HAL
open science

Systematic reduction of complex tropospheric chemical mechanisms using sensitivity and time-scale analyses

L. E. Whitehouse, A. S. Tomlin, M. J. Pilling

To cite this version:

L. E. Whitehouse, A. S. Tomlin, M. J. Pilling. Systematic reduction of complex tropospheric chemical mechanisms using sensitivity and time-scale analyses. Atmospheric Chemistry and Physics Discussions, 2004, 4 (4), pp.3721-3783. hal-00301336

HAL Id: hal-00301336

https://hal.science/hal-00301336v1

Submitted on 18 Jun 2008

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Atmos. Chem. Phys. Discuss., 4, 3721–3783, 2004
www.atmos-chem-phys.org/acpd/4/3721/
SRef-ID: 1680-7375/acpd/2004-4-3721
© European Geosciences Union 2004

Atmospheric Chemistry and Physics Discussions

© EGU 2004

Systematic reduction of complex tropospheric chemical mechanisms using sensitivity and time-scale analyses

L. E. Whitehouse[1], A. S. Tomlin[1], and M. J. Pilling[2]

[1]Energy and Resources Research Institute, University of Leeds, Leeds LS2 9JT, United Kingdom
[2]School of Chemistry, University of Leeds, Leeds LS2 9JT, United Kingdom

Received: 1 April 2004 – Accepted: 12 May 2004 – Published: 8 July 2004

Correspondence to: A. S. Tomlin (fueast@leeds.ac.uk)

Abstract

Explicit mechanisms describing the complex degradation pathways of atmospheric volatile organic compounds (VOCs) are important, since they allow the study of the contribution of individual VOCS to secondary pollutant formation. They are computationally expensive to solve however, since they contain large numbers of species and a wide range of time-scales causing stiffness in the resulting equation systems. This paper and the following companion paper describe the application of systematic and automated methods for reducing such complex mechanisms, whilst maintaining the accuracy of the model with respect to important species and features. The methods are demonstrated via application to version 2 of the Leeds Master Chemical Mechanism. The methods of local concentration sensitivity analysis and overall rate sensitivity analysis proved to be efficient and capable of removing the majority of redundant reactions and species in the scheme across a wide range of conditions relevant to the polluted troposphere. The application of principal component analysis of the rate sensitivity matrix was computationally expensive due to its use of the decomposition of very large matrices, and did not produce significant reduction over and above the other sensitivity methods. The use of the quasi-steady state approximation (QSSA) proved to be an extremely successful method of removing the fast time-scales within the system, as demonstrated by a local perturbation analysis at each stage of reduction. QSSA species were automatically selected via the calculation of instantaneous QSSA errors based on user-selected tolerances. The application of the QSSA led to the removal of a large number of alkoxy radicals and excited Criegee bi-radicals via reaction lumping. The resulting reduced mechanism was shown to reproduce the concentration profiles of the important species selected from the full mechanism over a wide range of conditions, including those outside of which the reduced mechanism was generated. As a result of a reduction in the number of species in the scheme of a factor of 2, and a reduction in stiffness, the computational time required for simulations was reduced by a factor of 4 when compared to the full scheme.

ACPD

4, 3721–3783, 2004

© EGU 2004

1. Introduction

A large number of volatile organic compounds (VOCs) contribute to the production of secondary pollutants in the lower atmosphere. The relative contribution of VOCs to the formation of specific secondary pollutants varies from one compound to another by virtue of differences in their reactivity and structure. Explicit mechanisms describing complex sets of degradation pathways are therefore required in order to evaluate the individual contributions of each VOC to secondary pollutant generation. The description of such complex pathways however, requires a large amount of mechanistic detail and leads to chemical schemes with very large numbers of species and reactions. Such explicit mechanisms are not easily employed within complex 3-D atmospheric dispersion models, particularly within the Eulerian framework, since large numbers of coupled species within each grid cell are computationally extremely expensive to solve. In recent years however, a number of techniques have been developed which aim to reduce large explicit mechanisms to more compact chemical representations in a systematic way, such that the accuracy of the reduced scheme is not substantially degraded when compared to the full scheme. Such methods include species lumping, sensitivity analysis, computational singular perturbation theory, low dimensional manifolds and repromodelling (Tomlin et al., 1998; Lam and Goussis, 1988; Maas and Pope, 1992; Turányi et al., 1993).

The systematic application of a number of these techniques may lead to chemical representations that accurately describe the behaviour predicted by fully explicit schemes, but at a reduced computational cost. For extremely complex schemes an additional requirement is that such reduction methods can be applied in an automatic way according to well defined principles. This paper, and the following companion paper, attempt to present such a systematic and automated reduction procedure that may be applied to any explicit atmospheric chemical mechanism. The procedure involves the application of a range of sensitivity analysis techniques followed by the lumping of the reaction mechanism based on the underlying chemical timescales. The methods are il-

© EGU 2004

lustrated via application to version 2 of the Master Chemical Mechanism MCM, (Jenkin et al., 1997). Section 2 of the paper gives a brief introduction to the MCM. Section 3 describes the formulation of the trajectory simulations that are used to define the conditions under which the reduced mechanisms have been derived and compared against the full mechanism. Section 4 describes the application of sensitivity based methods for reduction of the mechanism and Sect. 5 the application of time-scale based methods such as the Quasi-Steady State Approximation (QSSA). In Sect. 6 the application of perturbation methods for assessing the number of variables required to represent the dynamics of the full system is presented, along with an analysis of the numbers of slow, intermediate and fast variables at each stage of the reduction. In Sect. 7, comparisons between the full and reduced mechanisms are shown for a range of conditions. In Sect. 8, a discussion of the methods and results is given. Final conclusions are drawn in Sect. 9.

2. The Master Chemical Mechanism

The Protocol for the Master Chemical Mechanism (MCM) was developed in Jenkin et al. (1997) for investigating the production of ozone from a large range of VOCs. The mechanism was designed to describe the complete tropospheric oxidation of 124 VOCs and to accommodate a full range of VOC:NO_x ratios. The VOCs which are degraded in this mechanism were selected on the basis of available emissions data and provide approximately 91% mass coverage of the emissions of uniquely identifiable chemical species (Rudd, 1995). The MCM was designed to be applied in areas where detailed chemistry is required, e.g. in predicting the generation of intermediates such as multifunctional carbonyls, hydroperoxides and nitrates etc. for which field data are now becoming available. The protocol is summarised in Fig. 1.

The methodology is to start with the compounds with the lowest number of carbon atoms, e.g. CH_4, CH_3OH, HCHO, and work upwards. Each VOC is treated separately and its degradation is continued until a species is formed which has already appeared

© EGU 2004

and been treated in an earlier (low C number) VOC scheme. The reactions included in this protocol are classified as OH radical initiation reactions for all VOCs together with O_3 initiation reactions, NO_3 radical initiation reactions and initiation by photolysis where appropriate. Reaction of organic radicals, peroxy radicals, oxy radicals, Criegee bi-radicals and reactions of degradation products are also included, see Fig. 1.

The detailed kinetic modelling of these chemically complex processes requires the construction of a comprehensive mechanism comprising all the possible elementary reactions including the respective rate coefficients. The availability of kinetic and mechanistic data relevant to the oxidation of VOCs has increased significantly in recent years and various aspects of the tropospheric chemistry of organic compounds have been reviewed extensively in Atkinson (1990, 1994); Roberts (1990); Lightfoot et al. (1992). This database can be used to construct detailed degradation schemes for a range of organic compounds. To ensure a consistent approach in developing the mechanisms, a set of detailed criteria for defining the degradation schemes has been formulated by drawing upon all current sources of information. Experimental rate data are used where possible, and where no data are available comparison with similar compounds is used. Failing this, well founded assumptions, approximations and estimations are made based on this experimental database.

The rate coefficients for reactions of OH with various organic compounds were taken from either the IUPAC kinetic evaluations in Atkinson et al. (1992) or from Atkinson (1994, 1989). In the absence of experimental data, rate coefficients appropriate to 298K were estimated by the Structure Activity Relationship (SAR) method devised by Atkinson in Atkinson (1987) and updated in Kwok and Atkinson (1995) for alkanes, carbonyls, alcohols, ethers and chlorinated hydrocarbons, and by the method from Peeters et al. (1994) for alkenes and dienes. In order to reduce the complexity of the mechanism, a surrogate species RO_2 has been defined in the MCM as the sum of all peroxy radicals. The overall rate coefficients for the reactions of individual peroxy radicals with NO, NO_2, NO_3, HO_2 and the permutation reactions between RO_2s were evaluated based on experimental data in Lightfoot et al. (1992). Parametric forms are

© EGU 2004

based on those recommended in Jenkin et al. (1997).

Photolysis rate parameters depend on the solar zenith angle. Based on two parameter calculations made in Hough (1988) for clear sky conditions at an altitude of 0.5km, a formula was determined in Jenkin et al. (1997) to express the variation of the photolysis rate J with the solar zenith angle θ,

$$J = I\,(\cos\theta)^m\, e^{-n\sec\theta}, \tag{1}$$

where the three parameters (I, m, n) are optimised by fitting measured data (Jenkin et al., 1997). The solar zenith angle, θ, is calculated as a function of the time of day (represented by the local hour angle, LHA), the time of year (represented by the solar declination angle DEC) and the latitude (LAT) (Derwent and Jenkin, 1990):

$$\cos(\theta) = \cos(\mathrm{LHA})\cos(\mathrm{DEC})\cos(\mathrm{LAT}) \tag{2}$$

$$+\sin(\mathrm{DEC}) + \sin(\mathrm{LAT}). \tag{3}$$

Initial studies of ozone formation, in comparison with a previous chemical mechanism (Derwent et al., 1996), were found to give good agreement. More recent studies have also been carried out by evaluating the MCM against smog chamber and observational data in Jenkin et al. (1999). The overall results are very encouraging, and problem areas have been identified, which provide the impetus for further reviews and updates (Pilling et al., 1999).

The fully explicit MCMv2.0 was chosen for the investigation of reduction techniques in this work. The full mechanism contains 3487 species and 10763 reactions. The aim is to generate a set of mechanisms of different levels of reduction, and therefore different accuracies, to enable a choice of mechanisms available to potential users depending on the level of detail and accuracy required. Clearly the larger the mechanism the more computationally expensive it will be to solve. At the culmination of the reductions a comparison of accuracy and computational speed will be presented for each reduced mechanism.

Since the application of reduction methods in this work, an updated version of the MCMv3. has been developed. Major improvements have been made to the reaction

© EGU 2004

mechanisms of the aromatic compounds (Saunders et al., 2003; Jenkin et al., 2003). These are an important class of ozone producing VOCs. Revision of the photolysis rates has been carried out, especially with respect to multifunctional carbonyl products formed in the aromatic systems and the MCM has been expanded to include the important biogenic compounds α and β pinene. Although the work in this project was carried out on MCM v2.0, it is still relevant as the advances made in the new version are mainly in the aromatics with which we were not greatly concerned in this instance. Furthermore this work presents the development of systematic methodologies that can be generally and automatically applied and would therefore be relevant to the reduction of other mechanisms including any updated versions of the MCM.

3. Simulation conditions for mechanism reduction

3.1. Design of scenarios

For the purposes of applying the reduction methods, the MCM will be studied within many box model simulations, since this is computationally less expensive than a full grid calculation which would not be easily achievable using the full scheme. The selection of the box model scenarios is designed to ensure that the reduction of the MCM is valid over a wide range of urban conditions. The chemical rate equations of the system can be represented in the form

$$\frac{dc_i}{dt} = f(\boldsymbol{c}, \boldsymbol{k}) + E_i, \tag{4}$$

where $\boldsymbol{c}$ is the vector of the species concentrations, $\boldsymbol{k}$ is the rate coefficient vector and E_i is the emission into the box for species i.

The range of concentrations chosen to represent a typical urban atmosphere were taken from measurements in Birmingham during the PUMA campaign (PUMA, 2001), from Middlesbrough using data from the AURN site (AEAT-ENV) and from the AURN

© EGU 2004

site at Marylbone Road in London. From these data it was assumed that total NO_x varies between 0 ppb and 500 ppb, with a VOC/NO_x ratio within the range 0.25 to 8. A range of 33 box model scenarios was designed with appropriate emissions and initial conditions to simulate this range of concentrations. The full set of 94 trajectories was obtained by using three different initial conditions for ozone. The selected conditions of each trajectory are given in Table 1.

3.2. Selection of emissions profiles

Using data from the National Atmospheric Emissions Inventory (NAEI) and the West Midlands Inventory, relative emissions factors for VOCs in the mechanism have been calculated for a typical urban region. Each database contains eight source categories to which all emissions are ascribed. The percentage emission fractions for all the primary VOCs in the MCM for the National Emissions Inventory for the UK was obtained from NAEI (2000) for the different categories shown in Table 3. The relative sectors show different contributions for urban areas such as the West Midlands compared to the UK totals. In urban areas road transport has a much higher contribution than within the national average. Sources from solvent use are also higher in the West Midlands than the national average, with industry and process sources much lower.

Unfortunately, the data for emissions for speciated VOCs was only available for the UK inventory, so it was necessary to use the comparison between the relative amount for each emissions source to calculate the contributions in the West Midlands for each VOC within the MCM. Each UK total per emission source for a species was multiplied by the relative contribution of that source in the West Midlands. This gives the relative contribution of that species for a particular emission source category in the West Midlands. If these are summed across all emissions sources the total relative emission per species can be calculated for the West Midlands. This means that the total relative

© EGU 2004

contribution of a VOC to the total emission is given by:

$$\left(\sum_{\text{emissions sources}} \text{West Midland rel. contribution} \times \text{rel. national contribution of source} \right) \quad (5)$$

The full set of emissions factors used for the MCM box model simulations resulting from this procedure are given in Table 2. Total VOC and NO_x emissions are selected for each trajectory and these are given in Table 1. These values were selected in order to provide the range of VOC:NO_x required to present an appropriate span of conditions found in UK urban areas. In Table 1 it can be seen that some trajectory conditions do not have a number and were therefore not used in this investigation. Some of the maximum values generated by these trajectories lie outside those limits prescribed earlier and were therefore not used within the mechanism reduction strategy. Table 4 gives the initial conditions of those species with a non-zero initial concentration.

