelling content of the Concentration-Based Estimates of Deposition (CBED) fields (Fowler et al., 2004), these should not be used for the operational evaluation. Attention should be directed to the Defra acid rain, ammonia, nitric acid and particle composition monitoring network data. Analyses should be presented on a daily, twice-monthly and monthly basis, preferably.

Some acid deposition and eutrophication models eg. FRAME, use, or have used in the past, scaling to the CBED fields to assess critical loads exceedances. Scaling should be removed in completing the present comparisons with acid rain monitoring network data. Some statistical models present seasonal and annual results only and lack the ability to deal with daily, twice-monthly or monthly comparisons. In such cases, it is recommended that comparisons with monitoring network data should be carried out for up to 6 meteorological years to cover the year-by-year variability in rainfall patterns. We recommend focussing on two recent years with quite different annual meteorology (i.e. 2006, a wet year and 2003 a warm and dry year).

How is good agreement in one part of the model or region of the model domain to be set against poor agreement elsewhere? What significance is given to the situations where the FAC2 and NMB criteria are not met? What do they view as the overall level of model performance and is this an improvement, or not, on previous model versions and other evaluations of their model in the published literature? Comments should be made on whether model performance varies by season and by meteorological year.

76. The information required here for urban NO_2 models includes:

What are the FAC2 and NMB metrics for the comparison of model NO_2 and NO_x concentrations with observations for the chosen time periods? These may cover a field campaign or a pollution episode. They need to extend to at least one pollution year, if possible. Explanations need to be given concerning the choice of the time period, the selection of sites and monitoring networks. Comparisons with observations should be provided for as many sites as is practical and for as many time periods as is feasible. The modelling teams should explain how good agreement in one part of the model or region of the model domain is to be set against poor agreement elsewhere. They should comment on whether the overall level of model performance is better for NO_2 or NO_x. They should explain what significance they give to the situations where the FAC2 and NMB criteria are not met. Comments should be made, where possible, on whether the performance metrics are different for night-time *vs* daytime, winter *vs* summer and rural and suburban *vs* urban background sites. They should give a statement of what they view as the overall level of model performance and should comment on whether this is an improvement, or not, on previous model versions and other evaluations of their model in the published literature. The metrics should be calculated on the basis of the total concentrations (modelled + background vs measured). The metrics should be calculated on the basis of the modelled contributions (modelled vs measured-background), provided that the modelled contributions are larger than the total uncertainty in the background concentrations and the measurements.

77. The information required here for urban PM models includes:

What are the FAC2 and NMB metrics for the comparison of model $PM_{2.5}$ and PM_{10} with observations for the chosen time periods? These may cover a field campaign or a pollution episode. They need to extend to at least one pollution year, if possible. Explanations need to be given concerning the choice of the time period, the selection of sites and monitoring networks. Comparisons with observations should be provided for as many sites as is practical (see Annex A) and for as many time periods as is feasible. An explicit statement is required concerning the PM components that have been included in the model and observed concentrations. The metrics should be calculated on the basis of the total concentrations (modelled + background vs measured). The metrics should also be calculated on the basis of the modelled contributions (modelled vs measured-background), provided that the modelled contributions are larger than the total uncertainty in the background concentrations and the measurements. Operational evaluation of urban PM models should be extended to the individual PM components such as elemental or black carbon, organic carbon and secondary inorganic and organic species, if possible.

78. An air quality forecast model must satisfy the operational and diagnostic evaluation criteria described in this section when provided with suitable meteorological data but not necessarily forecast data. In addition, an air quality forecast model must pass a test against persistence (tomorrow is the same as today) using forecast meteorological data.

For each day when an air quality forecast is issued, the following data should be collected to complete an operational evaluation test based on persistence, where t_0 is hour of the day for which the forecast is issued :

a. the forecast air pollutant concentration, F_{t0+24}, for 24 hours in advance for a particular pollutant, time of day and averaging time period,

b. the observed air pollutant concentration, O_{t0}, on the day the forecast was issued, for the same time of day and averaging time period,

c. the observed air pollutant concentrations, O_{t0+24}, actually observed 24 hours on, for the same time of day and averaging time period.

The FAC2 and NMB metrics are then calculated for the set of paired values: F_{t0+24} and O_{t0+24} and O_{t0} and O_{t0+24}. The FAC2 and NMB evaluation metrics for F_{t0+24} and O_{t0+24} should be significantly better than those for O_{t0} and O_{t0+24}, showing improvement over the assumption of persistence. Additional operational evaluation procedures and metrics for air quality forecast models are given in Agnew et al. (2007).

5.4 Diagnostic Evaluation

79. Diagnostic evaluation examines the ability of an air quality model to represent the effect of emission changes on air pollutant concentrations (US EPA, 2007). However, because there may be no observed responses to compare against, model responses must be compared against our own theoretical expectations or against those of other models.

80. Diagnostic evaluation of ground-level ozone models is based on ‘indicator species ratios’. Sillman (1999) and Sillman and He, (2002) have shown how our understanding of ozone formation can be employed to characterise whether ozone formation is VOC- or NO_x-limited or somewhat in between these extremes. That is to say, it is possible to categorise from first principles whether ozone concentrations should respond more strongly to reductions in VOC emissions or NO_x emissions. Comparing model sensitivities to VOC or NO_x emission reductions with our theoretical expectations provides a means of evaluating model performance.

What are the responses of the mid-afternoon model ozone concentrations to 30% across-the-board reductions in the emissions of man-made VOC emissions and of NO_x emissions, separately, on a range of days and locations within the modelling domain?

Taking a range of locations, the responses (base case – control case) of the mid-afternoon concentrations are paired up to ascertain which is the greater, the response to VOC reductions or the response to NO_x reductions. Since some model NO_x responses may be negative, that is, ozone may increase in some locations, this analysis must take the signs of the responses carefully into account. Responses are +ve if ozone decreases or –ve if ozone increases, following the precursor emission reduction. A map or a table is then used to delineate where the model response to the 30% VOC reduction is greater than that to the 30% NO_x reduction and vice versa. A location where the VOC response is greater than the NO_x response is VOC-limited and one where the NO_x response is greater than the VOC response is NO_x-limited.

What are the model calculated indicator species ratios in the base case model run during the mid-afternoons for the chosen days and rural or suburban locations within the model domain?

The indicator species ratios are $[O_3]/[NO_z]$, where $[NO_z] = [HNO_3] + [PAN] + [\text{particulate nitrate}]$, $[O_3]/[HNO_3]$ and $[H_2O_2]/[HNO_3]$ and these are calculated in the model for the chosen locations and times of the day.

What is the spatial correspondence between the regions of the model domain allocated to VOC- or NO_x-limited regimes and to the indicator species ratios?

If the model is behaving according to our theoretical expectations, then the regions allocated to the VOC-limited regime should have $[O_3]/[NO_z]$ indicator ratios less than a threshold value and those allocated to the NO_x-limited regime should have $[O_3]/[NO_z]$ greater than this threshold. The other indicator species ratios should similarly delineate VOC-limited and NO_x-limited regimes but with different threshold values**.** It is recommended that an air quality model is considered acceptable if the spatial regions delineated by the VOC and NO_x responses are similar to those delineated by the indicator species ratios. Modelling teams should state the values of the thresholds and compare them with literature values.

This behaviour is illustrated graphically in Figure 5.1 which shows an indicator species analysis using the PTM model for Harwell, Oxfordshire for 28th July 2006. The plot shows a set of pairs of points for each trajectory model experiment showing the response to 30% reductions in VOC emissions (squares) and 30% reductions in NO_x emissions (plus signs) plotted against the ratio of the O_3 to NO_z concentrations in the base case. For $[O_3]/[NO_z]$ ratios much less than 20, the squares lie above the + signs showing that all these trajectory experiments are VOC-limited. For $[O_3]/[NO_z]$ ratios much greater than 20, the +signs lie above the squares showing that these trajectory experiments are NO_x-limited. This coherence between sorting the experiments by way of $[O_3]/[NO_z]$ ratios or NO_x and VOC reduction responses shows that the model is behaving according to our theoretical expectations.

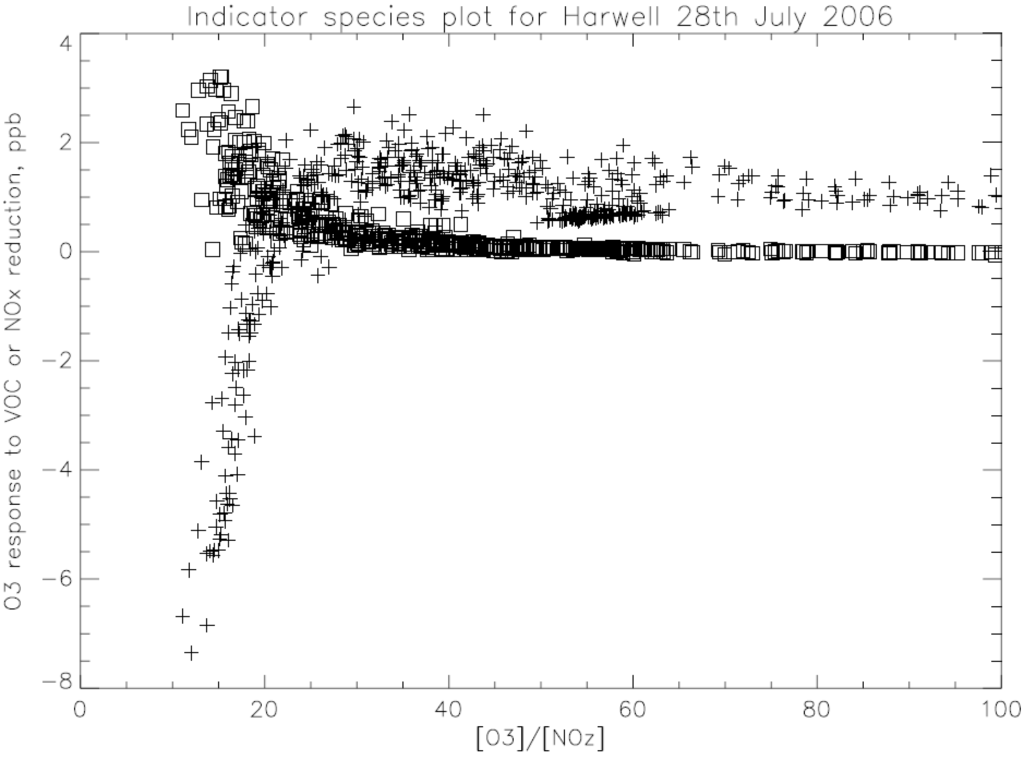


Figure 5.1. Indicator species analysis using the PTM model for Harwell, Oxfordshire for 28th July 2006. The plot shows a set of pairs of points for each trajectory model experiment showing the response (positive values indicate decreases in ozone, negative values indicate increases) to 30% reductions in VOC emissions (squares) and 30% reductions in NO_x emissions (plus signs) plotted against the ratio of the O_3 to NO_z concentrations in the base case. The results from 1000 trajectory cases are shown arriving at 15.00z. Positive values indicate decreases in ozone, negative values indicate increases in ozone.

There may be a range of model indicator species ratios for which the direction of control, whether VOC- or NO_x-limited is not clear-cut. When this occurs, it does not necessarily imply that our theoretical expectations are necessarily incorrect, merely unclear.

Ideally, this comparison between model responses and indicator species ratios could have been completed based on observations of the indicator species ratios: $[O_3]/[NO_z]$, $[O_3]/[HNO_3]$ and $[H_2O_2]/[HNO_3]$. Rarely if ever would there be any observations of HNO_3, PAN, particulate nitrate and H_2O_2 with which to complete the analysis based on observations alone within the UK.

81. Diagnostic evaluation of acid deposition and eutrophication models relies on the ability of models to reproduce the observed spatial patterns of wet deposition across the British isles. Previous assessments of acid rain and eutrophication in the UK have stressed the importance of the orographic enhancement of rainfall and acid deposition across the mountainous areas of the British Isles (NEGTAP, 2001). This enhancement of deposition is apparently not modelled by the EMEP model, the major policy tool used by policy-makers in Europe to construct air pollution strategies that address long-range transport of acidic and eutrophying substances. The result is that in the EMEP model assessments, the critical loads and levels set for the protection of ecosystems are generally exceeded in the south and east of the British Isles. In contrast, national assessments of critical loads and levels exceedances for sensitive ecosystems, focus attention on the upland regions of Wales, the Pennines, Dumfries and Galloway and the Highlands of Scotland.

In view of the likely importance of orographic enhancement in determining the location and magnitude of critical loads and levels exceedances for sensitive ecosystems, diagnostic evaluation focuses attention on the influence of altitude on deposition.

What are the annual wet depositions of SO_x, NO_y and NH_x along transects from north to south through Wales and through the Pennines and west to east through the Highlands of Scotland in the model and in the CBED deposition fields, using a spatial increment of 5 km?

The choice of deposition year is left to the modelling team, provided the necessary CBED deposition fields are available. Comparisons of transects for a number of years are preferable, particularly for the statistical models. Some acid deposition and eutrophication models eg. FRAME, use, or have used in the past, scaling to the CBED fields to assess critical loads exceedances. Scaling should be removed in completing the above comparisons with CBED fields. It is recommended that an air quality model is considered acceptable if it is able to reproduce the main salient features of the influence of orography on the deposition transects through Wales, the Pennines and the Highlands of Scotland as shown in the CBED wet deposition fields.

82. Diagnostic evaluation of urban NO_2 models examines the ability of an air quality model to represent the effect of pollutant emissions on the partitioning between the oxidant species. As environments become more polluted by NO_x emissions in moving from rural through to heavily-trafficked locations, systematic changes in the partitioning of the oxidants between NO_2 and $O_x=O_3+NO_2$ have been observed (Jenkin, 2004). The diagnostic evaluation of urban NO_2 models examines their ability to reproduce observations of oxidant partitioning and the results of other urban NO_2 models. It is recommended that an air quality model is considered acceptable if it is able to reproduce observed oxidant-partitioning relationships.

What are the main features of the scatter plots of model-derived mean NO, NO_2, O_3 and O_x concentrations plotted against mean NO_x concentrations for a number of locations and years compared with observations?

The modelling team should comment on whether the model is able to account correctly for the changes in oxidant partitioning at each site, whether rural, suburban, urban background, motorway, roadside and kerbside. They should comment on how the extrapolated background oxidant concentrations and the fractions of NO_x emissions appearing as direct NO_2 emissions, indicated in their analyses compare with the assumptions made in their modelling and expectations based on observations.

83. Diagnostic evaluation of urban PM models examines the ability of an air quality model to represent the effect of emission changes on $PM_{2.5}$ and PM_{10} concentrations. However, because there are no observed responses to compare against, model responses must be compared against our own theoretical expectations or against those of other models.

Diagnostic evaluation of secondary $PM_{2.5}$ and PM_{10} models is based on our understanding of gas-to-particle partitioning and an understanding of whether PM formation is ammonia- or strong acid-limiting. That is to say, it is possible to categorise whether PM concentrations should respond more strongly to reductions in NH_3 emissions or to SO_2 and NO_x emissions as first demonstrated by Ansari and Pandis (1998). Comparing model sensitivities to NH_3, SO_2 and NO_x emissions with our expectations and other model results, provides a means of evaluating model performance. It is recommended that an air quality model for secondary PM is considered acceptable if it indicates ammonia-limited or strong acid-limited responses to NH_3, SO_2 and NO_x emissions that are in line with theoretical expectations.

5.5 Retrospective Analyses

84. A retrospective analysis is intended to examine the ability of an air quality model to respond to emission changes by comparing recent trends in observed O_3, NO_2, NO_x, $PM_{2.5}$ and PM_{10} concentrations to the model-predicted trends over the same time period. The model is run for current episodes and episodes in one or more historical time periods using the emissions and meteorological data appropriate for each time period modelled. It may be difficult to obtain emissions and meteorological data that are consistent between the various model runs. Using inconsistent inputs will confound the interpretation of the predicted trend from the model.

85. Retrospective analysis is particularly important for acid deposition and eutrophication models. The question: “Can models reproduce measured trends in acid deposition?” needs to be answered. However, for the new Eulerian models, preparation of the necessary inputs, especially meteorological, remains a time consuming task. We recommend focus on two years: one recent (i.e. 2006) and one year corresponding to the start of the acid deposition monitoring network (i.e. 1988 or 1990). This will allow assessment of how models respond to changing emissions.

5.6 Pilot Demonstration of Model Performance Summary

86. The Model Performance Summary has been developed using R, a statistical and graphics package freely available as open-source software. Figure 5.2 presents an analysis of the monitoring data using the R-package OPENAIR (Carslaw and Ropkins, 2009). Appel (2007) has demonstrated a similar set of tools, AMET, for CMAS to evaluate meteorology and air quality data from WRF and CMAQ. Both OPENAIR and AMET used a MySQL database and R as the basis of the analytical environment.