3.3. Simulation conditions

Each simulation was run for a 72 h period starting at midnight with variations in both temperature and solar flux for a typical diurnal cycle in July at a latitude of 50° N. The temperature variation throughout the day was given by

$$T = 289.86 + 8.3 \sin\left(\left(7.2722 \times 10^{-5} t - 1.9635\right)\right) K, \quad (6)$$

where t is the time of day in seconds (Heard et al., 1998).

The trajectories selected all show a characteristic diurnal variation in ozone concentrations associated with species formed as result of photolytically initiated reactions. At sunset the reduction in photolytic activity and the reaction of ozone with NO leads to a reduction in the concentration of ozone.

Suitable time-points from the trajectories described above were used during the reduction procedure. These were selected over the second and third days of the simulations following the build up of intermediate compounds. Points were taken every 3 h from the 24 h point until the end of the third day.

ACPD
4, 3721–3783, 2004

© EGU 2004

This provides good data coverage over areas of swift growth and decay, but also does not neglect the night-time areas as shown in Fig. 2.

Three trajectories have been chosen to plot at each reduction stage so that the deviation in accuracy from the full scheme can be seen. These trajectories were selected as ones which typically represent realistic urban conditions with maximum ozone levels of less than 100 ppb and a variety of NO_x conditions from 300 ppb down to around 7 ppb.

3.4. Selection of important species

The following species were determined to be of interest and so were designated important species which should be accurately represented by the reduced schemes: O_3, NO, NO_2, HNO_3, NO_3, HONO, HNO_3, HO_2NO_2, HCHO, OH and the sum of all peroxy acyl nitrate (PAN) species, examined as total PAN. O_3, NO_x and total PAN were selected as they are major pollutants in urban environments and have a significant impact on human health and the environment. The remaining species are included as they are known to have extensive involvement in the mechanisms of the three major species or other important tropospheric processes.

4. The application of sensitivity analysis

Mechanism reduction techniques are well suited to situations where a large mechanism such as the MCM is to be applied to a specific set of environmental conditions. Reduction of the mechanism requires the identification of a smaller subset of reactions that can describe the chemistry of the mechanism to within a prescribed accuracy and therefore results in a less computationally expensive system. It is therefore a logical first step to reduce the number of chemical species, as the majority of computing time is spent solving their rate equations. Semi-automatic techniques based on sensitivity analysis have been used to derive greatly simplified chemical schemes which accu-

© EGU 2004

rately represent the behaviour of chosen key species (Turányi et al., 1993; Turányi, 1990b; Tomlin et al., 1992). As a result of the application of these techniques, redundant reactions and species can be removed from the mechanism as described in Tomlin et al. (1998); Heard et al. (1998); Zeng et al. (1997). In Heard et al. (1998) a number of systematic reduction techniques were applied to CBM-EX, an extended version of the carbon bond mechanism in order to ascertain their applicability to atmospheric systems. Similar methods were applied in Zeng et al. (1997) for the reduction of a subset of the MCM describing the tropospheric oxidation of butane. The systematic reduction of large kinetic systems using automated and slightly modified versions of such techniques will be briefly described in this section and subsequently applied to MCMv2.0.

4.1. Identification of redundant species

The aim of most kinetic modelling studies is to describe properly the concentration changes of certain species or key reaction features which are considered important (Turányi et al., 1993). The reduced model must generate values for the important species and features which are in agreement with those generated by the full scheme within a specified accuracy. It is necessary to identify the redundant species which can be excluded from the model, and necessary species which must be included in order to produce accurate results for the important species. The necessary species will be coupled to the important ones by significant reactions. Redundant species will be those which are not coupled in any significant way to the set of important and necessary species and so can be excluded from the model with only negligible effects on the important concentrations both directly and indirectly.

Redundant species can be identified via the investigation of the system Jacobian (Turányi, 1990b). A change in the concentration of a species affects the rates of those reactions in which this species is a reactant. A direct effect will be a change in the concentration in those species which are reactants or products of those reactions. The primary concentration changes then may cause a change in the concentration of other

© EGU 2004

species. A species may be considered redundant if its concentration change has no significant effect on the rate of production of important species. Following operator splitting the chemical rate equations form a system of ODEs:

$$\frac{d\boldsymbol{c}}{dt} = f(\boldsymbol{c}, \boldsymbol{k}), \qquad \boldsymbol{c}(0) = \boldsymbol{c}_0, \tag{7}$$

where $\boldsymbol{c}$ is the n dimensional species concentration vector, $\boldsymbol{k}$ is the vector of rate parameters. The influence of the change of concentration of species i on the rate of production of an N-membered group of important species is given by the sum of squares of normalised Jacobian elements,

$$B_i = \sum_{n=1}^{N} \left(\frac{\partial \ln f_n}{\partial \ln c_i} \right)^2 . \tag{8}$$

The higher the B_i value of a species, the greater its direct effect on important species. However, necessary species influence the concentration of important species through indirect coupling. Thus, the group of necessary species must be identified using an iterative method, where the highest ranked species on the basis of the B_i values are also considered in the summation for each subsequent iteration. Usually, a suitable threshold is chosen, often where a large gap in the spectrum exists, and the procedure repeated until no new species are admitted into the N-membered group. The redundant species are those which are not included in the summation after the final iteration at any of the selected reaction times. The major problem with this procedure is that the threshold limits for the B_i values are somewhat arbitrary for cases where no obvious gaps exist. For a complex mechanism like the MCM, with several thousand initial species, this problem arises and a continuous spectrum of B_i values is found for most conditions. Some modifications to previous methods were therefore required for reduction of the MCM.

In this work, after each iteration the top 10 species were added to the group of important species. The iterations were continued until there were 1500 species in the

ACPD
4, 3721–3783, 2004

© EGU 2004

group. A range of values was tested and 1500 was selected as a value at which a large number of species could be removed whilst continuing to preserve the accuracy of the model. The method therefore allows the identification of the most important 1500 species at each time point. The procedure was carried out at each of the chosen time-points for each trajectory. When this was complete all those species occurring in any of the lists of 1500 species were added to a master list of necessary species. Ultimately more than 1500 species were retained since any species which was important at even just one time-point of one trajectory must be kept in order to preserve the accuracy of the system over the wide range of trajectories selected. At this stage 396 species and 1224 reactions were removed from the mechanism. This reduced model will be referred to as Stage 1 in the following discussion.

4.2. Identifying redundant reactions through local rate sensitivities

In Turányi et al. (1989) a method was developed to identify redundant reactions based on the investigation of local rate sensitivities. The analysis measures the effect of a perturbation in a rate parameter on the rate of production of a necessary or important species (Heard et al., 1998). In Turányi et al. (1989) it has been shown that in such a localised first-order approach, the normalised rate sensitivities can be more efficiently used than concentration sensitivities, as algebraic expressions for the elements of the rate sensitivity matrix can be defined. The normalised rate sensitivity matrix F_{ij} can be calculated from:

$$F_{ij} = \frac{k_j}{f_i}\frac{\partial f_i}{\partial k_j}, \tag{9}$$

where f_i can be defined as,

$$f_i = \sum_j v_{ij} R_j. \tag{10}$$

ACPD
4, 3721–3783, 2004

© EGU 2004

Here v_{ij} is the stoichiometric coefficient of species i in reaction j and R_j is the rate of that reaction. Hence

$$F_{ij} = \frac{v_{ij} R_j}{f_i}. \tag{11}$$

4.3. Principal component analysis

The sensitivity matrix generates a very large amount of information that must be evaluated at many time-steps, as it is essentially a local calculation. Consequently, useful information is obtained from F_{ij} using either overall sensitivities or principal component analysis, which calculates the effect that changing a parameter such as a rate constant has on the concentration or rate of production of several or groups of species (Vajda et al., 1985; Tomlin et al., 1998). In this way principal component analysis provides a method for selecting a reaction set from the sensitivity data by examining reaction interactions and eliminating groups of reactions which do not significantly affect the concentration of important or necessary species rather than individual ones. Principal component analysis is achieved via the eigenvalue/eigenvector decomposition of the matrix $\tilde{\mathbf{F}}^T \tilde{\mathbf{F}}$. The eigenvectors reveal the related parameters, and the corresponding eigenvalues express the weight of these parameter groups (Vajda et al., 1985). This is usually carried out on a sub-matrix related to the important and necessary species rather than all the species of the mechanism. Reactions with large eigenvector coefficients associated with large eigenvalues are the most important reactions in the mechanism. By defining threshold values for significant magnitudes of both the eigenvalues and eigenvectors, the method provides an automatic way of deciding which of the reactions can be eliminated from the reduced model (Heard et al., 1998). Principal component analysis therefore provides a way of summarising the information contained in the sensitivity matrix, and at the same time reveals connections between parameters which would not be apparent when studying the original sensitivity matrix. The number of species included in the mechanism can also be reduced but only if the elimination of

© EGU 2004

reactions causes a particular species to also be eliminated. This is why for mechanism reduction the principal component analysis of the reaction system should be performed in tandem with the sensitivity analysis used to identify important and necessary species within scheme (Tomlin et al., 1998).

4.4. Application of rate sensitivity analysis to the MCM

4.4.1. The use of overall sensitivities

The computational requirements of interpreting the sensitivity information using principal component analysis were found to be high due the costs of manipulating a 9839 × 9839 matrix. In order to overcome this problem the use of an overall sensitivity measure was used prior to the application of principal components. In addition, it was necessary to exploit the sparsity of the system. It was possible via initial investigation of the matrix to determine the number of non-zero entries of the rate sensitivity matrix. These were marked in such a way that subsequent use of the matrix accessed only the entries with non-zero values. In addition, the matrix terms were normalised individually rather than normalising the Jacobian, in order to reduce the computational time of the method. Therefore the overall sensitivity measure is given by:

$$F_j = \sum_i \left(\frac{\partial f_i}{\partial k_j} \frac{k_j}{f_i} \right)^2 . \tag{12}$$

The user then selects a threshold and any reaction whose sensitivity is greater than this is retained within the mechanism. In this case a threshold of 1×10^9 was used over all time-points. Therefore, only those reactions that are not specifically selected to remain in the mechanism at any of the time-points in any of the trajectories may be safely removed.

A further 1429 reactions were found to be redundant following application of overall sensitivity methods. Over 95% of the reactions removed at this stage were peroxy-

© EGU 2004

peroxy reactions and in fact, following the application of overall sensitivities, no reactions of this type remained in the scheme, reflecting the polluted nature of the scenarios investigated, so that RO_2 + NO is fast. Note however, that many of the RO_2 + HO_2 reactions remain. This reduced model will be referred to as Stage 2 in the following discussion.

4.4.2. The use of a reduced set of important species

During Stage 3 redundant species are again identified but using a smaller set of important species. When the PAN species were examined it was found that only 49 of them contributed more than 0.1% to the concentration of total PAN. Any species contributing less than this was no longer considered to be important. This led to a reduced set of 58 important species. Again species which are necessary to this reduced set of important species are identified as described above and only 2454 species are considered to be necessary. At this stage the reaction groups relating to the formation and decomposition of minor PAN species were removed. In addition, reactions of OH with some minor partially oxygenated species (such as alcohols, acids, peracids and hydroperoxides) that generate acyl radical precursors were removed. Following this stage of reduction there are 2454 species and 6927 reactions remaining, showing that reactions describing the chemistry of relatively minor PAN species contribute a fairly large part of the MCM mechanism.

4.4.3. The application of principal component analysis to the MCM

The application of principal component analysis was then carried out for the 2454 necessary species identified in Stage 3. As the principal component analysis involves the decomposition of extremely large matrices it was not feasible to apply it to all 94 trajectories. A subset of 6 trajectories was therefore selected after examining the results from the overall sensitivity analysis. By looking at the reactions retained by each trajectory it was found that 95% of the reactions retained were selected by these 6 trajectories.

© EGU 2004

It was therefore considered that sufficient coverage of conditions relevant for reduction was achieved by carrying out the principal component analysis on this subset.

Commonly applied methods of principal component analysis used in for example the programme package KINAL (Turányi, 1990a), calculate the sensitivity matrix at each time-point and sum the entries over all time-points in the trajectory. The eigenvectors and eigenvalues are then evaluated for the summation matrices. In order to ensure that this did not compromise the accuracy of the results, eigenvalues and eigenvectors were calculated for the sensitivity matrix at each chosen time-point for one trajectory. It was found that carrying out the analysis in this way for one trajectory produced approximately the same results as using the matrix summed over all chosen time-points. Therefore it was accepted that the less time consuming method was valid. Eigenvalue and eigenvector thresholds of 10^{-4} and 0.13 respectively were selected after experimentation with higher thresholds was shown to lead to very large errors in the reduced system. The application of principal component analysis did not lead to significant further reductions in the mechanism and only a further 8 reactions were removed. We will return to this point in the discussion. This reduced model will be referred to as Stage 4 in the following discussion.