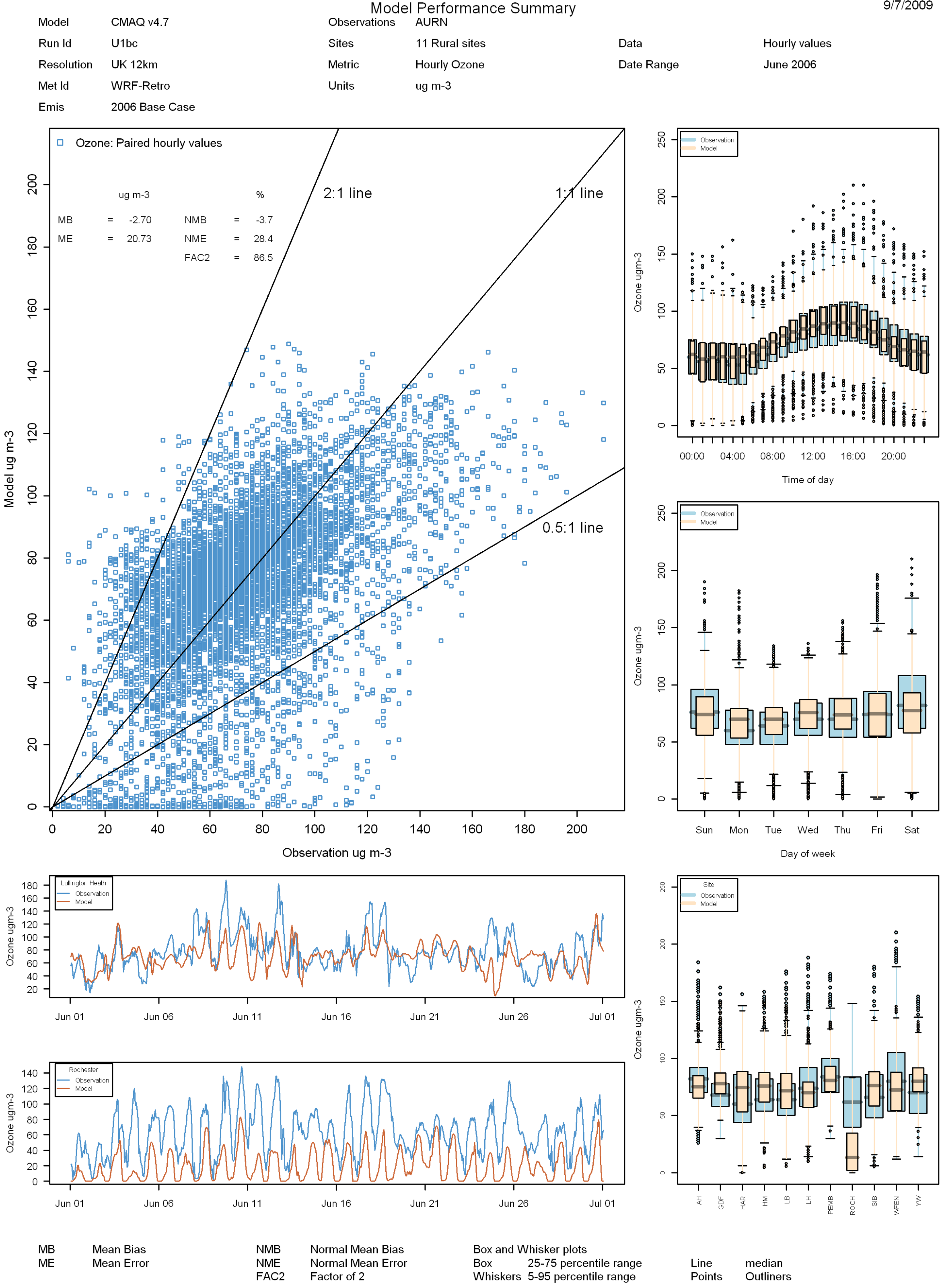
87. As an illustrative example, a pilot version of the Model Performance Summary, shown in Figure 5.2, consists of a series of plots demonstrating the main evaluation metrics for ground-level ozone. The header contains a brief summary of the model conditions and will cross-reference to the more detailed information provided by the modelling team in response to the Basic Information Questionnaire. Summary statistics in line with those recommended in Section 4 above are included together with the lines on the scatter plot marking the 0.5:1, 1:1 and 2:1 ratios. The box and whisker plots identify the performance for a series of different factors e.g. time of day, day of week, month, site, wind direction, and finally line plots of observations and model predictions. The Model Performance Summary is based on a selection of rural monitoring sites from around the country which are representative of general conditions. Supplementary analysis of sites not included in the summary should be presented separately, to give a fuller picture of how well the model performs more unusual conditions.

88. In this example Figure 5.2, hourly ozone observations from 11 AURN sites have been compared to values from a CMAQ simulation for June 2006. The selections of box and whisker plots reflects the nature of the data, i.e. for hourly data the diurnal profile has been included and as there are a number of sites these are also included. If it had been analysis of a single site the site plot would be redundant. The two line plots represent two sites in south-east England one performing well (Lullington Heath) and one not well (Rochester).

89. In this example, the model performs poorly when evaluated against the Rochester site observations. This can be seen clearly in the site box and whisker plot and in the scatter plot. This highlights the importance of selecting sites that are truly representative. If poor model performance is observed for a particular site and for a series of months, or across a series of different models, then location specific factors should be evaluated, and if deemed to be significant, then site should be excluded from the general evaluation.

90. Supplementary analysis of additional sites not included in the Model Performance Summary should be presented separately, to give a fuller picture of how well the model performs.

Figure 5.2 An example Model Performance Summary produced using ozone data from the CMAQ model for June 2006.



Model Performance Summary

9/7/2009

Model	CMAQ v4.7	Observations	AURN		
Run Id	U1bc	Sites	11 Rural sites	Data	Hourly values
Resolution	UK 12km	Metric	Hourly Ozone	Date Range	June 2006
Met Id	WRF-Retro	Units	ug m-3		
Emis	2006 Base Case				

		ug m-3			%
MB	=	-2.70	NMB	=	-3.7
ME	=	20.73	NME	=	28.4
			FAC2	=	86.5

MB	Mean Bias	NMB	Normal Mean Bias	Box and Whisker plots			
ME	Mean Error	NME	Normal Mean Error	Box	25-75 percentile range	Line	median
		FAC2	Factor of 2	Whiskers	5-95 percentile range	Points	Outliners

6 References

Agnew, P., Mittermaier, M.P., Coll, I, Vautard, R., Honore, C., Peuch, V.-H., Elbern, H., 2007. Evaluation of GEMS Regional Air Quality Forecasts. European Centre for Medium-range Weather Forecasts, Reading.

Amann, M, Cofala, J., Heyes, C., Klimont, Z., Mechler, R., Posch, M., Schöpp, W. (2004), RAINS REVIEW 2004. The RAINS model. Documentation of the model approach prepared for the RAINS peer review 2004, http://www.iiasa.ac.at/rains/review/review-full.pdf, International Institute for Applied Systems Analysis (IIASA), Laxenburg, Austria, February 2004. And see http://www.iiasa.ac.at/web-apps/apd/RainsWeb/

AMET (Atmospheric Model Evaluation Tool) available from www.cmascenter.org

Andersson-Sköld, Y. and Simpson, D., 1997. Comparison of the chemical schemes of the EMEP MSC-W and the IVL photochemical trajectory models. Norwegian Meteorological Institute, EMEP MSC-W Note 1/97.

Andersson-Sköld, Y., Simpson, D., 1999. Comparison of the chemical schemes of the EMEP MSC-W and the IVL photochemical trajectory models. Atmospheric Environment **33**, 1111-1129.

Ansari, A.S., Pandis, S.N., 1998. Response of inorganic PM to precursor concentrations. Environmental Science and Technology **32**, 2706-2714.

Appel, K.W., Gilliland, A.B., Sarwar, G., Gilliam, R.C., 2007. Evaluation of the Community Multiscale Air Quality (CMAQ) model version 4.5: Sensitivities impacting model performance. Part I – ozone. Atmospheric Environment **41**, 9603-9615.

AQEG, 2009. Ozone in the United Kingdom. Report of the Air Quality Expert Group, Defra, London, March 2009.
http://www.defra.gov.uk/environment/airquality/publications/ozone/pdf/aqeg-ozone-report.pdf

Beck, B., 2002. Model evaluation and performance. Encyclopaedia of envirometrics, Volume 3, 1275-1279, J. Wiley and Sons, New York, USA.

Beven, K.J., 1993. Prophecy, reality and uncertainty in distributed hydrological modelling. Advances in Water Resources 16, 41-51.

Beven, K., Freer, J., 2001. Equifinality, data assimilation, and uncertainty estimation in mechanistic modelling of complex environmental systems using the GLUE methodology. Journal of Hydrology **249**, 11-29.

Bloss, C., Wagner, V., Jenkin, M.E., Volkamer, R., Bloss, W.J., Lee, J.D., Heard, D.E., Wirtz, K., Martin-Reviejo, M., Rea G., Wenger, J.C. Pilling, M.J., 2005a. Development of a detailed chemical mechanism (MCMv3.1) for the atmospheric oxidation of aromatic hydrocarbons. Atmospheric Chemistry and Physics **5**, 641-644.

Bloss C., Wagner V., Bonzanini A., Jenkin M.E., Wirtz K., Martin-Reviejo M., Pilling, M.J., 2005b. Evaluation of detailed aromatic mechanisms (MCMv3 and MCMv3.1) against environmental chamber data. Atmospheric Chemistry and Physics, **5**, 623-639.

Bossmeyer, J., Brauers, T., Richter, C., Rohrer, F., Wegener, R., Wahner, A., 2006. Simulation chamber studies on the NO_3 chemistry of atmospheric aldehydes, Geophysical Research Letters **33**, L18810.

Britter, R., Schatzmann, 2007. Model evaluation guidance and protocol documentation. COST Action 732, Meteorological Institute, University of Hamburg, Hamburg, Germany.

Bush, T, Tsagatakis, I., King K., Passant, N., 2008. NAEI Emission Mapping Methodology 2006. Report of the National Atmospheric Emissions Inventory, October 2008.

Carslaw, D., Beevers, S., Ropkins, K., Bell, M., 2006. Detecting and quantifying aircraft and other on-airport contributions to ambient nitrogen oxides in the vicinity of a large international airport. Atmospheric Environment **40**, 5424-5434.

Carslaw, D.C., Ropkins, K., 2009. Open-source tools for analysing air pollution monitoring data. Institute for Transport Studies, University of Leeds, 5th January 2009.

Carter, W. P. L., 1990. A detailed mechanism for the gas-phase atmospheric reactions of organic compounds. Atmospheric Environment **24A**, 481-518.

Carter W.P.L., 2008. Development of the SAPRC-07 chemical mechanism and updated ozone reactivity scales. Final Report to the California Air Resources Board Contract No. 03-318.

Chang, J.C., Hanna, S.R., 2004. Air quality model performance evaluation. Meteorology and Atmospheric Physics **87**, 167-196.

Choudrie, S.L., Jackson, J., Watterson, J.D., Murrells, T., Passant, N., Thomson, A., Cardenas, L., Leech, A., Mobbs, D.C., Thistlethwaite, G., 2008. UK Greenhouse Gas Inventory 1990 to 2006: Annual Report for submission under the Framework Convention on Climate Change. AEA Report AEAT/ENV/R/2582. http://www.airquality.co.uk/archive/reports/cat07/0804161424_ukghgi-90-06_annexes_UNFCCCsubmission_150408.pdf

Collins, W.J., Stevenson, D.S., Johnson, C.E., Derwent, R.G., 1997. Tropospheric ozone in a global-scale three dimensional Lagrangian model and its response to NO_x emission controls. Journal of Atmospheric Chemistry **26**, 223–274.

Defra (2009). Local Air Quality Management Technical Guidance LAQM.TG(09). February 2009. http://www.defra.gov.uk/environment/airquality/local/guidance/pdf/tech-guidance-laqm-tg-09.pdf

De Meij, A., Thunis, P., Bessagnet, B., Cuvelier, C., 2009. The sensitivity of the CHIMERE model to emissions reduction scenarios on air quality in Northern Italy. Atmospheric Environment **43**, 1897–1907.

Dennis, R., Fox, T., Fuentes, M., Gilliland, A., Hanna, S., Hogefe, C., Irwin, J., Rao, S.T., Scheffe, R., Schere, K., Steyn, D., Venkatram, A., 2009. On the evaluation of regional-scale photocehmical air quality modelling systems. Atmospheric Modeling Division, National Exposure Research Laboratory, US Environmental Protection Agency, RTP, NC 27711, USA.

Derognat, C., Beekmann, M., Baeumle, M., Martin, D., Schmidt, H., 2003. Effect of biogenic volatile organic compound emissions on tropospheric chemistry during the Atmospheric Pollution Over the Paris Area (ESQUIF) campaign in the Ile-de-France region. Journal of Geophysical Research **108**, 8560.

Derwent, R.G., Collins, W.J., Johnson, C.E., Stevenson, D.S., 2001. Transient behaviour of tropospheric ozone precursors in a global 3-D CTM and their indirect greenhouse effects. Climate Change **49**, 463–487.

Derwent, R.G., Simmonds, P.G., O'Doherty, S., Stevenson, D.S., Collins, W.J., Sanderson, M.G., Johnson, C.E., Dentener, F., Cofala, J., Mechler, R., Amann, M., 2006. External influences on Europe's air quality: Baseline methane, carbon monoxide and ozone from 1990 and 2030 at Mace Head, Ireland. Atmospheric Environment **40**, 844-855.

Derwent, R.G., Jenkin, M.E., Passant, N.R., Pilling, M.J., 2007a. Photochemical ozone creation potentials (POCPs) for different emission sources of organic compounds under European conditions estimated with a Master Chemical Mechanism. Atmospheric Environment **41**, 2570-2579.

Derwent, R.G., Jenkin, M.E., Passant, N.R., Pilling M.J., 2007b. Reactivity-based strategies for photochemical ozone control in Europe. Environmental Science and Policy **10**, 445-453.

Dodge, M. C., 2000. Chemical oxidant mechanisms for air quality modeling: critical review. Atmospheric Environment **34**, 2103-2130.

Derwent, D., Jenkin, M., Passant, N., Pilling, M (2008). "*Up in the air*". Chemistry and Industry, 26 May 2008

Dore, C.J., Murrells, T.P., Passant, N.R., Hobson, M.M., Thistlethwaite, G., Wagner, A., Li, Y., Bush, T., King, K.R., Norris, J., Coleman, P.J., Walker, C., Stewart, R.A., Tsagatakis, I., Conolly, C., Brophy, N.C.J., Hann, M.R., 2008. UK Emissions of Air Pollutants 1970 to 2006. Report of the National Atmospheric Emissions Inventory. http://www.airquality.co.uk/archive/reports/cat07/0810291043_NAEI_2006_Report_Final_Version(3).pdf

Eliassen, A., Hov, Ø., Isaksen, I. S. A., Saltbones, J., Stordal, F., 1982. A Langrangian long-range transport model with atmospheric boundary layer chemistry. Journal of Applied Meteorology **21**, 1645–1661.

Emmerson, K.M., Evans, M.J., 2008. Comparison of tropospheric chemistry schemes for use within global models. Atmospheric Chemistry and Physics Discussions **8**, 19957-19987.

Entec (2005). Service Contract on Ship Emissions: Assignment, Abatement and Market-based Instruments. Task 1 - Preliminary Assignment of Ship Emissions to European Countries. Final Report to European Commission Directorate General Environment, Entec UK Limited, August 2005. http://ec.europa.eu/environment/air/pdf/task1_asign_report.pdf

Entec (2008). UK Ship Emissions Inventory. Final report for Defra. October 2008

Fowler, D., Smith, R.I., Sutton, M.A., Cape, J.N., Nemitz, E., Coyle, M., Muller, J., Milford, C., Famulari D., Anderson-Dunn, M., Tang, Y-S., Storeton –West, R., Crossley, A., Harvey, F., Twigg, M., Riedo, M., Loubet, B., Vincent, K., , Hayman G., Choularton, T. and Beswick, K., 2004. Acid Deposition Processes. Report EPG 1/3/166 to the Department for Environment, Food and Rural Affairs: May 2004.

Gery, M.W., Whitten, G.Z., Killus, J.P., Dodge, M.C., 1989. A photochemical kinetics mechanism for urban and regional scale computer modelling. Journal of Geophysical Research **94**, 12925-12956.

Gross, A., Stockwell, W.R., 2003. Comparison of the EMEP, RADM2 and RACM mechanisms. Journal of Atmospheric Chemistry **44**, 151–170.

Hov, Ø., Stordal., F., Eliassen, A., 1985. Photochemical oxidant control strategies in Europe: a 19 days' case study using a Lagrangian model with chemistry. Report NILU TR 5/85, Norwegian Institute for Air Research, Oslo Norway.

Hynes, R.G., Angove, D.E., Saunders, S.M., Haverd, V., Azzi, M., 2005. Evaluation of two MCM v3.1 alkene mechanisms using indoor environmental chamber data. Atmospheric Environment **39**, 7251-7262.

Irwin, J.S., Civerolo, K., Hogrefe, C., Appel, W., Foley, K., Swall, J., 2008. A procedure for inter-comparing the skill of regional-scale air quality model simulations of daily maximum 8-h ozone concentrations. Atmospheric Environment **42**, 5403-5412.

Jenkin, M.E., 2004. Analysis of sources and partitioning of oxidant in the UK. Part 1: The NO_x-dependence of annual mean concentrations of nitrogen dioxide and ozone. Atmospheric Environment **38**, 5117-5129.

Jenkin, M.E., Saunders S.M., Pilling M.J., 1997. The tropospheric degradation of volatile organic compounds: a protocol for mechanism development. Atmospheric Environment **31**, 81-104.

Jenkin, M.E., Murrells, T.P., Passant, N.R., 2000. The Temporal Dependence of Ozone Precursor Emissions: Estimation and Application. AEA Technology report for the Department of the Environment, Transport and the Regions. AEAT/R/ENV/0355 Issue 1. http://www.airquality.co.uk/archive/reports/empire/AEAT_ENV_0355_v2.pdf

Jenkin M.E., Saunders S.M., Wagner V. and Pilling M.J, 2003. Protocol for the development of the Master Chemical Mechanism, MCM v3 (Part B): tropospheric degradation of aromatic volatile organic compounds. Atmospheric Chemistry and Physics, 3, 181-193.

Jenkin, M.E., Watson L.A., Utembe, S.R., Shallcross, D.E., 2008. A Common Representative Intermediates (CRI) mechanism for VOC degradation. Part 1: Gas phase mechanism development. Atmospheric Environment **42**, 7185-7195.

Jonson, J.E., Simpson, D., Fagerli, H., Solberg, S., 2006. Can we explain the trends in European ozone levels? Atmospheric Chemistry and Physics **6**, 51–66.