5. The application of time-scale based methods

The atmospheric diffusion equation contains the physical processes, chemical transformation and deposition that together describe the time dependent behaviour of all the species in a system. Each process occurs over a particular time-scale, and within any particular model structure the physical processes are usually associated with a narrower range of time-scales than the chemical time-scales. Within a large mechanism these chemical time-scales can differ by many orders of magnitude (McRae et al., 1982). It is this very large range of time-scales that can lead to a chemical system being stiff and therefore computationally expensive to solve.

By examining the atmospheric chemical system using a dynamical systems ap-

ACPD
4, 3721–3783, 2004

© EGU 2004

proach, it can be shown that the longterm behaviour of the mechanism can be accurately represented by assuming that the faster time-scales equilibrate with respect to the slower ones and therefore the dimension of the system can be reduced. Various methods have exploited this time-scale separation for mechanism reduction as described in Tomlin et al. (1998). The methods include the quasi-steady-state approximation (QSSA), the slow manifold approach and the computational singular perturbation (CSP) method, each of which will be discussed briefly below. The methods are successful, since by reducing the number of fast variables within the system, both the dimension of the system (and therefore the number of equations which must be solved at each timestep) and its stiffness are reduced, thus reducing the computational burden of the model. The analysis of time-scales in atmospheric systems is therefore a valuable tool for mechanism reduction.

Figure 3 shows the range of time-scales encountered in the Leeds MCM, and some of the species possessing these time-scales.

Intermediate radical species, such as O or RO, are highly reactive and very short lived. They quickly reach their equilibrium concentrations with respect to the slower species over time-scales of the order of $10^{-6} - 10^{-9}$ s, e.g. O^1D has a chemical lifetime of 3.0×10^{-9} s. NO, CO and some organic species however are associated with much slower time-scales, around $10^0 - 10^7$ s. The very fast time-scales are responsible for the equilibrium processes and are utilised in the QSSA and partial-equilibrium approximations used in conventional mechanism reduction. Since we are mainly interested in modelling the long-term behaviour of an atmospheric system – over a period of hours or days – it is possible to decouple the fast time-scales and assume that they reach local equilibrium with respect to the slower species. This leads to the formation of slow manifold on which the dynamics of the reduced system is contained.

5.1. The evaluation of system timescales

For low-dimensional systems, the investigation of time-scales can be carried out through a non-dimensionalisation process often revealing small parameters within the

ACPD
4, 3721–3783, 2004

© EGU 2004

model equations from which time-scale separation can easily be identified. For larger systems such as those typically found in tropospheric chemical mechanisms, non-dimensionalisation may be impractical, and hence perturbation methods are generally used to investigate system dynamics. The evolution of a small disturbance in a non-linear system can be used to reduce the problem to one which is locally linear. The resulting set of linear equations is easier to solve and can provide information about local time-scales and the stability of the non-linear system. Here it is applied to a more general case of systems with time-scale separation. Useful information can therefore be derived from the analysis since the concentrations of the slow species will not change significantly over the time-scale of relaxation of the perturbed species.

The rate of change of concentration in a homogeneous reaction system was described in Eq. (7). Where operator splitting is used in a reactive flow problem this formulation describes the chemical reaction step and $f(\boldsymbol{c}, k)$ can be represented as a function of the reaction rates:

$$f_i(\boldsymbol{c}, \boldsymbol{k}) = \sum_j v_{ij} R_j. \tag{13}$$

Here i is the species number, v_{ij} is the stoichiometric coefficient of the ith species in the jth reaction and R_j is the jth reaction rate.

If the system of Eqs. (7) is disturbed at a particular point c^{f_x} with a small perturbation Δc_i so that

$$c_i = c_i^{fix} + \Delta c_i \tag{14}$$

and this is expanded as a Taylor series, then to first order, a set of linear equations can be defined describing the motion of the perturbation,

$$\frac{\partial \Delta c_i}{\partial t} = \left(\frac{\partial f_i}{\partial c_1}\right) \Delta c_1 + \left(\frac{\partial f_i}{\partial c_2}\right) \Delta c_2 + \ldots + \left(\frac{\partial f_i}{\partial c_n}\right) \Delta c_n. \tag{15}$$

The general solution to this equation can be written as

$$\Delta \boldsymbol{c}(t) = e^{\int_0^t \mathbf{J}(s) ds}, \tag{16}$$

ACPD
4, 3721–3783, 2004

© EGU 2004

where **J** is the Jacobian matrix which is given by

$$\mathbf{J} = \frac{\partial \boldsymbol{f}}{\partial \boldsymbol{c}} = \begin{pmatrix} \frac{\partial}{\partial c_1}\left(\frac{\partial c_1}{\partial t}\right) & \cdots & \frac{\partial}{\partial c_n}\left(\frac{\partial c_1}{\partial t}\right) \\ \vdots & & \vdots \\ \frac{\partial}{\partial c_1}\left(\frac{\partial c_n}{\partial t}\right) & \cdots & \frac{\partial}{\partial c_n}\left(\frac{\partial c_n}{\partial t}\right) \end{pmatrix}. \tag{17}$$

If the Jacobian does not change significantly over the time-scale of relaxation of the small perturbation then Eq. (16) reduces to,

$$\Delta \boldsymbol{c}(t) = e^{\mathbf{J}(t)t}\mathbf{A} \tag{18}$$

where **A** is a constant matrix. This is usually the case if there is sufficient time-scale separation between the fast relaxing mode and the slow mode (Tomlin et al., 2001). Equation (18) can be written in the form,

$$\Delta c_i(t) = C_1 e^{\lambda_1 t} + C_2 e^{\lambda_2 t} + \ldots + C_n e^{\lambda_n t}, \tag{19}$$

where C_i depends on the size and direction of the perturbation. The exponents λ_i are the eigenvalues of the Jacobian matrix calculated at the point c^{f_x}. Therefore, each time-scale of the local linear solution to the full equations has an eigenvalue associated with it. What Eq. (19) shows, is that the decay of a perturbation of an individual species may depend on several eigenvalues, and therefore time-scales, due to its coupling with other species. Perturbations which decay very rapidly are associated with large negative eigenvalues and therefore a very short time-scale. This approach is essentially local and the Jacobian will be time-dependent if there are any non-linear trajectories. Therefore the ordering of the time-scales may change as the reaction proceeds.

5.2. The quasi-steady-state approximation (QSSA)

5.2.1. Definition of the QSSA

The Quasi-Steady-State Approximation (QSSA) is perhaps the simplest method by which the system time-scales can be exploited for mechanism reduction. It equates

ACPD
4, 3721–3783, 2004

© EGU 2004

species with time-scales, and involves the identification of species which react on very short time-scales and can therefore be assumed to be in equilibrium with respect to the slower species. For fast species such as radicals, the rate of production and consumption by chemical reaction are comparable. It can therefore be assumed that the original equation system of Eq. (7), is reduced to a differential-algebraic system:

$$\frac{dc_i}{dt} = f_i^N \tag{20}$$

$$0 = f_i^Q \tag{21}$$

where N and Q represent the non-QSSA and QSSA species respectively (Tomlin et al., 1998). The number of differential equations that must be solved in the system is reduced, and since the fast time-scales have been eliminated, the stiffness of the resulting system is also reduced. The application of QSSA does not imply that the concentration of the QSSA species is constant over time but rather that f_i^Q is zero when c_i is equal to the QSSA concentration c_i^Q.

5.2.2. Selection of QSSA species via the calculation of instantaneous errors

Two available methods for identifying possible QSSA species based on systematic criteria are the use of singular perturbation theory developed by Lam and Goussis (1988) and the calculation of QSSA errors in Turányi et al. (1993). In this work we have utilised the latter method.

The difference in the QSSA species concentrations at an interval Δt after the point of application of the QSSA, as calculated from Eqs. (20) and (21), and Eq. (7), is the instantaneous error of the QSSA. This error represents a continuous perturbation of the calculated trajectory and causes an overall error in the concentrations of the non-QSSA species. This error feeds back to the QSSA species and so their overall error can be different from their instantaneous error. However, if the overall error of non-QSSA species is small, the overall error of QSSA species should be close to their instantaneous error (Turányi et al., 1993).

ACPD
4, 3721–3783, 2004

© EGU 2004

An expression for the instantaneous error Δc_i^s on application of the QSSA to a single species has been derived in Turányi et al. (1993) as:

$$\Delta c_i^s = \left| \frac{1}{J_{ii}} \frac{dc_i}{dt} \right| . \tag{22}$$

If the QSSA is to be applied to several species then the error is defined as:

$$\Delta c_i^g = \frac{1}{J_{ii}} \frac{dc_i}{dt} - \frac{1}{J_{ii}} \sum_{k \neq i} J_{ik} \Delta c_k^g \tag{23}$$

where Δc_i^g denotes the group error and the term that distinguishes Eqs. (23) from (22) represents the interactions between the instantaneous errors of the QSSA species. If the individual and group errors remain small then the species within the chosen group can be considered as valid QSSA species (Tomlin et al., 1998).

Conducting this analysis gives an initial estimation of errors associated with the QSSA group. However, during a simulation these errors may propagate, and the careful selection of QSSA species is therefore very important. A full estimation of the validity of QSSA over a simulation can most easily be determined by comparing with the results from the simulation conducted in the absence of the QSSA approximation using the error analysis to guide the choice of QSSA group members.

The atmospheric lifetime can also be used as a criterion for the selection of the QSSA species (Hesstvedt et al., 1978). The total removal rate of a species A in the atmosphere is expressed as,

$$R = [A] \sum_i k_i x_1 \tag{24}$$

where $[A]$ is the concentration of species A, k_i are the rate coefficients of the second order consuming reactions of A, and x_i are the concentrations of the other species

ACPD
4, 3721–3783, 2004

© EGU 2004

involved in the reaction i. This equation also implicitly includes any pseudo-first-order process (eg. photolysis). The atmospheric lifetime is defined as:

$$\tau_i = \frac{1}{\sum_i k_i x_i}. \tag{25}$$

This is then related to the diagonal Jacobian element,

$$\tau_i = -\frac{1}{J_{ii}}, \tag{26}$$

see McRae et al. (1982). From Eq. (22) it can be seen that the instantaneous error of a single QSSA species is equal to the product of its lifetime and its net production rate. Therefore a small QSSA error requires a short lifetime and/or slow rates of change for a species. A slow rate of change is therefore not essential for the application of QSSA as the error will be negligible if the lifetime of the species is small compared to the net production of the species.

5.3. Application of the QSSA to the MCM

Following the application of sensitivity analysis to the MCM, the reduced mechanism consists of 2454 species and 6919 reactions. Individual QSSA errors were calculated using Eq. (22). Any species having an absolute QSSA error of less than 3×10^{-2} at each time point in every trajectory was considered to be a QSSA species. The tolerance level for inclusion of species into the QSSA species group was selected by examination of the range of values and then locating a gap in the sequence of values. In total, 557 QSSA species were identified.

Those species with only one consuming reaction were removed and reactions combined in the following manner. If

$$\mathrm{D} + \mathrm{E} \longrightarrow \mathrm{A} + \mathrm{PROD}_2 \quad \text{rate coefficient} = k_1 \tag{27}$$

$$\mathrm{A} \longrightarrow \mathrm{C} + \mathrm{PROD}_1 \quad \text{rate coefficient} = k_2 \tag{28}$$

ACPD
4, 3721–3783, 2004

© EGU 2004

where A is the QSSA species, then A can be replaced in Eq. (27) by:

$$\mathrm{D + E \longrightarrow C + PROD_1 + PROD_2 \ \ rate\ coefficient} = k_1 . \tag{29}$$

343 species could be removed from the MCM in this way. Almost all of these species were alkoxy radicals with only one consuming reaction at the previous stage that was either decomposition or reaction with oxygen.

79 of the QSSA species had more than one parallel consuming reaction although the rates for these were not temperature and therefore time dependent. These reactions were removed as follows. If A is defined as a QSSA species and

$$\mathrm{A \longrightarrow B + PROD_1 \ \ rate\ coefficient} = k_1 \tag{30}$$

$$\mathrm{A \longrightarrow E + PROD_2 \ \ rate\ coefficient} = k_2 \tag{31}$$

$$\mathrm{C + D \longrightarrow A + PROD_3\ rate\ coefficient} = k_3 \tag{32}$$

then A is replaced in Reaction (32) by

$$\frac{k_1}{k_1 + k_2}(\mathrm{B + PROD_1}) + \frac{k_2}{k_1 + k_2}(\mathrm{E + PROD_2}) . \tag{33}$$

The majority of these species were excited Criegee biradicals, of which $[\dot{O}O\dot{C}(CH_3)CHO]^*$ is an example. It is formed from unsaturated dicarbonyls and the oxidation of alkylbenzenes and isoprene. The biradical is formed in the reactions,

$$\mathrm{CHOC(CH_3){=}CHC(O)C_2H_5 + O_3 \longrightarrow CHOC(O)C_2H_5 + [\dot{O}O\dot{C}(CH_3)CHO]^*} \tag{34}$$

$$\mathrm{CHOC(CH_3){=}CHC(O)CH_3 + O_3 \longrightarrow CHOC(O)CH_3 + [\dot{O}O\dot{C}(CH_3)CHO]^*} \tag{35}$$

$$\mathrm{CHOCH{=}C(OH_3)CHO + O_3 \longrightarrow CHO_2 + [\dot{O}O\dot{C}(CH_3)CHO]^*} \tag{36}$$

$$\mathrm{CHOC(CH_3){=}CHC_2H_4OH + O_3 \longrightarrow HOCH_2CHO + [\dot{O}O\dot{C}(CH_3)CHO]^*}, \tag{37}$$

ACPD
4, 3721–3783, 2004

© EGU 2004

and reacts via,

$$[\dot{O}O\dot{C}(CH_3)CHO]^* \longrightarrow OH + CO + CH_3O_3 \quad (38)$$

$$[\dot{O}O\dot{C}(CH_3)CHO]^* \longrightarrow \dot{O}O\dot{C}(CH_3)CHO, \quad (39)$$

where the rate coefficients of the equations are given in Table 5 and KDEC is defined as 10^6 s^{-1} (Pilling et al., 1999).

From Eqs. (38) and (39) $[\dot{O}O\dot{C}(CH_3)CHO]^*$ can be expressed as

$$[\dot{O}O\dot{C}(CH_3)CHO]^* = 0.89\,(OH + CO + CH_3CO_3) \quad (40)$$

$$+0.11\left(\dot{O}O\dot{C}(CH_3)CHO\right), \quad (41)$$

i.e. there is competition between decomposition to form OH and formation of the excited criegee biradical which primarily reacts with CO to form methylglyoxyl. The OH channel represents a significant radical formation route.

Substituting this into Eqs. (34) to (37) and removing Eqs. (38) and (39) the set of equations is reduced to

$$\begin{aligned} &CHOC(CH_3){=}CHC(O)C_2H_5 + O_3 \longrightarrow \\ &\quad CHOC(O)C_2H_5 + 0.89(OH + CO + CH_3CO_3) \\ &\quad +0.11(\dot{O}O\dot{C}(CH_3)CHO) \end{aligned} \quad (42)$$

$$\begin{aligned} &CHOC(CH_3){=}CHC(O)C_2H_5 + O_3 \longrightarrow \\ &\quad CHOC(O)C_2H_5 + 0.89(OH + CO + CH_3CO_3) \\ &\quad +0.11(\dot{O}O\dot{C}(CH_3)CHO) \end{aligned} \quad (43)$$

$$\begin{aligned} &CHOC(CH_3){=}CHC(O)CH_3 + O_3 \longrightarrow \\ &\quad CHOC(O)CH_3 + 0.89(OH + CO + CH_3CO_3) \\ &\quad +0.11(\dot{O}O\dot{C}(CH_3)CHO) \end{aligned} \quad (44)$$

$$\begin{aligned} &CHOCH{=}C(OH_3)CHO + O_3 \longrightarrow \\ &\quad CHO_2 + 0.89(OH + CO + CH_3CO_3) \end{aligned}$$

ACPD
4, 3721–3783, 2004

© EGU 2004

$$+0.11(\dot{O}O\dot{C}(CH_3)CHO) \quad (45)$$

$$CHOC(CH_3)=CHC_2H_4OH + O_3 \longrightarrow$$
$$HOCH_2CHO + 0.89(OH \;+\; CO + CH_3CO_3)$$
$$+0.11(\dot{O}O\dot{C}(CH_3)CHO), \quad (46)$$

where the rate coefficients remain as before.

A further 55 species can be removed in a similar manner to this, but as their rate coefficients are time-dependent the coefficient ratios within the lumped equations must be recalculated at each iteration, for example O^1D.

$$O_3 \longrightarrow O^1D \;\; k_1 = J(1) \quad (47)$$

$$O^1D \longrightarrow 2OH \; k_2 = 2.20 \times 10^{-10} \; H_2O \quad (48)$$

$$O^1D \longrightarrow O \; k_3 = 3.20 \times 10^{-11} \; O_2 e^{\frac{70}{T}}. \quad (49)$$

The expression for O^1D can be expressed as

$$\frac{k_2}{k_2 + k_3}(2[OH]) + \frac{k_3}{k_2 + k_3}([O]). \quad (50)$$

It can be seen that both k_2 and k_3 are time-dependent and therefore the ratios in the expression for O^1D must be recalculated after each iteration. Equation (47) will then become,

$$O_3 \longrightarrow \frac{k_2}{k_2 + k_3}(2[OH]) + \frac{k_3}{k_2 + k_3}([O]). \quad (51)$$

The 55 QSSA species in this group were mainly alkoxy radicals undergoing both decomposition and reaction with oxygen in the previous reduced scheme.

© EGU 2004

Sixty-six QSSA species remain which cannot be removed by any of the above methods. These could potentially be removed by creating explicit expressions for the removed species by setting their rate equations equal to zero. This was only implemented for one species, O, and resulted in an algebraic expression for O of the following form:

$$[O] = \frac{k_3[NO_3] + k_4[NO_2] + k_5[O_2] + k_6\frac{k_2}{k_1+k_2}[O_2]}{k_7[NO_2] + k_8[NO_2] + k_9[NO] + k_{10}[O_3] + k_{11}}, \tag{52}$$

where the rate expressions are given in Table 6. KMT01 and KMT02 both represent fall-off reactions of the following form: $k = \left(\frac{k_0 k_\infty}{k_0+k_\infty}\right) F$, where $\log F = \frac{\log F_c}{1+(\log(k_0/k_\infty))^2}$, (Atkinson et al., 1992).

This expression can be substituted for [O] in all equations, and therefore leads to the removal of one species. Unfortunately this method leads to highly complicated, and often nonlinear, algebraic expressions when used for other remaining QSSA species that may require solution using iterative methods. For this reason the method as only been applied to a single species at this stage, in order to avoid excessively complicating the mechanism before further lumping stages. An additional drawback for these highly coupled QSSA species is that their removal can only be achieved using complex algebraic expressions within the model code. The explicit description of the chemical model in kinetic form is lost at this stage of reduction as a result. The final reduced mechanism after the application of the QSSA will be addressed as Stage 5 in the following discussion.

6. The existence of slow manifolds within the model

6.1. Intrinsic Low Dimensional Manifold (ILDM) methods

The concept of the low-dimensional or slow manifold was first introduced to atmospheric systems by Lorenz (1986). It is particularly useful in homogeneous chemical

ACPD
4, 3721–3783, 2004

© EGU 2004

systems or systems employing operator splitting, since the slow manifold describes the long-term behaviour of the chemical system once the fast modes have died away. The QSSA represents a subset of such slow manifold based techniques since it assumes that fast species locally equilibrate. However, as demonstrated above, the QSSA assumes that time-scales are associated with individual or groups of QSSA species. In highly coupled systems this may not always be the case as complex functions of species may in fact create fast and slow time-scales. The concept of the ILDM therefore utilises a geometric approach to time-scale separation, which may in some cases lead to higher levels of reduction in the number of system variables than the QSSA.

Geometrically, the dynamics of the system in species phase space can be seen as reaction trajectories relaxing onto a series of increasingly lower-dimensional manifolds, which was interpreted in Roussel and Fraser (1991) as a "cascade through a nested hierarchy of smooth interfaces (inertial manifolds)". Scenarios with quite different initial conditions approach these attracting manifolds, and the slow time-scales govern the long term dynamics. The aim of mechanism reduction is to find the lowest dimension manifold – the intrinsic manifold – which will still represent the important long-term features of the reaction (Lowe and Tomlin, 2000), and to develop a model that provides a time-dependant representation of the slow dynamics. Several methods have been developed which allow the description of the system dynamics relating to the remaining intermediate modes. These include the use of look-up tables describing the slow dynamics, the QSSA, where algebraic functions for the non-intermediate modes are used, and repromodelling methods, where fitted models such as polynomial difference equations are formulated for the intermediate dynamics (Lowe and Tomlin, 2000).

6.2. Time-scale modes and species couplings

If the full chemical mechanism contains n_s species, then n_s coupled ODEs define the system, which therefore has n_s different time-scales associated with it. If the system is assumed to be in equilibrium with respect to its n_f fastest time-scales, these can be decoupled from the system. In addition, a number of species or combinations of species

© EGU 2004

may remain approximately constant over time. These form n_c conservation relations, of which element conservation in the mechanism is an example. The behaviour of the system can then be described independently from the conservation relations and the decoupled fast time-scales and hence by a smaller number of variables n_g, where $n_g = n_s - n_f - n_c$. Estimating the intrinsic dimension of the system is useful since it provides guidelines for the application of ILDM type methods. It also allows us to assess whether the QSSA has successfully removed all possible fast time-scales from the model.

In order to determine the intrinsic system dimension n_g, the system is first transformed into a new set of variables called time-scale modes which are each related to a single time-scale of the system. If $\lambda_1, \ldots, \lambda_n$ are distinct eigenvalues of $\mathbf{J}$ and $\mathbf{X}_r$ and $\mathbf{X}_l$ the matrices of its corresponding right and left eigenvectors then we define a new set of variables z where

$$z = \mathbf{X}_l^T \boldsymbol{c} \tag{53}$$

and

$$\frac{dz}{dt} = \mathbf{X}_l^T \boldsymbol{f}. \tag{54}$$

Conversely it follows that the inverse transformation matrix $\mathbf{X}_r$ exists and transforms the new variables back to the original species concentrations:

$$\boldsymbol{c} = \mathbf{X}_r z. \tag{55}$$

Using perturbation analysis (Tomlin et al., 2001) showed that the time-dependance of a perturbation in z is given by:

$$\boldsymbol{\Delta} z(t) = e^{diag(\lambda_1,\ldots,\lambda_n)t} \boldsymbol{\Delta} z_o. \tag{56}$$

The system will therefore respond to a perturbation in z according to single time-scales, each perturbation growing or decaying exponentially depending on the size and sign of the eigenvalues. The variables z can therefore be considered as the modes of the

© EGU 2004

system. Physically what this means is that the lumping matrix $\mathbf{X}_l^T$ shows us how each species contributes to the modes associated with each eigenvalue or system time-scale . By ordering the eigenvalues we can see which species are associated with the fast and slow modes of the system. This result is similar to CSP analysis and can allow us to identify species contributing to the fast decaying modes which can potentially be removed.

Assuming we can order the eigenvalues according to their values from large and negative to large and positive we can separate the system into slow, intermediate and fast modes. This can either be achieved by identifying large gaps in the spectrum of eigenvalues or by comparison of chemical time-scales with typical transport processes. We therefore partition the lumping matrix $\mathbf{X}_l^T$ into slow, intermediate and fast parts:

$$\mathbf{X}_l^T = \begin{bmatrix} \mathbf{X}_{lf}^T \\ \mathbf{X}_{li}^T \\ \mathbf{X}_{ls}^T \end{bmatrix}, \tag{57}$$

allowing the modes to be separated into fast:

$$z_f = \mathbf{X}_{lf}^T \boldsymbol{c} \tag{58}$$

intermediate

$$z_i = \mathbf{X}_{li}^T \boldsymbol{c} \tag{59}$$

and slow modes:

$$z_s = \mathbf{X}_{ls}^T \boldsymbol{c}. \tag{60}$$

In a similar way to the application of the QSSA, a model describing the long term dynamics of the system need only describe the intermediate modes since the fast ones can be assumed to be in local equilibrium and the slow ones assumed to represent some kind of conservation equation.

ACPD
4, 3721–3783, 2004

© EGU 2004

6.3. Estimating the underlying dimensionality of the trajectories

By ordering the eigenvalues as above we can consider individual modes i and investigate their collapse onto what is essentially an (n-1) dimensional manifold. It has been shown previously in Tomlin et al. (2001) that the equation describing the decay of each fast mode is given by:

$$\frac{dz_i}{dt} = \mathbf{X}^T_{l,\lambda_i}\frac{d\boldsymbol{c}}{dt} = \mathbf{X}^T_{l,\lambda_i}\boldsymbol{f}, \tag{61}$$

where $\mathbf{X}^T_{l,\lambda_i}$ is the transpose of the left eigenvector associated with eigenvalue λ_i. If the mode has collapsed then $\mathbf{X}^T_{l,\lambda_i}\boldsymbol{f}$=0 giving the equivalent of the Maas and Pope (1992) equation defining an (n-1) dimensional manifold following from the collapse of an individual mode z_i. Where the trajectory is close to the manifold (i.e. in the region of the fixed point about which the original expansion was made) we can estimate the normalised distance of an individual mode i from an (n-1) dimensional manifold using the equation:

$$\Delta\tilde{z} = \frac{\mathbf{X}^T_{l,\lambda_i}\boldsymbol{f}}{\lambda_i}\frac{1}{(\kappa + |\,\mathbf{X}^T_{l,\lambda_i}\boldsymbol{c}\,|)}, \tag{62}$$

where κ is a small parameter included to avoid division by zero.

Following the ideas of Roussel and Fraser (1991) we can imagine the system collapsing onto a cascade of manifolds of decreasing dimension with the fastest modes collapsing first and the slowest last. By estimating the distance of each mode from each (n-1) dimensional manifold and comparing against a tolerance parameter we can therefore determine at each time-point along the trajectory how many modes have effectively collapsed. It also follows that once the fastest mode has collapsed then the error of assuming an (n-2) dimensional manifold can be estimated by the distance of the next slowest mode from the manifold. If we assume n_f fast modes therefore, the error associated with the (n-n_f) dimensional slow manifold will be largely determined

ACPD
4, 3721–3783, 2004

© EGU 2004

by the distance of the z_{n_f} mode from the manifold although there may be some contribution from the faster modes.

The number of linear conservation relations which hold for the system can be determined by finding those modes whose eigenvalues are equal or close to zero. A very small eigenvalue signifies a slow rate of change and therefore signifies variables or combinations of variables which remain approximately constant over time. This does not imply that the concentration of specific species must remain constant over time but that certain species maintain a fixed relationship with other species throughout the calculation. This means that groups of species maybe conserved. Again, a user tolerance $ctol$ is defined and a conservation relation is defined for $|Re(\lambda(i))| \leq ctol$. The definition of manifold dimension is therefore given by

$$\text{dimension} = n_s - n_c - n_f. \tag{63}$$

The dimension therefore depends on selected user tolerances, the choice of which will affect the final accuracy of the model. In general there are usually gaps in the spectrum of eigenvalues and therefore $ctol$ might be chosen by first looking at how the spectrum changes over time and choosing an appropriate gap.