Kent, A. J., Grice, S. E., Stedman, J. R., Cooke, S., Bush, T. J., Vincent, K. J., Abbott, J., 2007. UK air quality modelling for annual reporting 2006 on ambient air quality assessment under Council Directives 96/62/EC, 1999/30/EC and 2000/69/EC. Report to the Department for Environment, Food and Rural Affairs, the Scottish Executive, Welsh Assembly Government and the Department of the Environment in Northern Ireland. AEA report. AEAT/ENV/R/2502, September 2007. http://www.airquality.co.uk/archive/reports/cat09/0807231621_dd12006mapsrep_v2.pdf

Kuhn, M., Builtjes, P., Poppe, D., Simpson, D., Stockwell, W., Andersson-Skold, Y., Baart, A., Das, M., Hov, O., Kirchner, F., Makar, P., Milford, J., Roemer, M., Ruhnke, R., Strand, A., Vogel, B., Vogel, H., 1998. Intercomparison of the gas-phase chemistry in several chemistry and transport models. Atmospheric Environment **32**, 693-709.

Lattuati, M., 1997. Contribution à l'étude du bilan de l'ozone troposphérique a` l'interface de l'Europe et de l'Atlantique Nord: modélisation lagrangienne et mesures en altitude. Thèse de sciences, Université Paris 6, France.

Lee, D.S., Kingdon, R.D., Jenkin, M.E., Garland, J.A., 2000. Modelling the atmospheric oxidised and reduced nitrogen budgets for the UK with a Lagrangian multi-layer long-range transport model. Environmental Modelling and Assessment **5**, 83-104.

Mattai, J and Hutchinson D, 2008. London Atmospheric Emissions Inventory 2004 Report, Greater London Authority, November 2008. http://www.london.gov.uk/mayor/environment/air_quality/docs/laei-2004-full-report-dec08.pdf

Monks, P.S., Blake, R.S., Borrell, P., 2007. Review of tools for modelling tropospheric ozone formation and assessing impacts on human health and ecosystems. A report to collate, evaluate and summarise information on tools for modelling ozone (O_3) formation and assessing impacts on human health and ecosystems prepared for the United Kingdom Department for Environment, Food and Rural Affairs (Defra).

Murrells, T., Cooke, S., Derwent, D., 2008. Performance of the Ozone Source Receptor Model against recommendations of the DEFRA Review on Tools for Modelling Tropospheric Ozone. AEA Report for Defra AEAT/ENV/R/2653, July 2008.

National Research Council, 2007. Models in environmental regulatory decision making. The National Academies Press, Washington DC, USA.

NEGTAP, 2001. Transboundary air pollution. Acidification, eutrophication and ground-level ozone in the UK. National Expert Group on Transboundary Air Pollution, Defra, London.

Olson, J., Prather, M., Berntsen, T., Carmichael, G., Chatfield, R., Connnell, P., Derwent, R., Horowitz, L., Jin, S., Kanakidou, Kasibhatla, P., Kotarthi, R., Kuhn, M., Law, K., Penner, J., Perliski, L., Sillman, S., Stordal, F., Thompson, A., Wild, O., 1997. Results from the Intergovernmental Panel on Climatic Change Photochemical Model Intercomparison (PhotoComp). Journal of Geophysical Research **102**, 5979–5991.

Oreskes, N., Shrader-Frechette, K., Belitz, K., 1994. Verification, validation, and confirmation of numerical models in the earth sciences. Science **263**, 641-646.

Passant, NR (2002), Speciation of UK Emissions of Non-Methane Volatile Organic Compounds, AEAT Report AEAT/ENV/R/0545, February 2002

Passant, NR (2003). Estimation of Uncertainties in the National Atmospheric Emissions Inventory. A paper produced for the Department for Environment, Food and Rural Affairs; the National Assembly of Wales; the Scottish Executive; and the Department of Environment in Northern Ireland. AEA Report AEAT/ENV/R/1039. January 2003.
http://www.airquality.co.uk/archive/reports/cat07/AEAT1039_finaldraft_v2.pdf

Pinho, P.G., Pio, C.A., Jenkin, M.E., 2005. Evaluation of isoprene degradation in the detailed tropospheric chemical mechanism, MCM v3, using environmental chamber data. Atmospheric Environment **39**, 1303-1322.

Pinho, P.G., Pio, C.A., Carter, W.P.L., Jenkin, M.E., 2006. Evaluation of alkene degradation in the detailed tropospheric chemistry mechanism, MCM v3, using environmental chamber data. Journal of Atmospheric Chemistry **55**, 55-79.

Pinho, P.G., Pio, C.A., Carter, W.P.L., Jenkin, M.E., 2007. Evaluation of α and β-pinene degradation in the detailed tropospheric chemistry mechanism, MCM v3.1, using environmental chamber data. Journal of Atmospheric Chemistry **57**, 171-202.

Pinho, P.G., Lemos, L.T., Pio, C. A., Evtyugina, M.G., Nunes, T.V., Jenkin, M.E., 2009. Detailed chemical analysis of regional-scale air pollution in western Portugal using an adapted version of MCM v3.1. Science of the Total Environment **407**, 2024-2038.

Righi, S., Lucialli, P., Pollini, E., 2009. Statistical and diagnostic evaluation of the ADMS-Urban model compared with an urban air quality monitoring network. Atmospheric Environment **43**, 3850-3857.

Russell, A., Dennis, R., 2000. NARSTO Critical review of photochemical models and modeling. Atmospheric Environment **34**, 2283-2324.

Saunders, S.M., Jenkin, M.E., Derwent, R.G., Pilling, M.J., 2003. Protocol for the development of the Master Chemical Mechanism, MCM v3 (Part A): tropospheric degradation of non-aromatic volatile organic compounds. Atmospheric Chemistry and Physics **3**, 161-180.

Schmidt, H., Derognat, C., Vautard, R., Beekmann, M., 2001. A comparison of simulated and observed ozone mixing ratios for the summer of 1998 in Western Europe. Atmospheric Environment **35**, 6277–6297.

Seaman, N.L., 2007. Evaluating the performance of meteorological processes within air quality modelling systems. Workshop on the Evaluation of Regional-scale Air Quality Modeling Systems, Research Triangle Park, North Carolina, USA.

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 chemistry and NO_x-VOC indicators. Journal of Geophysical Research **107**, D22, 4659, doi:10.1029/2001JD001123.

Simpson, D., Andersson-Skold, Y., Jenkin, M.E., 1993. Updaing the chemical scheme for the EMEP MSC-W oxidant model: current status. Norwegian Meteorological Institute, EMEP MSC-W Note 2/93, DNMI, Oslo, Norway.

Simpson, D., Fagerli, H., Jonson, J.E., Tsyro, S., Wind, P., 2003. Transboundary Acidification, Eutrophication and Ground Level Ozone in Europe. PART I: Unified EMEP Model Description. EMEP Report 1/2003, DNMI, Oslo, Norway.

Singles, R., Sutton, M.A., Weston, K.J., 1998. A multi-layer model to describe the atmospheric transport and deposition of ammonia in Great Britain. Atmospheric Environment **32**, 393-398.

Solberg S., Dye C., Walker, S.-E., Simpson, D., 2001. Long-term measurements and model calculations of formaldehyde at rural European monitoring sites. Atmospheric Environment **35**, 195-207.

Stockwell, W.R., Middleton, P., Chang, J.S., Tang, X., 1990. The second generation regional acid deposition model chemical mechanism for regional air quality modeling, Journal of Geophysical Research **95**, 16,343-16,367.

Stockwell, W.R., Kirchner, F., Kuhn, M., Seefeld, S., 1997. A new mechanism for regional atmospheric chemistry modelling, Journal of Geophysical Research **102**, 25,847-25,879.

Taylor, K.E., 2001. Summarizing multiple aspects of model performance in a single diagram. Journal of Geophysical Research **106**, 7183-7192.

UN ECE, 2004. Convention on Long-range Transboundary Air Pollution International Cooperative Programme on Modelling and Mapping of Critical Loads & Levels and Air Pollution Effects, Risks and Trends: 2004 manual on methodologies and criteria for modelling and mapping critical loads and levels and air pollution effects, risks and trends. http://www.icpmapping.org/

US EPA. AERMOD modelling system. United States Environment Protection Agency. http://www.epa.gov/scram001/dispersion_prefrec.htm#aermod

US EPA. AERMOD Deposition algorithms-Science Document (Revised draft) Environmental Protection Agency, March 2004.
http://www.epa.gov/ttn/scram/guidance/review/aermoddep/

US EPA . Science Algorithms of the EPA Models-3 Community Multiscale Air Quality (CMAQ) Modeling System. EPA/600/R-99/030, March 1999
http://www.epa.gov/asmdnerl/CMAQ/CMAQscienceDoc.html

USEPA (2003). AERMOD: latest features and evaluation results. EPA-454/R-03-003.
http://www.epa.gov/scram001/7thconf/aermod/aermod_mep.pdf

US EPA, 2007. Guidance on the use of models and other analyses for demonstrating attainment of air quality goals for ozone, $PM_{2.5}$ and regional haze. U.S. Environmental Protection Agency Report EPA-454/B-07-002, Research Triangle Park, North Carolina, USA.

Utembe, S.R., Jenkin, M.E., Derwent, R.G., Lewis, A.C., Hopkins, J.R., Hamilton, J.F., 2005. Modelling the ambient distribution of organic compounds during the August 2003 ozone episode in the southern UK. Faraday Discussions **130**, 311-326.

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., Szopa, S., Beekmann, M., Menut L., Hauglustaine D.A., Rouil L., Roemer, M., 2006. Are decadal anthropogenic emission reductions in Europe consistent with surface ozone observations? Geophysical Research Letters **33**, L13810.

Watson, L.A., Shallcross, D.E., Utembe, S.R., Jenkin, M.E., 2008. A Common Representative Intermediates (CRI) mechanism for VOC degradation. Part 2: Gas phase mechanism reduction. Atmospheric Environment **42**, 7196-7204.

Whitten, G.Z., Hogo, H., Killus, J.P., 1980. The carbon-bond mechanism: a condensed kinetic mechanism for photochemical smog. Environmental Science and Technology **14**, 690-700.

WMO, 2008. Overview of tools and methods for meteorological and air pollution mesoscale model evaluation andtraining. Joint Report of COST Action 728 and GURME. GAW Report No. 181. World Meteorological Organization, Geneva, Switzerland.

Yarwood, G., Rao, S., Yocke, M., Whitten, G., 2005. Updates to the carbon bond chemical mechanism: CB05. Final report to the U.S. EPA, RT-0400675. Available at www.camx.com.

Yu, S., Eder, B., Dennis, R., Chu, S.-H., Schwartz, S.E., 2006. New unbiased symmetric metrics for evaluation of air quality models. Atmospheric Science Letters **7**, 26-34.

Zádor, J., Wagner, V., Wirtz, K., Pilling M.J., 2005. Quantitative assessment of uncertainties for a model of tropospheric ethene oxidation using the European Photoreactor (EUPHORE). Atmospheric Environment **39**, 2805-2817.

Annexes

Annex A: Selection of Monitoring Network Data

Annex B: Choice of Chemical Mechanisms

Annex C: Deposition Parameterisations

Annex D: Emissions Inventory Data in Air Quality Models

Annex A: Selection of Monitoring Network Data

A.1 Introduction

Comparison of model predictions with observations is a critical task in the evaluation of air quality model performance. This Annex reviews the main air quality monitoring networks for the UK for ground-level ozone, acidification and eutrophication and the urban air pollutants: NO_2 and PM. Information is provided on the site classifications adopted and the levels of accuracy and precision that each network achieves.

The choice of the monitoring data required for model evaluation will depend upon both the temporal and spatial resolution of the model run. Automatic monitoring networks provide high temporal resolution data from an increasing number of fixed locations in urban and rural areas across the UK. These are often the most important for studies looking at the impact of air quality on human health. Non-automatic networks often provide a coarser temporal resolution, but may provide a much larger number of parameters, and tend to provide a wider picture of background concentrations for studies of ecosystems impacts. In all cases the use of fully ratified results from the UK national or European monitoring networks is recommended. However, if the model evaluation is required in near real-time, or at a very fine local scale, then either provisional national monitoring data or local data of a lesser or unknown data quality may sometimes be appropriate.

A.1 Data Quality

UK national and European monitoring networks are required to be operated to a high standard with approved monitoring equipment, and with data thoroughly scrutinised and ratified to meet the data quality objectives of European Air Quality Directives.

Annex 1 of The Directive 2008/50/EC of the European Parliament and of the Council, of 21 May 2008, on ambient air quality and cleaner air for Europe specifies the following:

Table A.1: Data quality objectives for ambient air quality assessment

	Sulphur dioxide, nitrogen dioxide and oxides of nitrogen and carbon monoxide	**Particulate matter (PM_{10}/$PM_{2,5}$) and lead**	**Ozone and related NO and NO_2**
Fixed measurements Uncertainty	15%	25%	15%

All European member states are required to provide evidence that their monitoring networks achieve these objectives. When carrying out model evaluation for these pollutants using ratified national monitoring data, these levels of uncertainty can therefore be assumed.

A.2 Site Location Types

It is important that data from suitable monitoring locations are used to validate a particular modelling study. e.g roadside or background monitoring data for an urban housing development study, or regional background concentrations for studying impacts on ecosystems. To this end a UK site classification scheme has been developed to ensure consistency of monitoring site descriptions. When using monitoring data for model evaluation the site location description should be provided to demonstrate that data from a suitable location have been identified.

The agreed UK site descriptions are as follows:

Table A.2: UK air quality monitoring site descriptions

1 - RURAL	
Description	An open countryside location, in an area of low population density distanced as far as possible from roads, populated and industrial areas.
Source Influences	Regional long-range transport, urban plume.
Objectives	Ecosystem impact studies. Assessing compliance with critical loads and levels for crops and vegetation. Investigating regional and long-range transport. Identification of ozone "hot spots"

2 - URBAN	
Description	URBAN
Source Influences	Vehicle, commercial, space heating.
Objectives	Identification of long-term urban trends.

3 - KERBSIDE	
Description	A site sampling within 1m of the kerb of a busy road.
Source Influences	Local traffic.
Objectives	Identifying vehicle pollution blackspots. Assessing worst case scenarios. Evaluating impacts of vehicle emission control technologies. Determining impacts of traffic planning/calming schemes.

4 - REMOTE	
Description	A site in open country, located in an isolated rural area, experiencing regional background pollutant concentrations for much of the time.
Source Influences	Regional/hemispheric background.
Objectives	Assessing unpolluted global or hemispheric background conditions. Long-range transport studies. Long-term baseline trend analysis.

5 - ROADSIDE	
Description	A site sampling between 1m of the kerbside of a busy road and the back of the pavement. Typically this will be within 5m of the road, but could be up to 15m.
Source Influences	Local traffic.
Objectives	Assessing worst case population exposure. Evaluating impacts of vehicle emission controls. Determining impacts of traffic planning/calming schemes.

6 - SUBURBAN	
Description	A location type situated in a residential area on the outskirts of a town or city.
Source Influences	Traffic, commercial, space heating, regional transport, urban plume downwind of a city.
Objectives	Traffic and land-use planning. Investigating urban plumes.

7 - URBAN BACKGROUND	
Description	An urban location distanced from sources and therefore broadly representative of city-wide background conditions e.g. urban residential areas.
Source Influences	Vehicle, commercial, space heating.
Objectives	Trend analysis. Urban planning. Traffic and land-use planning.

8 - URBAN CENTRE	
Description	An urban location representative of typical population exposure in towns or city centres e.g. pedestrian precincts and shopping areas.
Source Influences	Vehicle, commercial, space heating.
Objectives	Identification of long-term urban trends.

9 - URBAN INDUSTRIAL	
Description	An area where industrial sources make an important contribution to the total pollution burden. Intermediate. 20-30m from the kerb of a busy road.
Source Influences	Industrial, motor vehicles.
Objectives	Assessing local impacts on health and amenity. Process optimisation. Source attribution/identification. Providing model input data. Model development/validation. Local planning and plant authorization.

10 - INTERMEDIATE	
Description	20-30m from the kerb of a busy road
Source Influences	Vehicle, commercial, space heating.
Objectives	Identification of long-term urban trends.

11 - AIRPORT	
Description	Monitoring within the boundary of an airport perimeter.
Source Influences	Aircraft, vehicle, commercial, space heating.
Objectives	Determine air quality impact of airport.

A.3 Data Availabilty

Most monitoring data are now available relatively quickly and freely for download over the internet. The relevant websites are usually straightforward to find using the appropriate keywords in a suitable internet search engine. In the future locating data will become even simpler as the requirements of the European Environment Agency's Shared Environmental Information System (SEIS) and the EU INSPIRE Directive are implemented. These will ensure that all environmental monitoring data are catalogued and made available for sharing in a common file format.

In the meantime the following websites are recommended as providing the most comprehensive access to air quality monitoring data.

The UK Air Quality Archive at www.airquality.co.uk/archive/data_and_statistics.php provides an interactive on-line database of measured data and statistics for most of the UK air quality monitoring programmes. There are several options for either downloading pre-formatted annual data files, or going through a series of menus to request data from specific networks for selected locations and dates if required. This is the most comprehensive data source for results from national automatic monitoring networks.

Measurements of UK Pollutant Deposition can also be found through the website at www.uk-pollutantdeposition.ceh.ac.uk/networks.

Information on rural pollutant concentrations, deposition, and critical loads can be found on the UK Air Pollution Information System website at www.apis.ac.uk.

Regional measurements from automatic air quality monitoring networks are now also available for download through a number of well developed local websites and databases. Examples of these are:

- The London Air Quality Network – www.londonair.org.uk
- Scottish Air Quality – www.scottishairquality.co.uk
- Welsh Air Quality – www.welshairquality.co.uk
- Northern Ireland – www.airqualityni.co.uk
- Kent – www.kentair.org.uk
- Sussex – www.sussex-air.net
- Herts, Beds & Bucks – www.hertsbedsair.org.uk
- Heathrow Airwatch – www.heathrowairwatch.org.uk

Caution should be taken when using regional data since these monitors are not always operated to the same standard as national networks. In all cases the ratification status of the data should be clearly marked, and if provisional or regional quality data are used for model evaluation this must be highlighted.