6.4. Linking slow manifold distance to QSSA error

Equation (62) gives an estimation for $\Delta\tilde{\mathbf{z}}$. If the system is completely uncoupled the eigenvalues will be equal to the Jacobian diagonals and the set of eigenvectors will be equivalent to the identity matrix. So substituting,

$$\lambda_i = J_{ii} = -\frac{1}{\tau_i} \text{ and } \quad \mathbf{X}^T_{l,\lambda_i} = I \tag{64}$$

where τ_i is the chemical lifetime of species i, gives,

$$|\,\Delta z_i\,| = |\,f_i \tau_i\,| = |\,\Delta c_i^{qssa}\,| \tag{65}$$

ACPD
4, 3721–3783, 2004

© EGU 2004

This effectively demonstrates the equivalence of QSSA errors and the distance from the slow manifold for this special situation where no coupling exists between the fast variables.

6.5. Evaluation of time-scales remaining following the application of the QSSA

Examination of the range of time-scales remaining in a mechanism is a good way of identifying the sorts of methods which would lead to effective further reduction of the mechanism. For example, very little will be gained by applying ILDM methods to a system with very few fast time-scales remaining. The number of slow, fast and intermediate time-scales remaining at each stage of the reduction is presented for trajectory 7 after 36 h in Table 5. The table shows that following the application of the QSSA at Stage 5 of the reduction, most of the fast time-scales have been removed from the mechanism. This demonstrates that in this case the QSSA species as a group relate to the fast modes of the system. The large number of intermediate time-scales remaining implies that little further reduction would be achieved in this case by the use of an ILDM type method.

7. Comparison between full and reduced schemes

7.1. Comparison for original box model conditions

The number of species and reactions remaining at each stage of the reduction are presented in Table 6. The dimension of the system has been reduced by a factor of 2 at Stage 5 of the reduction with an associated speed up in simulation time of a factor of 4.

Comparisons between the full and reduced mechanisms for selected trajectories and important species can be seen in Figs. 4 to 9. The errors appear to be fairly small over a wide range of concentration conditions. Although a small accumulation in errors

© EGU 2004

occurs on the third day of simulation, these are still within an acceptable level.

7.2. Comparison of reduced mechanisms using a European trajectory model

Although generated under polluted urban conditions it is desirable that the reduced mechanisms are accurate over as wide a range of conditions as possible. In order to test this wider applicability, the production of eight species is presented over the course of a European scale trajectory using the 1D Harwell Trajectory model (Derwent et al., 1996). This trajectory follows an episode that developed over Europe, ending in Birmingham in order to coincide with the PUMA campaign measurements, and provides an opportunity to investigate how well the reduced mechanisms stand up outside of the conditions for which they were specifically developed.

The Harwell Trajectory model (Derwent and Nodop, 1986) is used to describe the chemical development of a parcel of boundary layer air. The emission terms are time and species dependent and depend on the location of the air parcel within the emissions grid. The emission terms are obtained by dividing the pollutant emission rate in mass per unit area per unit time by the boundary layer depth. The emission rates are taken from the following inventories,

- The EMEP emissions inventory is disaggregated over 150 × 150 km grid squares and represents the total biogenic emissions expressed in 100 tonnes per year as isoprene.
- The OECD-CORINAIR inventory expresses data in tonnes per year per 50×50 km grid-square with the VOCs split into non-methane hydrocarbons (NMVOCs) and methane. All data are summed over all categories and area and point sources.
- The UK NAEI uses 10×10 km disaggregated grid squares to represent emissions of CO, NO_x, SO_2 and NMVOCs over the UK.

The emissions are then assigned to individual species within the MCM according to the NAEI as described above. The results of this study made it possible to examine the

ACPD
4, 3721–3783, 2004

© EGU 2004

chemical development of air parcels travelling over Europe and examine the accuracy of the reduced mechanisms when compared with the trajectory simulated using the full mechanism.

This investigation was carried out using the full mechanism and reduced mechanisms at Stages 2, 3 and 5. It was necessary to restore the reaction

$$\begin{aligned} &CH_4 + OH \longrightarrow CH_3O_2 \\ &\text{rate coefficient} \;=\; 7.44 \times 10^{-18} T^2 \exp(-1361/T) \end{aligned} \tag{66}$$

to the reduced scheme, since it is important to include the oxidation of CH_4 for non-polluted conditions such as those found at the start of the trajectory. Although it reacts slowly with OH compared to other VOCs, when its background atmospheric concentration is high, it makes a major contribution to OH loss when the concentration of other VOCs is low.

The comparisons between full and reduced mechanisms are plotted in Figs. 10 and 11. In Fig. 10 the photochemical generation of ozone along the idealised trajectory shows a small deviation compared to the full mechanism. The greatest disparities occur at around 90 and 120 h, where the concentration of ozone is under predicted by 5% to 13%. The comparison of NO also exhibits a small deviation from the full mechanism, although less than that for ozone. On examination of the remaining figures a similar situation can be seen for each species. The largest jump in error occurs between Stage 2 and Stage 3, with an extremely small increase in error between Stage 3 and Stage 5 where the QSSA was applied. The worst results were found in the simulations of total PAN. The simulated concentrations rise steadily over the five-day trajectory exhibiting a clear diurnal afternoon peak. This rise coincides with increasing emissions of VOC into the atmosphere, and is perhaps not surprising since some of the minor PAN species were excluded from the list of important species at this stage.

The accuracy levels found here are slightly worse than the trends found using the scenarios over which the reduced mechanisms were developed, although still within acceptable limits. When conducting a mechanism reduction of this nature the only con-

© EGU 2004

ditions over which the mechanism can be assumed to be accurate are those for which the analysis on which the reductions were carried out was based. If the examination of trajectories over conditions outside those previously considered produced results as good as, or better than those obtained from the defined data set then this is a very positive piece of evidence. The more conditions over which these mechanisms can be demonstrated to be accurate the more basis there is for assuming that they can be used to model O_3, NO_x, PAN and VOCs within the troposphere under other un-tested conditions within the range considered to be typical for polluted environments.

8. Discussion

The first stages of the approach taken towards mechanism reduction in this work were derived from the protocol described in Heard et al. (1998) and Zeng et al. (1997). First, those species in which we are primarily interested were determined from the full MCMv2.0. which contains 10763 reactions and 3487 species (Pilling et al., 1999). Simulations were performed using the full mechanism for 94 box model simulations where initial concentrations and emissions were chosen to cover a wide range of urban conditions. Next, redundant species were identified and removed from the mechanism. At Stage 1, 396 species and 1224 reactions were removed from the mechanism leaving 9839 reactions.

Then, redundant reactions were removed by calculating the local rate sensitivity matrix and using overall sensitivity measures and principal component analysis as a way of interpreting the resulting matrix. Adaptations had to be made in order to cover the inherent problems of carrying out mechanism reduction on a system as large as the MCM. System sparsity was exploited where possible and methods were automated to a high degree. Using overall sensitivities the mechanism was reduced to 8410 reactions. All peroxy-peroxy radical reactions were removed during this stage, indicating that this loss mechanism for peroxy radicals is of much lower importance under urban and moderately polluted conditions than the reaction with inorganic nitrogen containing

© EGU 2004

species.

Before principal component analysis was carried out further species were removed using a second stage of identification of redundant species using a smaller set of important species. At this stage some minor PAN species were removed from the important species list leaving a set of 58 important species. This step was added as the large number of species still remaining in the system meant that the eigenvector eigenvalue analysis of the Jacobian, necessary to carry out principal component analysis, was too computationally expensive to be carried out for all the trajectories. Only 2454 species were considered to be necessary to model this reduced set of important species, involving 6927 reactions. The greatest deterioration in accuracy has occurred after this second removal of redundant species. This suggests that the PAN species removed as insignificant may have been influential when treated as a whole group. A potentially better method of reducing the number of PAN species would have been to define a species TOTALPAN where

$$\mathrm{TOTALPAN} = \sum_i \mathrm{PAN}_i \tag{67}$$

where PAN_i represents the 184 PAN species. Then a single rate equation

$$\frac{df(\mathrm{TOTALPAN})}{dk_i} \tag{68}$$

could have been examined. In this way those PAN species important to the production of TOTALPAN could have been identified and retained by the user, instead of defining them a priori. However, the errors still remained small for Stage 3 over quite a wide range of conditions.

In the fourth stage of reduction, principal component analysis on the sensitivity matrix was carried out. 2454 eigenvalues were retained as there were still 2454 necessary species present, signifying the maximum number of potentially relevant eigenvalues. An eigenvector tolerance of 0.13 was used. The large reduction at the previous stage did however mean that very little further reduction could be achieved using principal

© EGU 2004

component analysis without seriously compromising the accuracy of the mechanism. Despite extensive experimentation with potential thresholds, it was found that those which allowed accuracy to be maintained permitted the removal of only a further 8 reactions. This was potentially due to the fact that the MCM consists of many parallel reaction chains (Jenkin et al., 1997) which have very little cross coupling between them. This led to many reactions being removed using overall concentration sensitivity analysis leaving very few reactions remaining that were suitable for removal using principal component analysis. This is an important result for the future application of sensitivity analysis for the reduction of highly detailed tropospheric mechanisms, since avoiding the use of the computationally expensive principal component analysis could speed up the reduction process considerably.

The final stages of mechanism reduction involved identifying QSSA species and removing them by lumping reactions, and defining algebraic expressions for the QSSA species as functions of the non-QSSA species. The selection of QSSA species was found to lend itself very well to automation using local QSSA errors. Once a threshold had been selected a list of QSSA species was easily assembled from the data produced by each simulation. The species which were identified as being suitable for removal as QSSA species consist mainly of two species types. Most, were alkoxyradicals which decompose very quickly to produce an aldehyde and water. The remaining species were predominantly excited Criegee bi-radicals. One of the other species which has been identified as a QSSA species is O^1D, an excited oxygen atom, which having been formed by the photolysis of ozone, reacts very quickly with N_2, H_2O or O_2 to form OH or a ground state oxygen atom O^3P. Again, due to the very short life-time of this species, it was an obvious QSSA species as the intermediate steps which pass through O^1D can easily be lumped without impinging on the accuracy of the mechanism. Since this species is associated with extremely fast time-scales, its removal reduces the stiffness of the rate equation system considerably.

The application of the QSSA to the MCM is designed to remove fast time-scales from the mechanism, and the suitability of this technique for application to the MCM was in-

© EGU 2004

vestigated. It was shown that the number of fast time-scales has decreased from 478 before the removal of QSSA species to 6 after their removal. A small number of intermediate and slow time-scales were also removed. The decrease from 20% to 0.2% of the time-scales in the system being considered fast, shows a high level of success for removal species via the QSSA and a resulting decrease in dimension of the system. The application of the QSSA by simple reaction lumping for a very large number of species allows the basic structure of the chemical mechanism to be maintained giving significant advantages over other more geometric based techniques such as the ILDM for future kinetic studies. In this case, the large number of intermediate time-scales within the mechanism make the application of ILDM type methods impossible.

The removal of species from the mechanism leads to a decrease in the amount of computational time required to solve the system. This occurs by approximately a factor of n^2. In the Stage 5 reduced mechanism there are 1969 species, a decrease of a factor of approximately 2. In practice, the solver shows a speed up of approximately 4 making this assumption reasonably accurate. This factor of n^2 speed up could be considered to be due to the extensive use of $n \times n$ matrices in implicit integration routines.

The reduced mechanisms were shown to vary slightly in accuracy across the range of trajectories over which they were reduced. Most trajectories are well represented and the errors are less than 5% at all times. The reduced mechanism was also shown to be applicable over a long range trajectory calculation with the addition of only the reaction methane + OH that is relevant for unpolluted conditions. Some divergence was noted at the end of the 5 day trajectory but in general the performance of the reduced schemes was within acceptable limits, demonstrating potentially wide applicability.

8.1. Further reduction of the mechanism

At Stage 5 the mechanism contains 1969 species and 6168 reactions. Whilst the mechanism has been shown to be applicable over a wide range of conditions and can be computed 4 times faster than the full mechanism, it is still desirable to reduce it further. Conducting time-scale analysis on the system shows that there are large numbers of

© EGU 2004

intermediate time-scales still present in the reduced mechanism. The application of ILDM methods has been discounted and therefore other methods should be sought. An approach commonly used is species lumping, where a group of species can be represented in the mechanism by a single variable. In this way a smaller set of equations are needed to represent the dynamics of the chemical system and the number of intermediate time-scales would be reduced. The further application of lumping to the MCMv2.0. is described in the following companion paper.

9. Conclusions

Several methods for mechanism have been introduced and shown to be applicable using automated methods to complex tropospheric mechanisms such as the Master Chemical Mechanism. In particular, the methods of local concentration sensitivity analysis and overall rate sensitivity analysis proved to be efficient and capable of removing the majority of redundant reactions in the scheme across a wide range of conditions. The application of principal component analysis of the rate sensitivity matrix proved to be computationally very expensive due to its use of the decomposition of very large matrices. For a mechanism such as the MCM, where a large number of parallel reaction chains exist, its use did not prove to be effective in finding additional redundant reactions over and above the use of overall sensitivities.

The use of the QSSA also proved to be an extremely successful method of removing the fast time-scales within the system. The calculation of instantaneous QSSA errors was capable of automatically choosing potential groups of QSSA species based on user selected tolerances. Due to the fact that many of the QSSA species could be easily removed by reaction lumping, its application allowed a kinetic structure to be retained within the mechanism in contrast to other time-scale based methods such as the ILDM. Although they have proved successful within highly lumped schemes such as in Lowe and Tomlin (2000), due to the high number of intermediate time-scale remaining in MCM mechanism, the use of methods such as the ILDM would not be

© EGU 2004

possible without significant further reduction using other methods.

Analysis of the errors obtained by using the most reduced mechanism were determined to be less than 5% for many of the trajectories studied, and less than 10% for the majority. Those trajectories which showed disappointing levels of accuracy were in the very high NO_x and high VOC ranges of the conditions studied and would be considered to be at the extremes of conditions found in even polluted UK cities. Furthermore, it was found that with the addition of a reaction relevant to non-polluted environments, the reduced mechanism was capable of accurately predicting a trajectory describing the formation of a range of secondary species across Europe. The success of the mechanism under conditions well outside of those under which it was generated suggests that the reduced schemes are of quite wide applicability.

Finally, the use of the reduction techniques discussed here has allowed the development of a set of staged mechanisms with varying dimensions and accuracies to allow a user to choose the most appropriate level according to the desired application. An investigation was also carried out into the automation of reduction techniques such as the identification of redundant species, overall sensitivity analysis, principal component analysis, QSSA analysis and reaction lumping. This means that reduced mechanisms can be generated in a systematic way without the need for extensive manual analysis of data.

Acknowledgements. The authors would like to thank the NERC for financial support through the URGENT programme and D. Derwent, M. Jenkins and T. Turányi for useful discussions.

References

AEAT-ENV: http://www.aeat-env.com. 3727

Atkinson, R.: A structure-activity relationship for the estimation of rate constants for the gas-phase reactions of OH radicals with organic compounds, Int. J. Chem. Kin., 19, 799–828, 1987. 3725

© EGU 2004

Atkinson, R.: Kinetics and mechanisms of the gas-phase reactions of the hydroxyl radical with organic compounds, J. Phys. Chem. Ref. Data, Monograph 1, 1989. 3725
Atkinson, R.: Gas phase tropospheric chemistry of organic compounds – a review, Atm. Env., 24A, 1–41, 1990. 3725
Atkinson, R.: Gas-phase tropospheric chemistry of organic-compounds, J. Phys. Chem. Ref. Data, Monograph 2, 1994. 3725
Atkinson, R., Baulch, D. L., Cox, R. A., Hampson, J. A., Kerr, J. A., and Troe, J.: Evaluated kinetic and photochemical data for atmospheric chemistry: supplement IV, Atm. Env., 26, 1187–1230, 1992. 3725, 3747, 3770
Derwent, R. and Jenkin, M.: Hydrocarbons and long-range transport of ozone and PAN across Europe, Atm. Env., 25A, 1661–1678, 1990. 3726
Derwent, R. and Nodop, K.: Long range transport and deposition of acidic nitrogen species in North-West Europe, Nature, 324, 356–358, 1986. 3754
Derwent, R., Jenkin, M., and Saunders, S.: Photochemical ozone creation potentials for a large number of reactive hydrocarbons under European conditions, Atm. Env., 30, 181–199, 1996. 3726, 3754
Heard, A., Pilling, M., and Tomlin, A.: Mechanism reduction techniques applied to tropospheric chemistry, Atm. Env., 32, 1059–1073, 1998. 3729, 3731, 3733, 3734, 3756
Hesstvedt, H., Hov, O., and Isaksen, I.: Quasi-steady state assumption in air pollution modelling, Int. J. Chem. Kinet., 10, 971–994, 1978. 3742
Hough, A.: The calculation of photolysis rates for use in global tropospheric modelling studies., AERE Report R-13259 (HMSO), London, 1988. 3726
Jenkin, M., Saunders, S., and Pilling, M.: The tropospheric degradation of volatile organic compounds :A protocol for mechanism development, Atm. Env., 31, 81–104, 1997. 3724, 3726, 3758
Jenkin, M., Hayman, G., Derwent, R., Saunders, S., Carslaw, N., Pascoe, N., and Pilling, M.: Tropospheric chemistry modelling: Improvements to current models and application to policy issues, Report AEAT-4867/20150/R004, 1999. 3726
Jenkin, M., Saunders, S., Wagner, V., and Pilling, M.: Protocol for the development of the Master Chemical Mechanism, MCM v3 (Part B): tropospheric degradation of aromatic volatile organic compounds, Atm. Chem. Phys., 3, 181–193, 2003. 3727
Kwok, E. and Atkinson, R.: Estimation Of Hydroxyl Radical Reaction-Rate Constants For Gas-Phase Organic-Compounds Using A Structure-Reactivity Relationship – An Update, Atm.

© EGU 2004

Env., 29, 1685–1695, 1995. 3725
Lam, S. and Goussis, G.: Understanding complex chemical kinetics with computational singular perturbation, 22nd Symp. (Int.) Comb., 931–941, 1988. 3723, 3741
Lightfoot, P., Cox, R., Crowley, J., Estriau, M., Hayman, G., Jenkin, M., Moortgat, G., and Zabel, F.: Organic peroxy-radicals – kinetics, spectroscopy and tropospheric chemistry, Atm. Env., 26A, 1805–1964, 1992. 3725
Lorenz, E.: On the existence of a slow manifold, J. Atm. Sci., 43, 1547–1557, 1986. 3747
Lowe, R. and Tomlin, A.: Low-dimensional manifolds and reduced chemical models for tropospheric chemistry simulations, Atm. Env., 34, 2425–2436, 2000. 3748, 3760
Maas, U. and Pope, S.: Simplifying chemical kinetics:intrinsic low dimensional manifolds in composition space, Combust. Flame, 88, 239–264, 1992. 3723, 3751
McRae, G., Goodin, W. R., and Seinfeld, J.: Development of a second generation mathematical model for urban air pollution -I. Model formation, Atm. Env., 16, 679–696, 1982. 3737, 3743
NAEI: 1998 Emissions Inventory aeat/2/env/078 issue 1, 2000. 3728
Peeters, J., Boullart, W., and Van Hoeymissen, J.: Site-specific partial rate constants for OH addition to alkenes and dienes, Proc. EUROTRAC symp. 1994, Garmisch-Partenkirchen, Germany, April 1994, 110–114, 1994. 3725
Pilling, M., Saunders, S., Carslaw, N., Pascoe, S., Jenkin, M., and Derwent, D.: http://www.chem.leeds.ac.uk/Atmospheric/MCM/mcmproj.html, 1999. 3726, 3745, 3756, 3770
PUMA: http://badc.nerc.ac.uk/data/urgent/Projects-Doc/pumaco.html, 2001. 3727
Roberts, X.: The atmospheric chemistry of organic nitrates, Atm. Env., 24A, 243–287, 1990. 3725
Roussel, M. and Fraser, S.: On the geometry of transient relaxation, J. Chem. Phys., 94, 7106–7113, 1991. 3748, 3751
Rudd, H.: Emissions of volatile organic compounds from stationary sources in the United Kingdom: Speciation, Report AEA/CS/16419033/REMA-029, 1995. 3724
Saunders, S., Jenkin, M., Derwent, R., and Pilling, M.: Protocol for the development of the Master Chemical Mechanism, MCM v3 (Part A): tropospheric degradation of non-aromatic volatile organic compounds, Atm. Chem. Phys., 3, 161–180, 2003. 3727
Tomlin, A., Turányi, T., Merkin, J., and Brindley, J.: Mechanism reduction for the oscillatory oxidation of hydrogen: sensitivity and quasi-steady-state analyses, Combust. Flame, 91, 107–130, 1992. 3731
Tomlin, A., Turányi, T., and Pilling, M.: Autoignition and low temperature combustion of hy-

© EGU 2004

drocarbons, chap. Mathematical tools for construction, investigation and reduction of combustion mechanisms, Elesvier, Amsterdam, 1998. 3723, 3731, 3734, 3735, 3738, 3741, 3742

Tomlin, A., Whitehouse, L., Lowe, R., and Pilling, M.: Low dimensional manifolds in tropospheric chemical systems, Farad. Disc., 120, 125–146, 2001. 3740, 3749, 3751

Turányi, T.: KINAL: A program package for kinetic analysis of complex reaction mechanisms, Comp. Chem., 14, 253–254, 1990a. 3737

Turányi, T.: Reduction of large reaction mechanisms, New J. Chem., 14, 795–803, 1990b. 3731

Turányi, T., Bérces, T., and Vajda, S.: Reaction rate analysis of complex kinetic systems., Int. J. Chem. Kin., 21, 83–99, 1989. 3733

Turányi, T., Tomlin, A., and Pilling, M.: On the error of the quasi-steady-state approximation, J. Phys. Chem., 97, 163–172, 1993. 3723, 3731, 3741, 3742

Vajda, S., Valkó, P., and Turányi, T.: Principal component analysis of kinetic-models, Int. J. Chem. Kin., 17, 55–81, 1985. 3734

Zeng, G., Pilling, M., and Saunders, S.: Mechanism reduction for tropospheric chemistry – Butane oxidation, J. Chem. Soc. Farad. Trans., 93, 2937–2946, 1997. 3731, 3756

© EGU 2004

Table 1. Initial O_3 concentrations and desired maximum VOC concentrations and VOC to NO_x ratios for each trajectory scenario and the emissions rates used to achieve them.

initial O_3(ppb)			Desired conditions			
50	80	20	max VOC	$\frac{VOC}{NO_x}$	VOC emission rates	NO_x emission rates
trajectory number			(ppb)		(molecules $cm^{-3}s^{-1}$)	(molecules $cm^{-3}s^{-1}$)
	33	66	700	1.2	1.80×10^{8}	7.00×10^{7}
1	34	67	500	1.2	1.50×10^{8}	5.50×10^{7}
2	35	68	300	1.2	9.50×10^{7}	4.80×10^{7}
3	36	69	250	1.2	9.00×10^{7}	4.50×10^{7}
4	37	70	200	1	4.30×10^{7}	1.00×10^{7}
5	38	71	200	1.2	8.50×10^{7}	4.00×10^{7}
6	39	72	150	0.78	4.50×10^{7}	1.00×10^{7}
7	40	73	150	1	3.80×10^{7}	6.00×10^{5}
8	41		150	1.45	8.30×10^{7}	3.80×10^{7}
9	42	74	100	0.45	4.00×10^{7}	7.80×10^{6}
10	43	75	100	0.8	4.00×10^{7}	6.50×10^{6}
11	44		100	1.1	2.55×10^{7}	1.35×10^{7}
12	45		100	1.8	2.65×10^{7}	1.55×10^{7}
13	46	76	75	0.34	9.00×10^{6}	1.12×10^{6}
14	47		75	1	2.25×10^{7}	1.15×10^{7}
15	48	77	75	1.5	5.00×10^{7}	1.00×10^{7}
16	49	78	75	2.1	5.05×10^{7}	8.80×10^{6}
17	50	79	50	0.25	8.50×10^{6}	1.12×10^{6}
18	51	80	50	1	2.00×10^{7}	8.98×10^{6}
19	52	81	50	1.75	5.00×10^{7}	8.50×10^{6}
20	53	82	50	2.6	3.50×10^{7}	7.00×10^{6}
21	54	83	20	0.25	8.50×10^{6}	1.12×10^{6}
22	55	84	20	1	9.00×10^{6}	2.24×10^{6}
23	56	85	20	2	2.05×10^{7}	9.50×10^{6}
24	57	86	20	4.2	2.60×10^{7}	1.50×10^{7}
25	58	87	10	0.35	4.60×10^{6}	6.55×10^{5}
26	59	88	10	1	7.00×10^{6}	2.24×10^{6}
27	60	89	10	3	1.05×10^{7}	6.73×10^{6}
28	61	90	10	6	1.10×10^{7}	1.12×10^{7}
29	62	91	5	0.4	4.10×10^{6}	1.12×10^{6}
30	63	92	5	1	2.45×10^{6}	9.35×10^{5}
31	64	93	5	3	2.45×10^{6}	2.81×10^{6}
32	65	94	5	8	4.61×10^{6}	5.61×10^{6}

ACPD
4, 3721–3783, 2004

© EGU 2004

Table 2. Relative emissions factors for individual VOCs as calculated from Eq. (5) in Sect. 3.2.