On the European scale near-real-time automatic monitoring data are available from the EEA's OzoneWeb at http://www.eea.europa.eu/maps/ozone/map, and Airbase (http://air-climate.eionet.europa.eu/databases/airbase) provides the database of fully ratified data. Air quality data covering acidifying and eutrophying agents, O_3, PM and VOCs are available for rural, remote and mountain-top sites from the EMEP database at: http://tarantula.nilu.no/projects/ccc/emepdata.html.

A.4 Summary of Monitoring Networks

Within the UK and Europe monitoring data are generally split into two distinct types:

- *Automatic monitoring programmes* - which provide high temporal-resolution data (often to hourly or 15-minute level) available provisionally in near-real-time, and are usually ratified within 6-months of publication; and
- *Non-automatic monitoring programmes* – which provide lower temporal resolution data at the daily, monthly or quarterly level, often following the results of laboratory analysis. Results consequently take longer to publish, and are usually only published in their final ratified form 12-18 months following the measurements.

The spatial coverage of the monitoring programmes is specified by European Air Quality Directives. Most monitoring networks are now operated to achieve the level of minimum statutory compliance. However, in some cases additional monitoring is carried out for research purposes, often at the request of panels such as the UK Air Quality Expert Group (AQEG). A summary of the monitoring networks currently in operation in the UK is provided below. It can be assumed that parallel systems will be in operation across other European member states in order to fulfil their statutory monitoring requirements.

A.4.1 Automatic Networks

A.4.1.1 The Automatic Urban and Rural Network (AURN)

(Managed for Defra and the DAs by Bureau Veritas & Kings College ERG. Quality assured by AEA)

The AURN is the UK's largest automatic monitoring programme. It includes automatic air quality monitoring stations measuring oxides of nitrogen (NO_x), sulphur dioxide (SO_2), ozone (O_3), carbon monoxide (CO) and particles (PM_{10}). These are monitored on an hourly basis throughout the UK.

As of August 2008, the AURN consists of 120 monitoring sites. Of these, 61 are directly funded by Defra and the devolved administrations, whilst a further 59 affiliated sites are owned and operated by local authorities; eight of these sites are also in the London Air Quality Network (LAQN). The network has grown dramatically since it was first established in 1992.

The major objectives of the network are as follows:

Checking if statutory air quality standards and targets are met (e.g. EU Directives)
Informing the public about air quality
Providing information for local air quality review and assessments within the UK Air Quality Strategy
Identifying long-term trends in air pollution concentrations
Assessing the effectiveness of policies to control pollution

A number of organisations are involved in the day-to-day running of the network. Currently, the role of Central Management and Co-ordination Unit (CMCU) for the AURN is contracted to Bureau Veritas, whilst the Environmental Research Group (ERG) of King's College London has been appointed as Management Unit for the London Air Quality Network (LAQN). AEA undertakes the role of Quality Assurance and Control Unit (QA/QC Unit) for the entire AURN. The responsibility for operating individual monitoring sites is assigned to local organisations, such as local authority Environmental Health Officers with relevant experience in the field. Calibration gases for the network are supplied by Air Liquide Ltd and are provided with a UKAS certificate of calibration by AEA.

The techniques used for monitoring within the AURN are summarised below. These techniques represent the current state-of-the-art for automated monitoring networks and, with the exception of the automatic PM_{10} analysers, are the reference methods of measurement defined in the relevant EU Directives.

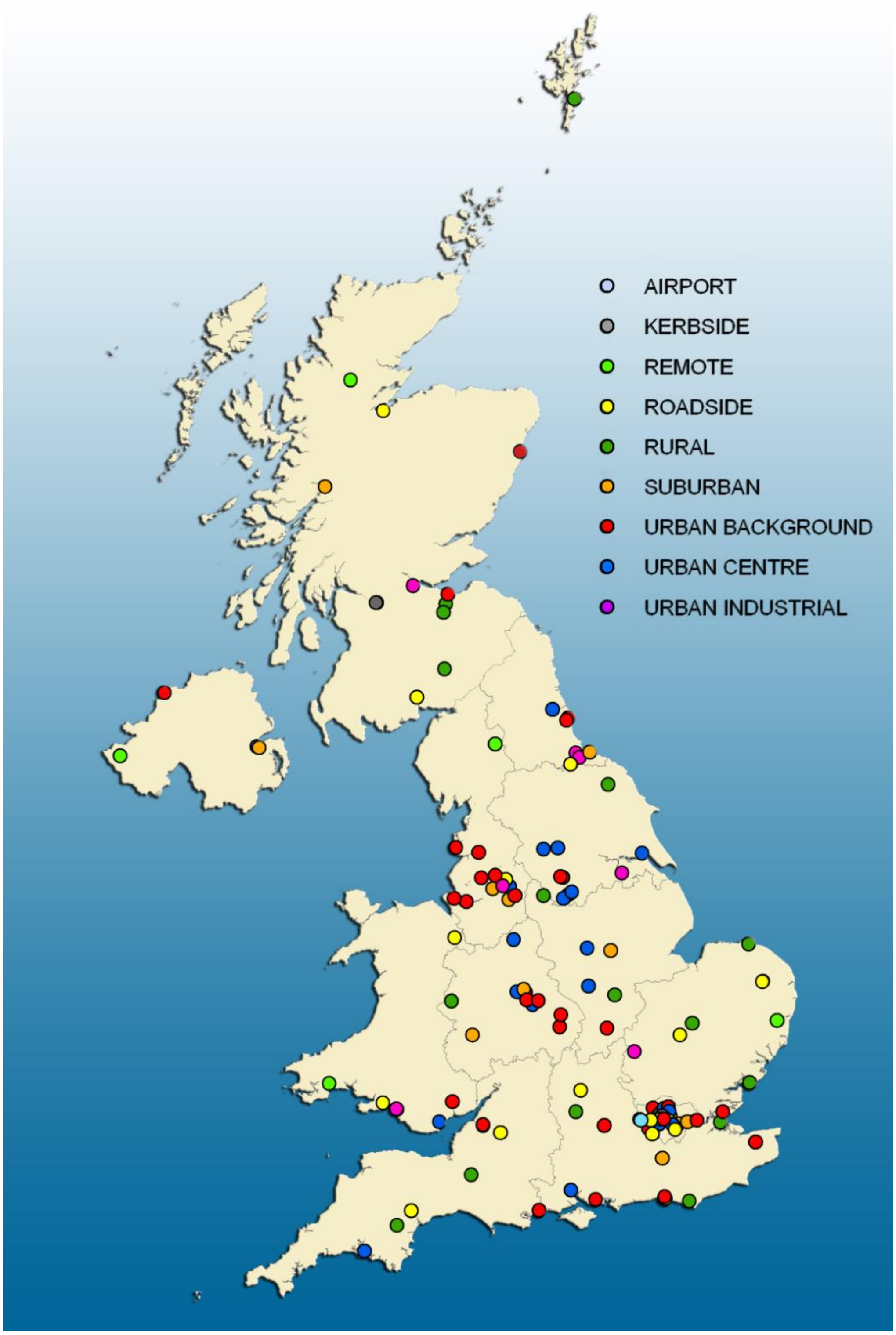
Table A.3 AURN Monitoring Techniques

O_3	UV absorption
NO/NO_x	Chemiluminescence
SO_2	UV fluorescence
CO	IR Absorption
PM_{10}	▶ Tapered Element Oscillating Microbalance ▶ Filter Dynamic Measurement System ▶ Beta Attenuation Monitor ▶ Gravimetric monitor

During 2007, Defra undertook a major review of the AURN, aimed at ensuring that it would continue to meet the UK's compliance monitoring needs in the most cost-effective way. For some pollutants, where ambient concentrations were well within EC Limit Values and AQS Objectives, monitoring was discontinued at some sites. This allowed the resources to be re-directed to locations and pollutants for which the priority was assessed to be higher. The spatial distribution of sites within the UK was also considered and, in some zones and agglomerations, (e.g. London) some sites were closed.

Carbon monoxide monitoring was significantly scaled down, being discontinued at 52 sites: this pollutant is well within EC Limit Values and AQS Objectives throughout the UK. Sulphur dioxide monitoring has also been decreased, with monitoring discontinued at 37 sites. Ozone monitoring was discontinued at 13 sites, PM_{10} particulate monitoring at 10 sites and NO_x and NO_2 monitoring ceased at 19 sites.

Figure A.1: Automatic monitoring stations in the National Automatic Urban and Rural Network (AURN) during 2007



In total, 20 AURN sites were discontinued in 2007. In many cases, however, they have not been shut down, but simply been de-affiliated from the AURN and continue to be operated by the relevant Local Authorities.

A.4.1.2 The Automatic Hydrocarbon Network
(Network managed and quality assured for Defra and the DAs by AEA)

Automatic hourly measurements of speciated hydrocarbons, made using an advanced automatic gas chromatograph (VOCAIR), commenced in the UK in 1991. By 1995, monitoring had expanded considerably with the formation of a 13-site dedicated network measuring 26 species continuously at urban, industrial and rural locations.

The focus in this ground-breaking measurement programme was two-fold: firstly to assess ambient concentrations of a range of Volatile Organic Compounds (VOCs) with significant photochemical oxidant formation potential, and secondly to measure two known genotoxic carcinogens (benzene and 1,3-butadiene) for comparison against emerging UK Air Quality Objectives. Data on these ‘air toxics’ were also regularly reported to the public.

The automatic hydrocarbon monitoring network, as originally constituted, used state-of-the-art measurement techniques, combined with advanced software techniques for signal processing and validation. It was the first network of its kind in the world. The Automatic Hydrocarbon Network operated successfully for 10 years before the programme was re-focussed, re-designed and simplified in 2002.

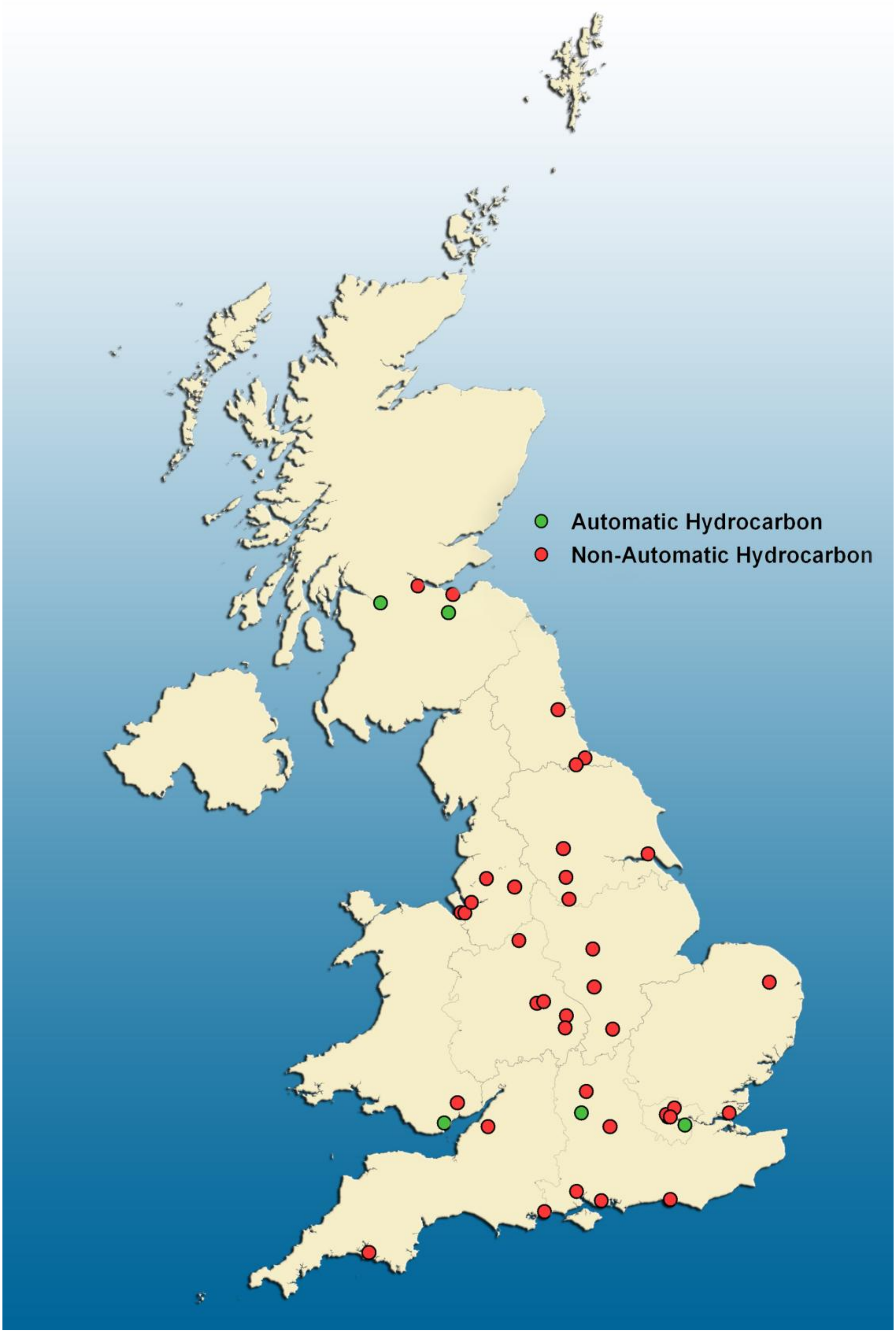
By the start of 2007, the UK Automatic Hydrocarbon Network consisted of six sites, located at Cardiff, Glasgow, Harwell, London Eltham, London Marylebone Road and Auchencorth Moss. Three of these sites – Cardiff, Glasgow and Harwell- utilise an Environnment VOC71M analyser configured to measure and report the concentrations of 1,3-butadiene, benzene, toluene, ethylbenzene, (m+p)-xylene and o-xylene.

Benzene data are used for comparison with the UK Air Quality Objectives and are also reported to the European Commission to fulfil requirements of the 2nd Daughter Directive (and the new Directive on ambient air quality and cleaner air for Europe (Directive 2008/50/EC, 21st May 2008), which supersedes it). 1,3-butadiene data were used for comparison with UK AQS Objectives.

Auchencorth Moss and the two London sites - London Eltham and London Marylebone Road - are fitted with automatic Perkin Elmer gas chromatographs measuring a wider range of VOCs, equivalent to that studied under the original measurement programme. These instruments are capable of measuring and reporting at least 27 hydrocarbons. In April 2007, a Perkin Elmer instrument was installed at Harwell as an upgrade to the existing VOC71M instrument.

During 2007, a strategic review of the UK monitoring networks resulted in the cessation of VOC monitoring at the Cardiff Centre monitoring site.

Figure A.2: UK monitoring stations for automatic and non-automatic hydrocarbons during 2007



A.4.2 Non-automatic networks

A.4.2.1 Deposition of acidic and eutrophying agents

The Acid Deposition Monitoring network (ADMN) was established in 1986 to monitor the composition of precipitation and hence to provide information on deposition of acidifying compounds in the United Kingdom. Its main emphasis has always been the assessment of potential impacts on UK ecosystems. Other measurements are also made within the programme - sulphur dioxide, nitrogen dioxide, particulate sulphate - to provide a more complete understanding of precipitation chemistry in the United Kingdom.

This network has evolved substantially over time. It was originally based on two sub-programmes- a ‘primary’ network providing high quality and high frequency data, which could be used to identify trends over time, and a ‘secondary’ network providing information on the spatial distribution of acid deposition in the UK. Originally, there were 9 primary and 59 secondary sampling sites. Subsequent changes made to the programme, including different measurement techniques, altered sampling frequencies and reductions to the number of monitoring sites, have made this distinction less clear cut.

In the UK, the Defra-funded National Ammonia Monitoring Network (NAMN, http://www.uk-pollutantdeposition.ceh.ac.uk/networks) was established in 1996 to quantify temporal and spatial changes in air concentrations and deposition in gaseous NH_3 and particulate NH_4^+ (included since 1999) on a long-term basis. The monitoring provides a baseline in reduced nitrogen (NH_x) species, which is necessary for examining responses to changes in the agricultural sector and to verify compliance with targets set by international agreements. Data from the network are also used to test the performance of an atmospheric chemistry and transport model (FRAME) that was developed at the same time with a special focus on NH_x, and to contribute to national Nitrogen (N) deposition estimates.

In 2007 there were 95 sites in the NAMN and the high spatial variability of ammonia concentrations demonstrates that this large number of sites is necessary. At 58 of these sites, where power is available, an active diffusion denuder methodology using the CEH DELTA (DEnuder for Long Term Atmospheric sampling) system is used to provide the main spatial and temporal patterns of NH_3 across the UK. Particulate NH_4^+, a secondary product is spatially more even and is monitored at a subset of DELTA sites.

The DELTA network is complemented by the implementation of a high sensitivity passive diffusion sampler, the ALPHA (Adapted Low-cost Passive High-Absorption) sampler at a further 49 sites to assess regional and local scale variability in air NH_3 concentrations in source regions. In the first phase of the network, the Gradko 3.5 cm membrane diffusion tube was used, but, owing to limitations in the sensitivity of the method (Limit of Detection = approx. 1-2 µg NH_3 m^{-3}), this was replaced by the new ALPHA sampler in the second phase of the network (since 2000). The ALPHA sampler was designed and developed specifically for monitoring ambient concentrations of NH_3, with a detection limit of around 0.02 µg NH_3 m^{-3}. To provide an ongoing validation of the ALPHA sampler, its performance is continuously assessed against the DELTA system at 9 sites within the network.