MCM name	species name	emissions ratio	MCM name	species name	emissions ratio
HCHO	METHANAL (FORMALDEHYDE)	7.20×10^{-3}	C2H6	ETHANE	6.52×10^{-3}
CH3OH	METHANOL	1.40×10^{-3}	C2H5OH	ETHANOL	3.69×10^{-3}
CH3CHO	ETHANAL (ACETALDEHYDE)	1.21×10^{-3}	NPROPOL	1-PROPANOL (N-PROPANOL)	5.47×10^{-2}
C3H8	PROPANE	1.91×10^{-2}	IPROPOL	2-PROPANOL (I-PROPANOL)	9.47×10^{-3}
C2H5CHO	PROPANAL (PROPRIONALDEHYDE)	1.38×10^{-3}	NBUTOL	1-BUTANOL (N-BUTANOL)	1.02×10^{-2}
CH3COCH3	PROPANONE (ACETONE)	1.22×10^{-2}	BUT2OL	2-BUTANOL (SEC-BUTANOL)	8.50×10^{-3}
NC4H10	N-BUTANE	6.21×10^{-2}	IBUTOL	2-METHYL-1-PROPANOL (I-BUTANOL)	2.48×10^{-3}
C3H7CHO	BUTANAL (BUTYRALDEHYDE)	7.24×10^{-4}	NC5H12	PENTANE (N-PENTANE)	3.72×10^{-2}
MEK	BUTANONE (METHYL ETHYL KETONE)	1.25×10^{-2}	C3H6	PROPENE (PROPYLENE)	1.44×10^{-2}
IC4H10	2-METHYL-PROPANE (I-BUTANE)	2.19×10^{-2}	IPEBOH	3-METHYL-2-BUTANOL	6.46×10^{-5}
IPRCHO	METHYLPROPANAL (I-BUTYRALDEHYDE)	5.95×10^{-5}	BUT1ENE	1-BUTENE	4.32×10^{-3}
TBUTOL	2-METHYL-2-PROPANOL (T-BUTANOL)	3.12×10^{-4}	MGLYOX	METYHL GLYOXAL	9.52×10^{-6}
C4H9CHO	PENTANAL (VALERALDEHYDE)	6.25×10^{-5}	MIBK	4-METHYL 2-PENTANONE (METHYL I-BUTYL KETONE)	8.56×10^{-3}
IC5H12	2-METHYLBUTANE (I-PENTANE)	3.25×10^{-2}	M22C4	2,2-DIMETHYLBUTANE	1.86×10^{-3}
MIPK	3-METHYL 2-BUTANONE (METHYL I-PROPYL KETONE)	3.76×10^{-7}	NC7H16	HEPTANE (N-HEPTANE)	5.31×10^{-3}
NEOP	2,2-DIMETHYLPROPANE (NEO-PENTANE)	1.32×10^{-4}	M3HEX	3-METHYLHEXANE	5.00×10^{-3}
NC6H14	HEXANE (N-HEXANE)	1.66×10^{-2}	NC10H22	DECANE (N-DECANE)	2.89×10^{-2}
HEX2ONE	2-HEXANONE (METHYL N-BUTYL KETONE)	4.22×10^{-4}	CHEX	CYCLOHEXANE	2.75×10^{-2}
M2PE	2-METHYLPENTANE	1.17×10^{-2}	ETHGLY	ETHANE-1,2-DIOL (ETHYLENE GLYCOL)	4.57×10^{-3}
GLYOX	GLYOXAL	1.22×10^{-5}	PROPACID	PROPANOIC ACID	6.59×10^{-7}
M3PE	3-METHYLPENTANE	7.91×10^{-3}	MEPROPENE	2-METHYLPROPENE (1-BUTENE, BUTYLENE)	2.41×10^{-3}
M23C4	2,3-DIMETHYLBUTANE	2.86×10^{-3}	TPENT2ENE	TRANS-2-PENTENE	2.90×10^{-3}
M2HEX	2-METHYLHEXANE	5.26×10^{-3}	ME3BUT1ENE	3-METHYL-1-BUTENE	1.65×10^{-3}
NC8H18	OCTANE (N-OCTANE)	2.15×10^{-2}	HEX1ENE	1-HEXENE	6.55×10^{-4}
NC9H20	NONANE (N-NONANE)	1.40×10^{-2}	C4H6	1-3 BUTADIENE	5.92×10^{-2}
NC11H24	HENDECANE (N-UNDECANE)	2.33×10^{-2}	C2H2	ETHYNE (ACETYLENE)	1.65×10^{-2}
NC12H26	DODECANE (N-DODECANE)	1.11×10^{-2}	CYHEXOL	CYCLOHEXANOL	2.78×10^{-4}
C2H4	ETHENE (ETHYLENE)	3.14×10^{-2}	DIETETHER	DIETHYL ETHER	1.30×10^{-4}
PROPGLY	PROPANE-1,2-DIOL (PROPYLENE GLYCOL)	9.92×10^{-5}	EOX2EOL	2-ETHOXY ETHANOL	2.95×10^{-3}
CH3CO2H	ETHANOIC ACID	4.02×10^{-4}	IPROACET	I-PROPYL ACETATE	4.68×10^{-3}
CBUT2ENE	CIS-2-BUTENE	4.37×10^{-3}	BUOX2ETOH	2-BUTOXY ETHANOL	5.31×10^{-3}
TBUT2ENE	TRANS-2-BUTENE	4.38×10^{-3}	NPROACET	N-PROPYL ACETATE	1.58×10^{-3}
PENT1ENE	1-PENTENE	3.38×10^{-3}	DMC	DIMETHYL CARBONATE	6.48×10^{-6}
CPENT2ENE	CIS-2-PENTENE	2.86×10^{-3}	CH2CL2	DICHLOROMETHANE (METHYLENE DICHLORIDE)	1.06×10^{-2}
ME2BUT1ENE	2-METHYL-1-BUTENE	1.36×10^{-3}	TCE	TETRA-CHLOROETHENE (PER-CHLOROETHYLENE)	7.53×10^{-3}
ME3BUOL	3-METHYL-1-BUTANOL (I-PENTANOL, I-AMYL ALCOHOL)	6.75×10^{-9}	CDICLETH	CIS-1,2-DICHLOROETHENE	4.49×10^{-5}
ME2BUT2ENE	2-METHYL-2-BUTENE	2.48×10^{-3}	TOLUENE	METHYLBENZENE (TOLUENE)	8.37×10^{-2}
CHEX2ENE	CIS-2-HEXENE	3.24×10^{-5}	OXYL	1,2-DIMETHYL BENZENE (O-XYLENE)	2.12×10^{-2}
THEX2ENE	TRANS-2-HEXENE	3.24×10^{-5}	EBENZ	ETHYL BENZENE	2.06×10^{-2}
C5H8	2-METHYL-1,3-BUTADIENE (ISOPRENE)	4.26×10^{-6}	IPBENZ	I-PROPYL BENZENE (CUMENE)	1.98×10^{-3}
APINENE	ALPHA-PINENE	3.67×10^{-4}	TM135B	1,3,5-TRIMETHYL BENZENE (MESITELENE)	6.30×10^{-3}
MIBKAOH	4-HYDROXY-4-METHYL-2-PENTANONE (DIACETONE ALCOHOL)	2.15×10^{-3}	PETHTOL	1-ETHYL 4-METHYL BENZENE (P-ETHYL TOLUENE)	6.93×10^{-3}
CH3OCH3	DIMETHYL ETHER	9.05×10^{-4}	STYRENE	ETHENYL BENZENE (STYRENE)	1.42×10^{-3}
CH3OCHO	METHYL FORMATE	3.93×10^{-7}	DIEK	3-PENTANONE (DIETHYL KETONE)	2.86×10^{-4}
ETHACET	ETHYL ACETATE	1.43×10^{-2}	BENZENE	BENZENE	2.17×10^{-2}
METHACET	METHYL ACETATE	1.07×10^{-4}	BENZAL	BENZENECARBONAL (BENZALDEHYDE)	8.40×10^{-4}
DIIPRETHER	DI I-PROPYL ETHER	2.35×10^{-5}	MXYL	1,3-DIMETHYL BENZENE (M-XYLENE)	3.78×10^{-2}
MO2EOL	2-METHOXY ETHANOL	1.88×10^{-9}	PXYL	1,4-DIMETHYL BENZENE (P-XYLENE)	2.54×10^{-2}
PR2OHMOX	1-METHOXY 2-PROPANOL	1.51×10^{-3}	PBENZ	N-PROPYL BENZENE	4.14×10^{-3}
BOX2PROL	1-BUTOXY 2-PROPANOL	3.98×10^{-4}	TM123B	1,2,3-TRIMETHYL BENZENE (HEMIMELLITENE)	5.50×10^{-3}
NBUTACET	N-BUTYL ACETATE	1.21×10^{-2}	TM124B	1,2,4-TRIMETHYL BENZENE (PSEUDOCUMENE)	1.82×10^{-2}
SBUTACET	S-BUTYL ACETATE	5.60×10^{-4}	OETHTOL	1-ETHYL 2-METHYL BENZENE (O-ETHYL TOLUENE)	4.99×10^{-3}
CH3CL	CHLOROMETHANE (METHYL CHLORIDE)	1.72×10^{-5}	METHTOL	1-ETHYL 3-METHYL BENZENE (M-ETHYL TOLUENE)	7.92×10^{-3}
CHCL3	TRICHLOROMETHANE (CHLOROFORM)	1.96×10^{-6}	DIME35EB	1,3-DIMETHYL 5-ETHYL BENZENE (3,5-DIMETHYL ETHYL BENZENE)	1.43×10^{-2}
CH3CCL3	1,1,1-TRICHLOROETHANE (METHYL CHLOROFORM)	1.55×10^{-2}	DIET35TOL	1,3-DIETHYL 5-METHYLBENZENE (3,5-DIETHYL TOLUENE)	2.51×10^{-3}
TRICLETH	TRI-CHLOROETHENE	2.10×10^{-2}	PHCOOH	BENZOIC ACID (PHENYL CARBOXYLIC ACID)	1.49×10^{-9}
TDICLETH	TRANS-1,2-DICHLOROETHENE	1.48×10^{-5}			

ACPD
4, 3721–3783, 2004

© EGU 2004

Table 3. Emissions fractions for given sources in the UK and West Midlands, RT – Road Transport, SU – Solvent Use, IP – Industry and Processes, FF – Fossil Fuel Distribution, DC – Domestic Combustion, PG – Power Generation, specifically power stations, OT – Other, NA – Natural.

Emission source	Category total kt C yr^{-1}	Category %	
	UK	UK	West Midlands
RT	527.3	26.9	45.9
SU	532.4	27.2	49.1
IP	298.3	15.2	0.3
FF	298.2	15.2	2.5
DC	40.1	2.1	0.5
PG	6.4	3.3×10^{-3}	0.0
OT	77.6	4.0	1.0
NA	178	9.1	0.7
total	1958.3	100	100

ACPD

4, 3721–3783, 2004

© EGU 2004

Table 4. Initial concentrations for species in MCM at 298 K.

Species	initial concentration (molecules cm^{-3})	(ppb)
NO	1×10^{9}	4.13×10^{-2}
NO_2	1×10^{9}	4.13×10^{-2}
HCHO	1×10^{11}	4.13×10^{1}

ACPD
4, 3721–3783, 2004

© EGU 2004

Table 5. Rate coefficients for example removing $[\dot{O}O\dot{C}(CH_3)CHO]^*$.

equation number	rate coefficient (molecule cm^{-3})s^{-1}
(34)–(36)	1.00×10^{-18}
(37)	1.20×10^{-17}
(38)	0.89 KDEC
(39)	0.11 KDEC

ACPD
4, 3721–3783, 2004

© EGU 2004

Table 6. Rate coefficients used in algebraic expression for O (Pilling et al., 1999; Atkinson et al., 1992), where M is a third body usually represented by [N_2].

	rate coefficient (molecules $cm^{-3})^{m-1} s^{-1}$
$k_3 =$	$1.85 \times 10^{-1} (\cos\theta)^{0.189} e^{-0.112 \sec\theta}$
$k_4 =$	$1.11 \times 10^{-2} (\cos\theta)^{0.397} e^{-0.183 \sec\theta}$
$k_5 =$	$5.22 \times 10^{-4} (\cos\theta)^{0.322} e^{-0.079 \sec\theta}$
$k_6 =$	$8.98 \times 10^{-5} (\cos\theta)^{1.436} e^{-0.936 \sec\theta}$
$k_7 =$	KMT02
	$k_0 = 9.0 \times 10^{-32} \left(\frac{T}{300}\right)^{-2.0} M$
	$k_\infty = 2.2 \times 10^{-11}$
	$F_c = e^{\frac{-T}{1300}}$
$k_8 =$	$6.50 \times 10^{-12} e^{\frac{120}{T}}$
$k_9 =$	KMT01
	$k_0 = 1.0 \times 10^{-31} \left(\frac{T}{300}\right)^{-1.6} M$
	$k_\infty = 3.0 \times 10^{-11} \left(\frac{T}{300}\right)^{0.3}$
	$F_c = e^{\frac{-T}{1850}}$
$k_{10} =$	$8.00 \times 10^{-12} e^{\frac{2060}{T}}$
$k_{11} =$	$6.00 \times 10^{-34} O_2 M \left(\frac{T}{300}\right)^{-2.3}$

ACPD
4, 3721–3783, 2004

© EGU 2004

Table 7. Number of lifetimes in time-scale ranges for each reduced mechanism.

range of life-times $\frac{1}{\lvert\lambda\rvert}$ (s)	number remaining			
	full	Stage 1	Stage 3	Stage 5
fast $< 1 \times 10^{-4}$	711	710	434	6
intermediate	2359	2175	1915	1863
slow $> 1 \times 10^{5} (\approx 1$ day)	416	205	105	100
total	3487	3091	2454	1969

ACPD

4, 3721–3783, 2004

© EGU 2004

Table 8. Reactions and species remaining at each stage.

stage of reduction	number of reactions remaining	% reactions remaining	number of species remaining	% species remaining
full	10763	100	3487	100
Stage 1	9839	91.4	3091	88.6
Stage 2	8410	78.1	3091	88.6
Stage 3	6927	64.4	2454	70.3
Stage 4	6919	64.2	2454	70.3
Stage 5	6168	57.3	1969	56.4

ACPD

4, 3721–3783, 2004

© EGU 2004

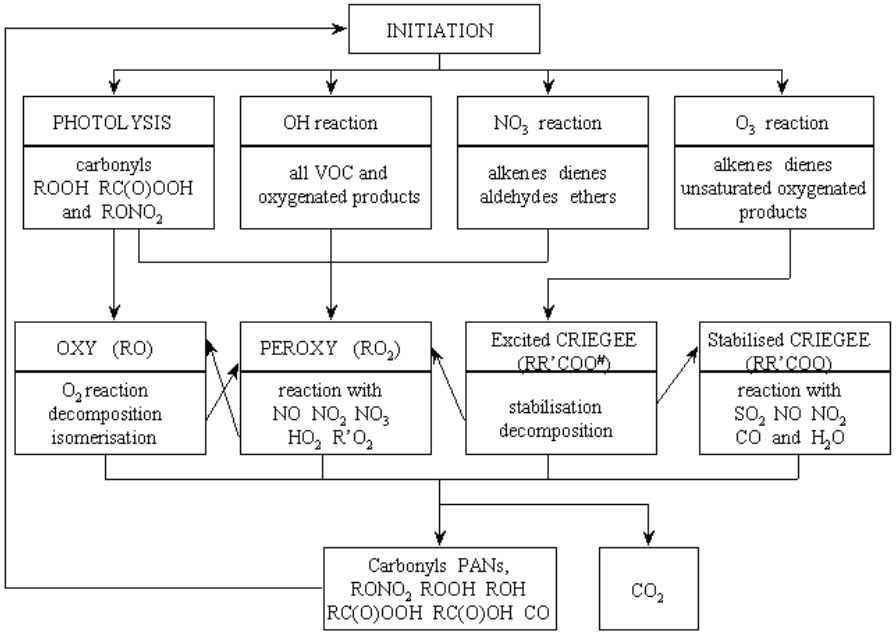

Fig. 1. The mechanism development of the tropospheric degradation of the VOCs.