The UK Nitric Acid Monitoring Network has been in operation since September 1999, providing data on nitric acid, particulate nitrate and other species. In the first phase of the network, monitoring was implemented at 12 sites using the CEH DELTA system, in an integrated fashion with the UK National Ammonia Monitoring Network (NAMN). An extension of the DELTA system at the NAMN sites was used to additionally sample gaseous HNO_3, SO_2, HCl and particulate NO_3^-, SO_4^{2-}, Cl^-, Na^+, Ca^{2+}, Mg^{2+}, in parallel with monthly sampling of NH_3 and NH_4^+ at the NAMN sites.

To improve on the national spatial coverage, the network was increased from 12 to 30 sites in the second phase of the network, starting January 2006. The new expanded network also replaced measurements of gaseous SO_2 and particulate SO_4^{2-} previously made under the Rural Sulphur Dioxide Monitoring Programme, which terminated at the end of 2005.

In 1999, seven new sites were established to monitor rainwater composition in ecologically sensitive areas. The UK Nitric Acid Monitoring Network was also introduced to the acid deposition contract in 1999; this provides monthly data on acid gases and particulates. Further changes in 2006 saw the exposure of triplicate nitrogen dioxide diffusion tubes at three sites and the installation of a wet-only daily precipitation collector at one site. In addition, sulphur dioxide measurements undertaken by the ADMN and Rural SO_2 Monitoring Network (part of the Acid Deposition Processes contract) were terminated at the end of 2005, being replaced by measurements made as part of the expanded nitric acid monitoring network.

In 2009, the Acid Deposition, Ammonia and Nitric Acid Monitoring Networks were merged into a single Eutrophying and Acidifying Pollutants UKEAP Network. This network covers the following measurements and sites, see Figure A.3:

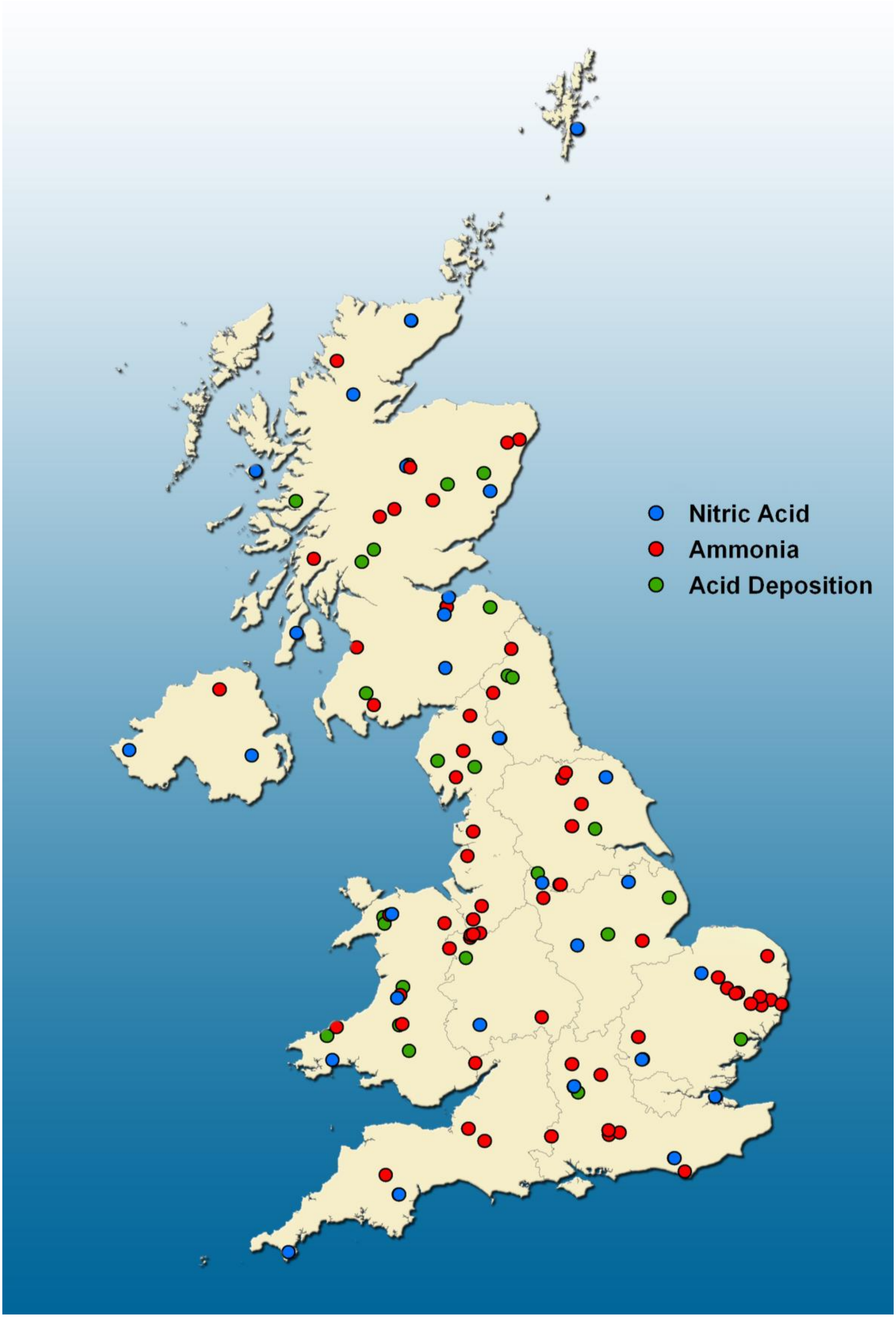
Table A.4 Eutrophying and Acidifying Pollutants UKEAP Network - site numbers & measured parameters

Precipitation Composition	– Rainwater sampling using a bulk collector on a ***fortnightly*** basis at 38 sites – Additional wet-only measurements on a ***daily*** basis at 1 site
Particulate Sulphate **Nitrogen Dioxide**	– Diffusion tube measurements on a ***monthly*** basis at 24 sites. Of these, 3 sites operate triplicate tubes.
Nitric acid, other acid gases and particulates	– Denuder (DELTA) measurements on a ***monthly*** basis at 30 sites

Data from the network are reported on the website http://www.uk-pollutantdeposition.ceh.ac.uk/networks. The aims of the network are to provide a long-term dataset of monthly speciated measurements of acid gases and aerosols that will be used to:

- Provide temporal and spatial patterns and trends, and compare results with dispersion models;
- Facilitate Pollution Climate mapping and assessment of Acid Deposition Processes;
- Contribute to mass closure from the measurements of several components of particulate matter (NO_3^-, SO_4^{2-}, Cl^-, Na^+, Mg^{2+} and Ca^{2+}) together with NH_4^+ from the closely integrated NAMN), which was one of the recommendations in the Department's Air Quality Expert Group's report on Particulate matter; and
- Calculate national and regional deposition budgets, especially in upland areas that are sensitive to acid deposition.

Figure A.3: Non-automatic monitoring stations for acid deposition, nitric acid and ammonia during 2007



A.4.2.4 The UK Black Smoke Network
(Network managed and operated for Defra and the DAs by the National Physical Laboratory)

Black Smoke measurements were the earliest systematic measurements of air pollution by particulate matter in the United Kingdom, with records dating back to the 1920s. Air is sampled through a filter and the darkness of the stain, measured by optical reflectance, is converted to a Black Smoke Index, given in units of µg m^{-3}.

In 2005, monitoring of Black Smoke ceased to be a regulatory requirement in the United Kingdom, with particulate matter being regulated as PM_{10} and $PM_{2.5}$. In that year, Defra commissioned an independent review of the UK urban network for measurement of Black Smoke, SO_2 and NO_2. Its recommendations for Black Smoke monitoring included continuing a network of about 20 sites, some of which should be at AURN locations.

The resulting new UK Black Smoke Network commenced operation on 1st September 2006. The design of the Network was planned to incorporate 11 existing Black Smoke sites and 10 Black Smoke samplers in AURN stations. Samplers were installed at the AURN sites between October 2006 and March 2007.

The design of the Network was planned to incorporate 11 existing Black Smoke sites from the old Smoke and SO_2 Network, and to install 10 Black Smoke samplers in AURN stations. The sites are listed below.

Black Smoke sampling uses the 8-port sampler that has historically been used in the UK network, based on the standard method BS 1747 Part 11, ISO 9835. The principle of the 8-port sampler method involves drawing air at a constant flow rate of around 1.4 l/min through a Whatman Number 1 cellulose filter, so that about 2 m^3 of air (at ambient conditions of temperature and pressure) is sampled for each daily sample. Suspended particulate matter is collected on the filter over an area determined by a choice of clamp – in this case with a one inch diameter - forming a dark stain. The inlet, an upturned funnel, is not specifically designed to be size selective, and has been shown in one study to collect the approximate size fraction $PM_{4.5}$.

The 8-port sampler is designed with eight pairs of filter clamps for weekly operation, providing daily sampling from a midnight-to-midnight basis. The timed eight-port valve is set to switch over at midnight to expose a fresh filter paper each day. Weekly visits are made to change filter papers and to record weekly sample volumes and flow rates.

Table A.5 Black smoke monitoring sites

Existing Sites	AURN sites
Strabane 2	Edinburgh St Leonard’s
Cardiff 12	Glasgow Centre
Halifax 17	Manchester Piccadilly
South Kirkby 1	Belfast Centre
Halesowen 8	Bradford
Sunderland 8	Stoke Centre
Norwich 7	North Kensington
Dunmurry 3	Birmingham Tyburn
Woolwich 9	Nottingham Centre
Bath 6	Folkestone, Kent Network
Marylebone Road	

Black smoke concentrations are then estimated by means of a reflectance measurement. The darkness of the stain is measured with an EEL M43D reflectometer, the reflectance being determined relative to a blank filter of the same type. The instrument uses a light bulb to give a broad band source that is reflected back from the smoke stain to a photo-sensitive element and produces a reading between 0% and 100% reflectance. NPL measures the reflectance of all of the filters sampled by the network, whereas, previously, reflectometry measurements were performed by local Councils using their own reflectometers.

The measured reflectance, sampled volume and filter area are used to calculate the concentration of dark particulate matter in the sampled air, as Black Smoke Index, with units of µg m^{-3}, using the relationship given in BS 1747: Part 2:

$$C = \frac{1}{V}\left(91679.22 - 3332.046R + 49.618884R^2 - 0.35329778R^3 + 0.0009863435R^4\right)$$

where:

C = concentration in µg/m^3
V = volume of sampled air in ft^3
R = reflectometer reading (%)

The above relationship is only valid for a one inch diameter filter clamp - an additional factor is required for other clamp sizes. This relationship is also only valid for values of R above 40%, which was true for all samples measured in the Network over recent years.

Annex B: Choice of Chemical Mechanisms

B.1 Introduction

A variety of chemical mechanisms have been used in models which have been applied to the study of ground-level ozone in Europe (e.g., see Monks et al., 2007). These mechanisms vary considerably in size and complexity, and the choice of chemical mechanism is usually determined by the type of model being applied and the specific aims of the study.

Explicit mechanisms, which are more fundamentally linked to elementary studies of atmospheric chemical reactions, can be used to provide a detailed representation of the chemistry and a stringent examination of the roles played by individual emitted VOCs in generating ozone and related secondary pollutants. Because of their explicit nature, however, such mechanisms typically contain several thousand species and reactions, such that their use is generally restricted to models containing comparatively simple representations of transport processes. “Condensed” or “reduced” mechanisms are therefore often used to describe the general features of ozone formation more economically, although they are less readily validated through the comparison of model predictions with observations of related secondary pollutants (e.g., individual PANs, aldehydes, ketones) or of the emitted VOC species. This is because the reduction methodologies usually involve a combination of: (i) a systematic reduction in the complexity of the chemistry for the considered suite of VOCs (i.e. “*lumped chemistry*”) and, (ii) the lumping of emissions so that the chemistry for one VOC can be used to represent that of a number of VOCs (i.e. “*lumped emissions*”).

Whether an applied mechanism is highly parameterized, highly explicit, or of intermediate complexity, its performance ideally needs to be evaluated and refined by comparing its predictions with experimental data. Where such evaluations have been carried out, they are usually made using either the results of environmental chamber experiments, or by comparing the predictions of a model, which includes the mechanism, against field measurements of appropriate species. Alternatively, mechanisms have been tested indirectly against experimental data through comparison of their performance with a mechanism which has been evaluated.

In the following paragraphs B.2 – B.8, the main chemical mechanisms that have been adopted for use in ground-level ozone modelling in Europe are briefly introduced, and principle approximations and simplifications that have been employed are described. Recommendations are given in B.9 concerning the choice of chemical mechanisms to be employed in model intercomparisons.

B.2 The Master Chemical Mechanism (MCM)

Mechanism details: The most recent version (MCM v3.1) treats the degradation of methane and 135 non-methane VOCs (http://mcm.leeds.ac.uk/MCM/). These include the major emitted anthropogenic species as listed in the UK NAEI; it also includes the representative biogenic species isoprene, two monoterpenes (α– and β–pinene) and one oxygenated VOCs (2-methyl-but-3-ene-2-ol). The resultant mechanism contains about 13,500 elementary reactions of 5,900 species.

General description: The MCM is a highly detailed explicit chemical mechanism which describes the gas-phase chemical processes involved in the atmospheric degradation of a large series of primary emitted VOCs. It provides a direct means of utilising published laboratory and theoretical data on the kinetics and mechanisms of elementary chemical reactions relevant to VOC oxidation in atmospheric models. However, because not all the reactions involved in atmospheric VOC chemistry have been, or indeed can be studied, the kinetics and products of a large number of unstudied chemical reactions are defined on the basis of the studied reactions of a smaller subset of similar chemical species, by analogy and with the use of structure-activity relationships (SARs) to estimate the otherwise unknown parameters. The mechanism construction methodology is documented in a series of published protocols (Jenkin et al., 1997; 2003; Saunders et al., 2003; Bloss et al., 2005a), which describe the rules upon which the development of the MCM is based. The MCM has also been used as a reference benchmark mechanism to assist the development and/or evaluation of some reduced mechanisms. It therefore provides a primary link in the transfer of experimental/theoretical knowledge, and a means by which mechanisms used in atmospheric models can be traceable to the elementary reaction studies.

The MCM protocol recognises that the rigorous application of a series of rules can lead to an unmanageably large number of reactions, particularly for larger VOCs, and strategic simplification measures are therefore applied to control the ultimate size of the mechanism. These fall into the following three general categories: (i) non-proliferation of low-probability reaction channels; (ii) simplified degradation of product classes deemed to be “minor”, or for which chemistry is poorly established; and (iii) parameterisation of the reactions of peroxy radicals (RO_2) with each other (of which there would otherwise be approximately 0.5 million in MCM v3.1).

Evaluation: Testing and evaluation of relevant portions of the MCM have been performed against data from a number of environmental chambers (Saunders et al., 2003; Bloss et al., 2005a, 2005b; Zádor et al., 2005; Pinho et al., 2005; 2006; 2007; Bossmeyer et al., 2006; Hynes et al. 2005). In this way, the MCM has been evaluated against over 300 chamber experiments. The MCM has also been evaluated against ambient observations of intermediates formed during VOC degradation in numerous studies (e.g., Saunders et al., 2003; Utembe et al., 2005; Pinho et al., 2009).

B.3 The Statewide Air Pollution Research Center (SAPRC) Mechanism

Mechanism details: The SAPRC mechanism has undergone a series of updates since the original version of Carter (1990). The most recent version (SAPRC-07) has a standard base mechanism consisting of about 100 species and 250 reactions (Carter, 2008), to which additional chemistry is appended, depending on the number of emitted VOCs, or VOC mixtures, being treated. The method has been applied to almost 1100 VOC species in reactivity assessments.

General description: The central component of the SAPRC mechanisms is a “standard base mechanism” which represents the reactions of the inorganic reactants, reactions of common organic oxidation products and radicals formed generally from VOC degradation, and the initial chemistry for a small series of more important VOCs (e.g., ethene and isoprene). The base mechanism contains a combination of explicit chemistry, and chemistry represented by lumped molecular species and chemical operators describing, for example, NO-to-NO_2 conversions resulting from peroxy radical reactions. The initial reaction sequences for the majority of emitted VOCs are not included in the base mechanism, but can be added to the mechanism as required for a particular application, using a mechanism generation and parameter estimation system. This chemistry either appears as explicit reactions for given individual VOCs, or as lumped model species for complex mixtures whose parameters are derived from the mixture of detailed model species they represent.

Evaluation: At all stages of development, the SAPRC mechanisms have been comprehensively evaluated by comparing predictions with results of more than 1500 environmental chamber experiments (see Carter, 2008; and references therein).

B.4 The Carbon Bond Mechanism (CBM)

Mechanism details: CBM has undergone a series of updates since the original version of Whitten et al. (1980). The most recent version (CBM-05) has a base mechanism consisting of 51 species and 156 reactions (Yarwood et al., 2005).

General description: CBM is comprised of a combination of explicit chemistry, surrogate approximations, and lumped or generalised chemistry which can be applied, in principle, to any emissions speciation. Selected emitted VOC species (e.g., ethene, formaldehyde and isoprene) are identified explicitly in the mechanism. In most other cases, the emitted VOCs are represented in terms of the series of chemical environments of the constituent carbon atoms within the molecule - the underlying concept being that similarly bonded carbon atoms react independently of the molecules in which they occur. Finally, some emitted VOCs are represented by another specific emitted species which falls into one of the previous two categories (i.e., conventional emissions lumping). The degradation chemistry is represented in terms of a series of surrogate product types, which are formed from the appropriate molecular substructures in the precursor VOC or product.

Evaluation: At all stages of development, CBM has been evaluated using the results of about 200 experiments performed in a number of environmental chambers (e.g., Gery et al., 1989; Yarwood et al., 2005).

B.5 The Regional Atmospheric Chemistry Mechanism (RACM)

Mechanism details: RACM was developed to describe oxidation chemistry in the lower atmosphere (Stockwell et al. 1997), as a further development of the regional acid deposition model version 2, RADM-2 (Stockwell et al., 1990). RACM treats the degradation of 16 representative anthropogenic VOCs, and 3 representative biogenic VOCs (isoprene, α–pinene and limonene). The resultant mechanism contains 77 species and 237 reactions.

General description: RACM treats the degradation of a combination of explicit emitted VOC species, and lumped emitted VOC species which represent species within a given class and OH-reactivity range. The lumped species are either generic species with weighted reaction parameters, based on the component VOC emissions distribution, or explicit species used to represent themselves and higher homologues of comparable reactivity within a VOC class. The latter case includes the anthropogenic species, toluene, m-xylene and m-cresol, and the biogenic species α–pinene and limonene. The subsequent degradation is represented by a combination of explicit chemistry, formation of surrogate product types, and the use of chemical operators describing, for example, NO-to-NO_2 conversions resulting from peroxy radical reactions (Stockwell et al. 1997).

Evaluation: RACM has been evaluated against environmental chamber data for selected VOCs and VOC mixtures (Stockwell et al., 1997), with its performance reviewed in relation to those of some other mechanisms by Dodge (2000). Its performance has also been compared with that of its precursor RADM-2 and the EMEP MSC-W mechanism for a wide range of ambient conditions, using box model scenarios (Gross and Stockwell, 2003).

B.6: The STOCHEM Mechanism

Mechanism details: The STOCHEM mechanism was originally reported by Collins et al. (1997), and has subsequently been updated as described by Derwent et al. (2001). The more recent version treats the degradation of methane and 11 non-methane VOCs, and consists of 70 species and 160 reactions.

General description: Originally developed for use in the global 3D chemistry-transport model of the same name, the STOCHEM mechanism makes use of a combination of lumped emissions and lumped (or selectively represented) chemistry to provide an economical treatment of VOC oxidation. Extensive mass-based lumping of VOC emissions is used to represent a detailed VOC speciation in terms of 11 species. For example, n-butane is used as a surrogate for alkanes with four or more carbon atoms, and o-xylene is used to represent emitted aromatic hydrocarbons with eight or more carbon atoms (Derwent et al., 2001). The degradation chemistry of the emitted species (and their products) is represented explicitly, but generally follows a simplified sequence of reactions in which the major oxidation pathway is used to represent the chemistry of all pathways. Further reductions are achieved through omission of some classes of reaction for selected species.

Evaluation: The STOCHEM mechanism has been evaluated primarily though comparison of its performance with those of other mechanisms for a range of ambient conditions, using box model scenarios (Olson et al.,1997); and through comparison with ambient data for ozone and other selected species when run in the STOCHEM model itself (e.g., Derwent et al., 2006), or in regional scale models such as the Ozone Source-Receptor Model, OSRM (Murrells et al., 2008).

B.7 The EMEP MSC-W Mechanism

Mechanism details: The original EMEP MSC-W mechanism was based on earlier chemical mechanisms developed by Eliassen et al. (1982), Hov et al. (1985) and Simpson et al. (1993) and has undergone a series of updates. The most recent version treats the degradation of methane and 11 non-methane VOCs. It consists of 79 species and 141 reactions, as described in the published mechanism comparison of Anderson-Sköld and Simpson (1999). Details of the latest version of the mechanism are given in Simpson et al., (2003).

General description: The EMEP MSC-W mechanism is highly reduced, making use of a combination of lumped emissions and lumped (or selectively represented) chemistry. Extensive mass-based lumping of VOC emissions allows the full speciation to be represented in terms of 11 species. For example, n-butane is used as a surrogate for alkanes with three or more carbon atoms, and o-xylene is used to represent all emitted aromatic species (Anderson-Sköld and Simpson, 1997). The degradation of the surrogate species (and their products) is represented explicitly, but generally follows a simplified sequence of reactions in which the major oxidation pathway is used to represent the chemistry of all pathways. Further reductions are achieved through omission of some classes of reaction for selected species.

The EMEP MSC-W mechanism has also been adapted/extended to produce the MELCHIOR and MELCHIOR-2 mechanisms, which have been applied exclusively in the CHIMERE model (e.g., Schmidt et al., 2001; Vautard et al., 2006; de Meij et al., 2009). MELCHIOR represents a parallel extension of early versions of the EMEP mechanism, consisting of 82 species and 338 reactions (Lattuati, 1997). The extensions focussed on updating the chemistry of the aromatic and biogenic species. MELCHIOR-2 is a subsequently reduced version of MELCHIOR, consisting of about 40 species and 120 reactions (Derognat et al., 2003).

Evaluation: The EMEP MSC-W mechanism has been evaluated primarily though comparison of its performance with those of other mechanisms for a range of ambient conditions, using box model scenarios (Anderson-Sköld and Simpson, 1997; 1999; Kuhn et al., 1998; Gross and Stockwell, 2003); and through comparison with ambient data for ozone and other selected species when run in the EMEP MSC-W model (e.g., Solberg et al., 2001; Jonson et al., 2006).

B.8 The Common Representative Intermediates (CRI) Mechanism

Mechanism details: The most recent full version (CRI v2) treats the degradation of methane and 115 non-methane VOCs (Jenkin et al., 2008). These include the major emitted anthropogenic species as listed in the UK NAEI; it also includes the representative biogenic species isoprene and two monoterpenes (α– and β–pinene). The resultant mechanism contains 1183 elementary reactions of 434 species. A series of further reduced versions are available the smallest of which (CRI v2-R5) treats the degradation of methane, 19 non-methane anthropogenic VOCs and the three biogenic VOCs indicated above. It contains 555 elementary reactions of 196 species.

The CRI method has also been used to provide a representation of the degradation of an additional 40 emitted VOCs in the UK NAEI speciation (mainly large alkanes, cycloalkanes and aromatics) as an annex to MCM v3.1, with the extended mechanism describing the degradation of 176 emitted VOCs (Derwent et al. 2007a; 2007b).

General description: CRI v2 was built up on a compound-by-compound basis, with the performance of its chemistry optimised for each compound in turn by comparison with that of MCM v3.1, using box model simulations (Jenkin et al., 2008). It is a lumped chemistry mechanism, in which lumped intermediates are defined on the basis of the theoretical maximum number of molecules of ozone formed from their complete subsequent degradation, which is based on the number of reactive (i.e., C-C and C-H) bonds the intermediates contain. Each of the series of generic intermediates is therefore used to represent a large set of species possessing the same index, as formed in the MCM. The reduced versions (including CRI v2-R5) were developed in conjunction with a systematic lumping strategy for anthropogenic emissions, in which the emitted VOCs were grouped on the basis of their chemical class and POCP value (Watson et al., 2008).

Evaluation: The performances of CRI v2 and its reduced variants (including CRI v2-R5) were tested during development against that of MCM v3.1 for a wide range of ambient conditions, using box model scenarios and simulations of a major field campaign (Jenkin et al., 2008; Watson et al., 2008). CRI v2-R5 has also been further evaluated in a recent mechanism intercomparison (Emmerson and Evans, 2008).

B.9 Recommendations Concerning the Choice of Chemical Mechanism

In the paragraphs above, a selection of chemical mechanisms has been laid out together with some supporting references and information. Choice of chemical mechanism may exert an important influence on the results obtained in most atmospheric chemistry modelling studies. It is recommended that in all modelling studies at least two chemical mechanisms are selected and the results are presented which show the sensitivity of the model output to the choice of chemical mechanism. This may well be extremely ambitious. Because of the importance given by Defra to the evaluation of model performance, it is furthermore recommended that the choice of chemical mechanism is restricted to include at least one of the mechanisms that have been extensively tested against laboratory data, that is: MCM (B.2), SAPRC (B.3), CBM (B.4), RACM (B.5) and CRI (B.8).

Annex C: Deposition Parameterisations

C.1 Introduction

This Annex briefly reviews the main dry and wet deposition parameterisations used in air quality models.

The Annex draws upon information provided in the documentation for a range of models including:

ADMS[1,2]
AERMOD[3]
EMEP Unified Model[4]
CMAQ Community Scale Model[5]
FRAME[6]
TRACK[7]

More details of the parameterisations can be obtained from the model documentation.

In any model intercomparison, it is important that modelling teams state which dry and wet parameterisations are being employed and where they have obtained the necessary meteorological and surface data to drive them.

Deposition parameterisations are used within air quality models for two separate but related purposes. These are:

- To assess the rate of removal of gaseous and particulate species from the air column by deposition processes in order to calculate the change in their concentrations with time or location.
- To assess the deposition flux to the ground, vegetation or other surfaces in order to assess the potential impact on ecosystems by reference to critical loads or stomatal flux-based critical levels.

Different deposition parameters may be used within the same model for these two purposes. For example, the deposition flux may be calculated using the methods specified in the ICP mapping manual[8] under the UNECE Convention on Long-range Transboundary Air Pollution while a different parameterisation is used internally within the model to calculate the rates of removal from the air column.

---

[1] CERC 2009. Modelling dry deposition. ADMS4 documentation P17/13F/09. http://www.cerc.co.uk/software/pubs/ADMS4TechSpec/P17_13.pdf

[2] CERC 2009. Modelling wet deposition. ADMS4 documentation P17/12D/09. http://www.cerc.co.uk/software/pubs/ADMS4TechSpec/P17_12.pdf

[3] US EPA 2004. AERMOD Deposition Algorithms – Science Document. http://www.epa.gov/scram001/7thconf/aermod/aer_scid.pdf

[4] David Simpson, Hilde Fagerli, Jan Eiof Jonson, Svetlana Tsyro and Peter Wind. Unified EMEP Model Description. http://emep.int/UniDoc/report.html

[5] Science Algorithms of the EPA Models-3 Community Multiscale Air Quality (CMAQ) Modeling System. EPA/600/R-99/030, March 1999. http://www.epa.gov/asmdnerl/CMAQ/CMAQscienceDoc.html

[6] Singles, R., M.A. Sutton & K.J. Weston (1998) A multi-layer model to describe the atmospheric transport and deposition of ammonia in Great Britain. Atmos. Environ., 32, 393-399.

[7] Lee, D.S., Kingdon, R.D, Jenkin, M.E. and Garland, J.A. (2000). Modelling the atmospheric oxidised and reduced nitrogen budgets for the UK with a Langrangian multi-layer long-range model. Environmental Modelling and Assessment, 5, 83-104.

[8] Manual on methodologies and criteria for modelling and mapping critical loads and levels and air pollution effects risks and trends. UNECE Convention on Long-range Transboundary Air Pollution. http://icpmapping.org/cms/zeigeBereich/11/manual-english.html

C.2 Dry Deposition

C.2.1 Deposition velocity

Many chemical species react with or are absorbed by elements of the earth's surface- the ground, water and vegetation. This dry deposition process is responsible for removing a substantial part of the chemical species from the atmosphere. Air quality models, in general, calculate the rate of dry deposition or the dry deposition flux, F_d, as the product of the difference in concentrations, $\Delta\chi$, between a reference height defined by the model and the surface, and a deposition velocity, V_d:

$$F_d = \Delta\chi \cdot V_d$$

In many cases, the concentration at the surface is effectively near zero because the chemical species is removed quickly at the surface, so that the dry deposition flux is given by:

$$F_d = \chi_r \cdot V_d$$

where χ_r is the concentration at the model reference height.

In other cases, it is convenient to assume that there is a linear relationship between the concentration at the reference height and the concentration at the surface. One example is where the chemical species diffuses away from a wetted surface at which the concentration in the air is in linear equilibrium (Henry's Law) with the aqueous solution below. Then the flux:

$$F_d = \chi_r \cdot V_d^{'}$$

where $V_d^{'}$ is the effective deposition velocity, taking into account the equilibrium relationship and the rate of diffusion into the aqueous solution. The assumption of linearity is only approximately true over small concentration ranges for many species that dissociate in water such as sulphur dioxide, ammonia and hydrogen chloride. Furthermore, the equilibrium is affected by other species that dissociate, including carbon dioxide.

In some cases, the concentration at the surface is greater than that at the reference height. The surface is then a net source rather than a net sink for the chemical species. This situation commonly arises for ammonia, nitric oxide and biogenic species such as terpenes and isoprenes. If the gross emission from the surface is independent of the concentration at the reference height then it can be treated as a separate emissions source, E so that the net deposition flux is:

$$F_d = \chi_r \cdot V_d - E$$

The use of the deposition velocity in this way assumes that the deposition of the chemical species is independent of other chemical species. It is necessary that the species should not react significantly during the time it takes to travel from the reference height to the surface. The simple deposition velocity approach may not be appropriate, depending on the vertical structure of the model, for estimating ozone deposition in urban areas where there is a substantial emission of nitric oxide near ground level from traffic. Similarly, the assumption may not be valid in rural areas for estimating nitric acid deposition where there are substantial emissions of ammonia from farm animals.

The simplest treatment of dry deposition in an air quality model is to assume that the deposition velocity for a particular chemical species is constant for a specific land cover type, (e.g. grassland, forest, sea). A slightly more refined approach is to assume that there is a defined seasonal or diurnal variation in the deposition velocity. More sophisticated models employ a electrical resistance analogue approach as described below.

C.2.2 Electrical resistance analogue

An electrical resistance analogue is generally used to represent the major components of resistance to mass transfer for gas deposition to vegetation and the ground. Figure C.1 shows a typical electrical resistance model, although there are some differences between models.

The main components are:

R_a, the aerodynamic resistance to vertical transport below a specified reference height above the surface,
R_b, the resistance of the nearly laminar sublayer of air in contact with surface elements, and
R_c, the resistance of the surface itself to uptake.

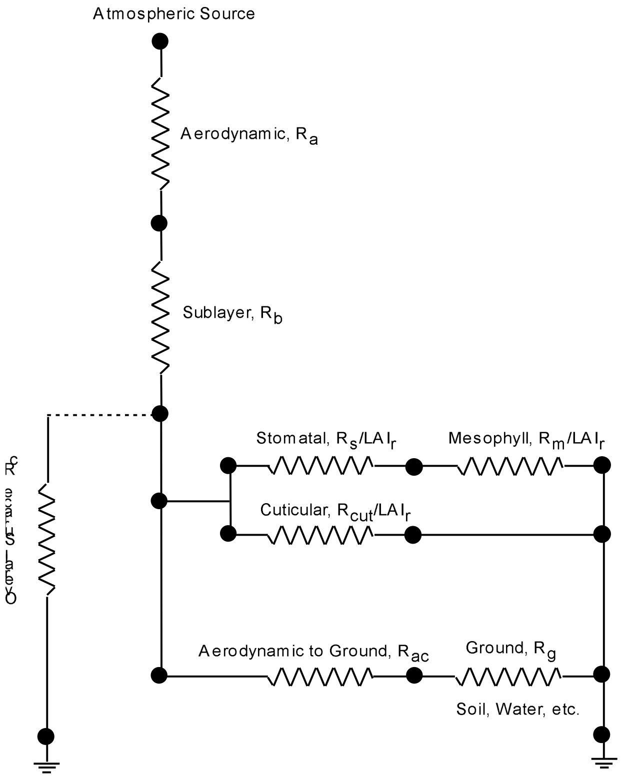
The dry deposition velocity is estimated from:

$V_d = (R_a + R_b + R_c)^{-1}$

The **aerodynamic resistance**, R_a depends on the atmospheric turbulence, which in turn is related to the wind speed, the surface roughness, the boundary layer height and the surface heat flux. The aerodynamic resistance is usually calculated using methods based on Monin-Obukhov similarity. The methods are well-established and have been validated experimentally.

The laminar sub layer resistance is related primarily to the intensity of turbulence in the turbulent boundary layer above the laminar sublayer and the physical properties of the air. Most models apply similar engineering formulae to estimate the laminar sublayer resistance.

Figure C.1. Scheme of major components of resistances used for gaseous dry deposition



Some models assume that the **surface resistance** is constant for a given depositing species and land cover type. Other models break the surface resistance down into smaller elements to represent the resistance through plant stomata, R_s, the resistance to uptake by the mesophyll cells, R_m, the resistance at the outer cuticle of the plant leaves, R_{cut}, the resistance in the plant canopy, R_{ac}, and the resistance at the ground surface, R_g.

The **resistance through the plant stomata**, R_a is the most thoroughly studied of the surface elements and many plant-specific laboratory studies have been carried out. Many of the experimental studies have been based on the transfer of water vapour through the stomata during transpiration and these have then been scaled to other chemical species based on the ratio of the diffusivities of the chemical species. It is not always clear whether the diffusivities in air or water vapour or self diffusivity have been used in calculating the ratios. In general the stomatal resistivity is calculated as

$$R_s = R_i \frac{1}{(f_1 f_2 f_3 f_4 f_5)}$$

where

R_i = minimum stomatal resistance,
f_1 = multiplicative scaling factors for solar irradiance (dimensionless),
f_2 = multiplicative scaling factor for soil moisture (dimensionless),
f_3 = multiplicative scaling factor for leaf humidity (dimensionless), and
f_4 = multiplicative scaling factor for temperature (dimensionless)
f_5 = multiplicative scaling factor for ozone(dimensionless)

The scaling factor $f_{1\text{-}5}$ are in the range 0-1.

The minimum stomatal resistance is the resistance corresponding to ideal growing conditions. It varies throughout the year to take account of the stage in the plant growing cycle, the leaf cover and the extent that the leaves are shaded by other leaves. The factors $f_{1\text{-}5}$ are stress factors that account for the reduction in conductivity during periods of plant stress. Not all models take account of all the stress factors.

The stress factor for irradiance can be calculated from the global position, the time of year and the cloud cover.

Calculating the stress factor for soil moisture requires information on the soil moisture content. The soil moisture content depends on the balance between accumulated rainfall and losses through evaporation, transpiration and run-off. Many models do not include the effects of soil moisture content because calculating the soil moisture content is difficult. However, the effect of soil water stress during dry summer periods can affect peak ozone concentrations.