ACPD
4, 3721–3783, 2004

© EGU 2004

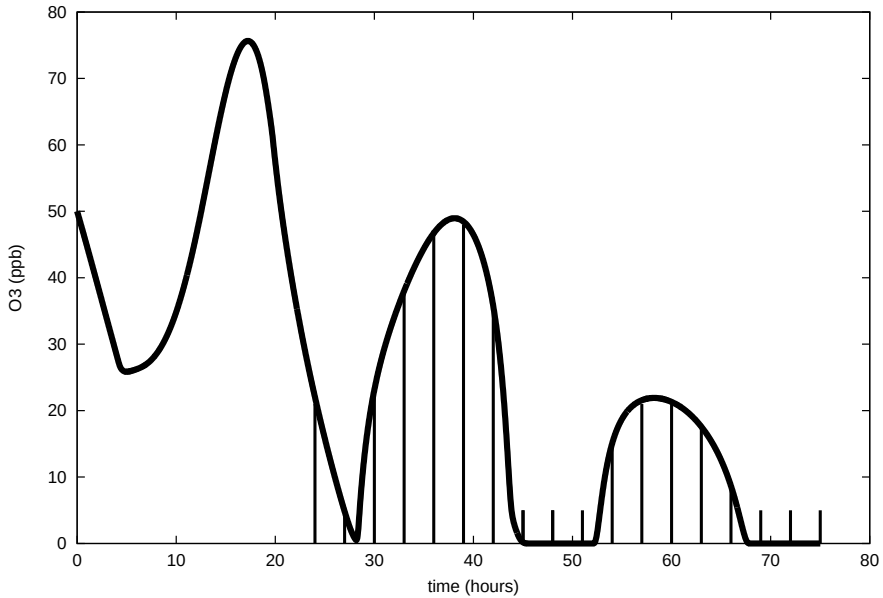

Fig. 2. Example trajectory indicating time-points used during mechanism reduction – the vertical lines indicate the locations of the selected time-points for all trajectories.

ACPD

4, 3721–3783, 2004

Systematic reduction of complex tropospheric mechanisms

L. E. Whitehouse et al.

© EGU 2004

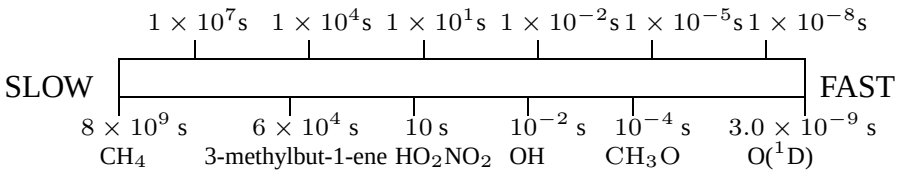

Fig. 3. Diagram showing the range of chemical lifetimes from fast to slow and examples of the lifetimes of species in the MCM.

ACPD
4, 3721–3783, 2004

© EGU 2004

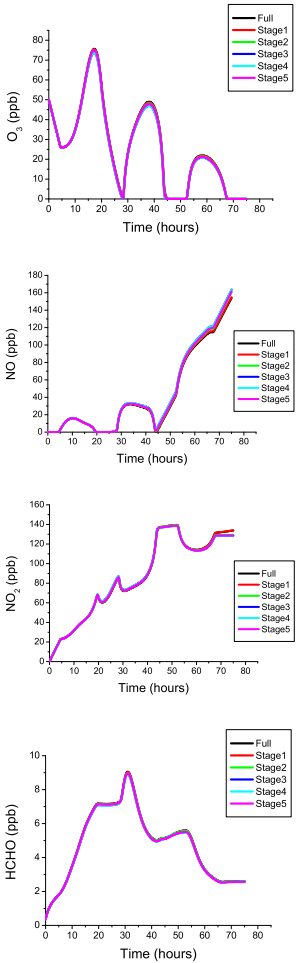

Fig. 4. Plot of Trajectory 7 to Stage 5 for O_3, NO, NO_2, HCHO.

ACPD
4, 3721–3783, 2004

Systematic reduction of complex tropospheric mechanisms

L. E. Whitehouse et al.

© EGU 2004

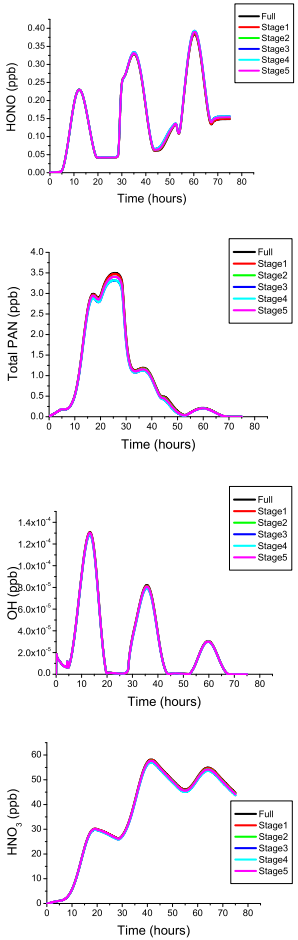

Fig. 5. Plot of Trajectory 7 to Stage 5 for HONO, total PAN, OH and HNO_3.

ACPD
4, 3721–3783, 2004

Systematic reduction of complex tropospheric mechanisms

L. E. Whitehouse et al.

© EGU 2004

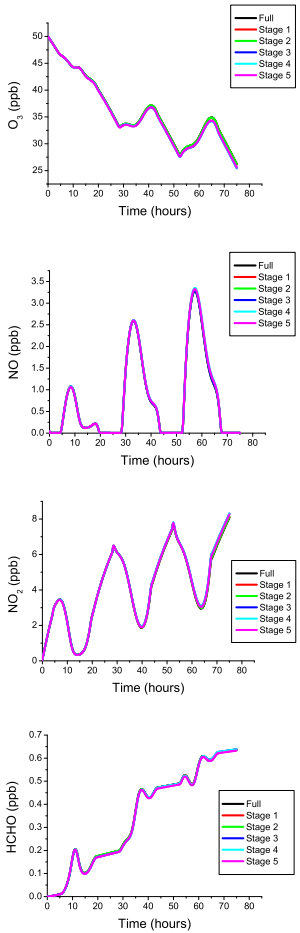

Fig. 6. Plot of Trajectory 25 to Stage 5 for O_3, NO, NO_2 HCHO.

ACPD
4, 3721–3783, 2004

Systematic reduction of complex tropospheric mechanisms

L. E. Whitehouse et al.

© EGU 2004

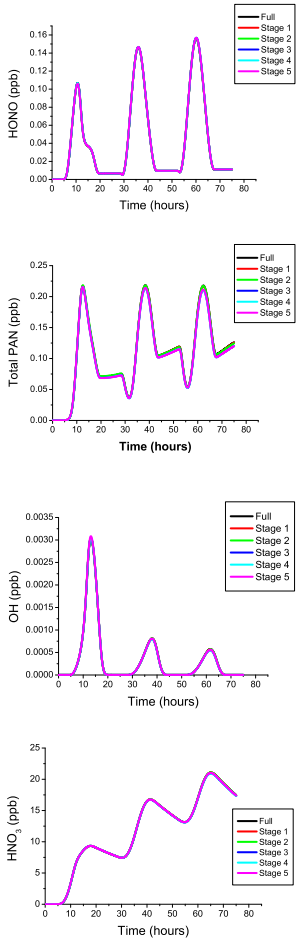

Fig. 7. Plot of Trajectory 25 to Stage 5 for HONO, total PAN, OH and HNO_3.

ACPD
4, 3721–3783, 2004

Systematic reduction of complex tropospheric mechanisms

L. E. Whitehouse et al.

© EGU 2004

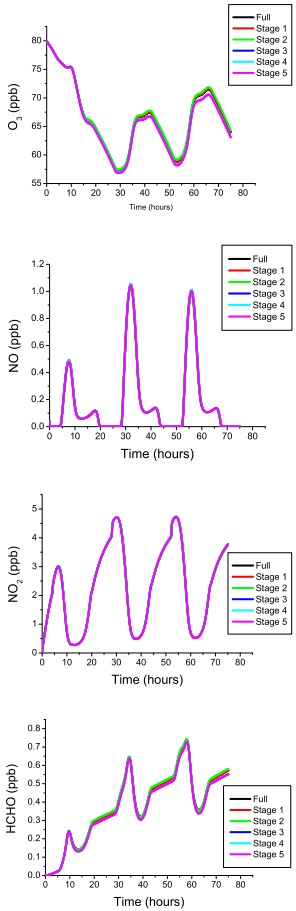

Fig. 8. Plot of Trajectory 63 to Stage 5 for O_3, NO, NO_2 HCHO.

ACPD
4, 3721–3783, 2004

Systematic reduction of complex tropospheric mechanisms

L. E. Whitehouse et al.

© EGU 2004

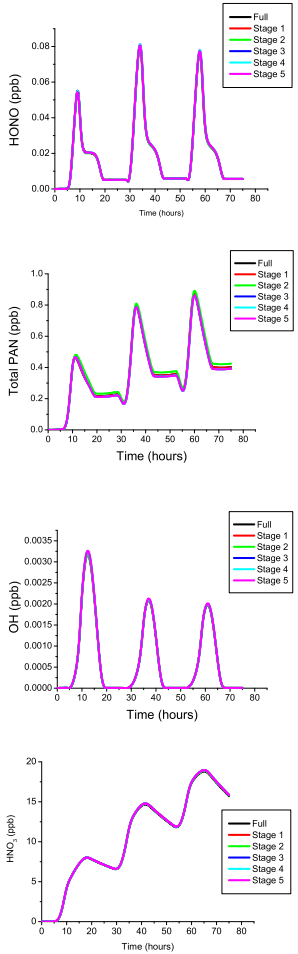

Fig. 9. Plot of Trajectory 63 to Stage 5 for HONO, total PAN, OH and HNO_3.

ACPD
4, 3721–3783, 2004

Systematic reduction of complex tropospheric mechanisms

L. E. Whitehouse et al.

© EGU 2004

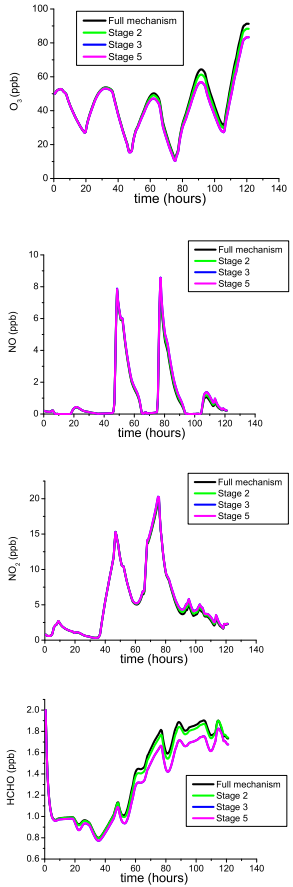

Fig. 10. Plots for PUMA trajectory run for O_3, NO, NO_2, HCHO.

ACPD
4, 3721–3783, 2004

Systematic reduction of complex tropospheric mechanisms

L. E. Whitehouse et al.

© EGU 2004

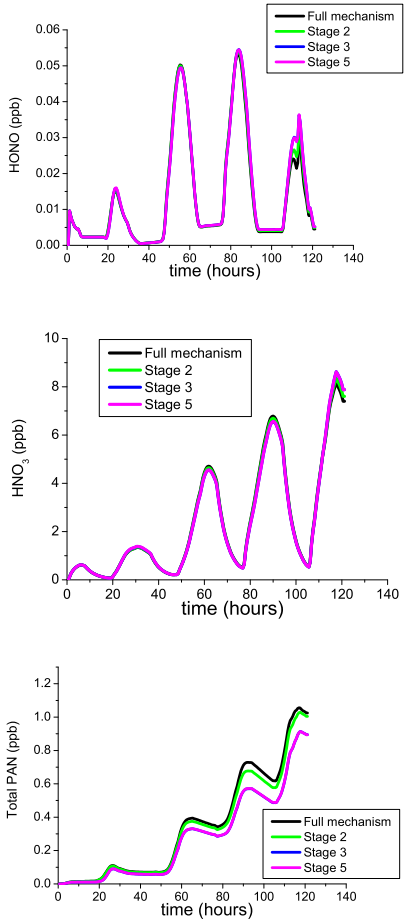

Fig. 11. Plots for PUMA trajectory run for HONO, HNO_3, total PAN.

ACPD
4, 3721–3783, 2004

Systematic reduction of complex tropospheric mechanisms

L. E. Whitehouse et al.

© EGU 2004