Calculating the stress factor for leaf humidity requires information on the humidity at leaf level. Usually meteorological models provide information at some reference height above ground but not at leaf level. It is then necessary to infer the humidity at leaf level taking into account the rate of evaporation and transpiration. An iterative calculation may be required.

The temperature scaling factor is usually calculated from the ambient temperature provided by the meteorological model.

Calculation of the ozone stress factor requires an estimate of the ground level ozone concentration. It is difficult to take this into account in an air quality model because the ozone concentration is usually one of the model outputs.

The stomatal resistance may be obtained directly as an output from some meteorological models because the rate of evapotranspiration from plants affects the weather. The air quality model does not then need to make a separate calculation. Some progress has been made to integrate meteorological models and air quality models. Detailed modelling of the effect of ozone stress on transpiration would require the use of an integrated model.

Empirical estimates of the **cutical, ground surface and mesophyll resistances** have been obtained for sulphur dioxide and ozone but data are limited for other species. The available models have tried to scale these data for other chemical species on the basis of their effective Henry's Law constants or judgement of their relative reactivity. However, for many substances, there has been little or no experimental validation. The cutical and ground surface resistances are considerably reduced for sulphur dioxide if the surface is wetted. For ozone, surface wetting reduces the area available for mass transfer and thus increases the cuticular resistance. Some meteorological models calculate when the leaf surface is wetted. The AERMOD-PRIME model provides an algorithm for calculating surface wetness based on the cloud cover and humidity. Others have applied simple rules of thumb based on observations.

The **aerodynamic resistance to ground or in-canopy resistance** is usually calculated using a simple formula from the height of the canopy, the density of the ground cover and the friction velocity. The experimental basis for this formula is very limited.

C.2.3 Dry deposition of particulate matter

The dry deposition velocity for particulate matter is typically calculated using the following resistance model:

$$V_{dp} = \frac{1}{R_a + R_p + R_a R_p V_g} + V_g$$

where

V_{dp} = deposition velocity for particles (m/s),
R_a = aerodynamic resistance (s/m),
R_p = quasilaminar sublayer resistance (s/m), and
V_g = gravitational settling velocity for particles (m/s).

The aerodynamic resistance is calculated using the algorithms used for gases.

The gravitational settling velocity is typically calculated using Stokes Law with a slip correction factor.

Various semi-empirical formulae are used in the models to predict the rate of diffusion of particles through the laminar sublayer. The formulae take into account the Brownian diffusivity of the particles and the intensity of turbulence in the air stream above the laminar layer. The EMEP Unified model (Simpson et al., 2003) uses alternative formulae for vegetated and non-vegetated surfaces because vegetated surfaces are more effective in removing particulate matter.

C.3 Wet Deposition

Many atmospheric reactions of pollutants occur primarily in the aqueous phase in clouds. Wet deposition then occurs when the clouds rain out. Rain falling through the lower atmosphere can wash out further some of the chemical species in the air. The wet deposition flux is given by:

$$F_{wg} = C_l r$$

where

F_{wg} = flux of pollutants by wet deposition
C_l = concentration of pollutant in the raindrops
r = the rate of rainfall

The concentration in the raindrops gradually increases as the drops fall eventually approaching an equilibrium concentration, determined by the solubility of the chemical species in the raindrop and the gaseous concentration of the pollutant.

For many gases of relatively low solubility, the concentration in the raindrop approaches the equilibrium value quite quickly, so that equilibrium is reached before the raindrop hits the ground.

For gases of high solubility, the equilibrium is reached very slowly: the rate of increase in the concentration in the raindrop is controlled by the rates of mass transfer in the gaseous phase to the surface of the raindrop, at the surface of the raindrop and in the liquid phase within the raindrop. In many cases, the gas phase mass transfer rate is the determining step and the rate is almost independent of the depositing species. For these gases, it is convenient to define a washout coefficient:

$$F_{wg} = C_g z_p \Lambda$$

where z_p is the height of the column of air through which the raindrop falls; and
C_g is the average concentration of the gas in the column; and Λ is the washout coefficient.

The washout coefficient is dependent on the rainfall rate:

$$\Lambda = ar^b$$

where the coefficients a and b are typically approximately 10^{-4} s^{-1} and 0.64 respectively.

Aerosols trapped within the cloud will be rained out in the same way as gaseous pollutants. The raindrops will wash out additional particulate material from the air below the cloud as they fall. The equivalent washout coefficient is:

$$\Lambda = \frac{3E_0 r}{2D_m}$$

where D_m is the particle diameter and E_0 is a dimensionless collision efficiency. Various semi-empirical relationships predict the collision efficiency.

Nitrogen dioxide and ozone are sufficiently insoluble that the raindrops are in equilibrium with the surrounding air at ground level. Ammonia, hydrogen chloride and nitric acid are sufficiently soluble that their deposition is controlled by the rate of mass transfer.

In the absence of other substances in solution, sulphur dioxide is sufficiently insoluble that it comes to equilibrium with the surrounding air at ground level. However, it can react with hydrogen peroxide dissolved in the raindrops so that its solubility is initially increased. The quantity of sulphur dioxide washed out is then limited by the initial concentration of hydrogen peroxide in the raindrop.

Wet deposition can be increased in hilly areas. There are two factors here. Firstly, the rainfall is greater in hilly areas near the coast because warm, wet air coming from the sea is cooled as it rises up the slope of the hills, forming clouds saturated with water and with the potential to form rain. This orographic enhancement can lead to substantially increased rainfall. Secondly, reactions of some pollutants, e.g. sulphur dioxide with atmospheric oxidants such as hydrogen peroxide and ozone are much faster in the aqueous phase so that the cloud droplets contain increased concentrations of oxidised sulphur. Aerosol particles such as ammonium nitrate and ammonium sulphate can act as cloud condensation nucleii and become dissolved in cloud droplets. These cloud droplets can then be washed out by rainfall from above the hill cloud. This is known as the seeder-feeder effect.

Annex D: Emissions Inventory Data in Air Quality Models

D.1 Introduction

All air quality models require emission rates to the atmosphere as input to the models. Emission rates come from, or are expected to come from, emission inventory compilations. In many cases these originate from institutes or organisations who are compiling national emission inventories and their primary objectives are to provide annual emission rates for different pollutants and source sectors following internationally agreed procedures and standards. This ensures countries meet their obligations in reporting emissions on a consistent basis to international bodies such as the UN Framework Convention on Climate Change (UNFCCC), the UN Economic Commission for Europe (UNECE) and the European Union. This allows progress of a country to be monitored towards achieving national emission reduction targets set under various Protocols and Conventions, such as the Kyoto Protocol for greenhouse gases, the Gothenburg Protocol to Abate Acidification, Eutrophication and Ground-Level Ozone, part of the Convention on Long-Range Transboundary Air Pollution (CLRTAP) and the EU's National Emissions Ceilings Directive (NECD). Emission inventory compilers must follow guidelines and rules which define what sources are and are not included in a country's emissions inventory and how emissions from literally hundreds of individual sources are grouped together into sectors.

When using a national emissions inventory in air quality models for a particular policy assessment, it is important that the modeller understands the definition of sectors if he or she intends to interpret the results or change the emissions in order to model a policy scenario.

The most accessible and complete emissions inventory covering the whole of the UK is the National Atmospheric Emissions Inventory (NAEI, www.naei.org.uk) which is updated on a yearly basis. Emissions inventories for other parts of Europe are provided through the European Monitoring and Evaluation Programme (EMEP, www.emep.int).

Air quality models require the spatial and temporal distribution of emission rates rather than just an annual emission rate for the whole country. They usually also require a more detailed chemical speciation of the pollutants provided by the inventory. Whilst these are available from the NAEI and EMEP, it needs to be recognised that by their very nature, these inventories generally use a top-down approach to spatially distribute national emission estimates calculated using national statistics and a variety of assumptions may need to be made to do this that do not take into account all the local variations in types of activities and emission factors. It follows that the spatial and temporal variability in emissions are not known to the same degree of certainty as the national annual estimates.

In some instances, models may be able to use local inventories, for example the London Atmospheric Emissions Inventory (LAEI, see Mattai and Hutchinson, 2008), but even these may need to make generalisations based on national data, for example that the age mix of road vehicles on roads in the local area is the same as the national average, simply because the local information is not available. There is the added difficulty of knowing whether local inventories are all constructed to the same standards using consistent methodologies and emission factors.

D.2 Choice of Emission Inventories and Version Control

In any model intercomparison, it is important that modellers state which inventories they are using (e.g. the NAEI, EMEP). If the objective is to compare the performance of models to assess their scientific basis, for example in the way it treats transport, meteorology, atmospheric chemistry or dry deposition, then all models should use the same emissions inventory to remove at least one degree of freedom in the model intercomparison.

The NAEI (for the UK emissions) and EMEP (for the rest of Europe) should be used wherever possible. The exception might be where the intercomparison was being applied to models calculating air quality in London in which case the common use of the LAEI would be acceptable.

Modellers must also state which version of inventories they are using. This is important because inventories are updated periodically, not only with the calendar year which the emissions refer to that reflect actual levels of activities and emission factors in that year, but also because of methodological changes. The NAEI provides an updated, but consistent time-series of emissions each year – the time-series reflects real changes in emissions for each year in the inventory, not changes in methods used. New spatially disaggregated emission maps are generated each year for the latest year in the time-series, but the NAEI does not re-map all earlier years' emissions.

Modellers may want an emissions map for an earlier year (say for 2003) that is consistent with the most recent emissions map produced by the NAEI (say for 2006). Rather than using a potential "out-of-date" map for 2003, modellers could apply some scaling factors to 2006 mapped emissions from a given source based on trends (from 2003 to 2006) in the national estimates in emissions from that source to ensure consistency with the latest inventory, though for major point sources, it might be possible to have available the actual emission rates for these sources in that year (2003) providing they are still deemed to be reliable.

Modellers should state not only the calendar year of the inventory (i.e. in what year the emissions refer to) to use in conjunction with, say, the meteorology of the same year, but what inventory version year it is based on. Modellers should state something like "Emissions are for 2003 based on the 2006 version of the NAEI" rather than simply "the 2003 NAEI". The version of EMEP data for European emissions should be given the same attention, i.e. the year of the emissions and the version of the inventory from which it is derived or the date from it which it is extracted from the EMEP website. These details are given in the NAEI and EMEP websites.

When models are being used to forecast air quality in future years, then they will need to use emission projections. Again, it is essential that modellers state which version of the NAEI projections are being used. These details are always provided by the NAEI in terms of what the inventory base year is, what energy projections are used (e.g. UEP32) and what DfT traffic projections are used.

Emission projections for sources in Europe can be taken from various sources on-line to apply to the EMEP gridded inventories, e.g. the EMEP site itself or the IIASA RAINS/GAINS projections might be used (Amann et al, 2004). The latter are periodically updated and are available for different emission scenarios, so it is vital that modellers state which version of the EMEP or IIASA projections are used by referring to the version names provided.

Natural sources of VOC emissions are not covered in national inventories and models usually calculate them internally using land-cover datasets, emission potentials and environmental correction factors that depend on meteorological conditions. These show great spatial and seasonal variability and differ from year to year depending on weather conditions. Modellers need to state the sources of biogenic emissions used or the land-cover, vegetation species, and meteorology data used and emission potentials and methodologies used to calculate them in their models.

Emissions from sporadic or irregular events such as forest fires or accidental releases are not usually captured in emission inventories. If these events are known to have occurred in the period of time being modelled, modellers should state that they have not been included or if they have been treated, then how.

D.3 Uncertainty in National Emissions Inventories

The previous sections have indicated how important it is that modellers state the **inventory sources and versions** used when models are compared against each other. Preferably, all models should use the same inventory versions. If comparing the results of air quality models against observations, it is also important to understand the **uncertainties** in the emissions inventories and what factors are contributing to them. This needs to be understood in terms of the national emissions inventory (i.e. the annual emission rates) and in terms of their spatial and temporal variability.

The NAEI reports uncertainty estimates in the annual rates of emissions that it reports at the national level. The NAEI makes quantitative estimates of uncertainties based on calculations using a direct Monte Carlo simulation technique corresponding to the IPCC Tier 2 approach recommended for greenhouse gases and also the methodology proposed in guidance produced by the UNECE Taskforce on Emission Inventories. This is explained in detail in the National Greenhouse Gas Inventory Report (Choudrie et al, 2008) and by Passant (2003). Taken from the 2006 version of the NAEI (Dore et al, 2008), Table D.1 shows the levels of uncertainty in total UK emissions of the main air quality pollutants.

Table D.1. Uncertainties in the estimates of UK emissions in 2006

	Estimated uncertainty (%)
NO_x	±10%
PM_{10}	-20 to +30%
NMVOCs	-9 to +10%
SO_2	±4%
NH_3	±20%
CO	-20 to +30%

These estimates, while relevant to the national totals that the UK has a legal requirement to report to the various international bodies, gives a rather false impression when it comes to uncertainties relevant to emission rates used in air quality models. The uncertainties in emission estimates for specific sectors and at specific locations and times will be considerably higher than this, but are much more difficult to quantify. There are several reasons for this and it helps to understand these reasons when considering uncertainties in emission inventories used for modelling a particular air quality issue, whether it be ground-level ozone, acidification and eutrophication or urban air quality, even if the uncertainties cannot be quantified. These will be considered in terms of the spatial variability, temporal variability and chemical speciation in the emissions.

D. 4 Spatial Variability in Emissions and their Uncertainties

The national emissions inventory is largely based on national statistics which are known with high levels of accuracy, for example the total consumption of coal, gas or petrol. Comprehensive traffic surveys conducted by DfT provide the total number of vehicle kilometres travelled by vehicles in the UK and are known with reasonably good levels of accuracy. The NAEI 1 km x 1 km maps are produced by spatially disaggregating national emissions using traffic flow data on individual roads and by using proxy statistics such as household and population census, land use and employment data. In reality, with the possible exception of traffic, levels of domestic, industrial and agricultural activity occurring in every square kilometre are not known with any accuracy. On the other hand, emissions from major point sources such as power stations and refineries can be spatially assigned with reasonably high levels of accuracy because we know the location of the source and the operators themselves may provide specific emissions data for their operations on the site, either from measurements or calculated from known activity levels at the site (e.g. fuel consumption).

In the case of traffic, although we may know the flows of vehicles on individual roads, even broken down by vehicle type, there is seldom information to break the flows down further by fuel type or engine size and age of vehicle. For example, the flow of cars along a road section may be known, but not the mix of petrol and diesel cars and the age distribution which determines the mix of Euro emission standards. These data can only be gathered at national level from licensing statistics.

The traffic emissions are mapped assuming the fleet mix is the same everywhere in the country, with no regional differences. This could actually be a significant source of uncertainty in the emissions used in urban air quality models where, for example, trends in roadside NO_2 concentrations at a given location have been interpreted in terms of trends in national traffic emissions when in fact the fleet mix in areas like central London have been evolving quite differently to the national fleet. Traffic data cannot separate cars from taxis, but relatively high levels of taxi activity, which mainly have diesel engines, are evident in central London while buses in the area are known to have been renewed or equipped with diesel particulate filters. The NAEI is currently incorporating some of these very specific traffic features for London, but area-specific activities like this in all other cities are not currently known.

What all this means is that the accuracy of the spatial distribution of emissions will vary from one pollutant to another according to the relative importance of point source emissions, which might be known reasonably well, and the line and more diffuse area sources such as domestic combustion and many fugitive sources which are known with less accuracy. So for a pollutant like SO_2, the spatial distribution may be known reasonably well because it is dominated by major point sources, but for VOCs and NH_3 which are dominated by more diffuse area sources, the distribution is known with much less accuracy.

In providing the 1 km x 1 km maps of UK emissions, the NAEI has considered the quality of the maps in terms of the contribution of point and area sources to mapped emission totals for each pollutant. These contributions are shown in Table D.2.

Table D.2. Contribution of point sources to UK mapped emission totals in the NAEI (2006). Taken from the report by Bush et al (2008).

Pollutant	Points sources (%)	Area sources (%)
CO	24%	76%
NH3	2%	98%
NMVOCs	20%	80%
NOx	32%	68%
PM_{10}	20%	80%
SO2	78%	22%

From this table it can be seen how 78% of the SO_2 emissions in the UK are from point sources, while only 2% of NH_3 emissions are from point sources, the remainder occurring from area sources.

Sources where the spatial disaggregation of pollutant emissions are particularly uncertain are:

- VOCs emissions from solvent use and from biogenic sources
- NO_x and SO_2 emissions from domestic combustion, off-road machinery and shipping
- PM emissions from domestic combustion, off-road machinery and shipping and activities such as construction, agriculture and other fugitive releases of dust
- NH_3 emissions from agricultural activities

Uncertainties in the spatial distribution of emissions are influenced not only by lack of knowledge about what activities are occurring in specific places, but also by the methods used to calculate the emissions. In many cases, emissions are calculated by combining an emission factor with an activity data and if you know where the activity is occurring, the emission can be estimated with the same level of certainty as can the emissions from the same source at national level. So for example, the method used to calculate emissions from domestic combustion is as valid for a particular grid square as for the whole of the UK provided the amount of domestic combustion activity (i.e. fuel consumed) is known.

In the case of road transport, this is not necessarily the case. The national inventory combines vehicle kilometre data with grammes per kilometre emission factors derived from functions that relate emission factor to average vehicle speed. Further methodologies are used to calculate cold start excess emissions and evaporative emissions from vehicles. All these methodologies are designed to do the job well for application to national scale inventories, but they can break down if applied at a much more local level. For example, using average speed-related emission functions to calculate emissions on the whole of a national road network is fine, but using them with average speed data to calculate emissions occurring at a specific road junction is not strictly valid because of the transient nature of the emission processes at the location resulting from the mix of acceleration, deceleration and idling conditions occurring there. For this microenvironment level, more complex modal emission models are required to account for the variation in emissions over different parts of the vehicle operational cycle. Local cold start emissions need information on local parking durations, ambient temperatures and lengths of trips. Thus, road transport emissions used in urban air quality models are to some extent compromised by the necessary simplifications and empiricism of the methodologies and emission factors used to calculate them.

Mapping emissions from certain sources such as off-road machinery and construction activities is particularly problematic because of the transient nature of the activities. The activity may occur in a certain area for a certain period of time, then move on. Agriculture dominates NH_3 emissions, but while the total number of livestock may be known, spatially distributing the emissions depends on knowledge of farming practices and livestock movements in different areas.

Spatially resolving the biogenic emissions of VOCs might appear to have the advantage that major forests are largely fixed positionally, but here the spatial pattern depends on understanding the classes of vegetation spread around the country and use of fields for different types of crops can change.

Emissions from shipping have until recently not been well-characterised, but are making an increasingly large contribution to emissions impacting on local and regional air quality and acidification. This is because fuels and emissions from ship engines have not been subject to the same level of control and regulation as have those from industry and road transport, so the relative contribution made by shipping to national NO_x and SO_2 emissions has been growing.

Part of the problem in defining an inventory for shipping emissions has been largely because it has been foreseen as an international emissions problem since it occurs across national boundaries and sea territories. Following national inventory reporting rules, emissions from international shipping are excluded from national totals, so this has meant that shipping has contributed a relatively small proportion to the emissions that individual countries report, yet of course the emissions are occurring, and contributing to local and regional air quality regardless of who is politically responsible for those emissions. The problem has required a centralised shipping inventory approach and the situation has been improving with shipping inventories developed by Entec for the European Commission and more recently for Defra (Entec, 2005, 2008). The problem with developing a spatially resolved inventory for shipping has been having reliable activity data indicating where different types of shipping movements occur. Entec has recently used actual shipping movement data held by Lloyd’s Registry to map out air quality pollutant emissions around UK waters so it is hoped this situation will be improved.

Although it is not possible to quantify the uncertainties in the spatial distribution of emissions in terms of confidence levels, the NAEI has developed a fairly sophisticated approach to provide an overall data quality confidence rating for each pollutant map (Bush et al, 2008). A rating index is assigned to each pollutant-source combination from 1 (highest quality) to 5 (lowest quality). Then, an overall 'confidence rating' can be calculated for each pollutant map thus:

$$\mathbf{R = \frac{\Sigma_{\mathit{i}} (E_{\mathit{i}} \,.\, R_{\mathit{i}})}{E_{\mathit{T}}}}$$

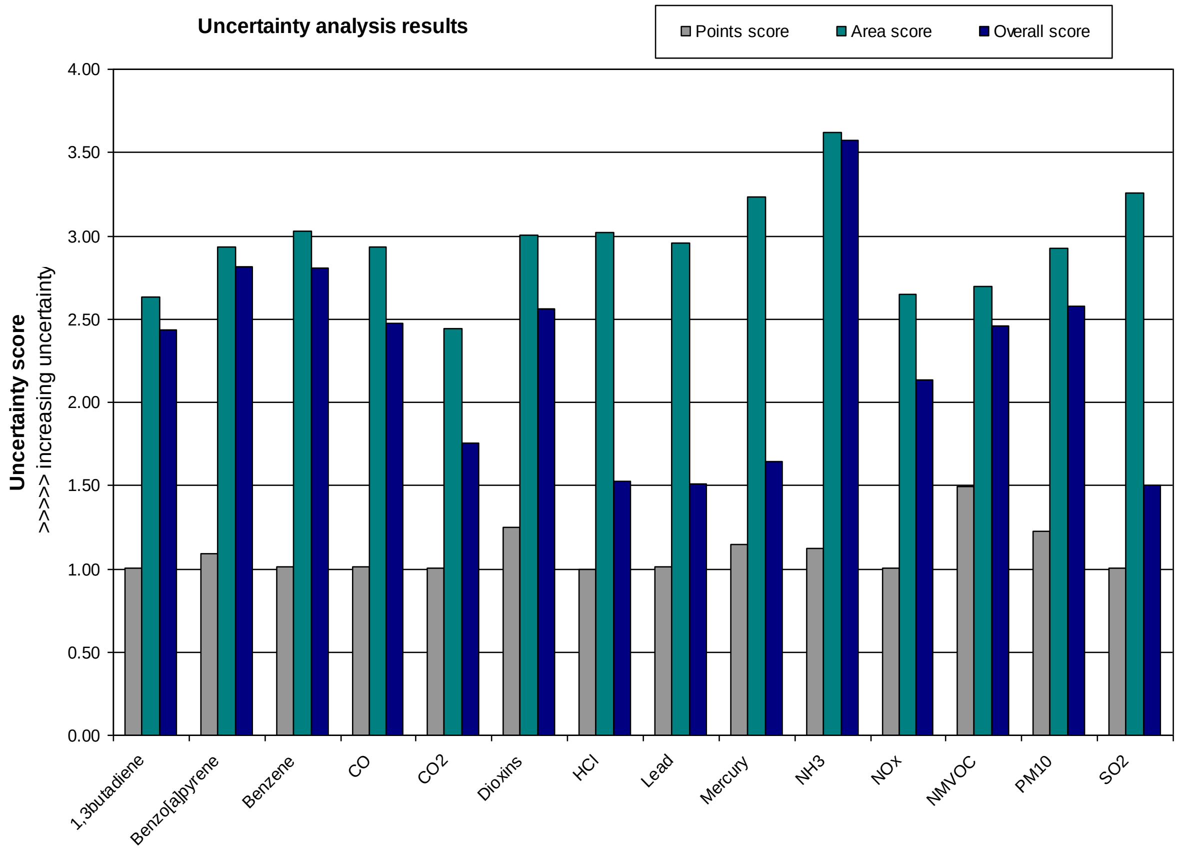
where R is the overall confidence rating for a given pollutant, E_i is the emission estimate for source *i*, R_i is the data quality ratings applied to the mapping of emissions for source *i* and E_T is the total emissions of the pollutant.

Some general rules of thumb were followed when choosing data quality ratings for mapping procedures. Point source data from the Pollution Inventory, industry or regulators are given a rating of 1 because the locations of emissions are 'known' precisely. Modelled point source data were given a quality rating of 2 to reflect the fact that, although all point sources are known, there is uncertainty regarding the distribution of emissions over these sources. Mapping of area sources is by use of various 'grids' which are spatially resolved data such as traffic flows, population or employee numbers. Quality ratings for each set of area/line sources were allocated following assessment of:

- The quality of the spatially resolved data used to make the grid; and
- The reliability of the grid as a measure of emissions from a source.

A rating was applied to each of the above and the mean is used as the data quality rating for that set of area source data. For example, a grid based on 2001 census population data has been allocated a rating of 2 since it is based on accurate census data which is generalised across the 1km grid resolution. The use of such a grid to map emissions from decorative paint use is considered appropriate and was assigned a rating of 1. The area source data for decorative paints therefore has an overall quality rating of 1.5. On the other hand, while a grid based on suburban land cover is also good quality and assigned a rating of 2, its use to map emissions from small scale waste burning (bonfires) is considered much less reliable and is given a rating of 4. Area source data for these emissions have an overall quality rating of 3. Figure D.1 shows the resulting confidence ratings for all the NAEI pollutant maps.

Figure D.1 Confidence ratings for mapping elements of the 2006 NAEI (taken from Bush et al, 2008).



From this, it can seen how the overall quality ranking is good for SO_2 at around 1.5, but poor for NH_3 at about 3.6. Among the key air quality pollutant emissions relevant to the modelling of ground level ozone, acidification and eutrophication and urban air quality, the ranking in order of decreasing mapping quality is:

$SO_2 > NO_x >$ NMVOCs $> PM_{10} > NH_3$.

Whilst not quantitative, this ranking system does at least allow the modeller to appreciate what performance might be expected of the model in terms of the quality of the spatially resolved emissions inventory used by the model.

The EMEP inventory for Europe is at a 50km resolution and it brings together emission inventories for many countries on a common basis. It is bound to have a more uncertain spatial representation than say the UK's NAEI.

D.5 Temporal Variability in Emissions and their Uncertainties

Emission inventories and their spatial distribution are only ever provided as annual emission rates, but this is clearly not enough for most air quality models which can require emissions on at least an hourly basis. Models get round this problem by using a temporal profile for each pollutant and main emission source that takes into account the relative variation in activity levels of anthropogenic sources with time throughout the course of a calendar year or for some sources, the variation in meteorological factors such as temperature if this affects emission rates.

The NAEI provides a default set of temporal profiles for each pollutant and 11 source sectors (SNAP 1 level) by time-of-day, day of the week and month of the year. The profiles, developed by Jenkin et al (2000), take into account the temporal variations in activity levels and the impact these have on emissions.

The profiles are known with varying degrees of confidence. The temporal variation in traffic exhaust emissions, for example, is known with reasonable confidence because the diurnal and day of the week pattern in traffic levels is well-characterised, although one might expect some variation on different road or area types, for example in central urban areas and rural motorways. The change in industrial combustion and process emissions with day of the week is probably known reasonably well. High quality data are available on power generation output by time-of-day, day of the week and month which helps define a profile for this sector. Domestic combustion is expected to follow a consistent diurnal and seasonal pattern, though one might expect year-to-year variability according to weather patterns.

The temporal profiles for some of the more diffuse sources are known with much less certainty, just as they are in terms of their spatial distribution. PM emissions from construction activities, for example, are likely to be intermittent and may start or cease at any time throughout the year. The temporal profile for VOC emissions from solvent use (industry and domestic) is likely to be more uncertain, although one might expect a seasonal pattern with higher emissions occurring in warmer months due to evaporation. Similarly, the profile for evaporative VOC emissions from vehicles and fuel handling is expected to be seasonal, but the temperature dependence of evaporative losses is difficult to quantify with any certainty. Agricultural practices are seasonal, so one would expect some seasonal dependence in NH_3 emissions, but again the temporal nature of the emissions is difficult to quantify and is likely to be variable from year-to-year and across different parts of the country.

Biogenic emissions of VOCs depend critically on meteorological conditions and the question here is how well the emission potentials and environmental correction factors combined with appropriately time-resolved meteorological data fields used in models are able to reproduce the variation in biogenic emissions with time-of-day and season, especially under extreme summer episode conditions. It has been estimated that biogenic emissions are uncertain by at least a factor of 4 (AQEG, 2009).

Even for point sources of emissions, where one might expect a fairly accurate annual emission rate following a general industry- or power generation-based temporal profile, it is impossible to know with confidence what emissions from a particular stack might be at every hour of the day, yet these variations could be important in a number of air quality model applications.

The EMEP inventory for Europe brings together emission inventories for many countries on a common basis. Whilst one would expect emissions from the same source across Europe to exhibit a common temporal profile, there are bound to be some differences due to different weather patterns and climatic conditions and industrial and agricultural practices.

D.6 Chemical Speciation in Emissions and their Uncertainties

Besides the spatial and temporal variation in emissions, models may also require a further breakdown by chemical species. Some of the pollutants reported in inventories are actually groups or families of individual chemical species, most notably, the VOCs and NO_x.

VOCs are made up of many hundreds of individual organic species each with different reactivity in the atmosphere. Models of ozone and organic aerosol formation require the breakdown of the reported VOC emissions into as many of these as possible and for each source sector. The NAEI provides a VOC speciation profile describing the relative mass proportions of individual VOCs emitted from each detailed source category. The speciation profile covers 664 individual or groups of VOCs emitted from around 250 different sources (Passant, 2002). The profile is known reasonably well for some sources like solvents because the manufacturer or user knows the chemical composition of the product. It is known less well for vehicle exhaust emissions, but is still expected to be adequate for the most abundant and reactive compounds. For other combustion sources and process emissions, the profile is known less well. However, when their photochemical ozone forming potential is taken into consideration, solvent use and road transport emissions of VOCs make the most significant contribution to episodic peak ozone concentrations (Derwent et al, 2008).

The NAEI profile is held constant each year in relative terms and applied to a new VOC inventory. The profile itself is only periodically updated. The constancy in the profile is a reasonable approximation for many sectors, but changes could occur for example as a result of solvent substitution, process changes such as in the chemicals and food and drink industries, or due to new fuel formulations (e.g. biofuels). This adds to the uncertainty associated with using the profile across a time-series, perhaps especially in forecasting the effect of future emissions on ozone formation.

The application of the NAEI speciation profile, which is UK-biased, to VOC sources in other countries (e.g. by combining with the EMEP inventory of VOCs) is questionable given potentially different processes, fuels and solvents used.

The overall uncertainty in the NAEI VOC speciation profile has not been quantified, however in the context of modelling ground-level ozone, it is probably not as significant as the uncertainties in the spatial and temporal variability of the emissions.

The separation of the NO_x inventory into NO and NO_2 is now understood to be important in modelling urban concentrations of ozone and NO_2. Primary NO_2 emissions influence the presence of local oxidant in urban areas and in recent years has been perturbed by increased diesel vehicle activity and exhaust abatement technologies introduced to curb vehicle emissions. Primary NO_2 fractions given as the mass fraction of NO_x directly emitted as NO_2 have been developed for a range of vehicle categories and other combustion sources of NO_x. These are used in the NAEI, but their uncertainties are large.

Inventories for emissions of primary particulate matter (PM) in the NAEI are broken down into various size classes ($PM_{2.5}$ and smaller size classes) for each emission source. They are usually derived from fractions relative to PM_{10}. They are not broken down into their chemical constituents. The uncertainties in the PM size fractions for many sources are high.

D.7 Concluding General Remarks Concerning Emission Inventories Used in Models

This Annex has so far only given a description of the issues concerning use of emission inventories in models in relation to identifying sources of emissions inventories and in relation to how the spatial and temporal patterns of the emission rates and the chemical speciation are expressed.

Modellers are ***users*** of emissions inventory information and it is seldom, if ever, the responsibility of the modeller to develop an inventory. Modellers can only use what emissions inventory information is available. This Protocol therefore advises modellers to specify the sources and versions of emissions inventory used in their models and specify what spatial and temporal representations of the emissions are used (e.g. which version of the NAEI gridded data) and if relevant, what VOC speciation profiles and primary NO_2 fractions are used.

It is not possible to provide any quantitative levels of uncertainty in the emissions other than in the national emission totals (expressed as an annual emission rate) and in terms of a qualitative ranking in the confidence levels of the spatial disaggregation of the emissions. However, the previous sections should give the modeller an appreciation of what the major sources of emissions uncertainty are relevant to the air quality problem being addressed, whether it is ground-level ozone, acidification and eutrophication or urban air quality, to help them understand and interpret the comparison of their modelled results with observations. In other words to answer the question “What in the emissions inventory could be causing our models to under or overpredict concentrations?” If all models systematically under- or overpredicted concentrations using the same emissions inventory information, this could indicate a systematic bias in the emissions, though this is by no means a certainty. This is in itself useful for the compilers of emissions inventories to know. Conversely, variances in the directional differences between modelled and observed concentrations among different models makes it difficult to draw any conclusions about the quality of the inventory.

Based on the information given so far in this Annex, a summary is given in the following sections of the sources of uncertainties in the particular emission inventory information relevant to each main modelling area.

D.7.1 Emissions information relevant to modelling ground-level ozone

The uncertainties relate to the emissions of NO_x and VOCs in the UK and Europe. Important issues are:

- Quantities of VOC emissions from road transport, fuel evaporation and solvent use;
- The spatial distribution of these emissions in the UK;
- The temporal variation of these emissions in the UK and Europe;
- The VOC chemical speciation of these sources;
- The quantities of biogenic VOC emissions in the UK and Europe and the spatial distribution and temporal variation;
- The quantities of NO_x emissions from sources in the UK and Europe, including shipping;
- The spatial distribution and temporal variation of these emissions in the UK and Europe; and
- NO_x emissions from traffic in urban areas and the fraction emitted as primary NO_2, their spatial and temporal variation, to account for the ozone decrement and oxidant partitioning in urban areas.

D.7.2 Emissions information relevant to modelling acidification and eutrophication

The uncertainties relate to the emissions of SO_2, NO_x and NH_3 in the UK and Europe. Important issues are:

- Quantities of NO_x and SO_2 from combustion sources in the UK and Europe, including shipping
- The spatial distribution and temporal variation of these emissions in the UK and Europe
- Quantities of NH_3 emission from all sources in the UK and Europe, especially agriculture
- The spatial and temporal variation of NH_3 emissions especially in the UK

D.7.3 Emissions information relevant to modelling urban air quality

Urban air quality is influenced by local sources of NO_x and PM adding to background concentrations of NO_2 and PM_{10} that are influenced by emissions on a larger scale. Models may take background concentrations from observations or other models and calculate the additional contribution from locally emitted sources. Alternatively, they could model the background and local contributions using a nested approach.

- Modelling background NO_2 will be sensitive to quantities of NO_x emissions from sources in the UK and Europe, including shipping, and their spatial and temporal variations
- Modelling background PM_{10} will be sensitive to quantities of PM, NO_x, SO_2 and NH_3 emissions from sources in the UK and Europe, including shipping, and their spatial and temporal variations

All urban air quality models will be sensitive to local NO_x and PM emissions, particularly from traffic sources. Hence, the key uncertainties to consider are:

- The estimation of NO_x emissions from local traffic sources, domestic combustion and off-road machinery
- The estimation of primary NO_2 emissions from local traffic
- The estimation of PM_{10} and $PM_{2.5}$ emissions from local traffic, including emissions from tyre and brake wear, road abrasion and from off-road machinery
- The treatment of road dust resuspension. This is a source that is not included in emission inventories and usually requires more complex dispersion models to estimate its contribution to airborne PM
- The emissions of PM from local domestic combustion, especially coal and biomass burning
- The emissions of PM from fugitive dust sources such as construction

Depending on whether there are major industrial processes, power stations, sea ports or airports in the locality, then emissions from these sources will also have a local influence.